MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE MOHAMED KHIDER BISKRA

Faculté des sciences exactes et
des sciences de la nature et de la vie

Département des Sciences de la matière

THESE
Présentée par

OUASSAF Mebarka

En vue de l'obtention du diplôme de :
DOCTORAT EN SCIENCES

Option :
Chimie informatique et pharmaceutique

Intitulée:

Contribution à la découverte de médicaments par une étude computationnelle de plusieurs séries de molécules hétérocycliques.

Soutenue le :11-04-2019

Devant la commission d'Examen

M. Ammar DIBI	Professeur	Université de Batna 1	Président
M. Salah BELAIDI	Professeur	Univ. Mohamed Khider- Biskra	Directeur de thèse
M. Nadjib MELKEMI	MC/A	Univ. Mohamed Khider- Biskra	Examinateur
M. Lazhar BOUCHLALEG	MC/A	Université de Batna 2	Examinateur

ACKNOWLEDGEMENT

This dissertation is

dedicated to the memory of my beloved

parents,

They are the reason for my success.

To my husband and my children.

To my siblings.

To all those who are dear to me.

CONTENTS

CHAPETR I: THE TRIAZOLE AND THE AROMATASE INHIBITION FOR BREAST CANCER TREATMENT

CHAPTER II: COMPUTATIONAL APPROACHES FOR DRUG DESIGN AND DISCOVERY

CHAPETR III: DRUG LIKENESS SCORING AND STRUCTURE ACTIVITY/PROPERTY RELATIONSHIPS

CHAPETR IV: QSAR MODEL FOR PREDICTING THE AROMATASE INHIBITION ACTIVITY OF 1.2.3 TRIAZOLE DERIVATIVES

CHAPETR V: MOLECULAR DOCKING STUDIES AND IN SILICO ADMET OF NEW SUBSTITUTED 1.2.3 TRIAZOLE DERIVATIVES FOR ANTI-BREAST CANCER ACTIVITY

LIST OF ABBREVIATIONS

- Absorption, Distribution, Metabolism, and Excretion _____ ADME
- Acquired Immune Deficiency Syndrome _____ AIDS
- Aromatase inhibitors _____ AIs
- Assisted Model Building with Energy Refinement _____ AMBER
- Austin Model 1 _____ AM1
- Blood brain barrier _____ BBB
- Becke, three-parameter, Lee-Yang-Parr _____ B3LYP
- Concordance correlation cofficient _____ CCC
- Computer-Assisted Drug Design _____ CADD
- Cross-Validation _____ CV
- Cytochrome P450 _____CyP450
- Density Functional Theory _____ DFT
- Human Ether-à-go-go-Related Gene_____ hERG
- Hormone Receptor-Positive _____ ER+
- Half maximal Inhibitory Concentration _____ IC50
- Hard and Soft, Acids and Bases theory _____HSAB
- Hartree-Fock _____HF
- Highest Occupied Molecular Orbital _____ Homo
- Ligand Efficiency _____LE
- Ligand lipophilicity efficiency _____ LipE
- Linear Regression _____LR
- Lowest Unoccupied Molecular Orbital _____ LUMO
- Molar Refractivity _____ MR
- Molecular electrostatic potential _____ MEP
- Molecular electrostatic surface map _____MESP
- Molecular dynamics _____ MD
- Molecular Weight _____ MW
- Møller-Plesset level 2 _____ MP2
- Molecular Mechanic 2_____MM2

- Multiple Linear Regression _____ MLR
- Number of Hydrogen-Bond Donors and Acceptors _____ NHBD and NHBA nrotb
- Number of Rotatable Bonds _____ nrotb
- Partial Least Squares _____ PLS
- partition coefficient octanol/water _____ logP
- Parameterized Model number 3 _____ PM3
- Pharmacokinetics _____ PK
- P-glycoprotein) _____ Pgp
- Polar Surface Area _____ PSA
- Predictive Residual Sum of the Squares _____ PRESS
- Protein Data Bank _____ PDB
- Quantitative Structure–Activity Relationship _____ QSAR
- Quantitative Structure–Property Relationship _____ QSPR
- Root-Mean Squared Error _____ RMSE
- Structure–Activity Relationships _____ SARs
- Topological polar surface area _____ TPSA
- Two-Dimensional or Three-Dimensional _____ (2D or 3D) QSAR

LIST OF FIGURE

CHAPETR I

CHAPETR II

CHAPETR III

CHAPETR IV

CHAPETR V

LIST OF TABLES

CHAPETR IV

CHAPETR V

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Cancer is the leading cause of death worldwide, accounting for an estimated 7.6 million deaths in 2008 [1]. Breast cancer is the most common type of cancer in women for both developed and developing countries. The effectiveness of anti-hormonal treatment of early breast cancer can be attributed to the fact that approximately two-thirds of breast tumors are hormone-dependent and require growth factors like estrogen to grow [2]. There is some different type of systemic therapy for breast cancer, one kind is hormonal therapy. Hormonal therapy can be given to women whose breast cancers test positive for estrogen to lower estrogen levels. Letrozole is a third generation of non-steroidal aromatase inhibitor – one class of hormonal therapy drugs- that was first introduces by Novartis to the market as Femara ® for the treatment of local or metastatic breast cancer [3-5].

It is well known that compounds having an aza-hetero ring, such as triazole show inhibitory activity against aromatase [6].

Aromatase is a cytochrome P-450 enzyme, which is responsible for the conversion of androgen to estrogen in the final step of the steroid biosynthesis cascade. Inhibition of this enzyme is therefore of practical importance in the treatment of estrogen-dependent diseases, for example breast cancer, cancer of the uterine body, and endometriosis. Like aromatase, several other steroid organic enzymes are also cytochrome P-450 [7].

Drugs are essential for the prevention and treatment of disease. Thus, ideal drugs are in great demand. But the process of Drug design is a tedious, time-consuming and cost intensive process. Thus, several approaches are required which collectively would form the basis of Computer Aided or In Silico Drug Designing. Use of computational methods in drug discovery and development process is nowadays gaining popularity, implementation and appreciation. Different terms are being applied to this area, including computer-aided drug design (CADD), computational drug design, computer-aided molecular design (CAMD), computer-aided molecular modeling (CAMM), rational drug design, In Silico drug design, computer-aided rational drug design. All the world's major pharmaceutical and biotechnology companies use computational design tools. At their lowest level the contributions represent the replacement of crude mechanical models by displays of structure which are a much more accurate reflection of molecular reality capable of demonstrating

motion and solvent effects. Beyond this, theoretical calculations permit the computation of binding free energies and other relevant molecular properties. [8]

At a broader sense drug designing is classified in to two major areas, the first is referred to as ligand-based drug design and the second, structure-based drug design.

Drug discovery may also require fundamental research into the biological and chemical nature of the diseased state. These and other aspects of drug design and discovery require input from specialists in many other fields and so medicinal chemists need to have outline knowledge of the relevant aspects of these fields.

Drug design explain:

1. Drug receptor interaction on the basis of various physic-chemical properties.
2. Relationship between biological activity and structure.
3. Modify the drug molecule according to the need.
4. The effect of the drug towards the biological responds by various processes. [9]

Our work is placed in the context of fundamental and original research of some heterocyclic compound and their derivatives, the main objective of this work is the application of different methods of molecular modeling to predict the chemical reactivity, physical property and biological activity expected in new molecules.

The memory structure, consisting of five chapters, has been conceptually divided into two different parts. On the one hand, the bibliographical background section, which is composed by Chapter 1 and 2. On the other hand, Chapter 3, Chapter 4 and Chapter 5 dedicated to applications and results, deepens specific practical applications. The contents of the chapters are briefly described:

The first chapter **the triazole and the aromatase inhibition for breast cancer treatment:** is divided into two parts: in the first part, we will be spread out, over general information concerning the pharmacological classification and properties of triazole. In the second part, we will present general information of the breast cancer disease and therapeutic options available. The importance of an antihormonal treatment with aromatase inhibitors is introduced.

Chapter 2 Computational Approaches for Drug Design and Discovery: introduces the state of the art in molecular modeling techniques, with the special focus on the structure-based drug discovery methods. It also gives an overview of the various algorithms and models that are used in *in silico* drug discovery with a particular spotlight on algorithms and scoring functions employed in this work (QSAR and docking molecular)

The third chapter **drug likeness scoring and structure activity/property relationships of 1,2,3-triazole derivatives as aromatase inhibitor:** comprises a structural, electronic and energetic study on triazole and its derivatives. In this chapter we have the results of a comparative study on three methods used in calculation, PM3 and Density functional theory DFT, and Ab initio/HF, thus, the substitution effect on the electronic and energetic parameters of the basic core of triazole. We will also present a qualitative study on the relation structure-properties of a bioactive series of triazole (work published in: Journal of Fundamental and Applied Sciences Vol 10, No 3 (2018)).

Chapter 4: **QSAR Model for Predicting the Aromatase Inhibition Activity of 1.2.3 triazole derivatives:** In this chapter, we establish a quantitative relationship between physiochemical properties and biological activity of a series of bioactive derivatives of 1,2,3- triazole (QSAR Model).

Chapter 05: **Molecular docking studies and in Silico ADMET of new substituted 1.2.3 triazole derivatives for anti-breast cancer activity:** A virtual receptor site model is proposed for the interaction between the triazole derivatives and the enzyme aromatase and Determination of PK parameters In order to identify new molecules likely to become drugs (work published in: Journal of Bio nanoscience, Volume 12, Number 1, 2018).

REFERENCE

1. Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians*, *61*(2), 69-90.

2. Miller, W. R. (2003). Aromatase inhibitors: mechanism of action and role in the treatment of breast cancer. *Seminars in Oncology*, *30*(4 Suppl 14), 3-11.

3. Yao, N., & Recht, A. (2015). The real difficulty in improving concordance of clinical practice with guidelines. *Annals of Translational Medicine*, *3*(6).

4. Crowley, E., Di Nicolantonio, F., Loupakis, F., & Bardelli, A. (2013). Liquid biopsy: monitoring cancer-genetics in the blood. *Nature Reviews. Clinical Oncology*, *10*(8), 472-484.

5. Caporuscio, F., Rastelli, G., Imbriano, C., & Del Rio, A. (2011). Structure-based design of potent aromatase inhibitors by high-throughput docking. *Journal of Medicinal Chemistry*, *54*(12), 4006-4017.

6. Duan, Y.-C., Zheng, Y.-C., Li, X.-C., Wang, M.-M., Ye, X.-W., Guan, Y.-Y., … Liu, H.-M. (2013). Design, synthesis and antiproliferative activity studies of novel 1,2,3-triazole-dithiocarbamate-urea hybrids. *European Journal of Medicinal Chemistry*, *64*, 99-110.

7. Okada M, Yoden T, Kawaminami E, ShimadaY, Kudoh M, Isomura Y, Shikama H Fujikura T, Chem. Pharm. Bull. 1996. 44,1871-1879.

8. Sushil K. S, E. Sharma and Y. Sharma, (2017). A review: Recent computational approaches in medicinal chemistry: Computer aided drug designing and delivery. *The Pharma Innovation Journal*; 6(5): 05-10.

9. O. Lokendra K, S. Rachana, B. M. Rani, (2013). Modern drug design with advancement in QSAR: A review. Int. J. Res. Bio Sciences 2 (1), (1-12).

CHAPTER I

THE TRIAZOLE AND THE AROMATASE INHIBITION FOR BREAST CANCER TREATMENT

1. INTRODUCTION

Breast cancer is considered to be the most common type of cancer and is the leading cause of cancer-related death in women accounting for an estimated 23% of new cases and 14% of all cancer deaths in 2008 [1]. It is widely accepted that the majority of breast cancers are hormone-dependent and that estrogen is a key mediator in the progression and metastasis of breast tumors. Particularly, for postmenopausal women it has been reported that the concentration of 17b-estradiol (E2) in breast tumor can be tenfold higher than those in plasma [2]. The high concentration of E2 in breast tumors could be attributed to increased uptake from plasma or in situ aromatization of androgens to estrogens. The latter is afforded by aromatase, an enzyme involved in the rate-limiting step of estrogen biosynthesis by catalyzing three consecutive hydroxylation reactions that aromatize C19 androgens to C18 estrogens. [3]

Two main strategies to control or block breast cancer progression include binding of the estrogen receptors with estrogen receptor antagonists (such as tamoxifen), and inhibiting the production of estrogen with aromatase inhibitors. [4].

Triazoles are common pharmacophore found in a diverse range of biologically active molecules due to their potential structural features (i.e., capability of hydrogen bonding, stable to metabolic degradation and less undesired effects) [5]. Based on the aromatase inhibitors, the triazole ring plays a pivotal role in chelation with heme iron [6].

This chapter is divided into two parts. In the first, we will talk about the triazole molecule (its properties, its manufacture, and its biological activity). In the second part we will discuss in depth the breast cancer and hormone therapy.

2. THE TRIAZOLE

2.1 Triazole ring

[1,2,3]-Triazoles are important five-membered nitrogen heterocycles that are member of azole. They are aromatic ring compounds similar to the azoles, pyrazole and imidazole, but with an additional nitrogen atom in the ring structure. Synonyms for these triazoles sometimes denote that a proton is attached in the 1-position, as for example, the naming 1H-1,2,3-triazole or 1,2,3-1H-triazole. The table below shows the proprieties of this compound. [7]

Table I.1 proprieties of 1.2.3 triazole

Properties of 1, 2, 3-triazole
Molecular formula $C_2H_3N_3$
Molar mass 69.0654
Boiling point 203 °C
Melting point 23-25 °C
Density 1.192
Appearance colourless liquid
Solubility in water very soluble
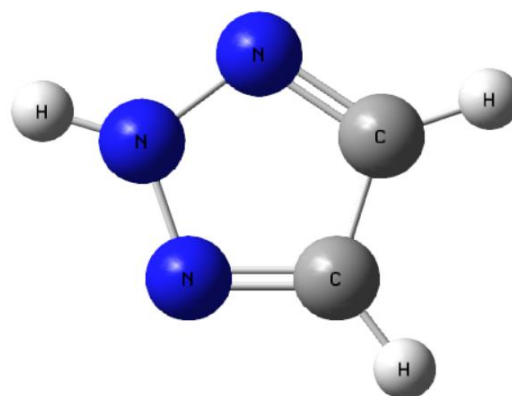
Generally, 1,2,3-triazoles are further subdivided into three main class, namely, monocyclic 1,2,3-triazoles, benzotriazoles and 1,2,3-triazolium salts as depicted in Figure I.1. [8] Monocyclic 1,2,3-triazoles and benzotriazoles are remarkably stable towards hydrolysis, oxidative/reductive conditions, and enzymatic degradation but reductive cleavage occurs under forcing conditions leading to the formation of triazolium salts. [9,10]
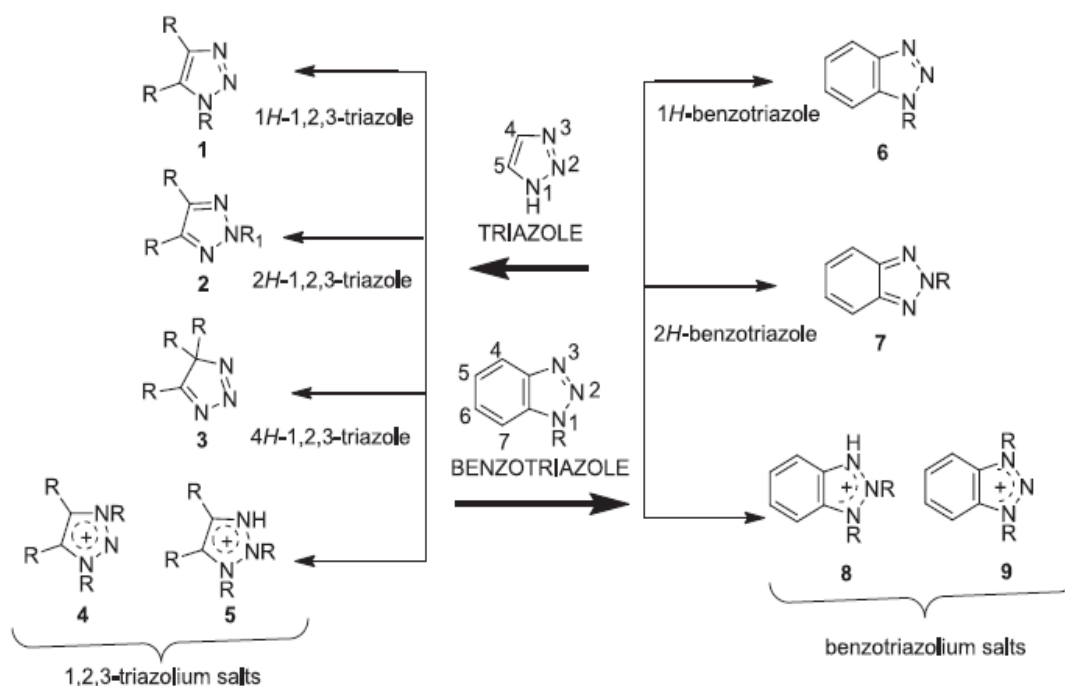
Figure I.1 classification of 1,2,3-triazole

The features possessed by the 1, 2, 3-triazoles make them pharmaceutically important molecules. They are stable to reduction and oxidation as well as to hydrolysis in acidic and basic conditions, which indicates their high aromatic stabilization. 1,2,3-triazoles have a high dipole moment (about 5 D) [10] and are able to participate actively in hydrogen bond formation as well as in dipole–dipole and π stacking interactions [11] which helps them in binding easily with the biological targets [12] and improves their solubility.

2.2 . Synthetic strategies

1,2,3-Triazole ring system has been a subject of intense research due to its versatile potential to interact with diverse biological systems.

In recent years, many synthetic methodologies have been developed for the synthesis of this ring system. The most popular reaction to produce the 1,2,3-triazole moiety is the 1,3-dipolar cycloaddition also known as Huisgen cycloaddition, between an azide and a terminal alkyne, under thermal conditions but was not initially applied much in organic synthesis owing to the poor regioselectivity (1,4- and 1,5-disubstituted 1,2,3-triazoles), low chemical yield and elevated temperatures [13]. A copper (I)- catalyzed version of azide–alkyne cycloaddition reaction (CuAAC), the click chemistry approach invented by

Sharpless resulted in the production of a large number of 1,4-disubstituted 1,2,3- triazoles in very high yields [14,15]. Over the past few decades, the various synthetic methodologies have received a lot of attention and offered new opportunities for medicinal chemists. The various metal free and metal catalyzed approaches explored for the preparation of triazole frame work are illustrated in Figs. I.2 and I.3. [16]
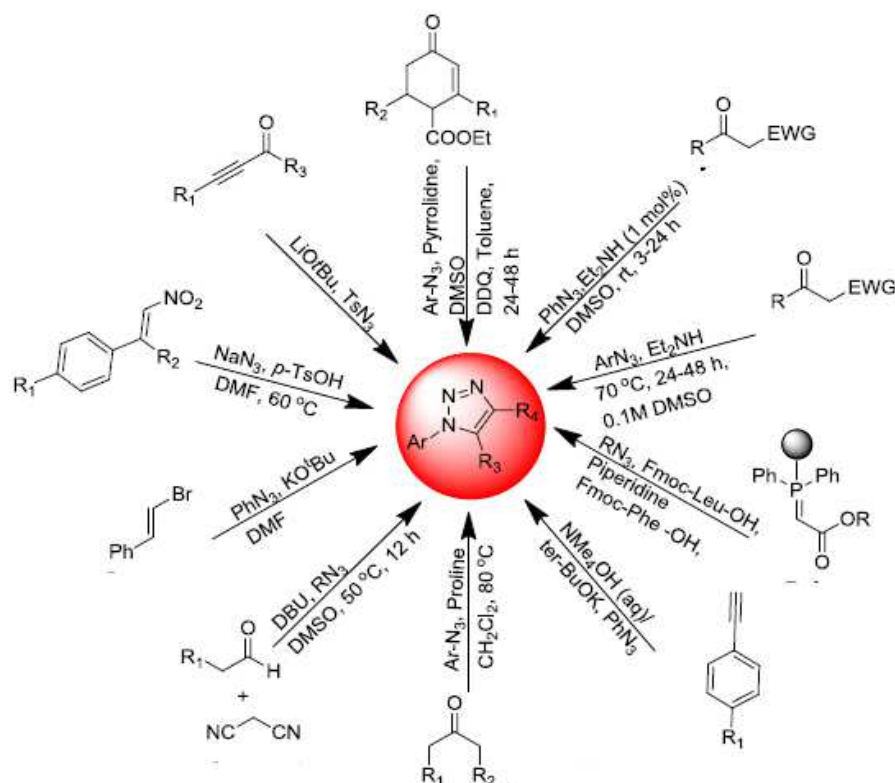
Figure I.2. A summary of various metal free approaches toward synthesis of 1,2,3-triazoles.
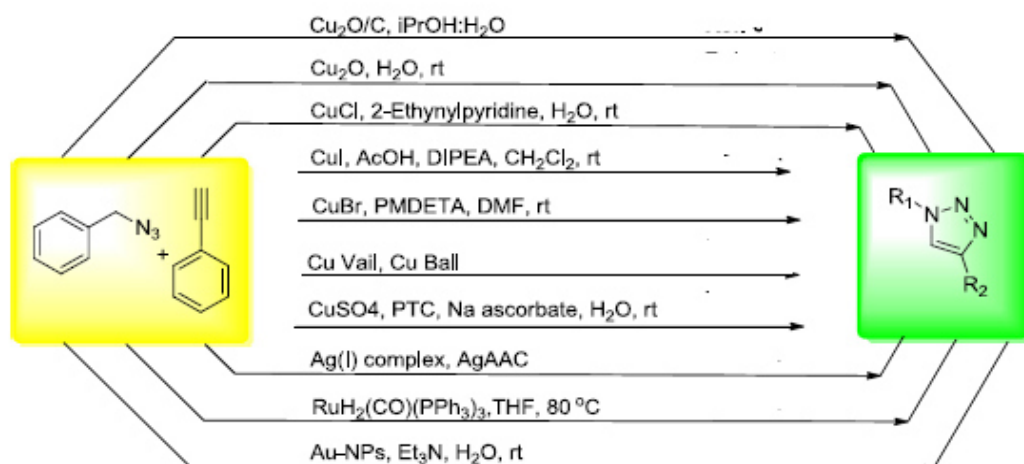
Figure. I 3. A summary of transition metal catalyzed synthesis of 1,2,3-triazoles

2.3 Biologically Active Derivatives

Due to their easy access and high enzymatic stability, benzotriazole and 1,2,3-triazole systems are frequently used as building blocks in drug design. Among antitumor agents, vorozole (structure 1) is a high-affinity competitive aromatase inhibitor, designed for inhibiting estrogen synthesis in patients with breast cancer [17-19]. Benzotriazole derivative 2 exhibits remarkable activity against leukemia, ovarian, renal, and lung cancers [20]. The structures may be complex, like compound 3[21], or simple, like compound 4 [22], both of them exhibiting anti-inflammatory activities, although based on different principles. Nucleoside analog 5 inhibits strongly helicase activity of hepatitis C virus [23], whereas compound 6 and several of its analogs show strong activity against respiratory syncytial virus (RSV) [24].

Very simple derivatives of benzotriazole with biological activity include 5,6 dimethylbenzotriazole, a very effective agent against cysts of Acanthamoeba castellanii [25], tetrabromobenzotriazole, which provides selective inhibition of protein kinase CK2 [26] and induces apoptosis of Jurkat cells [27], 1-salicylyl-4- methylbenzotriazole, potassium channel activator [28] and 1-isopropyl-1H-benzotriazole-4-carboxylic acid, a selective agonist of human orphan G-protein-coupled receptor GPR109b [29].Several 1,2,3-triazole derivatives have been designed to target G-protein-coupled receptors. Among them are neurokine NK1 antagonists 7 [30,31] and 8 [32], selective A3 adenosine receptor agonist 9 [33] and highly selective a1 adrenoreceptor antagonists [34]. Other1,2,3-triazole derivatives are of interest as inhibitors of some key enzymes: acetylcholinesterase [35], glycogen synthase kinase-3 [36], glycosidase [37], galectin-1 [38,39], and a-2,3-sialyltransferase [40].There are also 1,2,3-triazoles with antiviral [41,42], antibacterial [43,44], antithrombotic [45], or antiplatelet [46] activities. Some triazoles work as potassium channel activators [47], others as calcium signal transduction inhibitors [48]. 1,5- Diaryl-_2-1,2,3-triazolines are recognized anticonvulsant agents [49,50]. Among biologically active benzotriazoles are also inactivators of the severe acute respiratory syndrome 3CL protease [51], trichostatin suppressors [52], antagonists of the gonadotropin releasing hormone [53], and nonpeptide inhibitors of protein tyrosine phosphatase 1B [54].
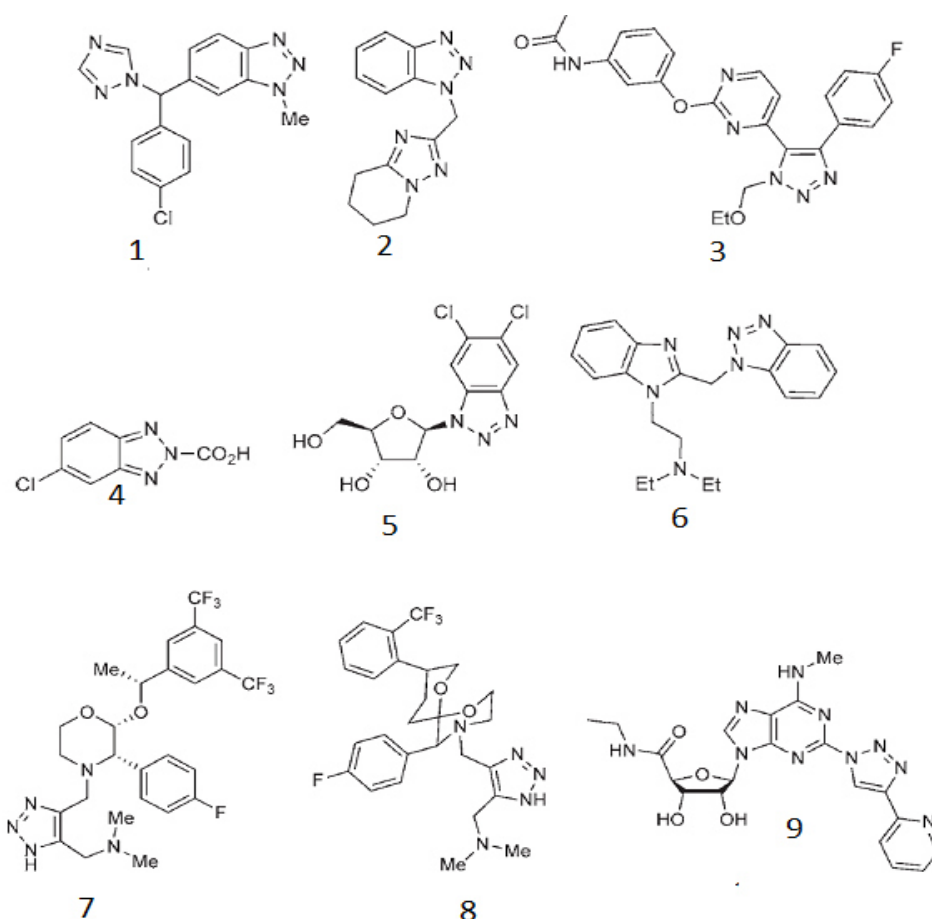
Figure I.4 Some Biologically active derivatives

2.4 Other Applications

Due to strong complexing affinities to copper and some other ions, benzotriazole and its derivatives have found wide application in anticorrosion formulations. Hundreds of patents covering this subject are registered each year. One of the major applications of such formulations is in electronics that include thiol passivation of copper interconnects during semiconductor manufacturing, grinding composition for polishing of semiconductor devices, corrosion-preventing agents for etching of insulator films in manufacture of semiconductor devices, cleaning solutions for electrohydrodynamic cleaning of semiconductors, components of polymer coatings for silver-plated circuits, and in dispersants for preparation of nickel-coated copper powder for electricity-conducting inks. Benzotriazole is also commonly used as an unticlogging agent in jet inks for forming high-quality images.

Anticorrosion abilities of benzotriazole and its derivatives are also widely utilized in fluids for all kind of machinery.

They are important antifriction–antiwear additives for engine oils, components of antirusting grease for aircraft, biodegradable lubricants for turbines, brake liquids based on polyoxyalkylene synthetic oils, metal corrosion inhibitors in aqueous coolants containing acetic acid and propylene glycol, grease for gas compressors for fuel cell systems, emulsifiable oil for preparation of noncombustible oil–water hydraulic emulsions for coal mining, environment protecting lubricating oil for refrigerators, antifreeze composition for diesel engines, and lubricating oil compositions for hot rolling aluminium plates. [55]

3. AROMATASE INHIBITION FOR BREAST CANCER TREATMENT

3.1 What Is Breast Cancer?

Breast cancer starts when cells in the breast begin to grow out of control. These cells usually form a tumor that can often be seen on an x-ray or felt as a lump. The tumor is malignant (cancer) if the cells can grow into (invade) surrounding tissues or spread (metastasize) to distant areas of the body (figureI.5). Breast cancer occurs almost entirely in women, but men can get breast cancer, too. Cells in nearly any part of the body can become cancer and can spread to other areas. [56]
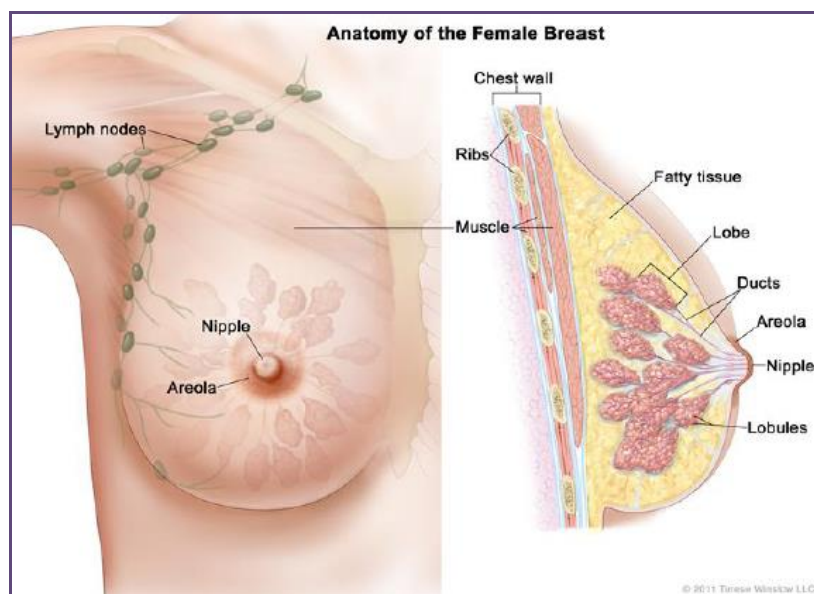
Figure I-5 Anatomy of the Female Breast.

3.2 Breast cancer molecular subtypes

In breast cancer, hormone receptors are the proteins located in and around breast cells. These receptors signal cells — both healthy and cancerous — to grow. In the case of breast cancer, the hormone receptors tell the cancer cells to grow uncontrollably, and a tumor result. Hormone receptors can interact with estrogen or progesterone. Estrogen receptors are the most common. This is why ER-positive is the most common form of breast cancer.[57]

Profiling the biologic make-up of breast cancer tissue can help doctors determine how well a patient will respond to different cancer treatments and how aggressive the cancer might be. Researchers are focused on three subtypes—hormone receptor-positive, HER2-positive, and triple-negative breast cancer
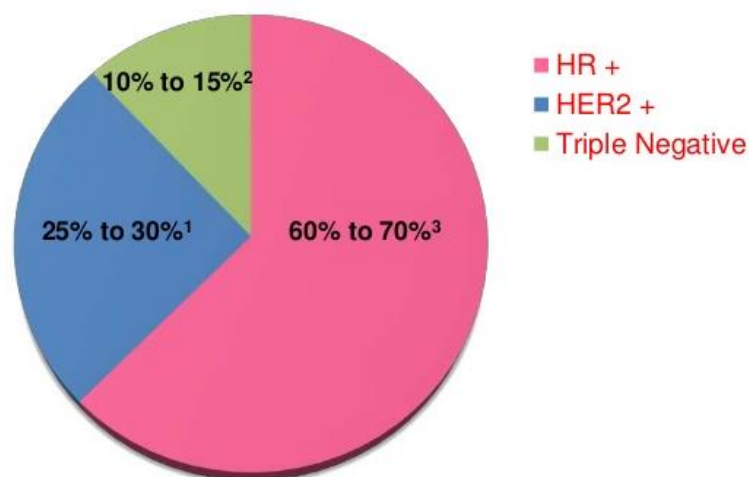
Figure 3-6 Breast cancer molecular subtypes

3.2.1 Hormone Receptor-Positive Breast Cancer

About 80% of all breast cancers are "ER-positive." (figure I.6) That means the cancer cells grow in response to the hormone estrogen. About 65% of these are also "PR-positive." They grow in response to another hormone, progesterone.

If your breast cancer has a significant number of receptors for either estrogen or progesterone, it's considered hormone-receptor positive.

Tumors that are ER/PR-positive are much more likely to respond to hormone therapy than tumors that are ER/PR-negative.

You may have hormone therapy after surgery, chemotherapy, and radiation are finished. These treatments can help prevent a return of the disease by blocking the effects of estrogen. They do this in one of several ways.
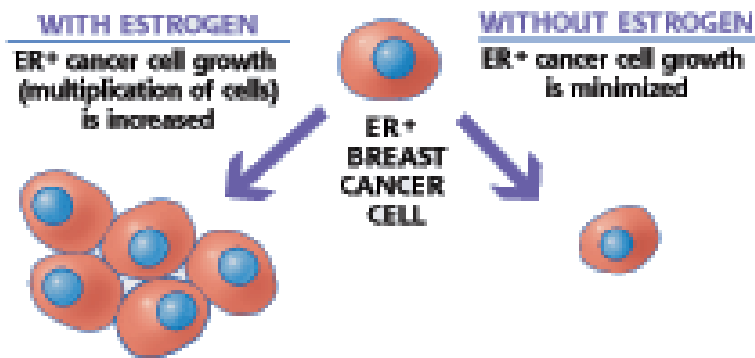
Figure 3.7 ER+ breast cancer cell

3.2.2 HER2-Positive Breast Cancer

In about 20% of breast cancers, the cells make too much of a protein known as HER2. These cancers tend to be aggressive and fast-growing.

For women with HER2-positive breast cancers, the drug trastuzumab (Herceptin) has been shown to dramatically reduce the risk of the cancer coming back. It's standard treatment to give this medication along with chemotherapy after surgery to people with breast cancer that's spread to other areas. It can also be used for early-stage breast cancer. But there is a small but real risk of heart damage and possible lung damage. Scientists are still studying how long women should take this medication for the greatest benefit.

3.2.3 Triple-Negative Breast Cancer

Some breast cancers (between 10% and 20%) are known as "triple negative" because they don't have estrogen and progesterone receptors and don't overexpress the HER2 protein. Many breast cancers associated with the gene BRCA1 are triple negative. [58]

3.3 Treatment options

If breast cancer is not treated, the cancer cells in the breast will keep growing. They can spread to other parts of the body, such as bones, the liver or the lungs. This is called secondary breast cancer. Over time, these cancer cells can stop some organs in your body from working, or lead to other life-threatening problems.

Breast cancer is treated in several ways. It depends on the kind of breast cancer and how far it has spread. People with breast cancer often get more than one kind of treatment. [59]

The Table I.2 summarizes these types.

Table I.2 treatment options

Treatments for breast cancer	
Surgery	An operation to take out the cancer and some of the healthy tissue around it
Chemotherapy	Medicine to kill cancer cells in the breast and other parts of the body
Radiotherapy	Treating the area where the cancer was found with radiation (X-rays)
Hormone therapy	Medicine to stop hormone receptor-positive breast cancer growing
Targeted therapy	Medicine to stop certain types of breast cancer, such as HER2-positive breast cancer, growing

3.3.1 Hormone therapy for breast cancer

Hormone therapy to treat breast cancer uses drugs or treatments to lower levels or block the action of female sex hormones (estrogen and progesterone) in a woman's body. This helps slow the growth of many breast cancers.

Hormone therapy makes cancer less likely to return after breast cancer surgery. It also slows the growth of breast cancer that has spread to other parts of the body.

Hormone therapy can also be used to help prevent cancer in women at high risk for breast cancer.

Figure 3-8 Estrogen hormone level

Hormone therapy can work in two ways:

- By blocking the estrogen from acting on cancer cells
- By lowering estrogen levels in a woman's body [62].

3.3.2 Drugs That Block Estrogen

Some drugs work by blocking estrogen from causing cancer cells to grow. [62].

- Tamoxifen (Nolvadex)
- Toremifene (Fareston)
- Fulvestrant (Faslodex)

3.3.3 Drugs That Lower Estrogen Levels

Some drugs, called aromatase inhibitors (AIs), stop the body from making estrogen in tissues such as fat and skin. But these drugs do not work to make the ovaries stop making estrogen. For this reason, they are used mainly to lower estrogen levels in women who have been through menopause (postmenopausal). Their ovaries no longer make estrogen.

Premenopausal women can take AIs if they are also taking drugs that stop their ovaries from making estrogen. [60]

Aromatase inhibitors include:

- Anastrozole (Arimidex)
- Letrozole (Femara)
- Exemestane (Aromasin)

3.4 What is an aromatase inhibitor and how does it work?

Aromatase (CYP19) is an enzyme of the cytochrome P450 superfamily that catalyzes the final and rate-limiting step of the conversion of androgens, testosterone and androstenedione, into estrogens, estradiol and estrone, respectively.

Aromatase inhibitors and inactivators interfere with the body's ability to produce estrogen from androgens by suppressing aromatase enzyme activity Before menopause, ovarian aromatase is responsible for the majority of circulating estrogen and is exquisitely sensitive to changes in luteinising hormone (LH). Following menopause, aromatase in fat and muscle may be responsible for much of the circulating estrogen. Aromatase in highly estrogen-sensitive tissues, such as the breast, uterus, vagina, bone, brain, heart and blood vessels, provides local estrogen in an autocrine fashion. [61] .

3.5 The overall structure of aromatase

A ribbon diagram of the overall crystal structure of human placental aromatase is shown in Figure. I.4 The tertiary structure consists of 12 major helices (labeled A through L) and 10 strands (numbered 1 through 10) distributed into 1 major and 3 minor sheets, and follows the characteristic cytochrome P450 fold. The bound androstenedione molecule at the heme distal site, the active site of the enzyme, and shown within its unbiased electron density, makes two hydrogen bond-forming contacts the 3-keto and 17- keto oxygens with Asp309 side chain and Met374 backbone amide, respectively (Figure.I.9). The major _-sheet is a mixed 4-stranded sheet that begins near the amino terminus (first two strands are 1:83–88 and 2:93–97) but ends in two strands from the carboxyl terminal half of the polypeptide chain (3:373–376 and 6:393–396).

A feature somewhat special to the aromatase structure is that the amino-terminal residues 47–50, which makes one backbone hydrogen bond with _1, adds an extra _-strand-like element to this sheet. Each of the three minor sheets consists of two anti-parallel strands scattered over the polypeptide chain (sheet2: 4:381–383 and 5:386–388; sheet3: 8:473–475 and 9:479–481; sheet 4: 7:458–461 and 10:491–494). Of the 12 major helices, the lengths, locations and orientations of helices I (293–324), F (210–227), G (242–267), H (278–287), C (138–152), D (155–174), E (187–205), J (326–341), K (354–366) and L (440–455) are similar to those found in most of the cytochrome P450s. Other helices, namely A(57–68), A (69–80), B (100–109), B(119–126), G(232–236), H(271–

274), J(346–349), K(398–404), and K (414–418) are 1–4 turns long and have more variability among P450s in terms of their locations, lengths and orientations. Forinstance, when compared with two human P450s 3A4 and 2D6 that aromatase has the closest resemblance to (~14–18% sequence identity), the helix Ain aromatase is longer than a similar one in 3A4 and is not seen in 2D6. The other notable difference in the secondary structures between aromatase and 3A4 is that the helix F in 3A4 is separated into two shorter helices by a stretch of polypeptide.

As discussed below, this region of the structure contributes significantly to the constitution of the active site. Another difference is that the G-helix in aromatase is at least 1 turn longer than those in 3A4 and 2D6. The F-helix–loop–G-helix region, in general, appears to be different in different P450s. With the helix G_ in the middle, the loop is tighter in aromatase than in either 3A4 or 2D6, both of which have longer intervening loops.
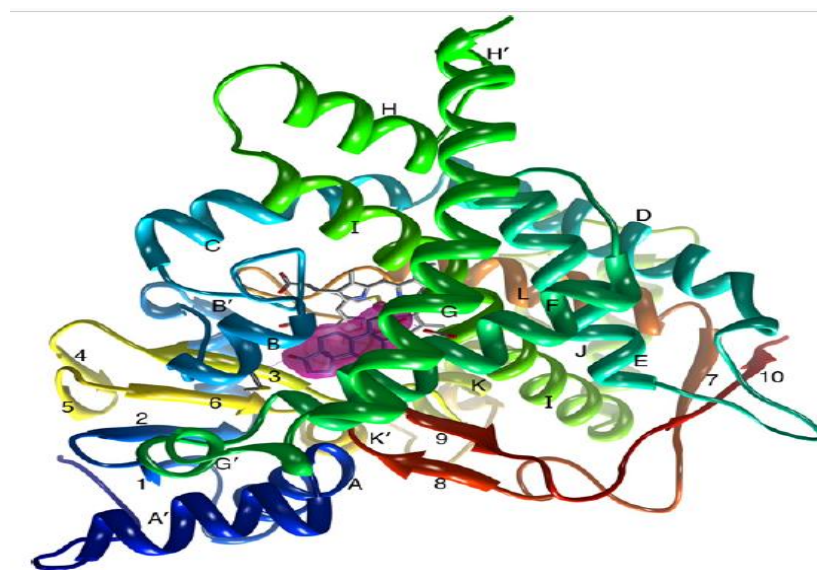
Figure. I.9 A ribbon diagram showing the overall structure of human placental aromatase.

A striking feature of the tertiary aromatase structure is that long loops interconnect well-defined secondary structure elements, in general agreement with other P450 structures. One example is the polypeptide between the 2-turn helix Kand helix L. This stretch of 35 residue polypeptides (405–439), devoid of much secondary structure except the 1-turn helix K, contributes the all-important cysteine ligand (Cys437) to the heme iron.

Other examples of long loops are between helices Band C, 7 and 8, and 9 and 10, all of which either contribute active site residues or have roles in the scaffolding of functionally important elements. Although these loops have little intra-loop interactions through hydrogen bonding, they stabilize by interacting with other structural elements and are, thus, well-defined in the electron density map. Another feature common to all cytochrome P450s is the ligation of the heme group via its propionate moieties by arginine and tryptophan side chains through ionic and hydrogen bonding interactions. These side residues in aromatase are Arg115, Trp141, Arg145, Arg375, and Arg435, homologous to those in 3A4, 2D6 and others. [63]

3.6 Types of aromatase inhibitors

AIs are classified as first, second, or third generation according to the specificity and potency with which they inhibit the aromatase enzyme. They are further subclassified as type 1 or type 2 inhibitors, according to the reversibility of their inhibitory activity (Table I.5).

Type 1 inhibitors, steroidal analogues of androstenedione, irreversibly inhibit the aromatase enzyme by covalently binding to it, thus earning the name "suicidal inhibitors." Permanent inactivation persists after discontinuation of the drug until the peripheral tissues synthesize new enzymes.

In contrast, nonsteroidal type 2 inhibitors bind reversibly to the aromatase enzyme, resulting in competitive inhibition [64]. Third-generation AIs (i.e., anastrozole, letrozole, and exemestane) are the most potent, most selective, and least toxic AIs known today and can reduce serum estrogen by more than 95%. In addition, their pharmacokinetic properties (a half-life of approximately 48 hours for anastrozole and letrozole and 27 hours for exemestane) allow for a once-daily dosing schedule [65].

Their selective inhibitory properties allow their use without the need for supplemental corticosteroidal or mineralocorticoid supplementation, as is the case with the nonspecific AI aminoglutethimide.

Several clinical trials have evaluated the efficacy and safety of these agents. Here, we review the published literature regarding these trials and summarize advances in the hormonal treatment of breast cancer.

Table I.3. Type of aromatase inhibitor

Generation	Type 1 (Steroidal Inactivator)	Type 2 (Nonsteroidal Inhibitor)
First	None	Aminoglutethimide
Second	Formestane	Fadrozole Rogletimide
Third	Exemestane (Aromasin)	Anastrozole (Arimidex) Letrozole (Femara) Vorozole

3.7 Mechanisms of action

The AIs are classified into two types: (*a*) type I, suicidal or noncompetitive inhibitors; and (*b*) type II, competitive inhibitors [66,67]. Type I inhibitors are steroidal compounds, and type II inhibitors are nonsteroidal drugs. Both types mimic normal substrates (androgens), competing with the substrate for access to the binding site on the enzyme. After initial binding, the next step differs for the two types: once a noncompetitive inhibitor has bound, the enzyme initiates its typical sequence of hydroxylation, but hydroxylation produces an unbreakable covalent bond between the inhibitor and the enzyme protein. Enzyme activity is thus permanently blocked; even if all unattached inhibitor is removed, enzyme activity can only be restored by new enzyme synthesis. Competitive inhibitors reversibly bind to the active enzyme site, and either no enzyme activity is triggered, or it is without effect. The inhibitor can disassociate from the binding site, allowing renewed competition between the inhibitor and the substrate for binding to the site. As a result, the effectiveness of competitive inhibitors depends on the relative concentrations and affinities of the inhibitor and the substrate. Continued activity requires constant presence of inhibitor.

To compete for binding to the active site, both competitive and noncompetitive inhibitors must necessarily share important structural features with the endogenous substrate. Noncompetitive inhibitors must also share structural features with androgens, allowing them to interact with the catalytic residual on the enzyme protein. This renders them inherently selective. By contrast, most competitive inhibitors interact with the heme iron, a common feature of all cytochrome P450 enzymes. Some may also bind to the highly conserved oxygen binding site in addition to the substrate binding site. Thus, unless the specificity of a competitive inhibitor is reinforced through other structural features, it may block the activity of a variety of cytochrome P450 enzymes, as does aminoglutethimide.

Both anastrozole and letrozole are type II nonsteroidal AIs, whereas exemestane has a steroidal structure and is classified as a type I AI, also known as an aromatase inactivator because it irreversibly binds with and permanently inactivates the enzyme. The clinical relevance of these differences in mechanism of action, if any, remains to be established. [68]
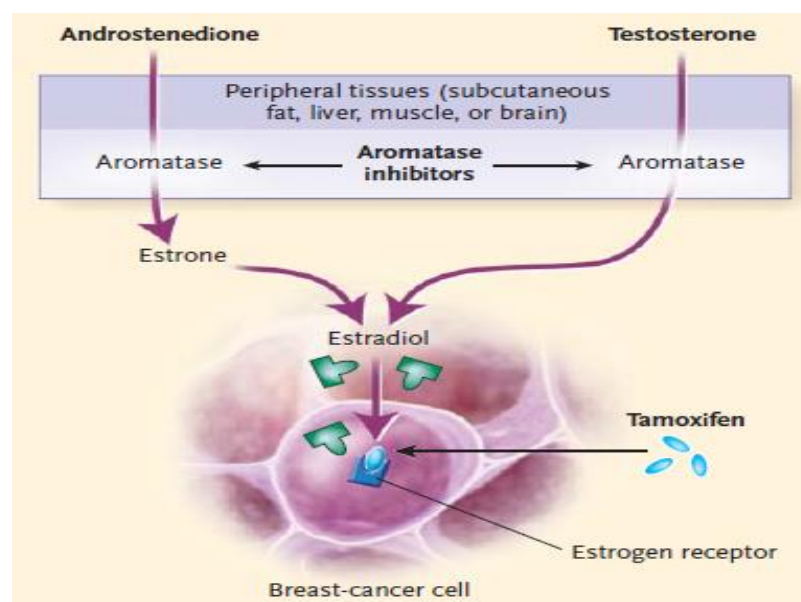
Figure I.9 Mechanism of Action of Aromatase Inhibitors

4. REFERENCE

1. Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians*, *61*(2), 69-90.

2. Van Landeghem, A. A., Poortman, J., Nabuurs, M., & Thijssen, J. H. (1985). Endogenous concentration and subcellular distribution of estrogens in normal and malignant human breast tissue. *Cancer Research*, *45*(6), 2900-2906.

3. Yue, W., Wang, J. P., Hamilton, C. J., Demers, L. M., & Santen, R. J. (1998). In situ aromatization enhances breast tumor estradiol levels and cellular proliferation. *Cancer Research*, *58*(5), 927-932.

4. Favia, A. D., Nicolotti, O., Stefanachi, A., Leonetti, F., & Carotti, A. (2013). Computational methods for the design of potent aromatase inhibitors. *Expert Opinion on Drug Discovery*, *8*(4), 395-409.

5. Agalave, S. G., Maujan, S. R., & Pore, V. S. (2011). Click chemistry: 1,2,3-triazoles as pharmacophores. *Chemistry, an Asian Journal*, *6*(10), 2696-2718.

6. Neves, M. A. C., Dinis, T. C. P., Colombo, G., & Sá e Melo, M. L. (2009). Fast Three Dimensional Pharmacophore Virtual Screening of New Potent Non-Steroid Aromatase Inhibitors. *Journal of Medicinal Chemistry*, *52*(1), 143-150.

7. Sandip Ravindra Kale. Multicomponent reactions catalyzed by Hydrotalcites and Hydroxyapatites. Thesis submitted to Institute of Chemical Technology, Mumbai.

8. Huisgen R, Guenter S, Leander M. 1,3-Dipolar cycloadditions. XXXII. Kinetics of the addition of organic azides to carbon–carbon multiple bonds. Chem Ber 1967;100:2494–2507.

9. Rostovtsev, V. V., Green, L. G., Fokin, V. V., & Sharpless, K. B. (2002). A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective « ligation » of azides and terminal alkynes. *Angewandte Chemie (International Ed. in English)*, *41*(14), 2596-2599.

10. Dörner, S., & Westermann, B. (2005). A short route for the synthesis of "sweet" macrocycles via a click-dimerization–ring-closing metathesis approach. *Chemical Communications*, *0*(22), 2852-2854.

11. Qi, X., & Ready, J. M. (2007). Copper-Promoted Cyloaddition of Diazocarbonyl Compounds and Acetylides. *Angewandte Chemie (International ed. in English)*, *46*(18), 3242-3244.

12. Löwdin, P.-O. (1955). Quantum Theory of Many-Particle Systems. I. Physical Interpretations by Means of Density Matrices, Natural Spin-Orbitals, and Convergence Problems in the Method of Configurational Interaction. *Physical Review*, *97*(6), 1474-1489.

13. R. Huisgen, S. Guenter, M. Leander, 1.3-Dipolare cycloadditionen, XXXII. Kinetik der additionenorganischerazide an CC-mehrfachbindungen, Chem. Ber. 100 2494–2507. (1967).

14. Kolb, H. C., Finn, M. G., & Sharpless, K. B. (2001). Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie (International Ed. in English)*, *40*(11), 2004-2021.

15. H.C. Kolb, M.G. Finn, K.B. Sharpless, *Angew. Chem. Int. Ed*. 113 2056–2075. (2001).

16. Dheer, D., Singh, V., & Shankar, R. (2017). Medicinal attributes of 1,2,3-triazoles: Current developments. *Bioorganic Chemistry*, *71*, 30-54.

17. Dowsett, M., Harper-Wynne, C., Boeddinghaus, I., Salter, J., Hills, M., Dixon, M., … Smith, I. (2001). HER-2 amplification impedes the antiproliferative effects of hormone therapy in estrogen receptor-positive primary breast cancer. *Cancer Research*, *61*(23), 8452-8458.

18. Archer, C. D., Parton, M., Smith, I. E., Ellis, P. A., Salter, J., Ashley, S., … Dowsett, M. (2003). Early changes in apoptosis and proliferation following primary chemotherapy for breast cancer. *British Journal of Cancer*, *89*(6), 1035-1041.

19. Kirilovas, D., Naessen, T., Bergström, M., Bonasera, T. A., Bergström-Pettermann, E., Holte, J., … Långström, B. (2003). Effects of androgens on aromatase activity and 11C-vorozole binding in granulosa cells in vitro. *Acta Obstetricia Et Gynecologica Scandinavica*, *82*(3), 209-215.

20. Akhtar, T., Hameed, S., Al-Masoudi, N. A., & Khan, K. M. (2007). Synthesis and anti-HIV activity of new chiral 1,2,4-triazoles and 1,3,4-thiadiazoles. *Heteroatom Chemistry*, *18*(3), 316-322.

21. Tullis, J. S., VanRens, J. C., Natchus, M. G., Clark, M. P., De, B., Hsieh, L. C., & Janusz, M. J. (2003). The development of new triazole based inhibitors of tumor

necrosis factor-alpha (TNF-alpha) production. *Bioorganic & Medicinal Chemistry Letters*, *13*(10), 1665-1668.

22. Cha, J. H., Son, M. H., Min, S.-J., Cho, Y. S., & Lee, J. K. (2011). 3-Hydroxy-2-[(2E)-1-(2-hydroxy-6-oxocyclohex-1-en-1-yl)-3-(2-methoxyphenyl)prop-2-en-1-yl]cyclohex-2-en-1-one. *Acta Crystallographica Section E: Structure Reports Online*, *67*(Pt 10), o2739.

23. Borowski, P., Mueller, O., Niebuhr, A., Kalitzky, M., Hwang, L. H., Schmitz, H., … Kulikowsk, T. (2000). ATP-binding domain of NTPase/helicase as a target for hepatitis C antiviral therapy. *Acta Biochimica Polonica*, *47*(1), 173-180.

24. Kopańska, K., Najda, A., Zebrowska, J., Chomicz, L., Piekarczyk, J., Myjak, P., & Bretner, M. (2004). Synthesis and activity of 1H-benzimidazole and 1H-benzotriazole derivatives as inhibitors of Acanthamoeba castellanii. *Bioorganic & Medicinal Chemistry*, *12*(10), 2617-2624.

25. Battistutta, R., De Moliner, E., Sarno, S., Zanotti, G., & Pinna, L. A. (2001). Structural features underlying selective inhibition of protein kinase CK2 by ATP site-directed tetrabromo-2-benzotriazole. *Protein Science : A Publication of the Protein Society*, *10*(11), 2200-2206.

26. Sarno, S., de Moliner, E., Ruzzene, M., Pagano, M. A., Battistutta, R., Bain, J., … Pinna, L. A. (2003). Biochemical and three-dimensional-structural study of the specific inhibition of protein kinase CK2 by [5-oxo-5,6-dihydroindolo-(1,2-a)quinazolin-7-yl]acetic acid (IQA). *Biochemical Journal*, *374*(Pt 3), 639-646.

27. Luo, Y., & Hu, Y. (2003). A Novel and Efficient Strategy for the Preparation of 2,4-Disubstituted-1,2,3-triazoles. *Synthetic Communications*, *33*(20), 3513-3517.

28. Semple, G., Skinner, P. J., Cherrier, M. C., Webb, P. J., Sage, C. R., Tamura, S. Y., … Connolly, D. T. (2006). 1-Alkyl-benzotriazole-5-carboxylic acids are highly selective agonists of the human orphan G-protein-coupled receptor GPR109b. *Journal of Medicinal Chemistry*, *49*(4), 1227-1230.

29. T. Harrison, A. P. Owens, B. J. Williams, C. J. Swain, A. Williams, E. J. Carlson, W. Rycroft, F. D. Tattersall, M. A. Cascieri, G. G. Chicchi, S. Sadowski, N. M. J. Rupniak, and R. J. Hargreaves, J. Med. Chem., 44, 4296. , 2001.

30. Duffy, R. A., Varty, G. B., Morgan, C. A., & Lachowicz, J. E. (2002). Correlation of neurokinin (NK) 1 receptor occupancy in gerbil striatum with behavioral effects of NK1 antagonists. *The Journal of Pharmacology and Experimental Therapeutics*, *301*(2), 536-542.

31. Seward, E. M., Carlson, E., Harrison, T., Haworth, K. R., Herbert, R., Kelleher, F. J., … Williams, B. J. (2002). Spirocyclic NK(1) antagonists I: [4.5] and [5.5]-spiroketals. *Bioorganic & Medicinal Chemistry Letters*, *12*(18), 2515-2518.

32. Melman, A., Gao, Z.-G., Kumar, D., Wan, T. C., Gizewski, E., Auchampach, J. A., & Jacobson, K. A. (2008). Design of (N)-Methanocarba Adenosine 5′-Uronamides as Species-Independent A3 Receptor-Selective Agonists. *Bioorganic & medicinal chemistry letters*, *18*(9), 2813-2819.

33. Mayer, P., Brunel, P., Chaplain, C., Piedecoq, C., Calmel, F., Schambel, P., … Imbert, T. (2000). New substituted 1-(2,3-dihydrobenzo[1, 4]dioxin-2-ylmethyl)piperidin-4-yl derivatives with alpha(2)-adrenoceptor antagonist activity. *Journal of Medicinal Chemistry*, *43*(20), 3653-3664.

34. Manetsch, R., Krasiński, A., Radić, Z., Raushel, J., Taylor, P., Sharpless, K. B., & Kolb, H. C. (2004). In situ click chemistry: enzyme inhibitors made to their own specifications. *Journal of the American Chemical Society*, *126*(40), 12809-12818.

35. Lin, J. R. (2010). Bis(5-methyl-1-phenyl-1H-1,2,3-triazole-4-carboxylic acid) monohydrate. *Acta Crystallographica Section E: Structure Reports Online*, *66*(Pt 8), 1967.

36. Aguilar, M., Díaz-Pérez, P., García-Moreno, M. I., & García Fernández, J. M. (2008). Synthesis and Biological Evaluation of Guanidine-Type Iminosugars. *The Journal of Organic Chemistry*, *73*(5), 1995-1998.

37. Tejler, J., Tullberg, E., Frejd, T., Leffler, H., & Nilsson, U. J. (2006). Synthesis of multivalent lactose derivatives by 1,3-dipolar cycloadditions: selective galectin-1 inhibition. *Carbohydrate Research*, *341*(10), 1353-1362.

38. Giguère, D., Patnam, R., Bellefleur, M.-A., St-Pierre, C., Sato, S., & Roy, R. (2006). Carbohydrate triazoles and isoxazoles as inhibitors of galectins-1 and -3. *Chemical Communications (Cambridge, England)*, (22), 2379-2381.

39. Chang, K.-H., Lee, L., Chen, J., & Li, W.-S. (2006). Lithocholic acid analogues, new and potent alpha-2,3-sialyltransferase inhibitors. *Chemical Communications (Cambridge, England)*, (6), 629-631.

40. Zhang, X., Hsung, R. P., & Li, H. (2007). A triazole-templated ring-closing metathesis for constructing novel fused and bridged triazoles. *Chemical Communications*, *0*(23), 2420-2422.

41. Zhang, X., Hsung, R. P., & Li, H. (2007). A triazole-templated ring-closing metathesis for constructing novel fused and bridged triazoles. *Chemical Communications*, *0*(23), 2420-2422.

42. Shibl, M. F., Elroby, S. A. K., & Hilal, R. H. (2011). Solvent and substituent effects on the electronic structures of triazoles: computational study. *Molecular Simulation*, *37*(1), 11-17. Weidner-Wells, M. A., Werblood, H. M., Goldschmidt, R., Bush, K., Foleno, B. D., Hilliard, J. J., … Macielag, M. J. (2004). The synthesis and antimicrobial evaluation of a new series of isoxazolinyl oxazolidinones. *Bioorganic & Medicinal Chemistry Letters*, *14*(12), 3069-3072.

43. Rehse, K. (1970). [The Otto-reaction of strychnine]. *Archiv Der Pharmazie Und Berichte Der Deutschen Pharmazeutischen Gesellschaft*, *303*(6), 518-524.

44. Jordão, A. K., Ferreira, V. F., Souza, T. M. L., de Souza Faria, G. G., Machado, V., Abrantes, J. L., … Cunha, A. C. (2011). Synthesis and anti-HSV-1 activity of new 1,2,3-triazole derivatives. *Bioorganic & Medicinal Chemistry*, *19*(6), 1860-1865.

45. Biagi, G., Calderone, V., Giorgi, I., Livi, O., Scartoni, V., Baragatti, B., & Martinotti, E. (2000). 5-(4'-Substituted-2'-nitroanilino)-1,2,3-triazoles as new potential potassium channel activators. I. *European Journal of Medicinal Chemistry*, *35*(7-8), 715-720.

46. Berlin, J., Tutsch, K. D., Hutson, P., Cleary, J., Rago, R. P., Arzoomanian, R. Z., … Wilding, G. (1997). Phase I clinical and pharmacokinetic study of oral carboxyamidotriazole, a signal transduction inhibitor. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *15*(2), 781-789.

47. Berlin, J., Tutsch, K. D., Arzoomanian, R. Z., Alberti, D., Binger, K., Feierabend, C., … Wilding, G. (2002). Phase I and pharmacokinetic study of a micronized formulation of carboxyamidotriazole, a calcium signal transduction inhibitor: toxicity, bioavailability and the effect of food. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *8*(1), 86-94.

48. H. Matloubi, A. Shafiee, N. Saemian, G. Shirvani, and F. J. Daha, J. Label. Compd. Radiopharm, 47, 31, 2004.

49. C. Y. Wu, K. Y. King, C. J. Kuo, J. M. Fang, Y. T. Wu, M. Y. Ho, C. L. Liao, J. J. Shie, P. H. Liang, and C. H. Wong, Chem. Biol, 13, 261, 2006.

50. Cha, J. H., Son, M. H., Min, S.-J., Cho, Y. S., & Lee, J. K. (2011). 3-Hydroxy-2-[(2E)-1-(2-hydroxy-6-oxocyclohex-1-en-1-yl)-3-(2-methoxyphenyl)prop-2-en-1-

yl]cyclohex-2-en-1-one. *Acta Crystallographica Section E: Structure Reports Online*, *67*(Pt 10), o2739.

51. S. Raghavan, Z. Yang, R. T. Mosley, W. A. Schleif, L. Gabryelski, D. B. Olsen, M. Stahlhut, L. C. Kuo, E. A. Emini, K. T. Chapman, and J. R. Tata, Bioorg. Med. Chem. Lett, 12, 2855, 2002.

52. Xie, J., & Seto, C. T. (2007). A Two Stage Click-Based Library of Protein Tyrosine Phosphatase Inhibitors. *Bioorganic & medicinal chemistry*, *15*(1), 458-473.

53. Product Class 13: 1,2,3-Triazoles. (2004). In *Category 2, Hetarenes and Related Ring Systems* (2004e éd.). Thieme Verlag.

54. .American Cancer Society . www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html

55. Breast Cancer Network Australia. www.bcna.org.au

56. ER-Positive Breast Cancer: Prognosis, Life Expectancy, and More. (2018, septembre 6). https://www.healthline.com/health/breast-cancer/er-positive-prognosis-life-expectancy

57. WebMD Medical Reference Reviewed by Laura J. Martin, MD on November 11, 2017

58. Aromatase inhibitors in breast cancer The discovery of new compounds by computational designand biochemical evaluation. Marco André Coelho dasNeves. Faculdade de Farmácia. Universidade de Coimbra. 2008

59. Riggs, B. L., Khosla, S., & Melton, L. J. (2002). Sex steroids and the construction and conservation of the adult skeleton. *Endocrine Reviews*, *23*(3), 279-302.

60. Ghosh, D., Griswold, J., Erman, M., & Pangborn, W. (2009). Structural basis for androgen specificity and oestrogen synthesis in human aromatase. Nature, 457(7226), 219-223.

61. Goss, P. E., & Strasser, K. (2001). Aromatase inhibitors in the treatment and prevention of breast cancer. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *19*(3), 881-894.

62. Chakraborty, S., Ganti, A. K., Marr, A., & Batra, S. K. (2010). Lung cancer in women: role of estrogens. *Expert review of respiratory medicine*, *4*(4), 509-518.

63. Lønning, P. E. (1998). Pharmacological profiles of exemestane and formestane, steroidal aromatase inhibitors used for treatment of postmenopausal breast cancer. *Breast Cancer Research and Treatment*, *49 Suppl 1*, S45-52; discussion S73-77.

64. Choksi, P., Williams, M. E., Kidwel, K. M., Stella, J., Soyster, M., Hanauer, D., & Van Poznak, C. (2016). Adjuvant Aromatase Inhibitors in Early Breast Cancer May Not Increase the Risk of Falls. *Journal of Bone Reports & Recommendations*, *2*(2).

65. Howell, A., & Dowsett, M. (1997). Recent advances in endocrine therapy of breast cancer. *BMJ : British Medical Journal*, *315*(7112), 863-866.

66. Buzdar, A., & Howell, A. (2001). Advances in Aromatase Inhibition: Clinical Efficacy and Tolerability in the Treatment of Breast Cancer. *Clinical Cancer Research*, *7*(9), 2620-2635.

67. Goss P. E., Strasser K. (2001). Aromatase inhibitors in the treatment and prevention of breast cancer. *J. Clin. Oncol.*19(3)*:* 881-894.

68. Buzdar, A. U., Robertson, J. F. R., Eiermann, W., & Nabholtz, J.-M. (2002). An overview of the pharmacology and pharmacokinetics of the newer generation aromatase inhibitors anastrozole, letrozole, and exemestane. *Cancer*, *95*(9), 2006-2016.

CHAPTRE II

COMPUTATIONAL APPROACHES FOR DRUG DESIGN AND DISCOVERY

1. INTRODUCTION

Drug design is an integrated developing discipline which portends an era of tailored drug. It involves study of effects of biologically active compounds on the basis of molecular interactions in terms of molecular structure or its physicochemical properties involved. Its studies processes by which the drugs produce their effects, how they react with the protoplasm to elicit a particular pharmacological effect or response, how they are modified or detoxified, metabolized or eliminated by the organism. Currently two major molecular modeling strategies are employed in drug design process, ligand-based drug design and structure-based drug design. Drug discovery is constantly reevaluating itself in order to advance in speed, efficiency, and quality and thus remain successful. [1]

Studies of drug databases showed that successful drugs tend to have 'drug-like properties. Drug likeness, when viewed at the *in vivo* level, is thought of in terms of PK and safety. These complex i*n vivo* properties result from an interaction of physicochemical and structural properties, such as solubility, permeability and stability, which are studied *in vitro*. These properties are, in turn, dictated by fundamental molecular properties, such as molecular weight, hydrogen bonding and polarity, which are studied *in silico*. [2]

2. COMPUTER-AIDED DRUG DESIGN BASIC PRINCIPLES

CADD refers to the application of informatics to the discovery, design and optimization of biologically active compounds. Some of the most commonly used techniques will be described shortly.

2.1. Quantum Mechanics and Molecular Mechanics

There are two different approaches to compute the energy of a molecule. First, quantum mechanics, a procedure based on first principles. In this approach, nuclei are arranged in the space and the corresponding electrons are spread all over the system in a Continuous electronic density and computed by solving the Schrödinger equation. When chemical reactions do not need to be simulated, classical mechanics can describe the behavior of a Bimolecular system. This mathematical model is known as molecular mechanics, and can be used to compute the energy of systems containing a large number of atoms, such as molecules or complex systems of biochemical and biomedical interest.

In contrast to quantum mechanics, molecular mechanics ignore electrons and compute the energy of a system only as a function of the nuclear positions. Then, it is possible to take into account nan implicit way the electronic component of the system by adequate parameterization of the potential energy function. The set of equations and parameters which define the potential surface of a molecule is called force field. [3]

2.2. Force Fields

In molecular mechanics the electrons and nuclei of the atoms are not explicitly included in the calculations. Molecular mechanics considers a molecule to be a collection of masses interacting with each other through harmonic forces. Thus, the atoms in molecules are treated as ball of different sizes and flavors joined together by springs of variable strength and equilibrium distances (bonds). This simplification allows using molecular mechanics as a fast-computational model that can be applied to molecules of any size.

In the course of a calculation the total energy is minimized with respect to the atomic coordinates, and it consists of a sum of different contributions that compute the deviations from equilibrium of bond lengths, angles, torsions and non-bonded interactions:

$$E_{tot} = E_{str} + E_{bend} + E_{tors} + E_{vdw} + E_{elec} + ...$$

Where:

E_{tot} is the total energy of the molecule,

E_{str} is the bond-stretching energy term,

E_{bend} is the angle-bending energy term,

E_{tors} is the torsional energy term,

E_{vdw} is the van der Waals energy term,

E_{elec} is the electrostatic energy term.

The equilibrium values of bond lengths and bond angles are the corresponding force constants used in the potential energy function in the force field and it defines a set known as force field parameters.

Each deviation from these equilibrium values will result in increasing total energy of the molecule. So, the total energy is a measure of intramolecular strain relative to a hypothetical molecule with an ideal geometry of equilibrium. By itself the total energy has no strict physical meaning, but differences in total energy between two different conformations of the same molecule can be compared. [4,5]

2.3. Molecular dynamics

Molecular dynamics (MD) is a computational technique in which successive configurations of a system are generated by integrating the Newton's law of motion. The result is a trajectory that describes how the positions and velocities of particles in the system vary with time [6] MD simulations of biological macromolecules provide atomic detail on the internal motions of

these systems. Constant improvements in the methodology and computational power, extended the use of molecular dynamics studies to larger systems including, for example, explicit solvent and/or membrane environment, greater conformational changes and longer time scales. [7]

Drug design applications of molecular dynamics simulations include the estimation of free energies of binding, prediction of target selectivity, generation of multiple conformations for flexible docking, development of dynamic protein-based pharmacophores, metabolism prediction and refinement of protein homology models. [8]

2.4. Energy-Minimizing Procedures

Energy minimization methods can be divided into different classes depending on the order of the derivative used for locating a minimum on the energy surface. Zero order methods are those that only use the energy function to identify the regions of low energy through a grid search procedure. The most well-known method of this kind is the SIMPLEX method. Within first-derivative techniques, there are several procedures like the steepest descent method or the conjugate gradient method that make use of the gradient of the function. Second-derivative methods, like the Newton-Raphson algorithm make use of the hessian to locate minima. [9,10]

2.4.1. Steepest Descent Method

In the steepest descent method, the minimizer computes numerically the first derivative of the energy function to find a minimum. The energy is calculated for the initial geometry and then again after one of the atoms has been moved in a small increment in one of the directions of the coordinate system. This process is repeated for all atoms which finally are moved to a new position downhill on the energy surface. The procedure stops when a predetermined threshold condition is fulfilled. The optimization process is slow near the minimum, and consequently, the steepest descent method is often used for structures far from the minimum as a first, rough and introductory run followed by a subsequent minimization employing a more advanced algorithm like the conjugate gradient.

2.4.2. Conjugate Gradient Method

The conjugate gradient algorithm accumulates the information about the function from one iteration to the next. With this proceeding the reverse of the progress made in an earlier iteration can be avoided. For each minimization step the gradient is calculated and used as additional information for computing the new direction vector of the minimization procedure. Thus, each successive step refines the direction towards the minimum. The computational effort and the storage requirements are greater than for steepest descent, but conjugate gradients are the method of choice for larger systems. The greater total computational expense and the longer time per iteration is more than compensated by more efficient convergence to the minimum achieved by conjugate gradients. [11] As a summary, the choice of the minimization method depends on two factors: the size of the system and the current state of the optimization. For structures far from minimum, as a general rule, the steepest descent method is often the best minimizer to use for 100-1000 iterations. The minimization can be completed to convergence with conjugate gradients. There are several ways in molecular minimization to define convergence criteria. In non-gradient minimizers only, the increments in the energy and the coordinates can be taken to judge the quality of the actual geometry of the molecular system. In all gradient minimizers, however, atomic gradients are used for this purpose. The best procedure in this respect is to calculate the root mean square gradients of the forces on each atom of a molecule. The value chosen as a maximum derivative will depend on the objective of the minimization. If a simple relaxation of a strained molecule is desired, rough convergence criterions like a maximum derivative

of 0.1 kcal mol-1Å-1 is sufficient while for other cases convergence to a maximum derivative less than 0.001 kcal mol-1Å-1 is required to find a final minimum. [12]

3. DRUG-LIKE PROPERTIES

Since 2001, when the "property-based design" concept was presented by van deWaterbeemd et al., [13] there has been an increased focus on its application in drug design and discovery [14–15].

The relationship between chemical structure and physicochemical properties has attracted the attention of medicinal chemists as a new drug research and development strategy that complements the structure-activity relationships in the progress of drug design and discovery.

The chemical structure of a drug influences its physicochemical properties, and the physicochemical properties of a drug molecule [16], such as MW, lipophilicity, aqueous solubility (S), permeability, acid-base ionization constant (pKa), HBD and HBA, ROT and PSA, can be changed by modifying its structure. Further, these physicochemical properties of drug molecule influence its ADME/Tox properties (drug-like properties), such as metabolic stability, plasma stability, P-glycoprotein (Pgp) extrusion, serum albumin binding, cytochrome P450 (CyP450) inhibition, human Ether-à-go-go-Related Gene (hERG) inhibition, the ability to across the blood brain barrier (BBB), pharmacokinetics (PK) and toxicity. Ultimately, these physicochemical properties and the ADME/Tox properties of a drug molecule affect its pharmacodynamics activity. [17]

Table II-1 Drug-like properties.

Structural properties	Physicochemical properties	Biochemical properties	Pharmacokinetics (PK) and toxicity
– Hydrogen bonding	– Solubility	– Metabolism (phases I and II)	– Clearance
– Polar surface area	– Permeability		– Half-life
– Lipophilicity	– Chemical stability	– Protein and tissue binding	– Bioavailability
– Shape			– Drug–drug interaction
– Molecular weight		– Transport (uptake, efflux)	
– Reactivity			– LD50
– pKa			

4. RULES FOR DRUG DISCOVERY

The fastest method for evaluating the drug-like properties of a compound is to apply "rules."

Rules are a set of guidelines for the structural properties of compounds that have a higher probability of being well absorbed after oral dosing. The values for the properties associated with rules are quickly counted from examination of the structure or calculated using software that is widely available.

4.1 Lipinski Rules (Oral drug properties)

Lipinski's rule of five also known as the Pfizer's rule of five or simply the Rule of five (RO5) is a rule of thumb to evaluate drug likeness or determine if a chemical compound with a certain pharmacological or biological activity has properties that would make it a likely orally active drug in humans. The rule was formulated by Christopher A. Lipinski in 1997. [18]

The impact of these rules in the field has been very high. This acceptance can be attributed to many factors: [17]

- The rules are easy, fast, and have no cost to use.
- The "5" mnemonic makes the rules easy to remember.
- The rules are intuitively evident to medicinal chemists.
- The rules are a widely used standard benchmark.
- The rules are based on solid research, documentation, and rationale.
- The rules work effectively.

Lipinski's rule states that, in general, an orally active drug has no more than one violation of the following criteria:

1. Not more than 5 hydrogen bond donors (nitrogen or oxygen atoms with one or more hydrogen atoms).
2. Not more than 10 hydrogen bond acceptors (nitrogen or oxygen atoms).
3. A molecular mass less than 500 daltons.
4. An octanol-water partition coefficient log *P* not greater than 5.

Note that all numbers are multiples of five, which is the origin of the rule's name. [19]

4.2 Veber Rules

Additional rules were proposed by Veber et al. They studied structural properties that increase oral bioavailability in rats. They concluded that molecular flexibility, polar surface area (PSA), and hydrogen bond count are important determinants of oral bioavailability.

Rotatable bonds can be counted manually or using software. PSA is calculated using software and is closely related to hydrogen bonding. [20]

Veber rules for good oral bioavailability in rats are as follows:

- ≤10 rotatable bonds
- ≤140 Å2 PSA, or≤12 total hydrogen bonds (acceptors plus donors)

5. STRATEGIES OF IN SILICO DESIGN

In silico drug design can be applied by either of two strategies of design depending on the knowledge of the target, presence of the primary sequence and 3D structure. These two strategies are:

5.1 Ligand-based drug design

In the absence of the receptor 3D information, lead identification and optimization depend on available pharmacologically relevant agents and their bioactivities. The computational approaches include QSAR, pharmacophore modeling, and database mining, [21] In this work, we will use QSAR.

5.1.1 Quantitative structure-activity Relationship

Is the process by which chemical structure is quantitatively correlated with a well-defined process, such as biological activity or chemical reactivity. For example, biological activity can be expressed quantitatively as in the concentration of a substance required to give a certain biological response. Additionally, when physicochemical properties or structures are expressed by numbers, one can form a mathematical relationship, or quantitative structure-activity relationship, between the two.

The mathematical expression can then be used to predict the biological response of other chemical structures. QSAR's most general mathematical form is

Activity = *f* (physiochemical properties and/or structural properties)

That the three key components required for the development of a QSAR model are:

- Some measure of the activity for a group of chemicals in a biological or environmental system – toxicological endpoint.
- A description of the physicochemical properties and/or structure for this group of chemicals.
- Molecular descriptors.
- A form of statistical relationship to link activity and descriptors [22]

5.1.2 The data to model

The modeler, and user of a model, must consider the data to model. Data should, ideally, be of high quality, meaning they are reliable and consistent across the data set to be modelled. The definition of data quality is, at best, subjective and is likely to be different for any effect, endpoint or property.

Therefore, the modeler or user should determine whether the data are performed in a standard manner, to a recognized protocol, and if they are taken from single or multiple laboratories. [23]

5.1.3 Molecular Descriptors

Despite great advances in the field of drug design, the use of descriptors to define the molecular structure of biologically active compounds is the main method utilized to discover new lead molecules. Descriptors are the chemical characteristic of a molecule in numerical form, used for QSAR/QSPR studies. Figure II.1 depicts the basic definition of these descriptors. [24]

Figure II.1 *Molecular Descriptors*

5.1.4 Statistical methods used in QSAR analysis

Statistical methods are an essential component of QSAR work. They help to build models, estimate a model's predictive abilities, and find relationships and correlations among variables and activities. A suitable statistical method coupled with a variable selection method allows analysis of this data in order to establish a QSAR model with the subset of descriptors that are most statistically significant in determining the biological activity. The statistical method can be broadly divided in to two: linear and non-linear method. In statistics a correlation is established between dependent variables (biological activity) and independent variables (physiochemical properties or molecular descriptor). The liner method fits a line between the selected descriptor and activity as compared to non-linear method which fit a curved between the selected descriptor and activity. The statistical method to build QSAR model is decided based on the type of biological activity data. Following are commonly used statistical methods. [25]:

1. Principal component analysis (PCA)	9. Genetic function approximation (GFA)
2. Cluster analysis	10. Genetic partial least squares (GPLS)
3. Simple liner regression	11. Logistic regression
4. Multiple liner regression	12. K-Nearest Neighbor classification
5. Stepwise multiple liner regression	(KNN)
6. Principle component regression (PCR)	13. Neural Network
7. Continuum Regression	14. Discriminant analysis
8. Partial least squares (PLS)	15. Decision Trees
	16. SIMCA

In our study we used the MLR method.

- **Multiple Linear Regression** (MLR)

can be considered as an easy interpretable regression-based method, regression analysis correlates independent X variables or descriptors with dependent Y variables (biological data). The regression model assumes a linear relationship between *m* molecular descriptors and the response (biological activity) variable.

This relationship can be expressed with the single multiple-term linear equation:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + bmXm + e$$

The MLR analysis calculates the regression coefficients, bi, by minimizing the residuals, which quantify the deviations between the data (Y) and the model (Y'), as in the case of simple linear regression. [26]

5.1.5 Validation of the QSAR model

QSAR studies aimed to derive a model that is optimally active, means the model should provide a reliable estimate of the activity of new or untested compound similar to those in it. After development of the model it is necessary to test whether any data from the data set affect model extensively. This is done by using QSAR validation methods. Validation method is required to ensure model reliability, to ensure that the model is not due to chance factor. Various types of validation methods are used in the QSAR studies to predict

the accuracy of model or to estimate the validity or predictivity of the derived structure property model. [25]

- **Internal validation parameters**

Internal validation uses the dataset from which the model is built and checks for internal stability. Cross-Validation (CV) technique is widely employed as an internal validation method of statistical models (23-26). Usually, one compound of the set is extracted each time, and then the model is recalculated using as training set the n-1 (where n is number of compounds) remaining compounds, so that the biological activity value for the extracted compound is predicted once for all compounds. This process is repeated n times for all the compounds of the initial set, thus obtaining a prediction for each object. This process referred as leave-one-out (LOO) method. [25]

This validation is done using the data that created the model. The various internal validation parameters invoked in this study are presented thus;

Table 5-2 Statistical parameters for cross-validation

	Statistic Definition	Formula
R^2_{cv}	the square of the correlation coefficient [27]	$$R^2_{CV} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)} = 1 - \frac{PRESS}{SSY}$$ Yi the experimental property $\hat{y}_i$; the predicted property $\bar{y}_i$ the mean experimental
R^2_{adj}	the square of the correlation coefficient Adjusted [18]	p = number of independent variables in the model $$R^2_{adj} = 1 - (1 - R^2)\frac{n-1}{n-p-1} = \frac{(n-1)R^2 - P}{n-p+1}$$
F	Variance Ratio [29]	$$F = \frac{\frac{\sum(\hat{y}_i - \bar{y}_i)^2}{p}}{\frac{\sum(y_i - \hat{y}_i)^2}{n-p-1}}$$
S	Standard error of estimate [29]	$$S = \frac{\sum(y_i - \hat{y}_i)^2}{n-p-1}$$
SSY	Total Sum of Squares (TSS) [30]	$$ssy = \sum(y_i - \bar{y}_i)$$
(PRESS)	Predictive Error Sum of Squares [30]	$$press = \sum(y_i - \hat{y}_i)^2$$

- **External validation**

Several authors have suggested that the only way to estimate the true predictive power of a QSAR model is to compare the predicted and observed activities of an (sufficiently large) external test set of compounds that were not used in the model development.

To estimate the predictive power of a QSAR model, Golbraikh and Tropsha recommended use of the following statistical characteristics of the test set [31,32]:

Table II.3 Statistical parameters for external validation

Statistic	Definition	Formula
$R^2_{pred\ EXT}$	**Predictive residual sum of squares(external validation) [33]**	$R^2_{pred} = 1 - \dfrac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)}$
Golbraikh and Tropsha's criteria		
$R^{\circ 2}$	**The squared correlation coefficient predicted versus observed [20]**	$R^{\circ 2} = 1 - \dfrac{\sum_{i=1}^{ntest}\left(\hat{y}_i - y_i^{r\circ}\right)^2}{\sum_{i=1}^{ntest}\left(\hat{y}_i - \bar{\hat{y}}\right)^2}, y_i^{r\circ} = K\hat{y}_i$
$R'^{\circ 2}$	**The squared correlation coefficient versus predicted activities**	$R'^{\circ 2} = 1 - \dfrac{\sum_{i=1}^{ntest}\left(y_i - \hat{y}_i^{r\circ}\right)^2}{\sum_{i=1}^{ntest}(y_i - \bar{y})^2}, \hat{y}_i^{r\circ} = K'y_i$
K	**The slopes of regression**	$K = \dfrac{\sum_{i=1}^{ntest} y_i\ \hat{y}_i}{\sum_{i=1}^{ntest} \hat{y}_i^2}$
K'	**The slopes of regression**	$K' = \dfrac{\sum_{i=1}^{ntest} y_i\ \hat{y}_i}{\sum_{i=1}^{ntest} y_i^2}$
Matrice for external validation		
r_m^2	**Closeness between the R² and R²₀ determination coefficients**	$r_m^2 = R^2\left(1 - \sqrt{|R^2 - R^{\circ 2}|}\right)$
$r_m'^2$		$r_m'^2 = R^2\left(1 - \sqrt{|R^2 - R'^{\circ 2}|}\right)$
$\overline{r_{m(test)}^2}$		$\overline{r_{m(test)}^2} = \dfrac{r_m^2 + r_m'^2}{2}$
Concordance correlation coefficient		
CCC	**Concordance correlation coefficient [34]**	$ccc = \dfrac{2\sum_{i=1}^{ntest}(y_i - \bar{y})\ (\hat{y}_i - \bar{\hat{y}})}{\sum_i^{n\ test}(y_i - \bar{y})^2 + \sum_i^{n\ test}(\hat{y}_i - \bar{\hat{y}})^2 + n_{test}(\bar{y} - }$

5.1.6 Evaluation of the model

A developed QSAR model can be accepted generally in QSAR studies when it can satisfy the following criterion :

- If correlation coefficient $R \geq 0.8$ (for *in vivo* data).
- If coefficient of determination $R^2 \geq 0.6$
- If the standard deviation s is not much larger than standard deviation of the biological data.
- If its F value indicate that overall significance level is better than 95%.
- If its confidence interval of all individual regression coefficients proves that they are justified at the 95% significance level.
- If cross-validated R^2 (Q^2) > 0.5
- If R^2 for external test set, R^2 pred > 0.6
- $(r^2 - r^2_0)/r^2 < 0.1$ and $0.85 \leq K \leq 1.15$, or $(r^2 - r'^2_0)/r^2 < 0.1$ and $0.85 \leq K' \leq 1.15$ (for test set).
- r^2_m (overall) and R^2p are ≥ 0.5 (or at least near 0.5).
- External validation parameters for a predictive model, have to satisfy a threshold value which is 0.85 for CCC. [35]

Equation has to be rejected

- If the above-mentioned statistical measures are not satisfied.
- If the number of the variables in the regression equation is unreasonably large.
- If standard deviation is smaller than error in the biological data. [36]

5.2 STRUCTURE-BASED DRUG DESIGN STRATEGIES SBDD

Structure-based design has played an important role in drug discovery and development. This approach requires the understanding of receptor–ligand interactions. If the target 3D structure is known, it can be used for the design of new ligands.

Herein we will focus on the most commonly used strategies: molecular docking. [37]

5.2.1 Definition of Molecular docking

Molecular docking studies are used to determine the interaction of two molecules and to find the best orientation of ligand which would form a complex with overall minimum energy. The small molecule, known as ligand usually fits within protein's cavity which is predicted by the search algorithm. This protein cavity becomes active when comes in contact with any external compounds and are thus called as active sites.

The results are analyzed by a statistical scoring function which converts interacting energy into numerical values called as the docking score; and also, the interacting energy is calculated. The 3D pose of the bound ligand can be visualized using different visualizing tools like Pymol, Rasmol etc which could help in inference of the best fit of ligand. Predicting the mode of protein-ligand interaction can assume the active site of the protein molecule and further help in protein annotation. [36]
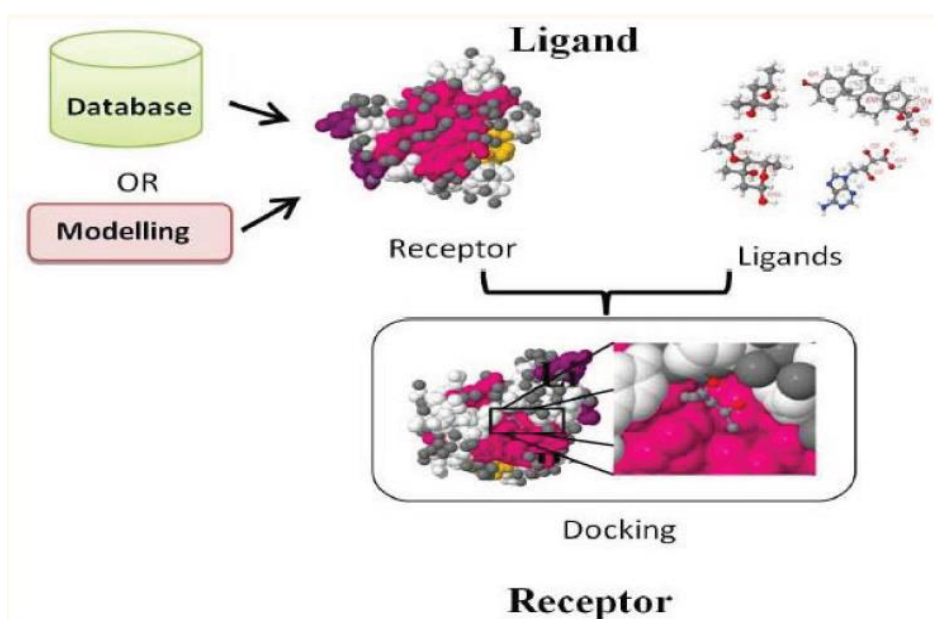
Figure II.2 Molecular docking flow chart

5.2.2 Different types of interactions

Interactions between particles can be defined as a consequence of forces between the molecules contained by the particles. These forces are divided into four categories:

- **Electrostatic forces** - Forces with electrostatic origin due to the charges residing in the matter. The most common interactions are charge-charge, charge-dipole and dipole-dipole.

- **Electrodynamics forces**-The most widely known is the Van der Waals interactions.

- **Steric forces** - Steric forces are generated when atoms in different molecules come into very close contact with one another and start affecting the reactivity of each other. The resulting forces can affect chemical reactions and the free energy of a system.

- **Solvent-related forces** - These are forces generated due to chemical reactions between the solvent and the protein or ligand. Examples are Hydrogen bonds (hydrophilic interactions) and hydrophobic interactions.
  A common characteristic of all these forces is their electromagnetic nature. Other physical factors - Conformational changes in the protein and the ligand are often necessary for successful docking. [37]

5.2.3 Types of docking

The following are majorly used type of docking are:

- **Lock and Key or Rigid Docking**: In rigid docking, both the internal geometry of the receptor and ligand is kept fixed during docking.

- **Induced fit or Flexible Docking**: In this model, the ligand is kept flexible and the energy for different conformations of the ligand fitting into the protein is calculated. Though more time consuming, this method can evaluate many different possible conformations which make it more reliable.[38]

5.2.4 Mechanics of docking

To perform a docking screen, the first requirement is a structure of the protein of interest usually the structure has been determined using a biophysical technique such as x-ray crystallography, or less often, NMR spectroscopy. This protein structure and a database of potential ligands serve as inputs to a docking program. The success of a docking program depends on two components: the search algorithm and the scoring function. [39]

Search algorithm: These algorithms determine all possible optimal conformations for a given complex (protein-protein, protein-ligand) in an environment i.e. the position and orientation of both molecules relative to each other. They can also calculate the energy of the resulting complex and of each individual interaction. [40]

The different types of algorithms that can be used for docking analysis are:

- Molecular dynamics
- Monte Carlo methods
- Genetic algorithms
- Fragment-based methods

- Point complementary methods
- Distance geometry methods
- Systematic searches

Scoring function: These are mathematical methods used to predict the strength of the noncovalent interaction called as binding affinity, between two molecules after they have been docked. Scoring functions have also been developed to predict the strength of other types of intermolecular interactions, for example between two proteins or between protein and DNA or protein and drug. These configurations are evaluated using scoring functions to distinguish the experimental binding modes from all other modes explored through the searching algorithm. [41]

For example:

• Empirical scoring function of Igemdock $Fitness = vdW + H_{bond} + Elec$

• Binding Energy $\Delta G_{bind} = \Delta G_{vdw} + \Delta G_{hbond} + \Delta G_{elect} + \Delta G_{conform} + \Delta G_{tor} + \Delta G_{sol}$

5.2.5 Major steps involved in mechanics of molecular docking

Molecular Docking is the process in which the intermolecular interaction between two molecules was studied in In-silico. In this process, the Macromolecule is the protein receptor. The micro molecule is the Ligand molecule which can be acted as an inhibitor. So, the Docking process involves the following steps:

Step I – preparation of protein:

Three-dimensional structure of the Protein should be retrieved from Protein data bank (PDB); afterward the retrieved structure should be pre-processed. This should admit removal of the water molecules from the cavity, stabilizing the charges, filling the missing residues, generation the side chains etc. according to the parameters available.

Step II – active site prediction:

After the preparation of protein, the active site of protein should be predicted. The receptor might possess lots of active sites merely the one of the concerns should be picked out. Mostly the water molecules and hetero atoms are removed if present.

Step III – preparation of ligand:

Ligands can be retrieved from several databases such as ZINC, Pub Chem or can be sketched applying Chem sketch tool.

Step IV- docking:

Ligand is docked against the protein and the interactions are analyzed. The scoring function gives score on the basis of best docked ligand complex is picked out. [42]

6. REFERENCE

1. Ekins, S., Mestres, J., & Testa, B. (2007). In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *British Journal of Pharmacology*, *152*(1), 9-20.

2. Smith, D. A., & Waterbeemd, H. van de. (1999). Pharmacokinetics and metabolism in early drug discovery. *Current Opinion in Chemical Biology*, *3*(4), 373-378.

3. Hagler, A. T., Lifson, S., & Dauber, P. (1979). Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 2. A benchmark for the objective comparison of alternative force fields. *Journal of the American Chemical Society*, *101*(18), 5122-5130.

4. Maple, J. R., Hwang, M.-J., Stockfisch, T. P., Dinur, U., Waldman, M., Ewig, C. S., & Hagler, A. T. (1994). Derivation of class II force fields. I. Methodology and quantum force field for the alkyl functional group and alkane molecules. *Journal of Computational Chemistry*, *15*(2), 162-182.

5. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., … Weiner, P. (1984). A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*, *106*(3), 765-784.

6. Hansson, T.; Oostenbrink, C.; van Gunsteren, W. F. (2002) Molecular dynamics simulations. *Curr. Opin. Struct. Biol.*, *12*, 190-196

7. Huo, S. H.; Wang, J. M.; Cieplak, P.; Kollman, P. A.; Kuntz, I. D. (**2002).** Molecular dynamics and free energy analyses of cathepsin D-inhibitor interactions: Insight into structure-based ligand design. *J. Med. Chem.*, *45*, 1412- 1419.

8. Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biology*, *9*, 71.

9. Ponder JW, Richards FM. (1987). An Efficient Newton-like Method for Molecular Mechanics Energy Minimization of Large Molecules. J. Comput. Chem.;8:10161024.

10. Schlick T, Overton MA. (1987). Powerful truncated Newton method for potential energy minimization. J. Comput. Chem.;8:1025-1039.

11. Fletcher R. (1980). Practical Methods of Optimization, Unconstrained Optimization. Wiley. New York, 1:102-118.

12. Fletcher R, Reeves CM.(1964).Function Minimization by Conjugate Gradients. Comput. J.;7:149-154

13. Bodor, N., & Buchwald, P. (2012). *Retrometabolic Drug Design and Targeting*. John Wiley & Sons.

14. Young, R. J., Campbell, M., Borthwick, A. D., Brown, D., Burns-Kurtis, C. L., Chan, C., … Zhou, P. (2006). Structure- and property-based design of factor Xa inhibitors: pyrrolidin-2-ones with acyclic alanyl amides as P4 motifs. *Bioorganic & Medicinal Chemistry Letters*, *16*(23), 5953-5957

15. Bemis, G. W., & Murcko, M. A. (1996). The properties of known drugs. 1. Molecular frameworks. *Journal of Medicinal Chemistry*, *39*(15), 2887-2893.

16. Mao, F., Ni, W., Xu, X., Wang, H., Wang, J., Ji, M., & Li, J. (2016). Chemical Structure-Related Drug-Like Criteria of Global Approved Drugs. *Molecules*, *21*(1), 75.

17. Kerns, E.H.; Di, L. Drug-Like Properties: Concepts, Structure Design and Methods: From ADME to Toxicity Optimization; Elsevier Inc.: Amsterdam, The Netherlands, 2008.

18. Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, *46*(1-3), 3-26.

19. Lipinski, C. A. (2004). Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today. Technologies*, *1*(4), 337-341.

20. Veber, D. F., Johnson, S. R., Cheng, H.-Y., Smith, B. R., Ward, K. W., & Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, *45*(12), 2615-2623.

21. Seetharama D. Satyanarayanajois (ed.), Drug Design and Discovery: Methods and Protocols, Methods in Molecular Biology, vol. 716

22. Ramírez, D. (2016). Computational Methods Applied to Rational Drug Design. *The Open Medicinal Chemistry Journal*, *10*, 7-20.

23. Recent Applications of Quantitative Structure-Activity Relationships in Drug Design Omar Deeb *Al-Quds University, Faculty of Pharmacy, Jerusalem Palestin.*

24. Naveen K. Mahobia. Roshan D. Patel,Naheed W. Sheikh, Sudarshan K. Singh1,Achal Mishra, Ravindra Dhardubey. (2010), Validation Method Used In Quantitative Structure Activity Relationship. Der Pharma Chemica, 2(5):260-271

25. Shao, J. (1993). Linear Model Selection by Cross-validation. *Journal of the American Statistical Association*, *88*(422), 486-494.

26. W. Wu; C. Zhang; W. Lin; Q. Chen; X. Guo; Y. Qian. PLoSONE, 2015, 10(3).

27. K. Brandon; O. Aline. Comprehensive R archive network (CRAN): http:// CRAN.R-project.org. Retrieved July 3 rd , 2015

28. K. Roy, K. Springer Briefs in Molecular Science. 2015

29. Golbraikh A, Tropsha A (2002) Beware of q2. J Mol Graph Model 20:269–276

30. Roy PP, Somnath P, IndraniMet al (2009) On two novel parameters for validation of predictive QSAR models. Molecules 14:1660–1701

31. Schuurmann G, Ebert RU, Chen J et al (2008) External validation and prediction employing the predictive squared correlation coefficients test set activity mean vs training set activity mean. J Chem Inf Model 48:2140–2145

32. Nicola Chirico N, Gramatica P (2011) .*J Chem Inf Model* 51(9):2320–2335

33. Abdulfatai, U., Uzairu, A., & Uba, S. (2016). Quantitative structure activity relationship study of anticonvulsant activity of α_substituted acetamido-N-benzylacetamide derivatives. *Cogent Chemistry*, *2*(1), 1166538.

34. Rarey, M., Wefing, S., & Lengauer, T. (1996). Placement of medium-sized molecular fragments into active sites of proteins. *Journal of Computer-Aided Molecular Design*, *10*(1), 41-54.

35. Sushma, B., Ch, V, Suresh., *(2012).* DOCKING-A Review. *Journal of Applicable Chemistry, 1 (2):167-173*

36. Rangaraju A. & Rao A. V., (2013) A REVIEW ON MOLECULAR DOCKING – Novel tool in drug design and analysis, Jour. Harmo. Res. Pharm., 2(4), 215-221.

37. Avinash Rangaraju, Dr.A.Veerabhadra Rao2 . (2012), Indole: the molecule of diverse pharmacological activities, Int. J. Drug Res. Tech., 2 (2), 189-197.

38. Manju, ..., R. K, ..., Gunjan, ..., & Anju, ... (2017). Review on introduction to molecular docking software technique in medicinal chemistry. *International Journal Of Drug Research And Technology, 2*(2), 8.

39. McMartin, C., & Bohacek, R. S. (1997). QXP: powerful, rapid computer algorithms for structure-based drug design. *Journal of Computer-Aided Molecular Design*, *11*(4), 333-344.

40. Schnecke, V., & Kuhn, L. A. (2000). Virtual screening with solvation and ligand-induced complementarity. *Perspectives in Drug Discovery and Design*, *20*(1), 171-190.

41. Mills, C. L., Beuning, P. J., & Ondrechen, M. J. (2015). Biochemical functional predictions for protein structures of unknown or uncertain function. *Computational and Structural Biotechnology Journal*, *13*, 182-191.

CHAPTRE III

DRUG LIKENESS SCORING AND STRUCTURE ACTIVITY/PROPERTY RELATIONSHIPS

1. INTRODUCTION

The 1,2,3- triazole can be divided in three subclasses depending on the position of the substituants in the ring. While 1H-1,2,3-triazole and 2H-1,2,3-triazole are aromatic compounds but their 4H-1,2,3-triazole isomers are not. This fact reflects the abundance of 1H- and 2H-1,2,3-triazoles and the scarcity of 4H-1,2,3-triazoles [1].

The two tautomeric forms: 1H- and 2H-1,2,3-triazoles derivatives are in equilibrium, both in solution and gas phase (Figure.III.1) [2].

2H-1,2,3-triazole   1H-1,2,3-triazole
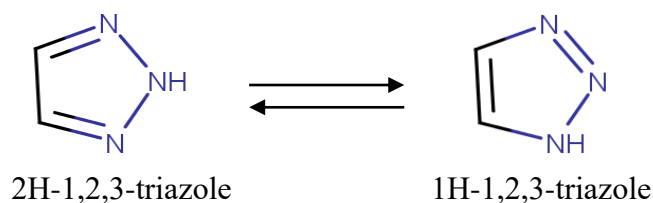
Figure.III.1 1,2,3-triazole tautomeric forms 2H-1,2,3-triazole and 1H-1,2,3-triazole (MarvinSketch 15.8.31)

It is well known that compounds having an aza-hetero ring, such as triazole show inhibitory activity against aromatase [3].

QSAR methods are attempts to correlate the molecular structure or properties derived from the molecular structure with a particular type of chemical, biochemical or biological activity. [4-6]

The present work reports ab initio and density functional results of molecular properties of 1H-1,2,3-triazole and 2H-1,2,3-triazole. Additionally, we applied DFT methods on some of 2H-1,2,3-triazole. We studied, at last, some of QSAR and drug likeness proprieties of 1,2,3 -triazole derivatives series reported in literature. On the other hand, calculated metrics aim to combine the potency with other parameters into a single metric which may be monitored during optimization. The earliest and most commonly applied metrics are the Ligand Efficiency (LE) and the Lipophilic Efficiency (LipE). Web based software was used to obtain parameter such as TPSA, nrotb and drug likeness.

2. MATERIAL AND METHODS

All calculations were performed using HyperChem 8.0.6 software [7] Gaussian 09 program package [8]; MarvinSketch 15.8.31 software [9], and Molinspiration online database, Molinspiration Cheminformatics [10].

The geometries of 2H-1,2,3-triazole and 1H-1,2,3-triazole, were fully optimized with ab initio/HF (6-31G$^+$, 6-31G^{++}(d,p), 6-311G^{++}(d,p))and DFT/B3LYP(6-311^{++}G (d, p)), integrated in Gaussian 09 program package. The calculation of QSAR properties is performed through the module QSAR Properties (HyperChem version 8.0.6), and allows the calculation of several properties commonly used in QSAR studies.

Molinspiration, and web-based software was used to obtain parameter such as TPSA (topological polar surface area), nrotb (number of rotatable bonds) and drug likeness. The calculated results are reported in the present work.

3. RESULTS AND DISCUSSION

3.1 Geometric and electronic structure of 1 H-1,2,3-triazole

The optimized geometrical parameters of 1H-1,2,3- triazole (Figure.III. 2) are obtained using ab-initio/*HF* and *DFT* methods, listed in (Table.III.1) and (Table III.2) with experimental results approximately similar to the theoretical results, regarding bond length and valence angle values [11].
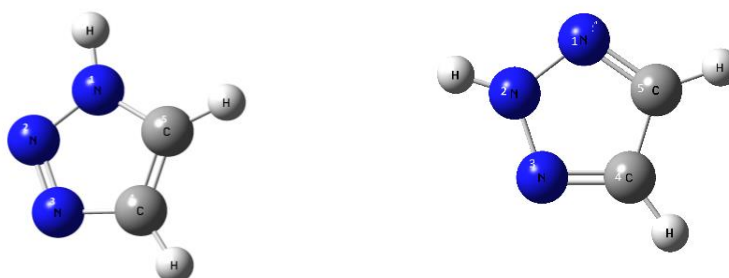
Figure.III.2 3D conformation of 1H and 2 H-1,2,3-triazole (Gauss View 3.0.9)

Table III.1 Bond lengths (angstrom) of 1H-1,2,3-triazole

Parameters	EXP[7]	Ab-initio/HF			DFT/B3LYP
		6-31G+	**6-31G++(d,p)**	**6-311G++(d,p)**	**6-311G++ (d,p)**
N1-N2	1.355	1.346	1.448	1.315	1.349
N2-N3	1.309	1.282	1.251	1.262	1.298
N3-C4	1.370	1.372	1.402	1.356	1.364
C4-C5	1.378	1.361	1.340	1.355	1.373
C5-N1	1.356	1.352	1.317	1.343	1.355

The theoretical dihedral angle values calculated by different methods are practically equal to zero degree explaining that the geometry of 1H-1,2,3- triazole is planar, and hence makes this conformation more stable. The charge densities calculated (Table III.3) by these methods are slightly different.

Table III.2. Calculated values, valence angles and dihedral angles of 1H-1,2,3-triazole

Parameters	EXP[19]	Ab initio/HF			DFT/B3LYP
		6-31G+	**6-31G++(d,p)**	**6-311G++(d,p)**	**6-311 G++ (d,p)**
N1-N2-N3	108.2	107.355	107.940	108.119	107.012
N2-N3-C4	108.2	109.529	108.636	109.270	108.692
N3-C4-C5	109.9	107.959	108.499	107.961	108.699
C4-C5-N1	104.4	104.328	107.654	103.521	103.560
C5-N1-N2	110.2	104.328	107.268	111.120	111.450
N1-N2-N3-C4	-	0	0	0	0
N2-N3-C4-C5	-	0	0	0	0
N3-C4-C5-N1	-	0	0	0	0
C4-C5-N1-N2	-	0	0	0	0
C5-N1-N2-N3	-	0	0	0	0

Table III.3. Net charge distribution for 1H-1,2,3-triazole.

Atoms	Ab initio/HF	DFT/B3LYP
	6-311G++(d .p)	**6-311++G(d,p)**
N1	-0.145	-0.051
N2	-0.025	-0.044
N3	-0.158	-0.158
C4	-0.220	-0.160
C5	-0.205	-0.210

3.2 Geometric and electronic structure of 2 H-1,2,3-triazole

The optimized geometrical parameters of 2 H-1,2,3-triazole by ab-initio/HF and DFT method listed in Table III.4 and Table III.5 are in accordance with numbering scheme given in Figure III.2.

Table III.4. Bond lengths (angstrom) of 2H-1,2,3-triazole

Parameters	EXP[19]	Ab initio/HF			DFT/B3LYP
		6-31G+	**6-31G++(d,p)**	**6-311G++(d,p)**	**6-311G++(d,p)**
N1-N2	1.323	1.326	1.303	1.301	1.326
N2-N3	1.323	1.326	1.303	1.301	1.326
N3-C4	1.347	1.319	1.308	1.306	1.333
C4-C5	1.401	1.411	1.404	1.405	1.406
C5-N1	1.347	1.319	1.308	1.306	1.333

Table III.5. Calculated values of valence angles and dihedral angles of 2H-1,2,3-triazole

| Parameters | EXP[19] | Ab initio/HF | | | DFT/B3LYP |
		6-31G$^+$	6-31G^{++}(d,p)	6-311G^{++}(d,p)	6-311^{++}G(d,p)
H5-C5-C4	130.8	121.897	121.742	130.349	130.293
H4-C4-C5	130.8	121.897	121.742	130.349	130.293
H2-N2-N3	-	122.698	122.001	122.019	121.094
N1-N2-N3	117.3	114.604	115.997	115.951	116.55
N2-N3-C4	-	104.569	104.084	104.152	103.111
N3-C4-C5	-	108.129	107.918	108.612	108.612
C4-C5-N1	-	108.129	107.918	108.612	108.612
C5-N1-N2	-	104.569	104.084	104.152	103.111
N1-N2-N3-C4	-	-0.003	-0.001	0.001	0.008
N2-N3-C4-C5	-	-0.005	0.0000	0.002	-0.022
N3-C4-C5-N1	-	0.012	0.002	0.002	0.030
C4-C5-N1-N2	-	0.001	0.000	-0.001	0.022
C5-N1-N2-N3	-	1.319	0.001	0.001	0.000

The efficiency of theoretical methods may be assessed by comparison with experimental results [12]. From these results a good correlation can be seen between the ab-initio/HF, and DFT for bond length. We can also note that charge densities calculated by ab-initio/HF are approximately similar to those calculated by the DFT method (Table III.6).

Table III.6. Net charge distribution for 2H-1,2,3-triazole.

| Atoms | *Ab initio/HF* | *DFT/B3LYP* |
	6-311G^{++}(d .p)	6-311^{++}G(d,p)
N1	-0.129	-0.128
N2	-0.063	-0.044
N3	-0.129	-0.128
C4	-0.188	-0.167
C5	-0.188	-0.167

The theoretical dihedral angle values calculated are practically equal to zero degree revealing that the geometry of 2H-1,2,3-triazole is planar, and hence makes this conformation more stable.

The efficiency of DFT/B3LYP method with 6-311++G(d,p) basis set may be scrutinized by comparison with the results obtained by ab-initio/HF method. A very good agreement between predicted geometries (bond lengths and bond angles) and corresponding experimental data, especially the DFT/B3LYP results can be observed. From that, we can say that the DFT method is more appropriate for further study on 1,2,3-triazole derivatives in other parts of this work.

To compare the stability of the two compounds 1H-1,2,3-triazoleand 2H-1,2,3-triazole, we calculated the HOMO (highest occupied molecular orbital), and the LUMO (lowest unoccupied molecular orbital) and their difference (ΔE) calculated by DFT/6-311G (d,p) by (Gaussian09) and heat of formation, calculated by (HyperChem 8.06) the results are reported in Table III.7.

Table III.7. *Calculated E_{HOMO}, E_{LUMO}, energy band gap (ΔE) and Heat of formation ΔH_f*

Nucleus	HOMO [au]	LUMO [au]	ΔE [au]	ΔH_f [kcal.mol^{-1}]
1H-1.2.3 triazole	-0.279	-0.023	0.256	67.813
2H-1.2.3 triazole	-0.289	-0.026	0.263	70.610

The calculations show the 2H tautomer as the more stable with a low reactivity. The difference of ΔE_{1H-2H} was -2.1kcal·mol^{-1}. These results agreed the results obtained by Tornkvist et al [13].

3.3 Molecular electrostatic potential

The molecular electrostatic potential is a well-established tool to explain the reactive behavior of a wide variety of chemical systems in both electrophilic and nucleophilic reactions. The study of biological recognition processes and hydrogen bonding interactions was performed to predict the reactive sites for electrophilic and nucleophilic attack for the investigated molecule [14]. The MEP was calculated at the DFT optimized geometry (Figure III.3).
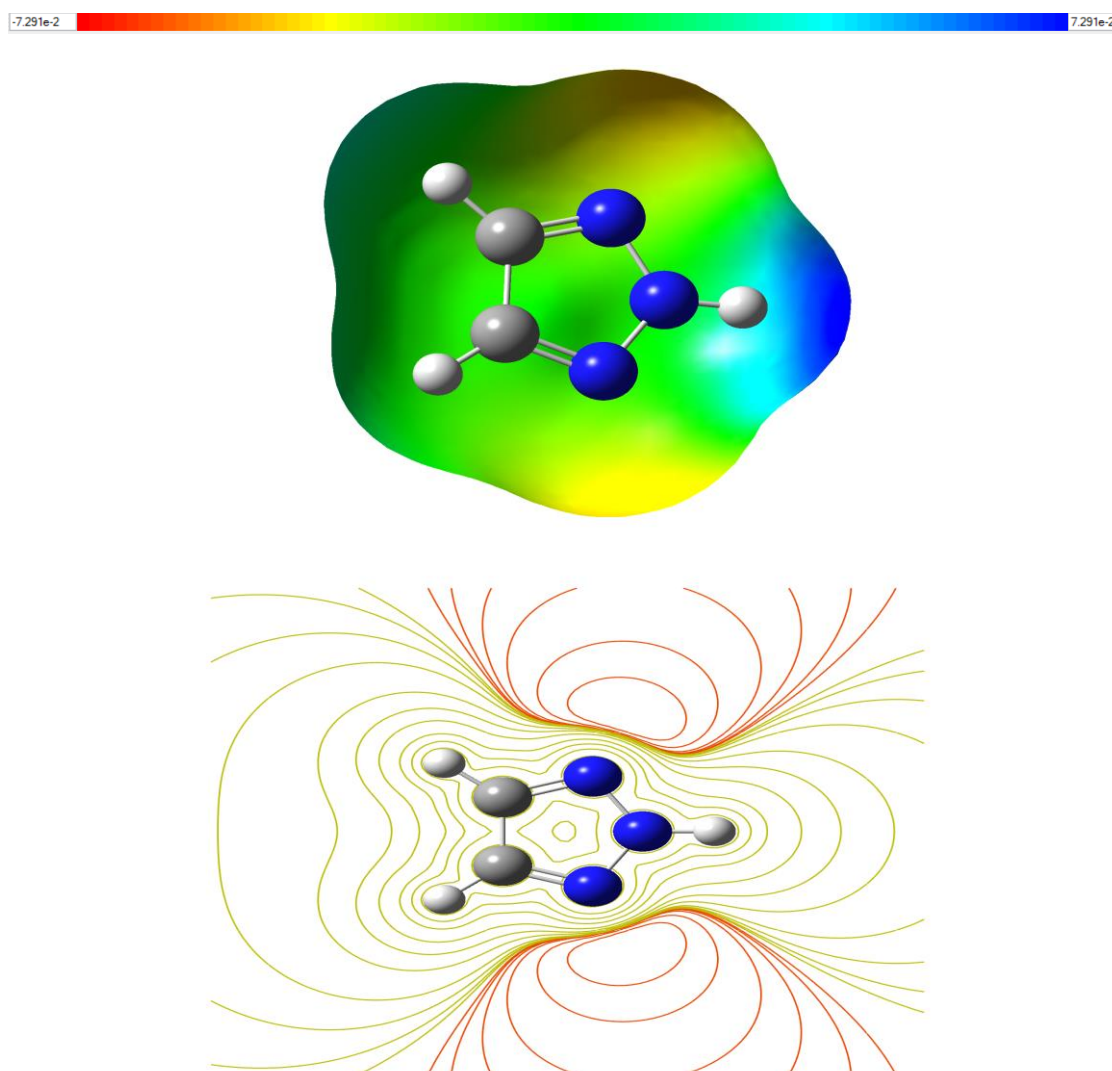
Figure. III.3 D MESP surface map and 2D MESP contour map for 2H 1,2,3 triazole (Gauss view 5.0.9)

The MEP is a useful property to study reactivity given that an approaching electrophile will be attracted to negative regions (where the electron distribution effect is dominant) [15]. The electrostatic potential values are represented by different colors. The positive, negative and neutral electrostatic potential regions of molecules are shown in terms of color grading. Generally, potential increases in the order red < orange < yellow < green < blue. The red color indicates the maximum negative region and the blue color represents the maximum positive region [16,17].

The MESP surface map (Figure III.3) for 2H-1,2,3-triazole shows that the maximum negative region is localized (yellow) over N1, N3 atoms and the maximum positive region is localized on NH group (bleu), indicating a possible site for nucleophilic attack. These sites give information about the region from which the compound can has intermolecular interactions. This predicted the most reactive site for both electrophilic and nucleophilic attack [18].

The green color situated in the middle between the yellow and blue regions and localized on CH groups explains the neutral electrostatic potential surface.

3.4 Substitution effect on 2H-1,2,3-triazole structure

To perceive the effect of the substitution, we have studied two series: the methyl and the ethyl group for the first series (an electron donor group), the cyanide and Chloride group for the second series (an electron attractor group) in position N2, C4 and C5 in the same series are given in Table III.8.

Table III.8. *Energies of 2H-1,2,3-triazole derivatives.*

2H 1.2.3 TRIAZOLE	HOMO [au]	LUMO [au]	ΔE [au]	μ [D]	ΔHf [kcal.mol⁻¹]
	-0.289	-0.026	0.263	0.217	70.610
Series 1					
T1	-0.272	-0.021	0.251	0.485	69.068
T2	-0.272	-0.018	0.252	0.700	61.168
T3	-0.272	-0.018	0.252	0.700	61.168
T4	-0.258	-0.015	0.243	0.343	59.688
T5	-0.258	-0.015	0.243	0.343	59.688
T6	-0.259	-0.011	0.248	0.908	51.918
T7	-0.270	-0.020	0.250	0.658	64.308
T8	-0.272	-0.020	0.251	0.398	56.550
T9	-0.272	-0.020	0.251	0.741	56.550
T10	-0.256	-0.016	0.240	0.741	50.581
T11	-0.256	-0.016	0.240	0.504	50.585
T12	-0.258	0.013	0.271	0.810	43.195
T13	-0.246	-0.010	0.236	0.508	50.195
T14	-0.245	-0.016	0.229	0.285	36.101
Series 2					
T`1	-0.285	-0.071	0.214	1.331	74.510
T`2	-0.281	0.038	0.243	1.217	65.360
T`3	-0.281	0.038	0.243	1.217	65.360
T`4	-0.281	-0.081	0.200	0.838	69.425
T`5	-0.279	-0.048	0.231	1.620	60.687
T`6	-0.281	-0.081	0.200	0.838	69.425
T`7	-0.311	-0.084	0.227	4.883	121.960
T`8	-0.308	-0.068	0.240	4.062	108.150
T`9	-0.308	-0.068	0.240	4.062	108.150
T`10	-0.329	-0.118	0.211	2.750	161.290
T`11	-0.325	-0.092	0.233	5.762	148.050
T`12	-0.329	-0.118	0.211	3.059	161.290
T`13	-0.343	-0.140	0.203	1.327	202.590
T`14	-0.279	-0.091	0.188	0.545	64.910

Heat of formation, dipole moment, HOMO (highest occupied molecular orbital), LUMO (lowest unoccupied molecular orbital) and their difference (ΔE) are reported for 2H-1,2,3-triazole and its derivatives in Table III.9 and Table III.10.

Table III.9. Mulliken charges of 2H-1,2,3-triazole and its derivatives (series 1)

Series1		T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
	N1	-0.12	-0.15	-0.15	-0.06	-0.06	-0.12	-0.31	-0.12	-0.05	-0.10	-0.10	-0.08	-0.05	-0.25
	N2	0.10	-0.1	-0.01	0.12	0.12	-0.01	0.27	0.02	0.02	0.30	0.30	-0.33	0.06	0.42
	N3	-0.12	-0.11	-0.11	-0.14	-0.14	-0.12	-0.31	-0.12	-0.12	-0.10	-0.10	-0.12	-0.05	-0.31
	C4	-0.17	-0.09	0.09	0.04	0.04	0.10	-0.2	0.14	-0.26	0.12	0.12	0.14	0.13	0.17
	C5	-0.17	0.09	-0.09	-0.06	-0.06	0.10	-0.2	-0.26	0.15	-0.24	-0.24	-0.23	0.10	0.02
Sub N2	C_1-methyl	-0.31	-	-	0.29	0.29	-	-	-	-	-	-	-	0.28	
	C1-ethyl	-	-	-	-	-	-	0.15	-	-	-0.27	-0.27	-	-	-0.17
	C2-ethyl	-	-	-	-	-	-	-0.65	-	-	-0.44	-0.44	-	-	-0.60
Sub C4	C_1-methyl	-	-0.70		-0.70	-	-0.61	-	-	-	-	-	-	-0.69	-
	C1-ethyl	-	-	-	-	-	-	-	-0.38	-	-0.42	-	-0.29	-	-0.48
	C2-ethyl	-	-	-	-	-	-	-	-0.51	-	-0.53	-	-0.63	-	-0.65
Sub C5	C_1-methyl	-		-0.70		-0.74	-0.61	-	-	-	-	-	-	-0.69	-
	C1-ethyl	-	-	-	-	-	-	-	-	-0.38		-0.42	-0.18	-	-0.48
	C2-ethyl	-	-	-	-	-	-	-	-	-0.51		-0.53	-0.74	-	-0.65

Table III.10. Mulliken charges of 2H-1,2,3-triazole and its derivatives (series 2).

	Series2	T'1	T'2	T'3	T'4	T'5	T'6	T'7	T'8	T'9	T'10	T'11	T'12	T'13	T'14
	N1	-0.09	-0.04	-0.09	-0.06	-0.01	-0.01	-0.08	-0.11	-0.12	-0.03	-0.10	-0.04	-0.01	0.04
	N2	0.34	-0.04	-0.04	0.35	-0.07	0.35	0.08	0.06	0.06	0.15	0.04	0.15	0.14	0.34
	N3	-0.09	-0.09	-0.04	-0.01	-0.01	-0.06	-0.08	-0.12	-0.11	-0.04	-0.10	-0.03	-0.01	0.04
	C4	-0.17	-0.37	-0.24	-0.36	-0.53	-0.31	-0.14	0.02	0.72	0.74	0.89	0.01	0.99	-0.56
	C5	-0.17	-0.24	-0.37	-0.31	-0.53	-0.36	-0.14	0.72	0.02	0.01	0.89	0.74	0.99	-0.56
Sub N2	Cl	0.13	-	-	-0.10	-	-0.10	-	-	-	-	-	-	-	-0.09
	C-cyano	-	-	-	-	-	-	0.22	-	-	0.22	-	0.22	0.22	-
	N-cyano	-	-	-	-	-	-	0.20	-	-	-0.18	-	-0.18	-0.16	-
Sub C4	Cl	-	0.31	-	-	0.40	0.31	-	-	-	-	-	-	-	0.39
	C-cyano	-	-	-	-	-	-	-	-	-0.91	-0.89	-0.83	-	-0.94	-
	N-cyano	-	-	-	-	-	-	-	-	-020	-0.18	-0.15	-	-0.14	-
Sub C5	Cl	-	-	0.31	0.31	0.40	-	-	-	-	-	-	-	-	0.39
	C-cyano	-	-	-	-	-	-	-	-0.91	-	-	-0.83	-0.89	-0.94	-
	N-cyano	-	-	-	-	-	-	-	-0.20	-	-	-0.15	-0.18	-0.14	-

Sub N Substitution is in the atom N2, Sub C4 Substitution is in the atom C4, Sub C5 Substitution is in the atom C5

Net atomic charges are also reported for the effect of the substitution on the electronic parameters and energy and their impact on the stability and chemical reactivity of molecule. Two series substituted molecule were studied (Figure III.4). This calculation is performed by DFT/B3LYP method using $6\text{-}311^{++}G$ (d,p) basis set.

The heat of formation decreased approximately to 2 and 7 kcal·mol-1 at each addition of methyl and ethyl groups respectively. Compounds T1, T4, T7 (nitrogen substituent) were the greatest values of the heat of formation compared to other derivatives.

N 2 substituted 1,2,3-triazoles are one special class of triazole derivatives [19]

Series1	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14
R1	CH_3	H	H	CH_3	CH_3	H	C_2H_5	H	H	C_2H_5	C_2H_5	H	CH_3	C_2H_5
R2	H	CH_3	H	CH_3	H	CH_3	H	C_2H_5	H	C_2H_5	H	C_2H_5	CH_3	C_2H_5
R3	H	H	CH_3	H	CH_3	CH_3	H	H	C_2H_5	H	C_2H_5	C_2H_5	CH_3	C_2H_5

Series2	T`1	T'2	T'3	T'4	T'5	T'6	T'7	T'8	T'9	T'10	T'11	T'12	T'13	T'14
R1	Cl	H	H	Cl	Cl	H	CN	H	H	CN	H	CN	CN	Cl
R2	H	Cl	H	Cl	H	Cl	H	CN	H	H	CN	CN	CN	Cl
R3	H	H	Cl	H	Cl	Cl	H	H	CN	CN	CN	H	CN	Cl

Figure III. 4. 2H-1,2,3-triazole systems (Marvin sketch 15.8.31)

The Highest occupied molecular orbital (HOMO) and the lowest un-occupied molecular orbital (LUMO) are very important parameters for quantum chemistry. These values help to exemplify the chemical reactivity and kinetic stability of the molecule. The HOMO represents the ability to donate an electron and the LUMO as electron acceptor represents the ability to obtain an electron.

In order to evaluate the energetic behavior of the title compound, the relatively high value of $\Delta E_{HOMO-LUMO}$ indicates that the title compound presents high chemical stability and low reactivity [20,21]. Reactive (HSAB principle: hard and soft acids and bases), Hard bases have highest-occupied molecular orbitals (HOMO) of low energy, and hard acids have lowest-unoccupied molecular orbitals (LUMO) of high energy [22,23].

In the mono-substituted Methyl group category, the compound T1 shows a maximum positive charge on nitrogen (0.1) leading to a nucleophilic substitution (Table III.9).

In the case of dimethyl and diethyl substitution of 2H-1,2,3-triazole in the N2 position compound T4 and T10 show maximum charge (0.12) and (0.30) respectively and smaller HOMO-LUMO energy gap (0.243) and (0.240) respectively (Table III.9) and leads to preferential site of nucleophilic attack. We also note that the methyl and ethyl substituent (donor effect) in (Table III.8) has the effect of increasing the energy of the HOMO, with little change in the LUMO.

From the results shown in Table 4, the 4,5-dimethyl1,2,3-triazole (compound T6) shows an important dipole moment value (0.908 D). It is the most soluble in polar solvents than other derivatives.

In the present work, we have studied cyanide and Chloride substituted 1,2,3-triazole along the same line of methyl and ethyl substituted 2 H-1,2,3-triazole for a comparative study.

The heat of formation increased approximately 4 and 51 kcal·mol^{-1} at each addition of Chloride and cyanide groups respectively.

In mono-substituted cyanide and chloride derivatives, 2-chloro-1,2,3-triazole (compound T'1) is predicted to be more chemically reactive than 4-Cyano-1,2,3-triazole,2-Cyano- 1,2,3-triazole and 4-Chloro- 1,2,3-triazole on the basis of least HOMO-LUMO energy gap (0.214).

Indi-substituted cyanide and chloride derivatives, dichloro 1,2,3-triazole (compounds T'4and T'6) is more reactive than dicyano 1,2,3-triazole (compound T'10). this is due to smaller HOMO-LUMO energy gap (0.200) (Table III.8).

The compound T'4 and T'6 is predicted to be the most reactive with smaller HOMO-LUMO energy gap of all 1,2,3-triazole systems (except the tri-substituted compounds).

The study of the effect of substitution on the 1.2.3 triazole base groups electron donor group and electron attractor group shows that by comparing the LUMO - HOMO gaps, the most chemically active are the tri-substituted. CompoundsT`14 and T14 show maximum charge in the N2 position (0.42) and (0.34) respectively. These results are in close agreement with the experiment, whereas the majority of tri-substituted 1,2,3-triazole have biological activity [24-26].

We note also that the cyanide and chloride substituent (attractor effect) lowers the energies of HOMO and LUMO. The influence on the energy of the LUMO is more important.

The effective atomic charges calculation which depicts the charges of every atom in the molecule distribution of positive and negative charges are vital to increase or decrease in bond length between the atoms, atomic charges, effect dipole moments, molecular polarizability, electronic structure, acidity-basicity behavior properties of molecular system and electrostatic potential surfaces [27,28].

The compound T`14 is predicted to be the most reactive with smaller HOMO-LUMO energy gap and with the most important positive charge on nitrogen N2. This Nitrogen is preferred for nucleophilic attack.

Based on our conclusions on the effect of substitution on the1,2,3-triazole molecule. We choose a series of triazole derivatives, having a biological activity.

3.5 Structure activity/property relationships of aromatase inhibitory activity of substituted 1,2,3-triazole

For the series of 1,2,3-triazole derivatives (Figure III.5) we have studied seven physicochemical properties with respect to their biological activities i.e, aromatase inhibitory activity [29]. The properties involved are: surface area grid (SAG), molar volume (V), hydration energy (HE), partition coefficient octanol/water (logP), molar refractivity (MR), polarizability (Pol) and molecular weight (MW). The results obtained using HyperChem 8.0.6 software are shown in Table III.11 and Table III.12. For example, figure III.6 shows the favored conformation in 3D of the compound 11.

Table III.11 . QSAR proprieties for 1,2,3-triazole derivatives

Compounds	Volume [Å³]	Surface area[Å²]	Hydratation Energy [Kcal.mol⁻¹]	Polarizability [Å³]	Refractivity [Å³]
1	748.96	465.70	-11.47	25.74	75.13
2	907.17	557.73	-10.40	31.25	88.93
3	1063.73	633.76	-17.14	41.98	123.41
4	836.96	502.85	-10.09	30.47	87.00
5	802.85	500.60	-13.67	29.90	90.01
6	889.17	547.10	-15.81	32.37	96.62
7	936.17	575.25	-15.48	34.30	101.34
8	950.02	582.33	-19.98	34.22	101.60
9	942.17	577.93	-14.68	34.20	100.90
10	911.32	561.11	-22.74	33.00	98.23
11	1097.37	663.18	-16.62	42.03	125.14
12	1089.77	632.87	-13.87	41.98	120.40
13	849.89	526.55	-6.26	29.39	243.35
14	1010.69	609.48	-14.54	37.25	108.53
15	787.34	481.85	-5.46	28.62	82.06
16	746.55	467.42	-9.53	28.04	85.03
17	835.52	517.36	-11.78	30.52	91.64
18	882.86	544.68	-10.49	32.35	95.92
19	855.60	523.83	-18.77	31.15	93.25
20	876.54	539.32	-11.40	32.44	96.36
21	889.10	549.66	-15.94	32.37	96.62
22	694.35	441.70	-7.35	23.89	70.15

Table III.12. *Drug likeness scoring for compounds*

Compounds	Lipinski rules					Veber rules			Ligand efficiency and Lipophilicity efficiency		
	Log P	MW	HBD	HBA	N VIOLATION	nrotb	TPSA	ABS	pIC50[41]	LE	LipE
1	1.61	226.28	4	0	0	4	54.51	90.19	*5.36*	0.44	3.75
2	2.80	268.36	4	0	0	7	54.51	90.19	*5.26*	0.36	2.46
3	3.60	327.39	5	0	0	4	78.30	81.98	*4.99*	0.24	1.39
4	2.30	266.35	4	0	0	3	54.51	90.19	*4.97*	0.40	2.67
5	1.42	260.30	4	0	0	3	54.51	90.19	*5.47*	0.38	4.05
6	0.83	290.32	5	0	0	5	63.74	87.00	*5.22*	0.36	4.39
7	0.61	324.77	5	0	0	5	63.74	87.00	*5.41*	0.36	4.8
8	0.55	315.33	6	0	0	5	87.54	78.79	*5.04*	0.32	4.49
9	0.90	304.35	5	0	0	5	63.74	87.00	*5.38*	0.32	4.48
10	0.20	306.32	6	0	0	6	72.98	83.82	*5.63*	0.35	5.83
11	1.44	366.42	5	0	0	6	63.74	87.00	*4.94*	0.30	3.5
12	3.11	367.45	5	0	0	4	78.30	81.98	*4.71*	0.23	1.6
13	3.08	243.35	3	0	0	7	30.72	98.40	*5.02*	0.25	1.94
14	2.42	327.39	5	0	0	5	78.30	81.98	*5.33*	0.41	2.91
15	2.85	241.34	3	0	0	3	30.72	98.40	*5.10*	0.28	2.25
16	1.69	235.29	3	0	0	3	30.72	98.40	*5.01*	0.38	3.32
17	1.11	265.31	5	0	0	6	39.95	95.21	*4.97*	0.41	3.86
18	1.26	279.34	4	0	0	5	39.95	95.21	*5.65*	0.39	4.39
19	0.08	281.31	5	0	0	6	49.19	92.02	*4.76*	0.31	4.68
20	0.88	299.76	4	0	0	5	39.95	95.21	*5.30*	0.33	4.42
21	0.83	290.32	5	0	0	5	63.74	87.00	*5.87*	0.39	5.04
22	1.89	201.27	3	0	0	4	30.72	98.40	4.90	0.31	3.01

QSAR (Quantitative Structure-Property Relationships) deals with relationship between physico-chemical properties & chemical structure, are based on the assumption that the structure of a molecule (i.e, its geometric, steric and electronic properties) must contain the features responsible for its physical, chemical, and biological properties and on the ability to represent the chemical by one or more numerical descriptors [30,31]. The relationships between the physicochemical properties of drugs and pharmacokinetic processes have been extensively studied. Some physicochemical parameters of drugs, such as lipophilicity, hydrogen-bonding capacity, molecular size and polar surface area, have proved to be useful for predicting passive transfer and permeation across biomembranes in ADME, but none has attracted as much interest in quantitative structure-permeation relationship (QSPeR) studies as lipophilicity [32].

Molecular volume determines transport characteristics of molecules, such as intestinal absorption or blood-brain barrier penetration [33]. Volume is therefore often used in QSAR studies to model molecular properties and biological activity.

Molecular weight (MW) is related to the size of the molecule. As molecular size increases, a larger cavity must be formed in water in order to solubilize the compound. Increasing MW reduces the compound concentration at the surface of the intestinal epithelium, thus reducing absorption. Increasing size also impedes passive diffusion through the tightly packed aliphatic side chains of the bilayer membrane [34].

Hydration energy is a key factor determining the stability of different molecular conformations in water solutions.

The molar refractivity is a steric parameter that is dependent on the spatial array of the aromatic ring in the synthesized compounds. The spatial arrangement also is necessary to study the interaction of the ligand with the receptor. Molar refractivity is related, not only to the volume of the molecules but also to the London dispersive forces that act in the drug receptor interaction [34]

The polarizability of an atom or molecule describes the response of its electron cloud to an external field. Polarizability appears in many formulas for low-energy processes involving the valence electrons of atoms or molecules. It is also widely used to describe the inductive and dispersive interactions of a molecule or molecular system, and it plays an important role in modeling many molecular properties and biological activities [35].

The values of polarizability are generally proportional to the values of surfaces and of volumes. We observe that polarizability data are generally proportional to refractivity, molecular volume and surface. Compound 11 shows the maximum value of both polarizability (42.03 $Å^3$) and refractivity (125.14$Å^3$). This compound has also high values of molecular weight (366.42), volume (1097.37 $Å^3$) and surface (663.18$Å^2$).

The most important hydration energy in the absolute value is that of the compound 10 (22.74 kcal·mol^{-1}) and the weakest is that of compound 15 (5.46 kcal·mol^{-1}) (Table III.11).

In fact, in the biological environments the polar molecules are surrounded by water molecules. Hence hydrogen bonds are established between a water molecule and these molecules.

Hydrophobic groups in 1,2,3- triazole derivatives induce a decrease of hydration energy. However, the lipophilicity increases proportionally with the hydrophobic features of substituent. As seen in Table III.9, the compound 10 is expected to have the highest hydrophilicity, whereas compound 15 will be most Lipophilic. This implies that these compounds will have poor permeability across cell membrane.
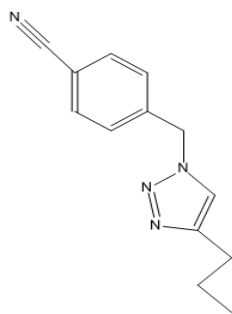
1
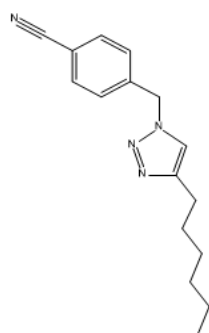
2
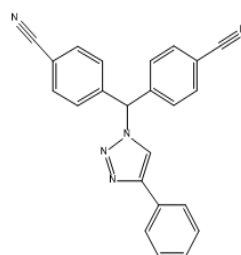
3
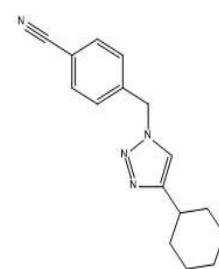
4
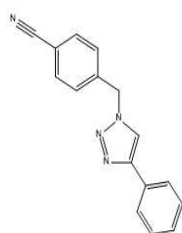
5
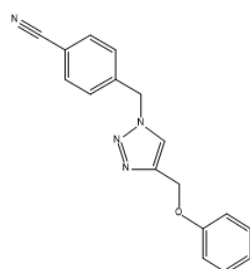
6
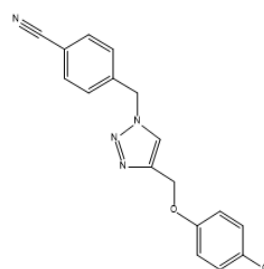
7
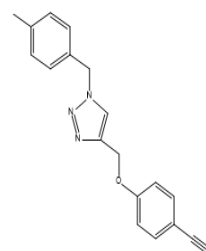
8

9

10

11

12

13

14

15

16

17   18   19   20

21   22
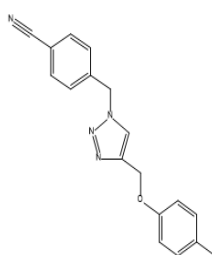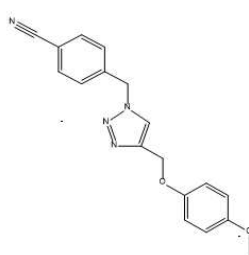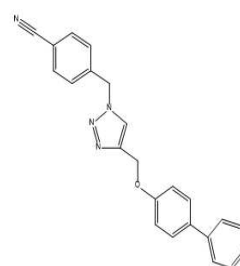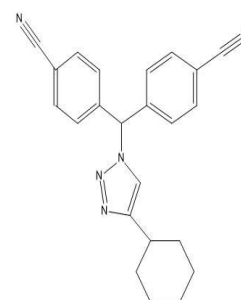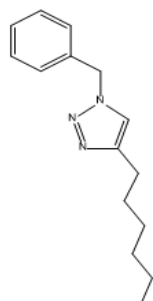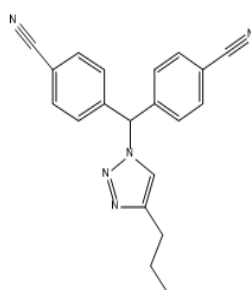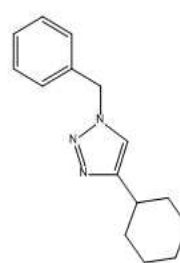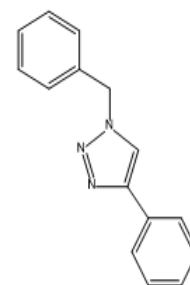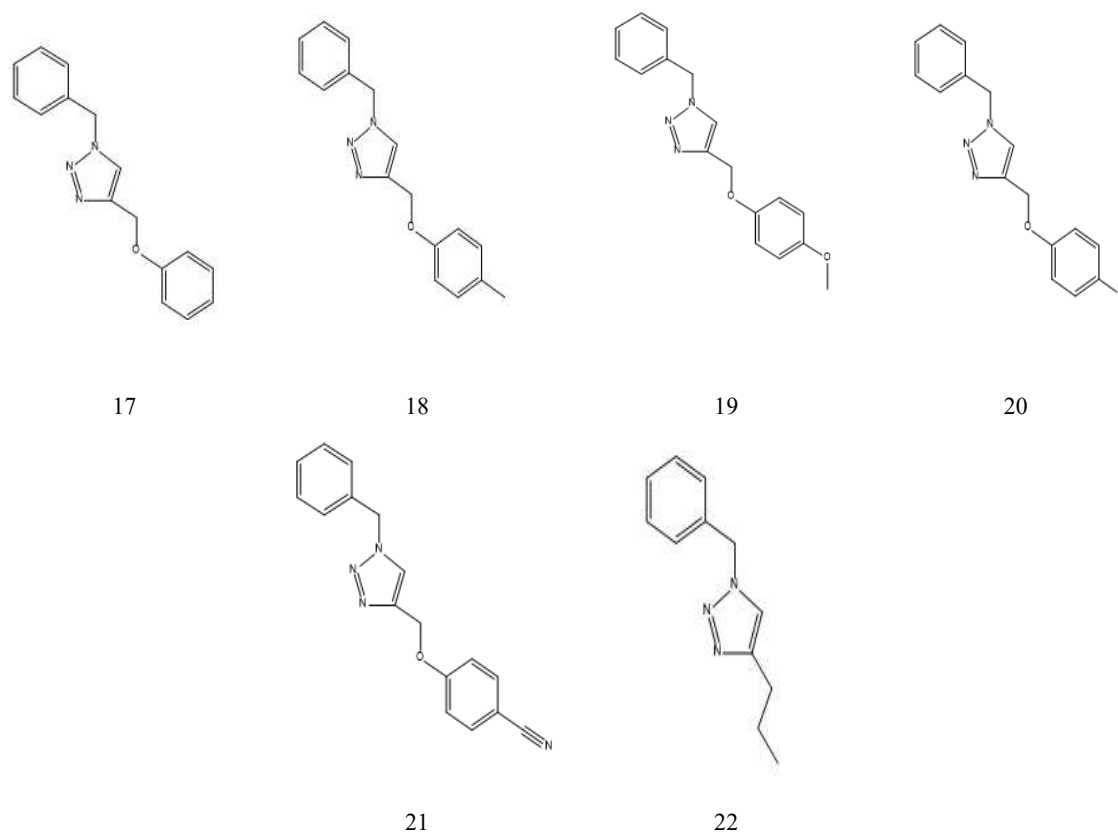
Figure III.5. 2D structures of 1,2,3-triazole derivatives
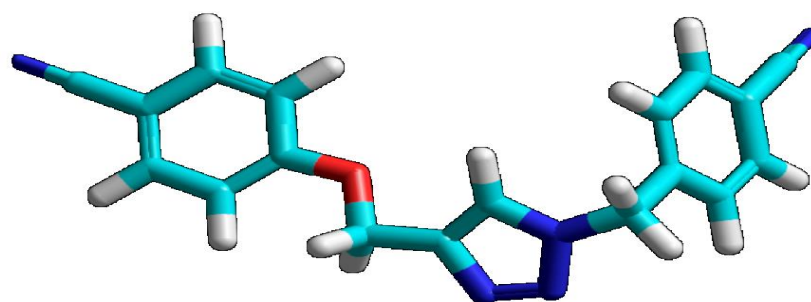
Figure. III.6. 3D conformation of compound 11 (HyperChem 8.03)

3.6 Drug-likeness properties of 1,2,3-triazole derivatives

Structures of all the selected 1,2,3-triazole derivatives in Figure III.5, were fed in the online Molinspiration software version (www.molinspiration.com) for calculation of molecular properties (number of hydrogen bond donors and acceptors, molecular weight) in Table III.12. For example, in figure 6 the favored conformation in 3D is of the compound 11.

Lipophilicity is usually measured by the partition coefficient (log P) of a compound in a single electrical form of molecular. It is a parameter that describes the partition equilibrium of a solute between water and an immiscible organic solvent. Log P is one criterion used in medicinal chemistry to assess the drug likeness of a given molecule, and is used to calculate lipophilic efficiency: a function of potency and log P that evaluate the quality of research compounds. For a given compound lipophilic efficiency is defined as the pIC50 (or pEC50) of interest minus the log P of the compound [36].

Molecular polar surface area (PSA), surface belonging to polar atoms, is a descriptor shown to correlate well with passive molecular transport through membranes and, therefore, allows prediction of transport properties of drugs. The calculation of PSA, however, is rather time-consuming because of the necessity to generate a reasonable 3D molecular geometry and the calculation of the surface it self.   Molecules with PSA values of 140 A^0 or more are expected to exhibit poor intestinal absorption [37]. TPSA was used to calculate the percentage of absorption (%ABS) according to the equation (1) [38].

$$\%ABS = 109 \pm 0.345 \times TPSA \qquad (1)$$

The LogP values for all compounds 1,2,3-triazole derivatives except 3 (which has the higher value 3.60) are in the field of optimal values (0< Log P <3) so we could say that these compounds have optimal biological activity (permeability, solubility). For a too high log P, the drug has low solubility and a too low Log P so it has difficulty penetrating the lipid membranes. Therefore the different properties which describe a molecular are so important for a drug's pharmacokinetics in the human body, including absorption, distribution, metabolism, and excretion "ADME" Components of the Lipinski's rule. [39]

However, there are some exceptions to this rule and a compound is likely to be orally active as long as it did not break more than one of its rules because some of orally active drugs such as atorvastatin, cyclosporin do not obey the rule of five. [39]

The calculation results show that all compounds meet the Lipinski rules. The suggesting log P of these compounds was found below 5 which means a good permeability across cell membrane: TPSA below 140 $Å^2$, *n* violations =1 or <0 it means compound easily bind to receptor, molecular mass <500, *n*rotb < 5, H-bond donors (HBD) ≤ 5 and H-bond acceptors (HBA) ≤ 10. [40]

In our case, the Lipinski and Veber rules are validated, therefore, theoretically, there would not have a problem with oral bioavailability for all chosen compounds.

TPSA of 1,2,3-triazole derivatives were found in the range of 87.54- 30.72 and were well below the 140 $Å^2$. We can observe obviously that all the title compounds (1–22) exhibited a great %ABS ranging from 78.79 to 98.40% indicating that these compounds should have good cellular plasmatic membrane permeability (Table III.12).

Ligand lipophilicity efficiency (LEE or LipE) is a ligand efficiency index that was first proposed by Leeson and Springthrope. LipE provides a straightforward and meaningful way to evaluate the quality of research compounds, linking potency and lipophilicity in an attempt to estimate drug-likeness, LipE attempts to maximize the minimally acceptable lipophilicity per unit of *in vitro* potency or more simply, to improve potency, while maintaining low lipophilicity.

$$LipE = pIC50 - \log P \qquad (2)$$

LipE is defined as the difference of log*P* (or log *D*) and the negative logarithm of a potency measure (pKd, pKi or pEC50) [41].

In addition, we have studied the Ligand Efficiency (LE) to penalize large compounds over small compounds with similar potency because larger compounds tend to have poorer physicochemical and ADME properties.

The LE metric was first defined by Andrews, which defines it as biological activity per molecular size.

$$LE = 1.4\,pIC50/N_H \qquad (3)$$

Where: N_H is the number of heavy atoms.

Ligand efficiency is a simple metric for assessing whether a ligand derives its potency from optimal fit with the target protein or simply by virtue of making many contacts.

It shows generally a dependency on ligand size i.e. Ligand efficiency drops dramatically when the size of the ligand increases [42-44].

We can see through the results in Table III.12 that compound 10 had the highest LipE and LE value of the data set and was deemed to be the most optimal compound.

4. CONCLUSION

The present work studied the molecular proprieties of 1,2,3-triazole. The DFT and ab-initio/*HF* methods can be used quite satisfactorily in predicting the chemical reactivity of the molecules and the effect of substitution of either donor or acceptor electron.

The calculations show the 2H tautomer to be the more stable with low reactivity, and difference of ΔE_{1H-2H}= -2.1kcal/mol.

The study of substitution on the 1.2.3-triazole base groups (electron donor group and electron attractor group) shows that by comparing the LUMO - HOMO gaps, the most chemically active compounds are found to be the trisubstituted T`14 and T14.1,2,3-triazole derivatives exhibited a great %ABS and thus reflecting a good cellular plasmatic membrane permeability.

Compound 19 in the series of 1,2,3- triazole derivatives, presents a low coefficient of division (log*P*), hence it is the most absorbent product.

Compound 10 has an important hydration energy leading to a better distribution in fabrics. Moreover, it had the highest LipE and LE values of the data set that's why it was deemed to be the most optimal compound and all the criteria of the Lipinski's rule are checked**.**

5. REFERENCES

1. Banday, A. H., Shameem, S. A., & Ganai, B. A. (2012). Antimicrobial studies of unsymmetrical bis-1,2,3-triazoles. Organic and Medicinal Chemistry Letters, 2(1), 13.

2. Modzelewska B B, Jagiello W E. (2000). Synthesis and biological activity of bis-1.2.4-triazole and bis-1.3.4-thiadiazole derivatives, Acta. Pol. Pharm. 57(3), 199–204.

3. Roger H, Gabriel E, Graham J, Rongliang L, Jacek G M, Maxwell M R. (2010). Synthesis of an antibacterial Compound Containing a 1,4-Substituted 1H-1.2.3-triazole. Org. Process. Res. Dev. 14 (1), 152-158.

4. Zhou, L., Amer, A., Korn, M., Burda, R., Balzarini, J., De Clercq, E., … Torrence, P. F. (2005). Synthesis and antiviral activities of 1,2,3-triazole functionalized thymidines: 1,3-dipolar cycloaddition for efficient regioselective diversity generation. *Antiviral Chemistry & Chemotherapy*, *16*(6), 375-383.

5. Tang, R., Jin, L., Mou, C., Yin, J., Bai, S., Hu, D., … Song, B. (2013). Synthesis, antifungal and antibacterial activity for novel amide derivatives containing a triazole moiety. *Chemistry Central Journal*, *7*(1), 30.

6. Ahmadi F, Rezayan M, Ghayahbashizadeh M S, Alipoiur E, Ostad S N, Vosooghi M, Reza khademi H, Amini M, (2014). Synthesis and Evaluation of Anti-inflammatory and Analgesic Activities of New 1.2.4-triazole Derivatives, J.Med.Chem. 11(1), 69-76.

7. HyperChem (Molecular Modeling System) Hypercube, Inc. USA, 2007.

8. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V.

Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, J. B. Farkas. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Wallingford, CT (2009).

9. Database, [http://www.molinspiration.com].

10. Marvin was used for drawing, displaying and characterizing chem- ical structures, substructures and reactions, Marvin 6.3.0, 2014, ChemAxon (http://www.chemaxon.com).

11. Begtrup M, Nielsen Cj. Nygaard L, Samdal S. (1988). The molecular structure and tautomer equilibrium of gaseous 1.2.3 triazole studied by microwave spectroscopy electron diffraction and abintio calculations, Acta Chem. Scand., A42,500-514.

12. Billes, F., Endrédi, H., & Keresztury, G. (2000). Vibrational spectroscopy of triazoles and tetrazole. *Journal of Molecular Structure: THEOCHEM*, *530*(1-2), 183-200.

13. Jbarah, A. A., Ihle, A., Banert, K., & Holze, R. (2006). The electrosorption of 1,2,3-triazole on gold as studied with surface-enhanced Raman spectroscopy. *Journal of Raman Spectroscopy*, *37*(1-3), 123-131.

14. Li, H.-B., Lu, N., Zhang, Q., Wang, Y., Feng, D., Chen, T., … Yu, P. (2018).: Electric-field control of ferromagnetism through oxygen ion gating. *Nature Communications*, *9*(1), 580.

15. Düğdü, E., Ünver, Y., Ünlüer, D., Tanak, H., Sancak, K., Köysal, Y., & Işık, Ş. (2013). Synthesis, structural characterization and comparison of experimental and theoretical results by DFT level of molecular structure of 4-(4-methoxyphenethyl)-3,5-dimethyl-4H-1,2,4-triazole. *Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy*, *108*, 329-337.

16. Chandrasekar, S., Balachandran, V., Evans, H.-S., & Latha, A. (2015). Synthesis, crystal structures HOMO-LUMO analysis and DFT calculation of new complexes of p-substituted dibenzyltin chlorides and 1,10-phenanthroline. *Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy*, *143*, 136-146.

17. Salah, T., Belaidi, S., Nadjib, M., Ismail, D., & BOUGHDIRI, S. (2016). In silico investigation by conceptual DFT and molecular docking of antitrypanosomal

compounds for understanding cruzain inhibition. *Journal of Theoretical and Computational Chemistry*, *15*.

18. . Shankar Rao, Y. B., Prasad, M. V. S., Udaya Sri, N., & Veeraiah, V. (2016). Vibrational (FT-IR, FT-Raman) and UV-Visible spectroscopic studies, HOMO-LUMO, NBO, NLO and MEP analysis of Benzyl (imino (1H-pyrazol-1-yl) methyl) carbamate using DFT calculaions. *Journal of Molecular Structure*, *1108*, 567-582.

19. Wu, L., Guo, S., Wang, X., Guo, Z., Yao, G., Lin, Q., & Wu, M. (2015). ChemInform Abstract: Tandem Synthesis of 2-Aryl-1,2,3-triazoles from α-Arylhydrazonoketones with NH 4 OAc via Copper-Catalyzed Aerobic Oxidation. *Tetrahedron Letters*, *56*.

20. Rocha, M., Di Santo, A., Arias, J. M., Gil, D. M., & Ben Altabef, A. (2015). Ab-initio and DFT calculations on molecular structure, NBO, HOMO-LUMO study and a new vibrational analysis of 4-(Dimethylamino) Benzaldehyde. *Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy*, *136 Pt B*, 635-643.

21. Almi, Zineb & Belaidi, Salah & Nadjib, Melkemi & Salima, Boughdiri & Belkhiri, Lotfi. (2016). Structure Activity Relationship and Quantitative Structure-Activity Relationships Modeling of Cyto-Toxicity of Phenothiazine Derivatives. Quantum Matter. 5. 10.1166.

22. Atkins P.W, Physical Chemistry, 3[th]edition, Oxford University Press ,Oxford,2001 p.p138-141.

23. Bakhshi H, Yeganeh H,. Mehdipour-Ataei,S. Solouk A, Irani S. Polyurethane Coatings Derived from 1,2,3-Triazole -Functionalized Soybean Oil-Based Polyols: Studying their Physical, Mechanical, Thermal, and biologicl Properties, 2013.Macromolecules 46 (19), 7777–7788

24. Haider S,. Alam MS, Hamid H, Shafi S, Nargotra A, Mahajan P, Nazreen S,. MKalle A, Kharbanda C, Ali Y, Alam A, Panda AK, Eur J .Med .Chem.2013,70,579-588. doi: 10.1016/j.ejmech.2013.10.032

25. Oliva CG, Jagerovic N,Goya P, Alkorta I, Elguero J, Cuberes R, Dordal A. , (2010). N-Substituted-1,2,3-triazoles. Synthesis, characterization and evaluation as cannabinoid ligandsARKIVOC, 2, 127-147

26. Belaidi S, Mazri R, Mellaoui M, Kerassa A, Belaidi H. (2014). Electronic Structure and

Effect of Methyl Substitution in Oxazole and Thiazole by Quantum Chemical Calculations, Res J Pharm Biol Chem Sci, 5(3),811-818

27. Harkati D, Belaidi S, Kerassa A ,Gherraf N, Quantum Matter, 2016.5, 36-44. doi 10.1166/qm.2016.1252

28. Hamadache, M., Benkortbi, O., Hanini, S., Amrane, A., Khaouane, L., & Si Moussa, C. (2016). A Quantitative Structure Activity Relationship for acute oral toxicity of pesticides on rats: Validation, domain of application and prediction. *Journal of Hazardous Materials*, *303*, 28-40

29. Belaidi S, Salah T, Melkemi N, Sinha L ,Prasad O, J.Comput. Theor.Nanosci, 2015.12, 2421-2427. doi 10.1166/jctn.2015.4042

30. Xiangli L, Bernard L ,Alfred T,. Pharm. Res.2011. 28,962–977. doi 0.1007/s11095-010-0303-7

31. Raevsky O A. (2014). Physicochemical Descriptors in Property-Based Drug Design, Mini Rev. Med. Chem, 4, 1041-1052

32. Wang, J., Xie, X.-Q., Hou, T., & Xu, X. (2007). Fast Approaches for Molecular Polarizability Calculations. *The Journal of Physical Chemistry A*, *111*(20), 4443-4448.

33. Pajouhesh H, Lenz G R. (2005). Medicinal Chemical Properties of Successful Central Nervous System Drugs, J. Neurosci, 2(4), 541–553

34. Ertl, P., Rohde, B., & Selzer, P. (2000). Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. *Journal of Medicinal Chemistry*, *43*(20), 3714-3717.

35. Arnott J A, Kumar R, Planey S L. (2013). Lipophilicity Indices for Drug Development, J. Appl. Biopharm, 1, 31-36

36. Viswanadhan V N, Ghose A. K. Revankar G R, Robins R K. (1989). Atomic physicochemical parameters for three-dimensional structure directed quantitative structure-activity relationships, J.Chem. Inf. Model, 29(3), 163-172.

37. Kerassa A, Belaidi S, Lanez T,.(2016). Computational Study of Structure-Property Relationships for 1,2,4-Oxadiazole-5-Amine DerivativesQuantum Matter, 5, 45-52

38. Leeson, P. D., & Springthorpe, B. (2007). The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature Reviews. Drug Discovery*, *6*(11), 881-890.

39. Lipinski C. A, Lombardo F, Dominy BW, Feeny P J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug . Delivery. Rev,23, 3-25

40. Veber D F, Johnson S R, Cheng H Y, Smith B R, Ward K W, Kopple K D. (2002). Molecular properties that influence J. Med. Chem,.45, 2615-2623

41. Keserü, G. M., & Makara, G. M. (2009). The influence of lead discovery strategies on the properties of drug candidates. *Nature Reviews Drug Discovery*, *8*(3), 203-212.

42. Leeson, P. D., & Springthorpe, B. (2007). The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature Reviews. Drug Discovery*, *6*(11), 881-890.

43. Andrew L. Hopkins, Colin R.Groom. Alexander A. Volume 9, Issue 10, 15 May 2004, Pages 430-431. Drug Discovery Today.

44. Leeson, P. D., & Springthorpe, B. (2007). The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature Reviews. Drug Discovery*, *6*(11), 881-890.

CHAPTRE IV

QSAR MODEL FOR PREDICTING THE AROMATASE INHIBITION ACTIVITY OF 1.2.3 TRIAZOLE DERIVATIVES

1. INTRODUCTION

Breast cancer is a kind of malignant tumor for women, [1] which accounts about 30% incidences for all malignant tumors in different age groups of women. In recent years, the mortality rate of breast cancer also shows an increasing trend and it has become one of the major causes of cancer death in female. [2]

A great majority of breast cancers is hormone-dependent [3] and it is widely accepted that estrogen plays an important role in the genesis and evolution of breast tumors. [4] Aromatase (CYP19) a cytochrome P450 enzyme, is responsible for the conversion of androgens including androstenedione and testosterone into estrogens, [5] therefore it is considered as a particularly attractive target for inhibition in the endocrine treatment of hormone-dependent breast cancer. Non-steroidal aromatase inhibitors, such as aminoglutethimide [6] anastrozole (Arimidex™) [7] and letrozole (Femara™).[8] Competitively inhibit the enzymatic activity of aromatase in a reversible manner and play an important role in the endocrine treatment for hormone-dependent breast cancers.

Among the AIs (inhibitory aromatase). letrozole and anastrozole both containing 1.2.4-triazole ring, were approved by the Food and Drug Administration (FDA) and using as the first-line therapy in the treatment of breast cancer in postmenopausal women since they have been shown to be superior to tamoxifen. [[9, 10] Based on the AIs, the triazole ring plays a pivotal role in chelation with heme iron. [11,12]

The results revealed that 1.2.3-trizole analog of letrozole showed equipotent activity to the parent compound.

In last decades, quantitative Structure-Activity Relationships (QSAR), [13] have been applied in many areas enabling to prevent time consuming and cost during the analysis of biological activities of interest. The main hypothesis involved in any QSAR is the assumption that the variation of the behavior of chemical compounds, as expressed by any experimentally measured biological or physicochemical property, can be correlated with numerical entities related to some aspect of the chemical structure termed molecular descriptors. [14-16]

Descriptors are generally used to describe different characteristics/ attributes of the chemical structure in order to yield information about the activity/property being studied.

Herein, a series of 1-substituted mono- and bis-benzonitrile or phenyl analogs of 1.2.3-triazole letrozole are employed for QSAR modeling of the aromatase inhibitory activity. A diverse set of quantum chemical and molecular descriptors were employed to provide numerical description of the investigated compounds, using multiple linear regressions (MLR).

2. MATERIALS AND METHODS

2.1 Data set

A series of thirty molecules belonging to 1.2.3-triazole derivatives have aromatase inhibitory activity, were taken from literature. [17] the studied compounds were randomly divided into training set (twenty-four compounds) and test set (six compounds). Training and test set compounds are represented in (figure.IV.1). These compounds in the series were sketched using ChemDraw module which is available in ChemOffice.
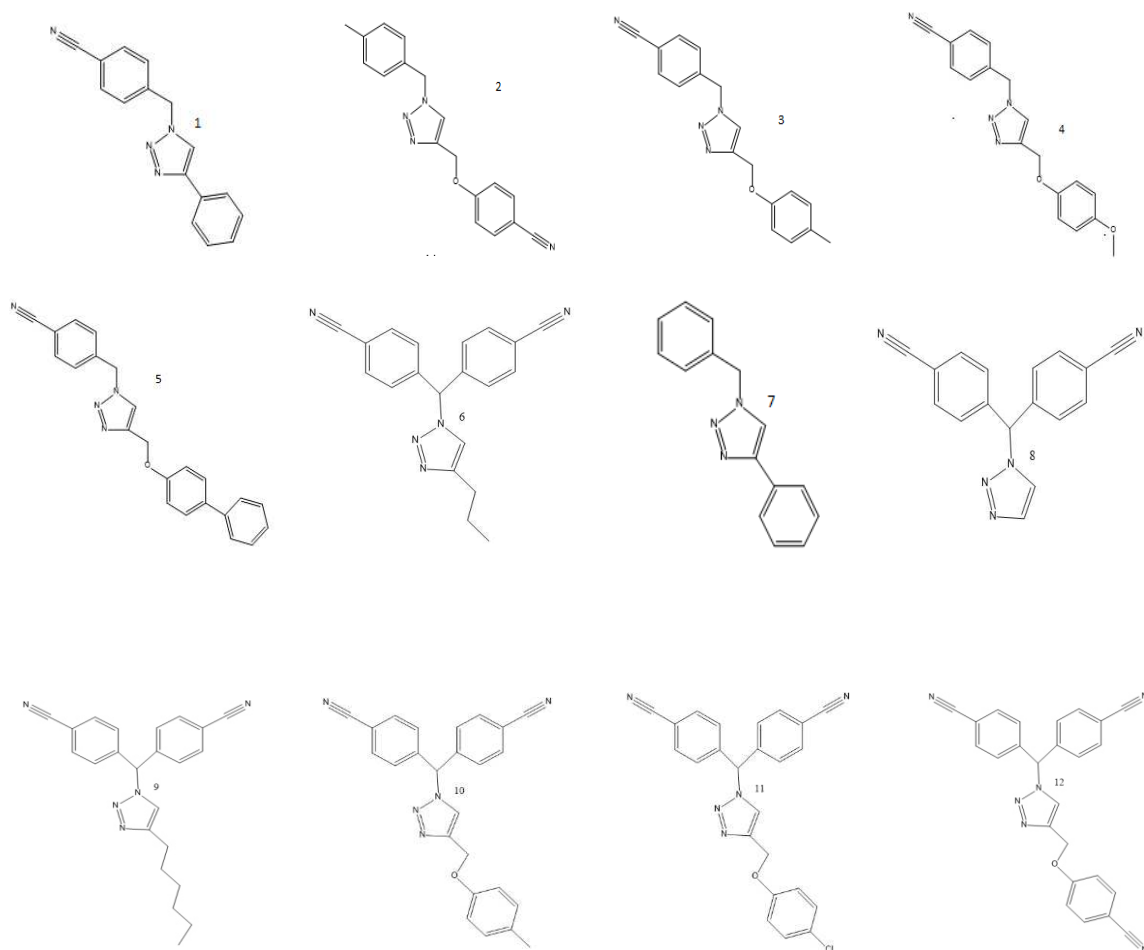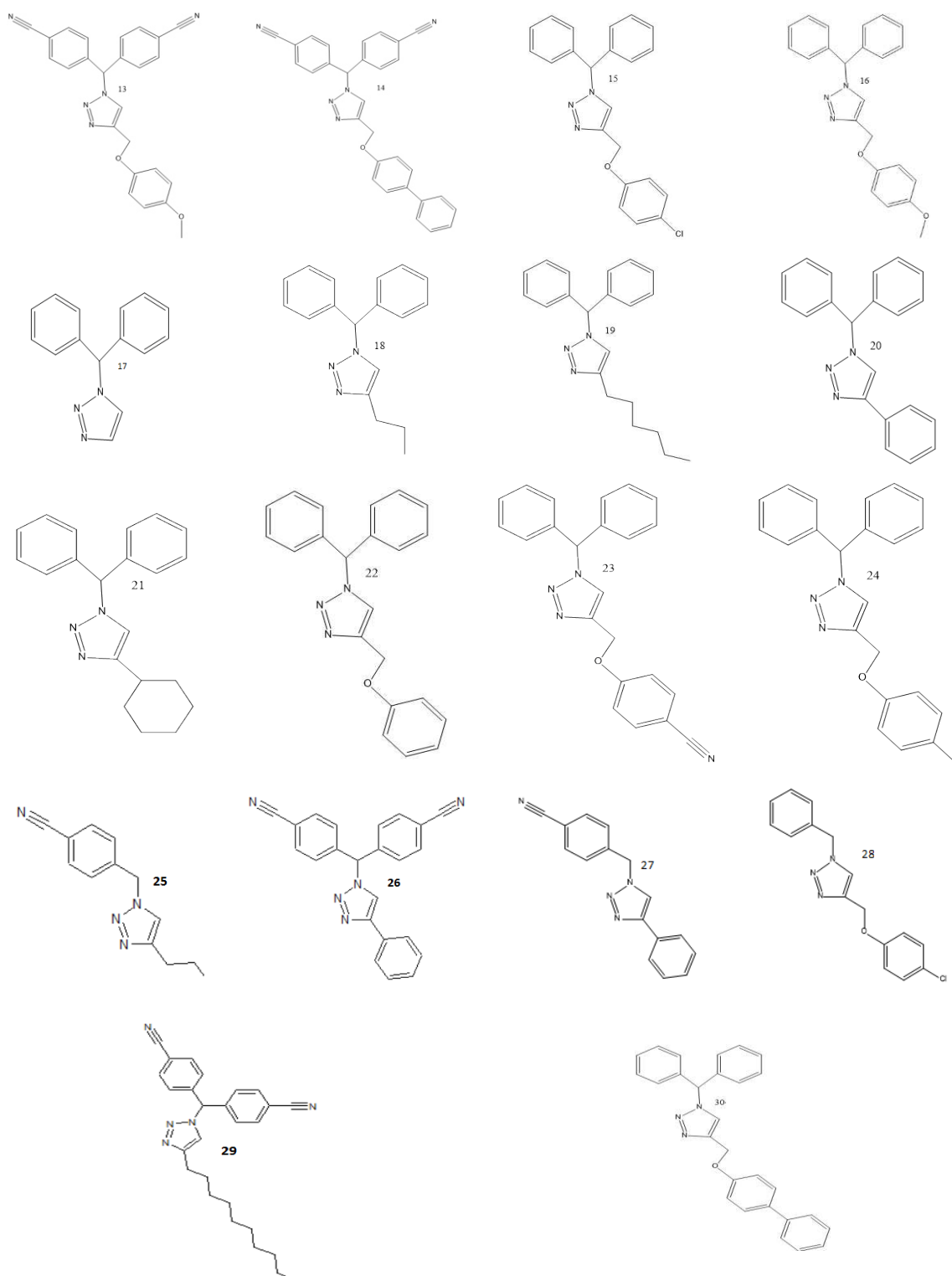
Fig.IV.1. *Chemical structures of the 1.2.3 triazole derivatives.*

2.2 Descriptors generation

Firstly, the thirty investigated molecules were pre-optimized by the Molecular Mechanics Force Field (MM+) included in HyperChem version 8.03 package. [19] After that, the resulted minimized structures were further refined using the semi empirical PM3 Hamiltonian implemented also in HyperChem. We chose a gradient norm limit of 0.01kcal/Å for the geometry optimization.

QSAR properties module from HyperChem 8.03 was used to calculate physical and chemical proprieties of a series of thirty 1.2.3-triazole derivatives: the molar polarizability (Pol), the molar refractivity (MR), partition coefficient octanol/water (log P), hydration energy (HE) , molar volume (MV) , Surface area grid (SAG) and molar weight (MW); these properties are described in (table IV.1).

Then, these 1.2.3 triazoles were reoptimized by using Gaussian 09 program package. [20] At the density functional theory level (DFT) using Becke's three-parameter Lee-Yang-Parr (B3LYP) With the 6-311G (d. p) basis set, this theory was used to calculate a number of electronic descriptors such as; HOMO and LUMO energies and atomic net charges (qN1, qN2,qN3,qC4 and qC5).(table IV.2)

Molinspiration web [22] based software was used to obtain TPSA parameter (topological polar surface area) which was used to calculate the percentage of absorption (%ABS) according to the equation [23]:

$$\%ABS \ = \ 109 \ \pm \ 0.345 \times TPSA$$

Also, we calculated the Ligand Efficiency (LE) according to the equation:

$$LE = 1.4 pIC50/NH$$

Where, NH is the number of heavy atoms [24]

Table VI.1. *Physicochemical descriptors.*

	S (A°²)	M(uma)	EH(kcal/mol)	pol(A°³)	ref(A°³)	TPSA	ABS	LE	Log P
1	458.7	246.27	-14.22	28.06	85.85	87.5	90.19	0.380	0.89
2	510.0	287.28	-21.48	30.55	91.44	96.7	78.79	0.320	1.11
3	507.0	276.30	-16.11	30.53	90.74	87.5	87.00	0.320	1.54
4	520.1	292.30	-18.99	31.17	92.83	49.1	83.82	0.350	0.39
5	516.9	338.37	-18.15	38.36	114.9	30.7	87.00	0.300	1.99
6	621.7	258.43	-5.25	34.90	98.20	30.7	81.98	0.410	4.14
7	450.9	227.31	-6.36	26.78	77.86	30.7	98.400	0.280	2.06
8	469.2	251.29	-11.96	28.68	85.76	63.7	95.210	0.390	1.82
9	520.5	285.31	-17.81	93.62	31.75	78.3	81.987	0.515	1.52
10	609.4	327.39	-14.54	108.5	37.25	78.3	81.987	0.265	2.42
11	822.9	425.52	-11.87	140.7	50.10	87.5	78.799	0.231	5.19
12	632.8	367.45	-13.87	120.4	41.98	87.5	78.799	0.216	3.11
13	633.7	361.40	-17.14	123.4	41.41	111.	70.591	0.233	2.22
14	660.4	377.40	-18.88	124.0	42.05	96.7	75.614	0.215	2.71
15	683.0	411.85	-18.50	128.7	43.97	87.5	78.799	0.189	2.49
16	702.0	407.430	-20.50	130.4	44.52	49.1	92.029	0.241	1.72
17	747.0	453.500	-20.86	152.5	51.74	30.7	98.402	0.415	3.32
18	454.7	235.29	83.66	28.04	0.331	30.7	98.402	0.331	2.08
19	544.1	277.370	98.57	33.55	0.282	30.7	98.402	0.282	2.97
20	629.8	319.450	112.3	39.05	0.285	30.7	98.402	0.285	4.16
21	570.5	311.390	113.4	37.70	0.284	30.7	98.402	0.284	2.78
22	588.6	317.430	110.4	38.28	0.243	39.9	95.217	0.243	3.66
23	614.8	361.830	118.7	40.27	0.239	63.7	87.010	0.239	3.05
24	627.5	352.400	119.0	40.19	0.237	39.9	95.217	0.237	2.99
25t	442.2	212.25	23.91	70.97	0.44	78.3	81.98	0.44	1.09
26t	653.1	310.44	36.78	103.1	0.24	39.9	95.21	0.24	3.86
27t	701.0	369.47	42.76	122.3	0.21	78.3	81.98	0.21	3.6
26t	642.4	375.86	42.1	124.7	0.25	39.9	95.21	0.25	1.97
29t	549.6	290.32	32.37	96.62	0.373	63.4	87	0.373	0.83
30t	627.8	341.41	40.18	120.1	0.28	39.9	95.217	0.28	1.71

Table IV.2. *Electronic descriptors*

	E_{LUMO}	qN_1	qN_2	qC_4	qC_5
1	-0.06	-0.49	-0.001	0.01	0.21
2	-0.09	-0.49	-0.012	0.05	0.22
3	-0.06	-0.49	-0.017	0.06	0.22
4	-0.06	-0.49	-0.017	0.28	0.22
5	-0.08	-0.70	0.019	0.27	0.25
6	-0.07	-0.48	-0.023	0.10	0.18
7	-0.01	-0.49	-0.016	0.11	0.17
8	-0.01	-0.49	0.023	0.08	0.17
9	-0.07	-0.68	0.007	0.28	0.22
10	-0.07	-0.49	-0.023	0.10	0.18
11	-0.07	-0.48	-0.019	0.36	0.18
12	-0.07	-0.49	-0.010	0.09	0.20
13	-0.07	-0.49	-0.006	0.09	0.21
14	-0.07	-0.49	-0.015	0.09	0.20
15	-0.07	-0.49	-0.015	0.09	0.21
16	-0.02	-0.48	-0.007	0.09	0.28
17	-0.01	-0.48	-0.019	0.09	0.18
18	-0.02	-0.49	-0.003	0.018	0.21
19	-0.01	-0.49	-0.013	0.11	0.17
20	-0.02	-0.47	-0.010	0.36	0.18
21	-0.02	-0.47	-0.005	0.36	0.10
22	-0.03	-0.52	0.034	0.04	0.22
23	-0.03	-0.48	-0.002	0.09	0.20
24	-0.02	-0.47	-0.010	0.36	0.18
25t	-0.25	-0.5	-0.005	-0.29	0.01
26t	-0.26	-0.49	0.017	-0.06	0225
27t	-0.07	-0.49	-0.02	0.101	0.18
26t	-0.02	-0.48	-0.001	0.098	0.20
29t	-0.04	-0.49	-0.011	0.05	0.27
30t	-0.06	-0.02	-0.558	-0.00	0.20

2.3 Regression analysis

Multiple linear regression analysis of molecular descriptors was carried out using the stepwise strategy in SPSS version 19 for Windows. [25]

2.4 Validation of the QSAR model

Testing the stability predictive power and generalization ability of the models is a very important step in QSAR study, as for the validation of predictive power of a QSAR model. two basic principles (internal validation and external validation) are used.

3. RESULTS AND DISCUSSION

In the present study we tried to developed the best QSAR model to explain the correlations between the physicochemical parameters and the biological activities IC50 values of 1.2.3 triazole derivatives with aromatase inhibitory effects.

The full linear equation for the prediction of the inhibitory IC50 activity is the following equation

$$p(IC50) = 2.726 + 0.005S - 0.366logP - 0.026ref + 6.894LE - 6.604qN1 + 0.01TPSA$$
$$+ 0.057EH + 9.431E_{LUMO} - 5.333qN2 + 0.722qC4 - 0.026ABS$$
$$+ 0.006M - 0.02POL - 2.661qC5$$

$n = 24$; $R = 0.991$; $R^2 = 0.982$; $S = 0.149$; $F = 34.569$; $Q = 6.651$

The significant equation consists of 14 descriptors; polarizability (Pol), the molar refractivity (ref), partition coefficient octanol/water (log P), hydration energy (HE), Surface area grid (S), molar weight (M), TPSA (topological polar surface area), the percentage of absorption (%ABS), Ligand efficiency (LE), E_{LUMO} and atomic net charges (qN_1, qN_2, qC_4, qC_5).

The F-value has found to be statistically significant at 95 % level since the calculated F value is higher as compared to tabulated value.

The positive value of quality factor (Q) for this QSAR's model suggests its high predictive power and lack of over fitting.

The positive coefficient of hydration energy and negative Log P indicates that the hydrophilic derivatives give a good biological activity.

From the equation, we can see any increase in the molecular surface causes an increase of the biological activity, which results in increased surface of contact between the ligand

and the receptor the same for the molecular weight. The positive coefficients of MW explain that any decrease molecular weight of the compounds causes a decrease in the biological activity.

It can be observed that high coefficients of Ligand efficiency LE**.** Thus, high LE lead to increasing aromatase inhibitory activity high LE prefers compounds that gain to escape the affinity-biased selection and optimization towards larger ligands. The focus should be directed towards the generation of compounds that use their atoms most efficiently.

In model, the positive coefficient of TPSA indicates that the substituents which increase molecular polar surface area will lead to increased activity. This relates to the molecular transport through membranes Suggested that a decrease in the permeability might decrease the activity.

It can be observed that high coefficients of atomic charges on atoms N2, C4 and C5 (qN2, qC4 and qC5 respectively). Thus, high negative charges lead to increasing aromatase inhibitory activity.

The charges allowed a physical explanation and electronic molecular properties contributing to aromatase inhibitory potency as the electronic character related directly to the electron distribution of interacting molecule at the site active.

Once the equation is obtained, it is important to determine its reliability and significance. The validation of the equation is done by cross-validation /leave-one out method. The results are shown below.

Table IV.3. Cross-validation parameter.

PRESS	SSY	PRESS/SSY	SPRESS	R^2_{cv}	R^2_{adj}
0.202	11.072	0.018	0.149	0.982	0.953

Also, for reasonable QSAR model the PRESS/SSY ratio should be lower than 0.4 [26]. The data presented in (Table.IV.3) indicate that for the developed model this ratio is 0.018.

Our result of R^2_{cv} and R^2_{adj} for this QSAR model has been to be 0.982 and 0.953 respectively. The high value of R^2_{cv} and R^2_{adj} are essential criteria for the best qualification of the QSAR model.

One can also use the S_{press} parameter that reflects the error changes predictions. Developed QSAR models have low values of S_{press} (<0.200) indicating that the model has small residual value between observed and predicted biological activities. We can see from the Table IV.3 that all residual values less than twice of standard error of estimate (0.149) therefore There, aren't any outliers.

In order to confirm our results, we have estimated the aromatase inhibitory activity pIC50 of training sets using the model expressed by equation. and compared them with the observed values. The data presented in (Table IV.4) show that the observed and predicted activities are very close each other.

The plots for this model show to be more convenient with $R^2 = 0.9914$. It indicates that the model can be successfully applied to predict the aromatase inhibitory activity of these compounds.

To investigate the presence of a systematic error in developing the QSAR models. The residuals of predicted values of the biological activity pIC50 were plotted against the experimental values as shown in (Figure.IV.3).

The propagation of the residuals on both sides of zero indicates that no systemic error exists. As suggested by Jalali-Heravi and Kyani. [27,28] It indicates that this model can be successfully applied to predict the aromatase inhibitory activity of this class of molecules.

The correlation matrix for pIC50 and selected descriptors to build the QSAR model is shown in (Table IV.4) shows that the distribution coefficient Log P and the Ligand efficiency LE are parameters important in the correlation between selected descriptors and the aromatase inhibitory activity.

The parameters used in the Model are almost Independent which can be seen from the correlation matrix.

Table IV.4 Experimental and predicted aromatase inhibitory activities (pIC50) of aromatase inhibitory activity (1-24) obtained from the model

Compound	pIC50 exp.	pIC50 pred.	Resid.	Compound	pIC50 exp.	pIC50 pred.	Resid.
1	5.470	5.599	-0.129	13	5.330	5.276	0.054
2	5.040	5.129	-0.089	14	4.920	4.903	0.017
3	5.380	5.227	0.153	15	4.870	4.782	0.088
4	5.630	5.652	-0.022	16	4.820	4.817	0.003
5	4.940	5.002	-0.062	17	5.340	5.298	0.042
6	5.330	5.309	0.021	18	4.960	4.993	-0.03
7	5.100	4.928	0.172	19	4.830	5.025	-0.19
8	5.650	5.655	-0.005	20	4.890	4.871	0.019
9	8.100	8.042	0.058	21	4.860	4.770	0.090
10	5.290	5.342	-0.052	22	4.510	4.416	0.094
11	4.960	5.039	-0.079	23	4.780	4.95754	-0.17
12	4.790	4.713	0.077	24	4.570	4.60468	-0.03
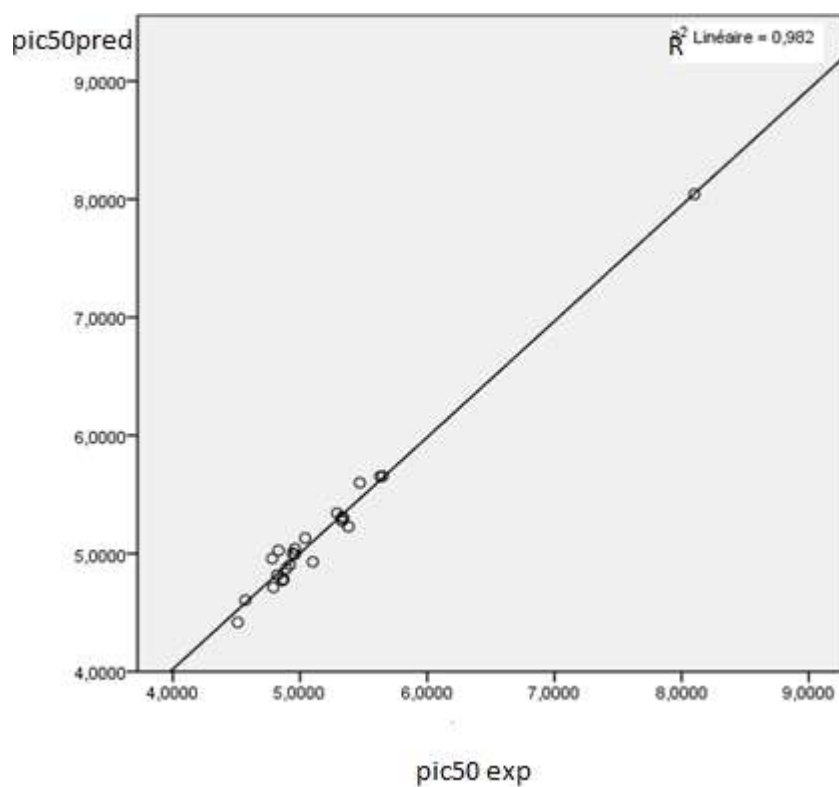
(a)
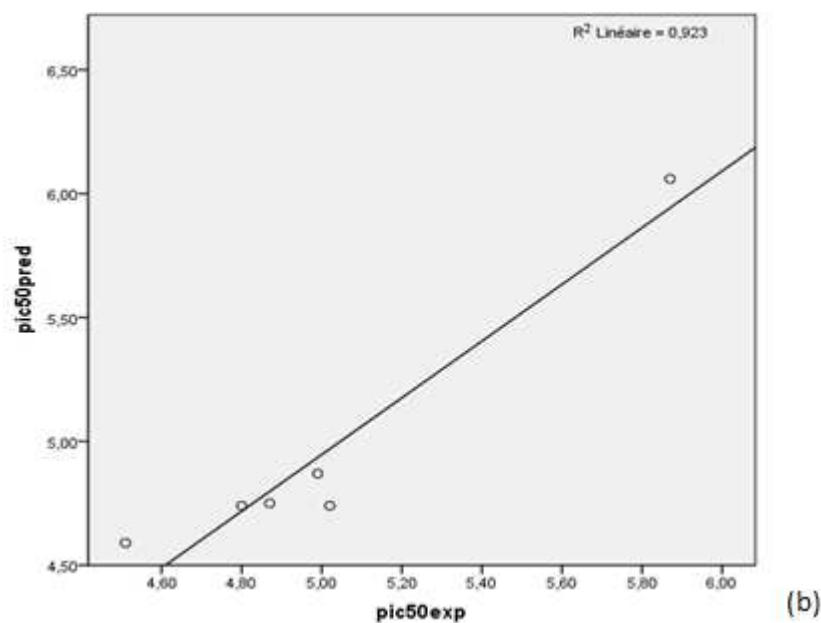
Figure. IV. 2. Scatter Plot between the Observed and Predicted Activities of Model of a-training set b- test set for the test set.
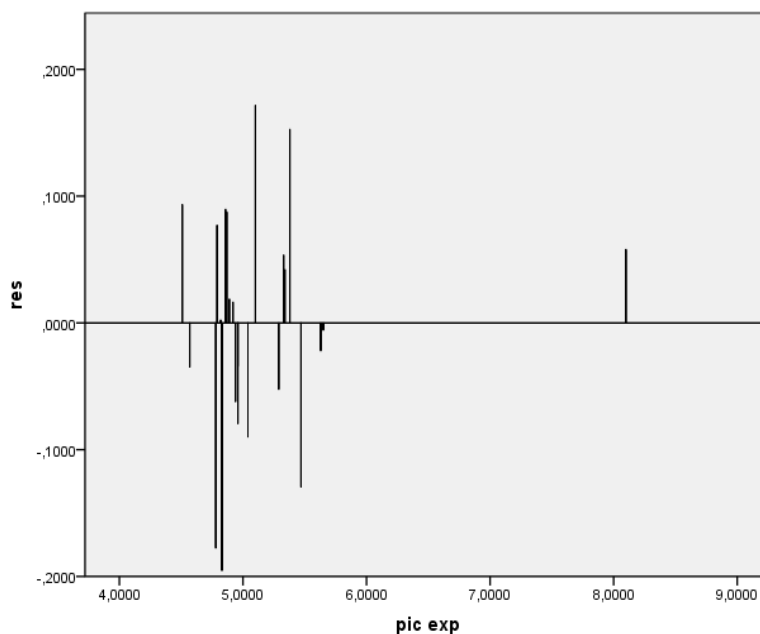
Figure IV.3. Plots of the residual values against the experimentally observed.

The proposed model passed all the tests for the predictive ability .

The results obtained show that the predicted values (Table IV.4) are very close to the observed values (figureIV.2.b) the value of R^2 is equal to 0.851 which confirms that model adequately describes the relationship between pIC50 predicted and observed model. Further the above QSAR model is confirmed its external predictability by predicting.

Table IV.5. *Correlation matrix of the fourteen selected descriptors*

	pIC50	S	M	logP	EH	Pol	Ref	LE	E_{LUMO}	qN1	qN2	qC4	qC5	TPSA	ABS
pIC50	1														
S	0.26	1													
M	0.26	0.90	1												
logp	0374	0.676	0.41	1											
EH	0.258	0.180	0.45	0.529	1										
Pol	0.224	0.800	0.80	0.570	0.07	1									
Ref	0.064	0.404	0.36	0.387	0.180	-0.83	1								
LE	0.776	0.413	0.45	0.314	0.063	-0.46	0.34	1							
E_{LUMO}	0.288	0.095	0.12	0.168	0.474	0.151	0.302	0.003	1						
qN1	0.555	0.267	0.09	0.230	0.22	0.202	0.21	-0.40	0.348	1					
qN2	0.094	0.356	0.20	0.125	0.14	-0.14	0.04	0.128	0.166	0.500	1				
qC4	0.1	0.27	0.2	0.297	0.07	0.199	0.07	-0.01	0.010	0.22	0.064	1			
qC5	0.071	0.044	0.16	0.472	0.55	-0.08	0.18	-0.06	0.335	0.397	0.267	-0.1	1		
TPSA	0.218	0.148	0.21	0.228	0.51	0.071	0.02	-0.19	0.700	0.06	0.143	-0.2	0.205	1	
ABS	0.20	0.2	0.3	0.072	0.48	-0.06	0.10	0.193	0.89	0.12	0.249	0.08	-0.27	0.812	1

Table IV.6 Observed and predicted activity of test compounds

Compounds	pIC50 exp.	pIC50 pred.
25t	4.99	4.87
26t	5.64	5.87
27t	4.8	4.74
28t	4.87	4.96
29t	5.87	6.1
30t	5	5.11

Table IV.7. Predictive power results for the external test set; Golbraikh and Tropsha criteria

Golbraikh and Tropsha's criteria							
R^2_{pred}	K	K'	$R^{°2}$	$R'^{°2}$	$\dfrac{R^2 - R^{°2}}{R^2}$	$\dfrac{R^2 - R'^{°2}}{R^2}$	$\left|R^2_0 - R'^2_°\right|$
0.851>0.6	0.982>0.85	1.017<1.15	0.972r is close to R^2	0.952 is close to R^2	-1.04<0.1	-0.11<0.1	0.02< 0.3

From the Table IV.5 it is obvious that the predicted responses of all the test compounds are in good agreement with their corresponding observed responses as well as ideal fit is attained produced by plotting a graph (Figure.IV.3) by correlating observed activity versus predicted activity of the test set compounds, the squared correlation coefficient is calculated as 0.923.

The predictive abilities of the best MLR was tested (Table IV.6) using the Golbraikh–Tropsha criteria and the R^2_{pred} test (see Model Validity section). All the calculated parameters indicated the model showed a good predictive power.

Analyzing the results of the external test set listed in Table IV.8, It could be observed that all the Golbraikh–Tropsha criteria were fulfilled.

r^2_m value of 0.556 whereas values of average r^2_m of 0.581 and $\Delta r^2_{m(test)}$ of 0.042 extend more efficient evidence of external predictability of the generated QSAR.(Table IV.8)

Once the QSAR model formulated and validated properly. Its utility is to predict the biological responses of the compounds which are generated by combinatorial deign and experimentally non-investigated.

Table IV.8 Validation characteristics of developed model according to r2m metrics and Concordance correlation coefficient

r_m^2 parameter				Concordance correlation coefficient
r_m^2	$r_m'^2$	$\Delta r_{m(test)}^2$	$\overline{r_{m(test)}^2}$	CCC
0.556>0.5	0.581	0.042<0.2	0.568>0.5	0.946>0.85

4. CONCLUSION

In this study SW-MLR was used to develop linear QSAR model for prediction of aromatase inhibitory activity of triazole derivatives. The built model displayed good correlations between the structure and activity of the studied compounds. The model was validated using LOO cross-validation and external test set. The built model has a good self- and external-predictive power. Based on QSAR model results coefficients of Ligand efficiency LE atomic net charges (qN1) and partition coefficient octanol/water (log P), were found to be important factors controlling aromatase inhibitor activity.

5. REFERENCES

1. Lézé, M.-P., Le Borgne, M., Marchand, P., Loquet, D., Kogler, M., Le Baut, G., … Hartmann, R. W. (2004). 2- and 3-[(aryl)(azolyl)methyl]indoles as potential non-steroidal aromatase inhibitors. *Journal of Enzyme Inhibition and Medicinal Chemistry*, *19*(6), 549-557.

2. Assi, H. A., Khoury, K. E., Dbouk, H., Khalil, L. E., Mouhieddine, T. H., & El Saghir, N. S. (2013). Epidemiology and prognosis of breast cancer in young women. *Journal of Thoracic Disease*, *5*(Suppl 1), S2-S8.

3. Daniel, A. R., Hagan, C. R., & Lange, C. A. (2011). Progesterone receptor action: defining a role in breast cancer. *Expert review of endocrinology & metabolism*, *6*(3), 359-369.

4. Reed, M. J. (1994). The role of aromatase in breast tumors. *Breast Cancer Research and Treatment*, *30*(1), 7-17.

5. Simpson, E. R., Mahendroo, M. S., Means, G. D., Kilgore, M. W., Hinshelwood, M. M., Graham-Lorence, S., … Michael, M. D. (1994). Aromatase cytochrome P450, the enzyme responsible for estrogen biosynthesis. *Endocrine Reviews*, *15*(3), 342-355.

6. Neves, M. A. C., Dinis, T. C. P., Colombo, G., & Melo, M. L. S. e. (2009). An efficient steroid pharmacophore-based strategy to identify new aromatase inhibitors. *European journal of medicinal chemistry*, *44*(10), 4121-4127.

7. Howell, A., Robertson, J. F. R., & Vergote, I. (2003). A review of the efficacy of anastrozole in postmenopausal women with advanced breast cancer with visceral metastases. *Breast Cancer Research and Treatment*, *82*(3), 215-222.

8. Simpson, D., Curran, M. P., & Perry, C. M. (2004). Letrozole: a review of its use in postmenopausal women with breast cancer. *Drugs*, *64*(11), 1213-1230.

9. Fornander, T., Hellström, A. C., & Moberger, B. (1993). Descriptive clinicopathologic study of 17 patients with endometrial cancer during or after adjuvant tamoxifen in early breast cancer. *Journal of the National Cancer Institute*, *85*(22), 1850-1855.

10. Jeong, H.-J., Shin, Y. G., Kim, I.-H., & Pezzuto, J. M. (1999). Inhibition of aromatase activity by flavonoids. *Archives of Pharmacal Research*, *22*(3), 309.

11. Agalave, S. G., Maujan, S. R., & Pore, V. S. (2011). Click Chemistry: 1,2,3-Triazoles as Pharmacophores. *Chemistry – An Asian Journal*, *6*(10), 2696-2718.

12. Worachartcheewan, A., Suvannang, N., Prachayasittikul, S., Prachayasittikul, V., & Nantasenamat, C. (2014). Probing the origins of aromatase inhibitory activity of disubstituted coumarins via QSAR and molecular docking. *EXCLI Journal*, *13*, 1259-1274.

13. C. Hansch and A. Leo , Exploring QSAR ,Fundamentals and Applications in Chemistry and Biology. *American Chemical Society*. Washington. D. C.**(1995).**

14. Aranda, J. F., Garro Martinez, J. C., Castro, E. A., & Duchowicz, P. R. (2016). Conformation-Independent QSPR Approach for the Soil Sorption Coefficient of Heterogeneous Compounds. *International Journal of Molecular Sciences*, *17*(8).

15. M. Karelson. Molecular Descriptors in QSAR/QSPR,Wiley-Interscience, New York**(2000).**

16. Almi, Z., Belaidi, S., Lanez, T., & Tchouar, N. (2014). Structure Activity Relationships, QSAR Modeling and Drug-like calculations of TP inhibition of 1,3,4-oxadiazoline-2-thione Derivatives. *International Letters of Chemistry, Physics and Astronomy*, *Vol. 18*.

17. Nantasenamat, C., Worachartcheewan, A., Prachayasittikul, S., Isarankura-Na-Ayudhya, C., & Prachayasittikul, V. (2013). QSAR modeling of aromatase inhibitory activity of 1-substituted 1,2,3-triazole analogs of letrozole. *European Journal of Medicinal Chemistry*, *69*, 99-114.

18. O. Deeb, Recent applications of quantitative structure-Activity relationships in drug design. In: Ekinci D. editor, Biochemistry genetics and molecular biology, *medicinal chemistry and drug design*, Rijeka, Croatia **(2012).**

19. HyperChem (Molecular Modeling System) Hypercube, Inc., 1115 NW, 4th Street, Gainesville, FL 32601, USA **(2007).**

20. Gaussian 09, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O.

Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian Inc., Wallingford, CT (**2010**)**.**

21. Ghose, A. K., & Crippen, G. M. (1987). Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure-activity relationships. 2. Modeling dispersive and hydrophobic interactions. *Journal of Chemical Information and Computer Sciences*, *27*(1), 21-35.

22. Database,(http://www.molinspiration.com).

23. Zhao, Y. H., Abraham, M. H., Le, J., Hersey, A., Luscombe, C. N., Beck, G., … Cooper, I. (2002). Rate-limited steps of human oral absorption and QSAR studies. *Pharmaceutical Research*, *19*(10), 1446-1457.

24. Andrews, P. R., Craik, D. J., & Martin, J. L. (1984). Functional group contributions to drug-receptor interactions. *Journal of Medicinal Chemistry*, *27*(12), 1648-1657.

25. SPSS software packages, SPSS Inc., 444 North Michigan Avenue, Suite 3000, Chicago, Illinoi, 60611, USA.

26. Reenu, & Vikas. (2015). Exploring the role of quantum chemical descriptors in modeling acute toxicity of diverse chemicals to Daphnia magna. *Journal of Molecular Graphics and Modelling*, *61*, 89-101.

27. Podunavac-Kuzmanović, S. O., Cvetković, D. D., & Barna, D. J. (2009). QSAR Analysis of 2-Amino or 2-Methyl-1-Substituted Benzimidazoles Against Pseudomonas aeruginosa. *International Journal of Molecular Sciences*, *10*(4), 1670-1682.

28. Heravi, M. J.; Kyani, A. (2004). Use of computer-assisted methods for the modeling of the retention time of a variety of volatile organic compounds: a PCA-MLR-ANN approach J. Chem. Inf. Comput. Sci., 44: 1328–1335.

CHAPTER V

MOLECULAR DOCKING STUDIES AND IN SILICO ADMET OF NEW SUBSTITUTED 1.2.3 TRIAZOLE DERIVATIVES FOR ANTI-BREAST CANCER ACTIVITY.

1. INTRODUCTION

This study is precisely on aromatase inhibitors as these inhibitors play a central role in the treatment of breast cancer [1]. Estrogen is important and promotes the growth and survival of normal and cancer epithelial cells. The chemical compound estrogen binds to the estrogen receptor (ER) and the activation of the cell progression takes place. The active estrogen receptor tends to bind with promoter gene present in the nucleus, which regulates the gene activity and translates the protein [2]. The activated receptor in turn binds to gene promoters in the nucleus and activates many other genes [3]. The activated gene products are thereby responsible for cell division, inhibition of cell death, new blood vessels formation and protease activity. Rapid expression of estrogen receptor is found at earlier stages of breast cancer. Nearly 70% cancer mainly depends on the over-expression of estrogen receptor [4,5].

these inhibitors have an azole hetero ring containing a sp^2 nitrogen atom which binds to the heme iron atom of aromatase to show activity inhibitory. as part of our search for new aza-hetrocyclic derivatives as aromatase inhibitors. [6,7]

The results revealed that 1,2,3-triazole analog of letrozole showed equipotent activity to the parent compound. In addition, the 1,4-disubstituted-1,2,3- triazole was shown to be the most potent compound (IC50 = 1.36 µM) among the tested 1,4-disubstituted-1,2,3-triazole series. However, the interaction mode of the 1,4-disubstituted-1,2,3-triazole series with the target enzyme remains to be explored. [8,9].

In the present work, the potential of the aromatase inhibitors, letrozole and 1.2.3 triazole Derivatives in the treatment of breast cancer has been analyzed using computational approach.

Molecular docking is a computational procedure that attempts to predict the noncovalent binding of a macromolecule (Receptor) obtained from data banks or MD simulations, etc. with a small molecule (Ligand) as a lead for further drug development. The lead candidates can be found using a docking algorithm that tries to identify the optimal binding mode of a small molecule to the active site of a macromolecular target.

The docking of aromatase inhibitors and 1.2.3 triazole with human placental aromatase cytochrome P450 reveals the role of amino acid, not only at the active site but also at the molecular or atomic level interaction of these compounds.

The ADMET properties of drugs, together with its pharmacological properties are conventionally viewed as a part of drug development. The best ligands after docking analysis were subjected to predict.

2. MATERIAL AND METHODS

2.1. Enzyme Structure

The X-ray crystal structures of human placental aromatase cytochrome P450 in complex with androstenedione (**4-ANDROSTENE-3-17-DIONE** $C_{19}H_{26}O_2$ AEMFNILZOJDQLW-QAGGRKNESA-N) (PDB ID:3EQM) [10]. were downloaded from RCSB Database (http://www.rcsb.org/pdb).

Computational analysis was carried out on chain A of 3EQM.

3EQM is a three-dimensional structure with EC Number: EC#: 1.14.14.14, chains (A), resolution 2.90 Å, and R-value 2.44.

After the download we simplified one of the two chains and we eliminated the water molecules and the cofactors

2.2 Ligand Structures

All ligand structures of 1,2,3-triazole derivatives [11] were optimized by using MM+ molecular modeling and the semi-empirical AM1 method, both of which are implemented in HyperChem 7.0 software (version 7.0, Hypercube, USA, http://www.hyper.com). For these calculations, the Polak–Ribiere conjugate gradient algorithm was employed, with the RMS gradient set to 0.0001 kcal/(Å.mol).

2.3 Cavity prediction

All cavities or the potential ligand binding site of 3EQM was predicted using MVD is given in Fig.V1. A cavity which has an important volume 304.64 $Å^3$ was predicted and used in our study. The binding site was set inside a restriction sphere of 15 radiuses with the center X: 12.67, Y: -3.46, Z: 14.09. The MolDock grid score was set with a grid resolution of 0.30.
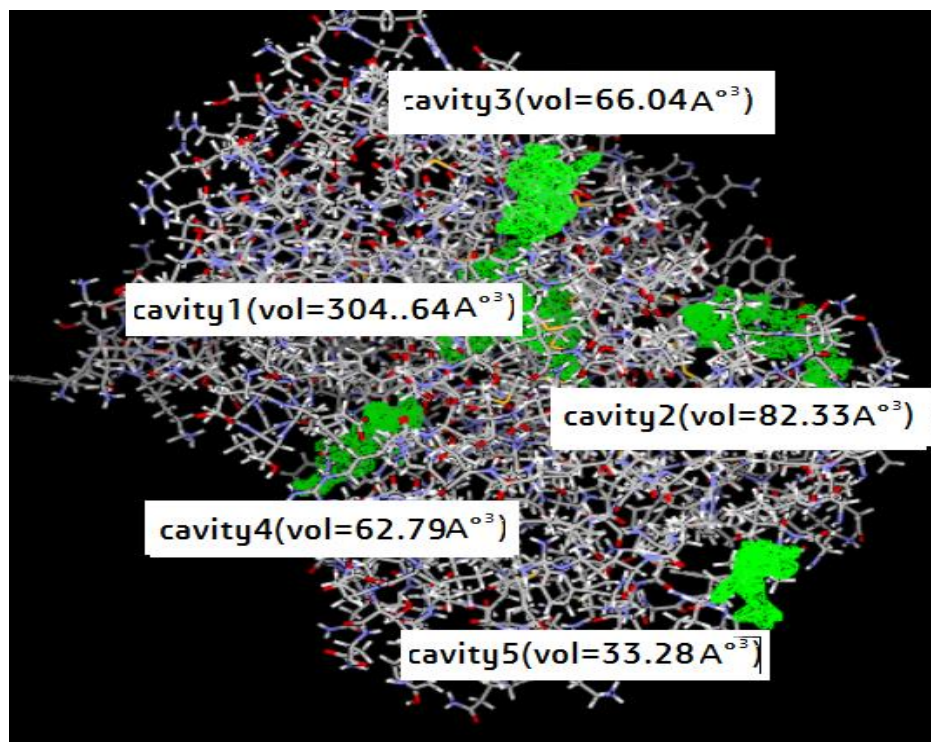
Figure.V.1. The five cavities MVD-detected cavities

2.4   Molecular docking simulation

In the present investigation, we make use of a docking algorithm called MolDock. MolDock is based on a new hybrid search algorithm, called guided differential evolution. The guided differential evolution algorithm combines the differential evolution optimization technique with a cavity prediction algorithm. The use of predicted cavities during the search process, allows for a fast and accurate identification of potential binding modes (poses).

We used MVD because it showed higher docking accuracy than other stages of the docking products (MVD: 87%, Glide: 82%, Surflex: 75%, FlexX: 58%) in the market [12,13]. Molecular docking technique was employed to dock some novel derivatives of 1,2,3- triazole designed L1-L29 (fig.V.3) against p450 receptor 3EQM using MVD to locate the interaction between various compounds and active site of aromatase.

MVD requires the receptor and ligand coordinates in either Mol2 or PDB format. Non polar hydrogen atoms were removed from the receptor file and their partial charges were added to the corresponding carbon atoms.
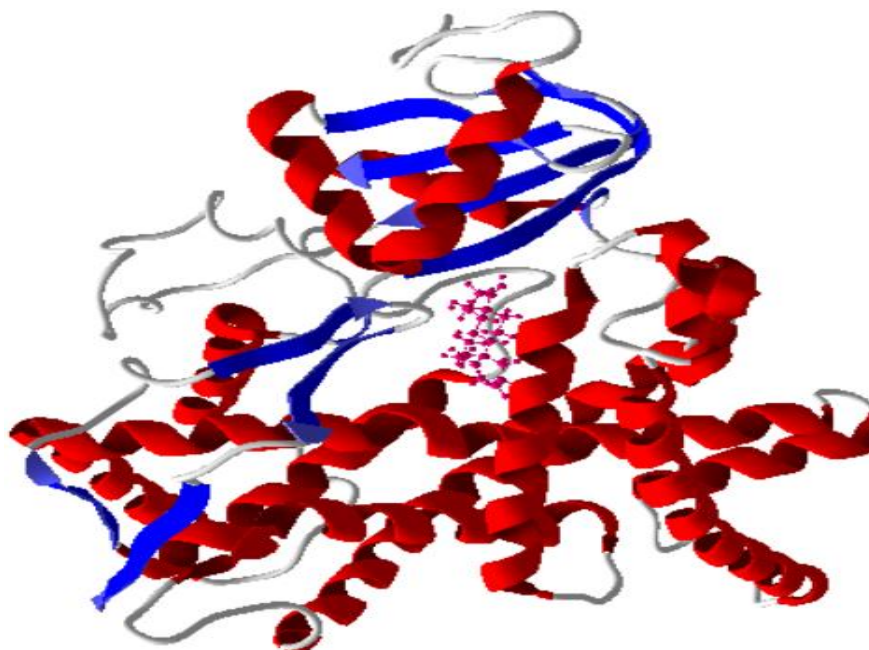
Figure V.2. Secondary structure of the target

In the docking simulations with MVD, the pose with the minimum MolDock score value was selected as the best solution, and for creating two dimensional representations of interactions between ligand and enzyme, LigPlot+ software was used.

Molecular docking was performed using MolDock docking engine of Molegro software [24]. The binding site was defined as a spherical region which encompasses all protein atoms within 15.0 A° of bound crystallographic ligand atom (dimensions X (85.67 A°), Y (51.14 A°), Z (43.73 A°) axes, respectively). Default settings were used for all the calculations.
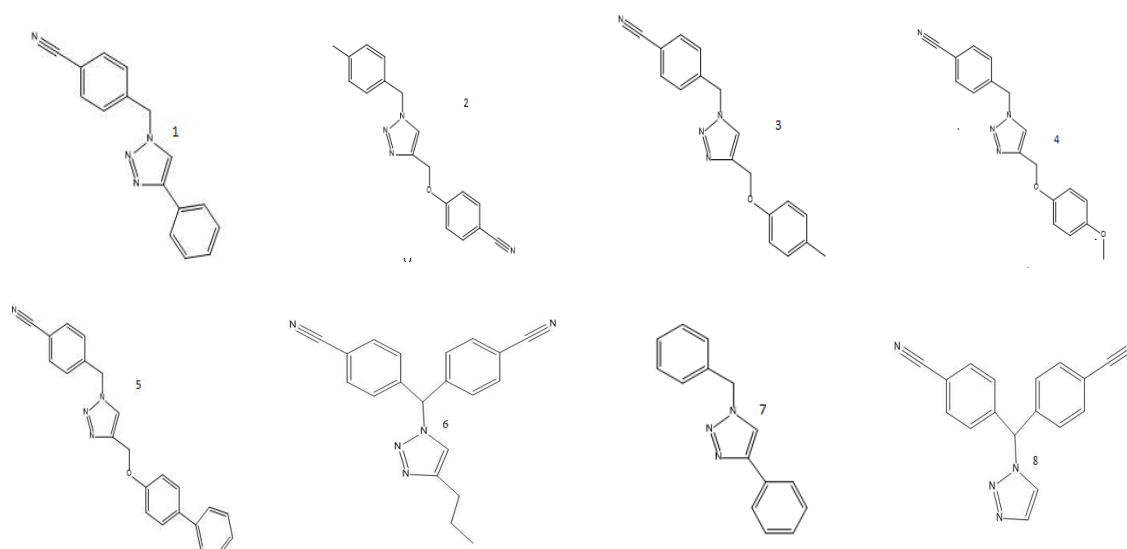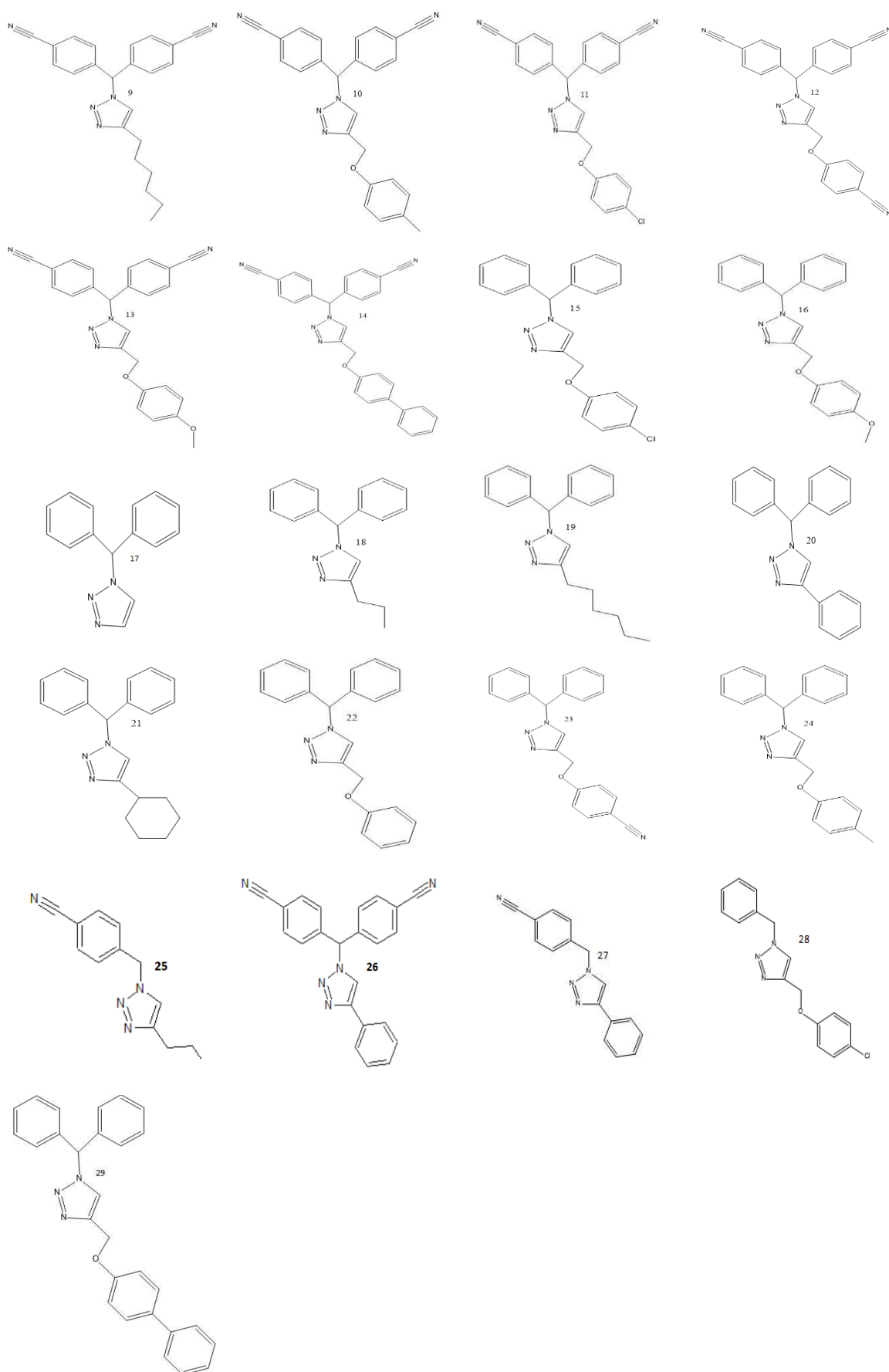
Figure.V.3. *Chemical structures of the 1.2.3 triazole derivatives.*

Docking was performed using a grid resolution of 0.3 A° and for each of the 10 independent runs; a maximum number of 1500 of interactions were executed on a single population of 50 individuals.

2.5 Molecular Property Prediction:

Molecular descriptors and drug likeliness properties of compounds were analyzed using the tool Molinspiration server (http://www.molinspiration.com), based on Lipinski Rules of five [14].

According to Lipinski's rule of five, poor absorption or permeation is more likely when there are more than 5 H-bond donors, 10 H-bond acceptors, the molecular weight is greater than 500 Da and the calculated LogP (CLogP) is greater than 5 (or MlogP>4.15)[16].. Moreover, good bioavailability is more likely for compounds with ≤10 rotatable bonds (nrotb) and total polar surface area (TPSA) of ≤140 Å (rule of Veber) [15].

2.6 Prediction of ADMET properties:

The pharmacokinetic properties such as Absorption, Distribution, Metabolism, Excretion and the Toxicity of the compounds can be predicted using the ADMETSAR (http:// www.admetexp.org) database. However, BBB (Blood-Brain Barrier) penetration, HIA (Human Intestinal Absorption), Caco-2 cell permeability and Ames test were calculated using the ADMETSAR.

In the ADMETSAR, web-based query tools incorporating a molecular build-in interface enable the database to be queried by SMILES and structural similarity search. It provides the latest and most comprehensive manually curated data for diverse chemicals associated with known ADMET profiles (ADMETSAR@LMMD) [16].

3. RESULTS AND DISCUSSION

3.1 Molecular docking studies

Molecular docking of 1,2,3-triazole derivatives to the aromatase enzyme was performed to investigate their binding interactions and to explore their binding modes. Moreover, our lead compound was also docked in order to investigate its binding pattern to the aromatase active site. The Protein-Ligand interaction plays a vital role in structural based drug design.

In this present study, we have identified 1,2,3-triazole derivatives, by means of molecular docking studies to be a novel and likely more potent inhibitor of aromatase than one of third-generation aromatase inhibitors (letrozole).

The results obtained show that all the ligands have interactions with the 3EQM at the level of the cavity 1, the FigureV.4. Explain Pose organize between human placental aromatase cytochrome P450 and most active compounds and Letrozole.
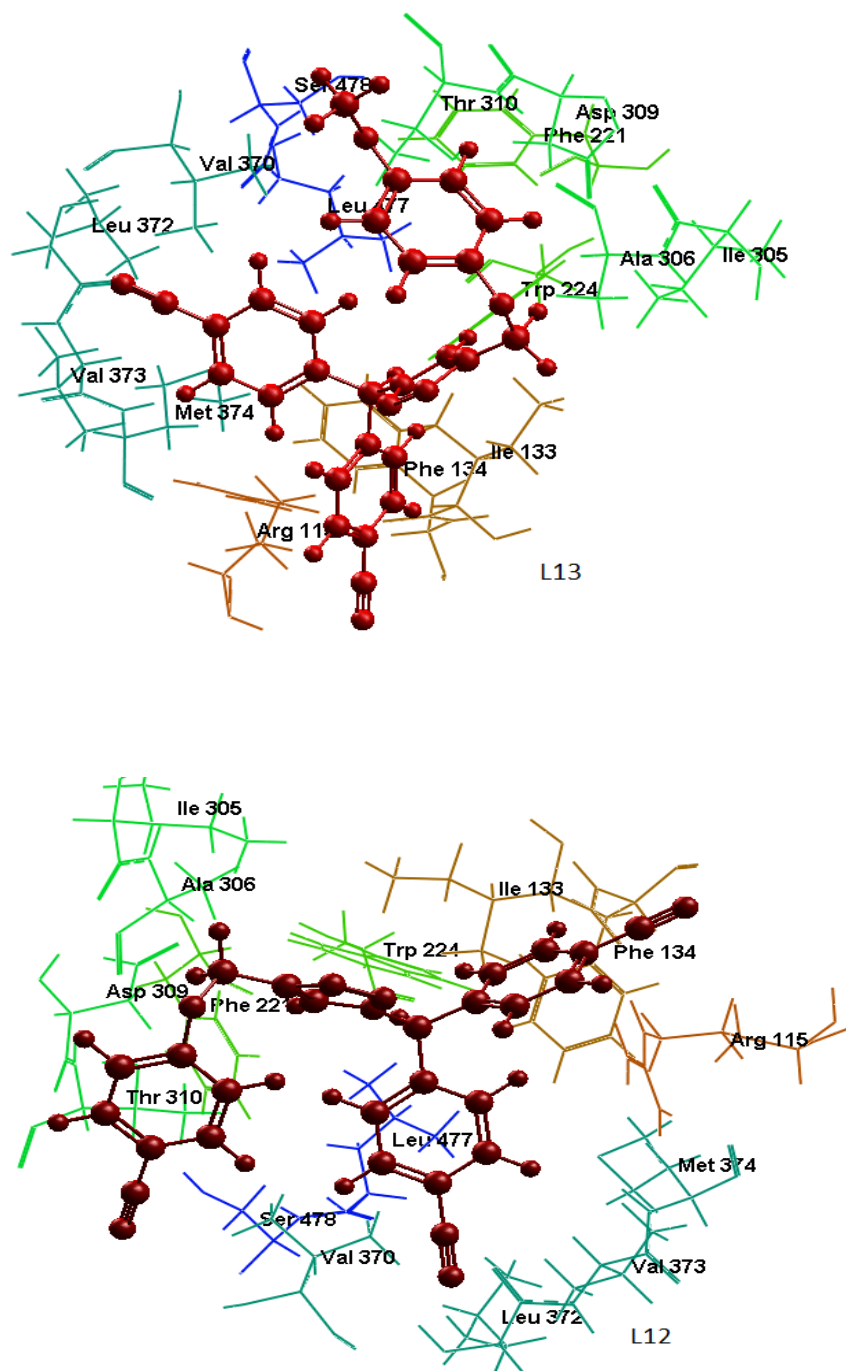

Letrozole

L14
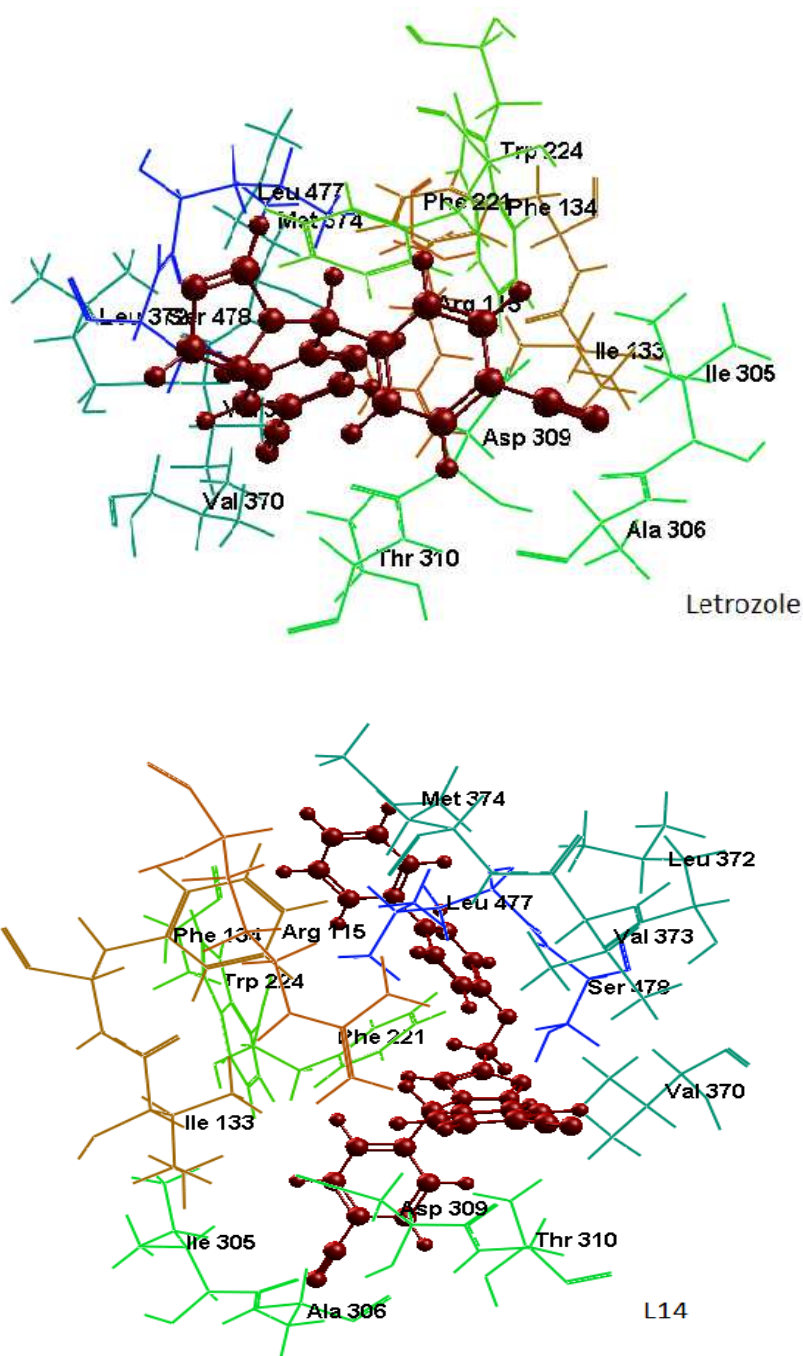
Figure.V.4 Pose organize between human placental aromatase cytochrome P450 and most active compounds and Letrozole

Table V.1.: Docking Results of 3EQM enzyme with of the compounds studied and Letrozole

Ligand	Mol Dock Score (Kcal/mol)	Interaction (Kcal/mol)	H-bond (Kcal/mol)	StericEnergy (Kcal/mol)
1	-114.854	-120.09	-2.44	3.32
2	-137.39	-141.36	-9.73	1.25
3	-142.30	-156.73	-5.02	11.26
4	-143.99	-151.37	-5.69	5.56
5	-138.84	-138.25	-6.11	-3.94
6	-134.71	-137.38	-3.56	0.69
7	-103.71	-110.53	-2.82	6.39
8	-118.15	-124.06	-5.57	3.96
9	-152.84	-155.29	-6.87	2.91
10	-152.71	-150.93	-1.74	6.57
11	-132.77	-146.23	-1.10	10.50
12	-165.8	-169	-5.08	1.66
13	-165.01	-178.04	-7.92	6.44
14	-179.89	-189.90	-5.22	6.94
15	-130.14	-140.23	-4.70	8.91
16	-141.25	-160.6	-5.70	15.11
17	-98.28	-106.75	-4.08	5.96
18	-109.72	-117.338	-6.88	6.9
19	-127.61	-134.97	-6.68	5.96
20	-122.64	-136.7	-5.40	12.90
21	-117.89	-133.05	-4.88	13.50
22	-127.59	-143.45	-6.94	14.02
23	-139.8	-152.16	-0.92	9.91
24	-131.99	-147.49	-6.57	9.05
25	-121.09	-120.45	-5.89	-2.36
26	-146.98	-156.76	-6.65	6.92
27	-121.91	-130.3	-6.61	-0.05
28	-152.2	-157.85	-5.85	3.64
29	-151.3	-179.28	-3.10	16.21
letrozole	-125.71	-130.48	-1.61	3.55

Based essentially on the comparison between the energies obtained using Mol Dock Score and the results obtained during molecular docking, it is noted that the energies of the complexes formed by ligands L14, L13, L12 Are lower compared to other complexes. -179.89, -165.8 and 165.01 kcal/mol respectively.

The energy map of 3EQM that might contribute in steric interaction favorable (green color), hydrogen acceptor favorable (turquoise color), hydrogen donor favorable (yellow color) and electrostatic potential of 3EQM (red and blue color) with the ligand viz. Letrozole, L14. L13. and L12 are shown in Fig V.5, respectively.
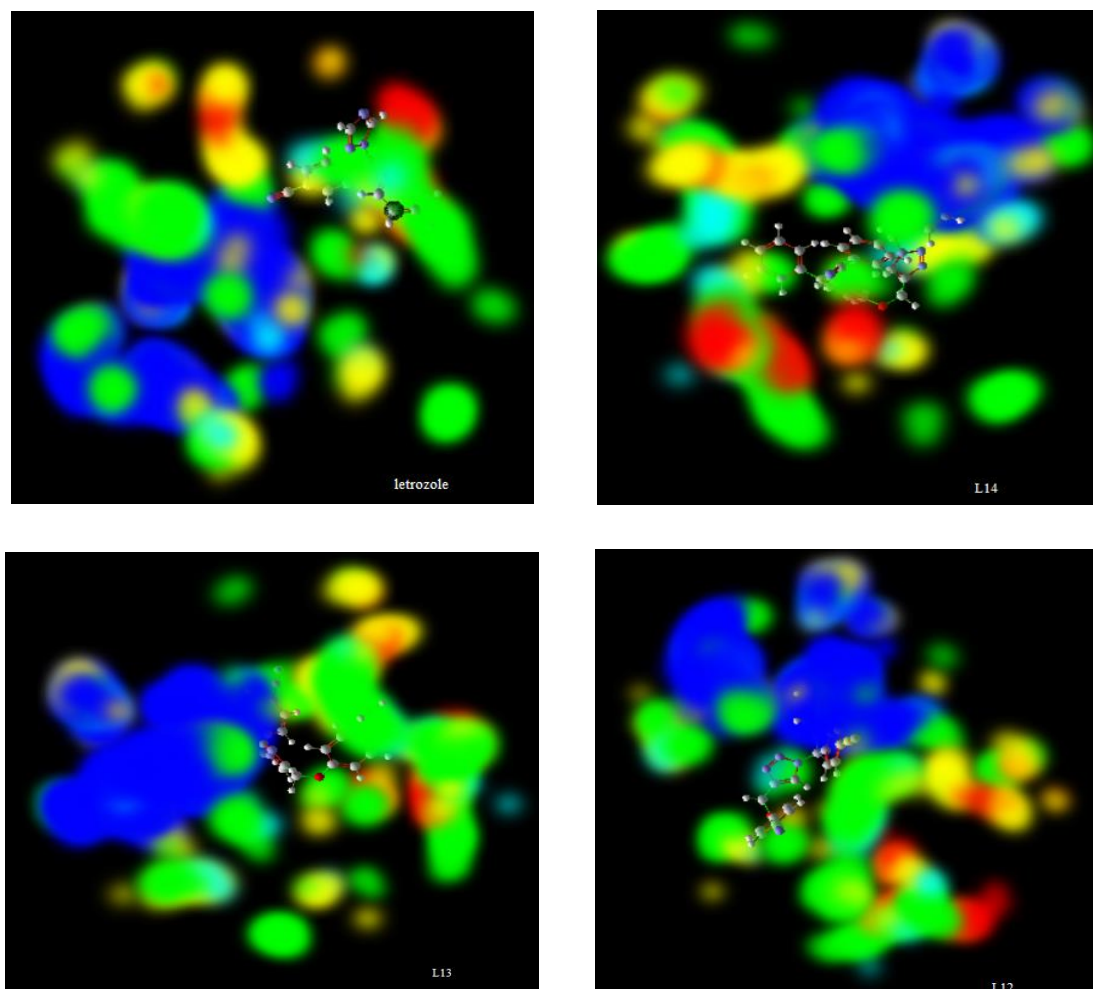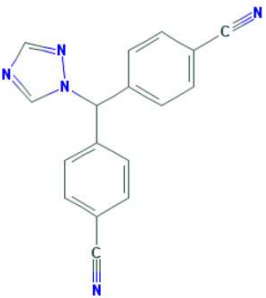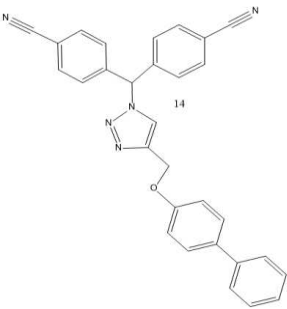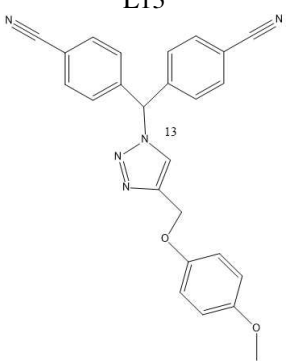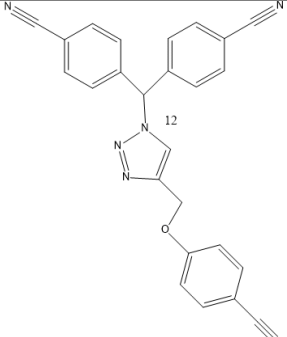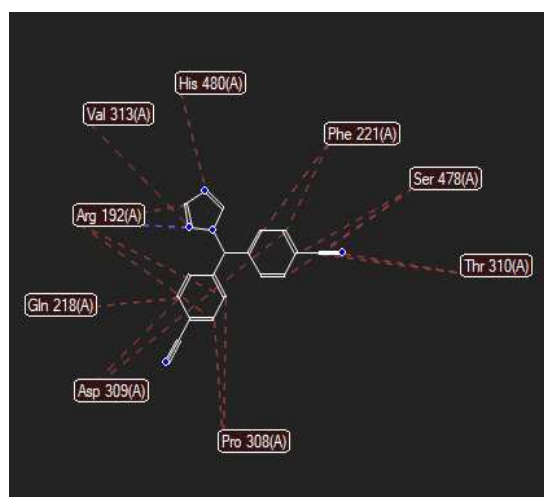

Figure.V.5 Energy map of most active compounds and Letrozole in the binding cavity of 3EQM

The main type of interaction for the letrozole atom control is the steric interaction. While The three compounds studied. The main interaction its hydrogen acceptor. Table V.2 illustrates this conclusion.
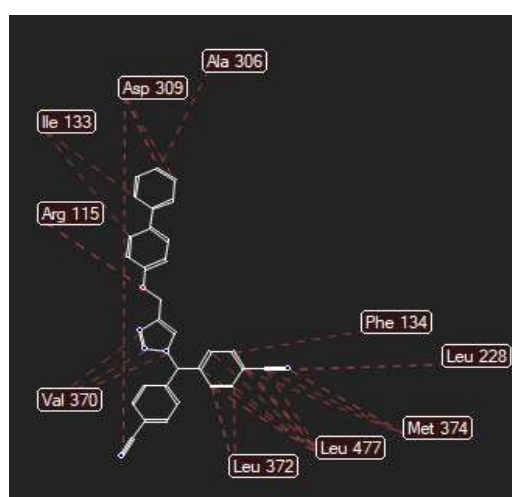
Table V.2. Summary of residues interacting with the aromatase inhibitors.

Interacting residues of receptor 3EQM			Interaction Distance(Aº)	Interaction Strength(kcal/mol)
letrozole	**h-bond**	arg192. ………..N	2.79	-1.17
	steric	Pro308………...C	3.16	0.84
		Arg192……........C	3.02	1.68
		Arg192…………C	2.65	3.92
		Arg192…………C	3.11	1.12
		Gln218…………C	3.07	1.41
		phe221………....C	3.09	1.30
		his480………….N	2.93	2.22
		val313………….C	3.11	1.16
		ASP309………...C	3.14	0.99
		ASP309………...C	3.20	0.63
		Ser478………….C	2.93	2.25
L14	**steric**	Val370………….C	2.89	2.48
		Ala306………….N	3.13	1.02
		Met374………….C	2.08	7.38
		Leu228………….N	3.12	1.10
		Arg115………….O	2.09	2.44
		Asp309………….N	2.86	2.66
		Asp309………….C	3.04	1.57
L13	**h-bond**	Arg192………….O	3.23	-0.74
		Met 374…………N	2.86	-2.50
		Arg115………….N	2.73	-1.90
	stric	Asp309…………..O	3.07	0.48
		Thr310…………..N	3.09	5.12
		Trp224…………. N	3.16	3.72
		ala306…………….N	2.69	2.07
L12	**H-bond**	Ser478………….O	2.83	-0.83
	steric	Thr310…………..N	3.16	0.83

Asp309……………C	2.95	2.13
Phe221……………N	3.14	0.95
Met374…………...N	2.70	2.50
Ala306……………N	2.95	2.15

letrozole

L14
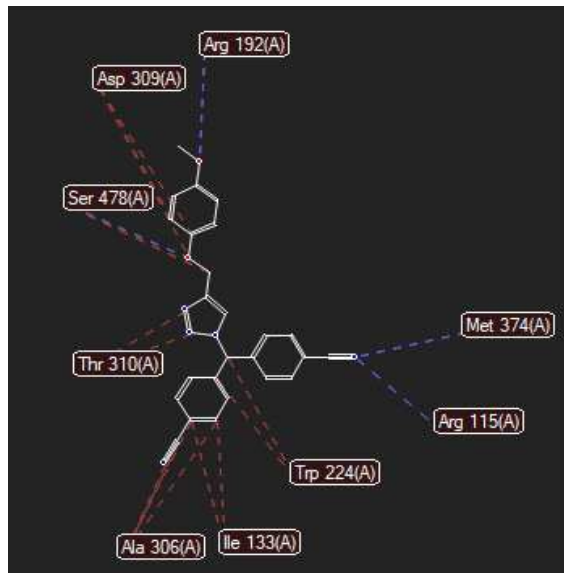
L13

L12

Figure.V.6 *Hydrogen bonding and steric interactions between aromatase receptor and most active compound*

The docked structure of L13 indicated that the CN group was extensively hydrogen bonded to the Arg115, and Met374 residues with bond distances of 3.14, and 3.34 A ° respectively (Table V.2.) (Fig V. 6). Another oxygen atom of compound L13, was also

involved in a hydrogen bond interaction with the Ser478 residue of the 3EQM. Additionally, Thr310, Ala306, Trp224, and Ile133 residues of the 3EQM displayed steric interactions with compound L13.
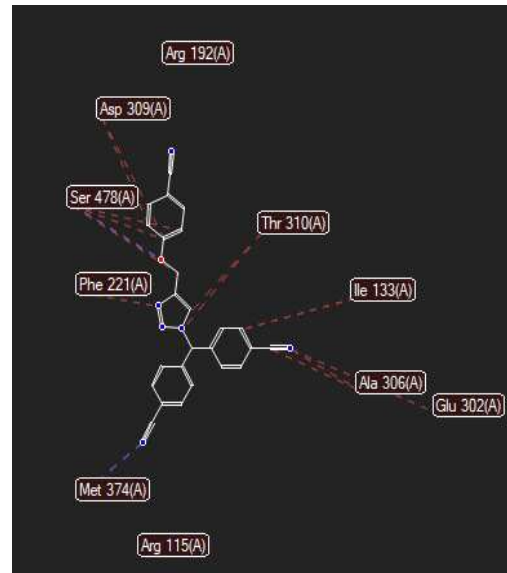
The stereo view of the docked complex indicated that both N atom (triazole cycle) of compound 12 interacted with the 3EQM by means of steric interactions (Fig.V. 6). The O atom in L12 was hydrogen bonded to the Ser 478 residues with bond distances of 2.83 A °tableV.2. Although, other H-bond interactions exist, these hydrogen bonds are relevant for inducing intrinsic activity towards highly selective and aromatase specific inhibitory property.

In particular two of the "rules" state that drugs should contain no more than 5 hydrogen bond donors and 10 acceptors. These rules arise because of the need to balance absorption and distribution properties with binding specificity within a relatively small drug molecule.

that two aromatic rings along with two hydrogen bond donors are important pharmacophores for strong aromatase inhibition. (compound 13) Moreover, the importance of hydrogen bonding in aromatase inhibition was suggested by Neves et al. [17].

The distances measured between the 3 ligands and the amino acids of the cavity 1 are cited in the table V.2.

The distances between the residues of the active site and the L14, L12, L13 vary between 2.09 Å and 3.20 Å, in this case it can be observed that according to Anne Imberty [18]., the interactions having distances between 2.5Å and 3.1Å Are considered as strong, those with distances between 3.1Å and 3.55Å are assumed to be average and when their distances are greater than 3.55Å, they are considered to be weak.
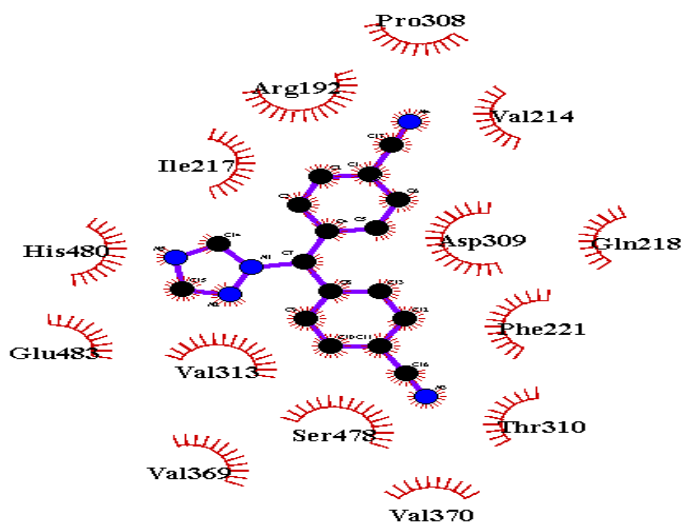
3eqm

LETRAZOLE

L14
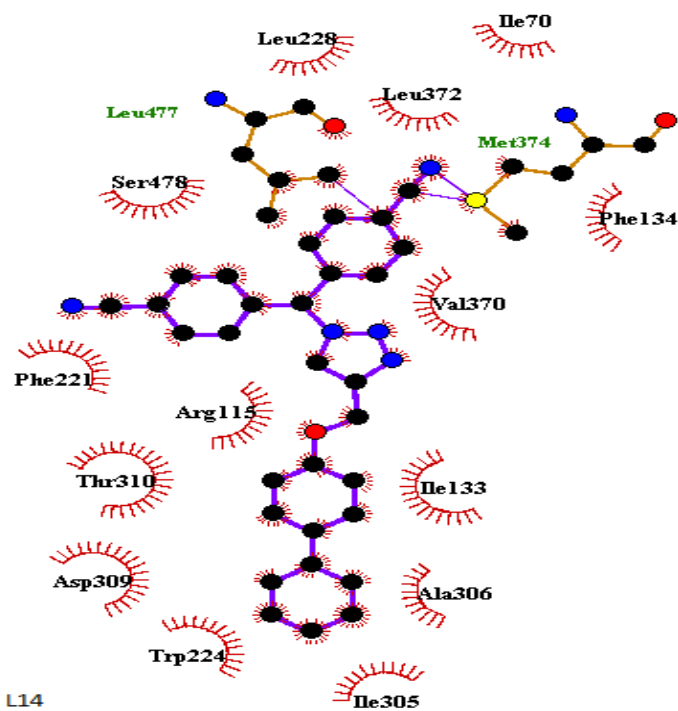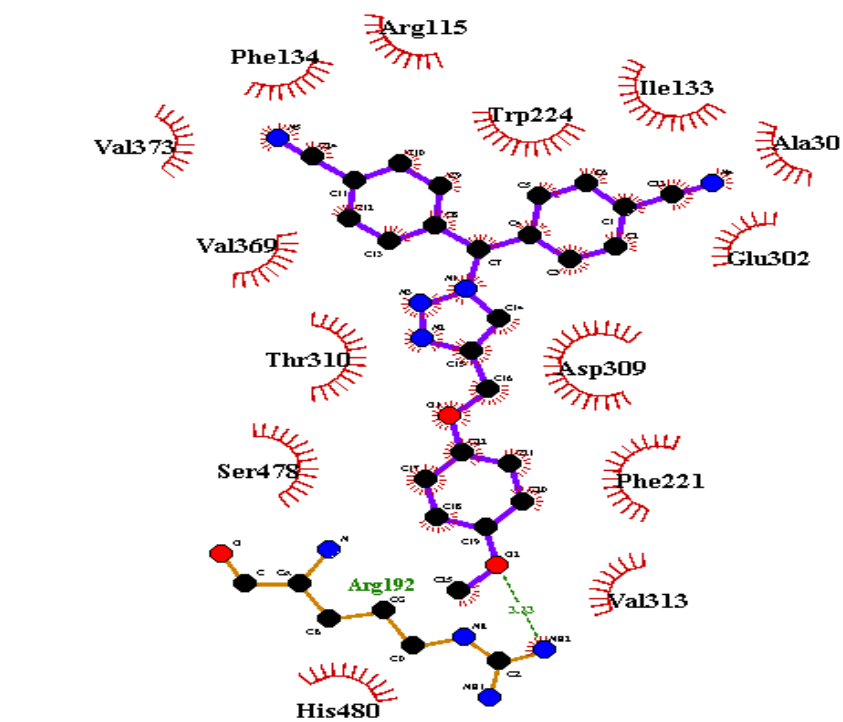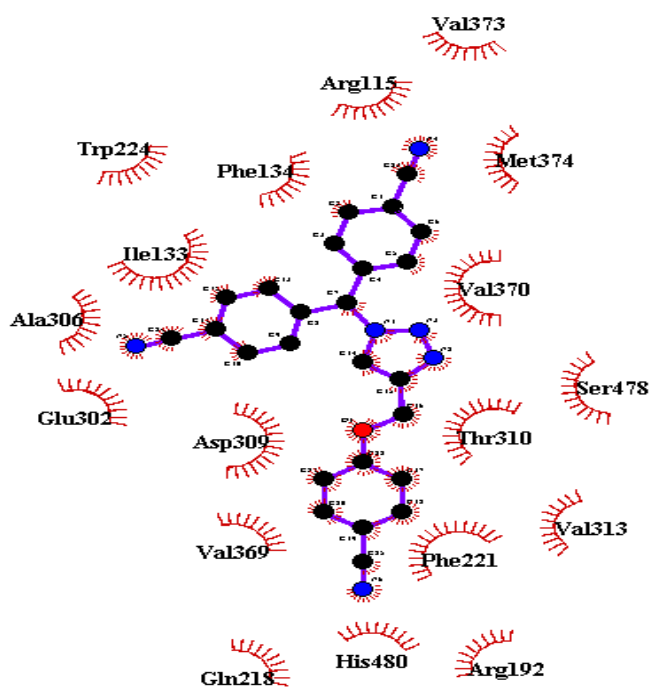
Figure V. 7: Ligplot + results showing the interactions of most active compounds and Letrozole with 3EQM.

To confirm our results, we have shown in the figure.V.7 the different interactions between the residues of the active site and the 4 ligands (fourth: letrozole ligand).

The most of amino acid residues in hydrophobic active site were involved in affinity hydrophobic bonding interactions of ligand [18]. (Figure V.8 et V.9). This figure shows hydrophobic and electrostatic bonding between the most compounds L14.L13.and L12 and letrozole with 3EQM.

According to the figure It was observed that Letrozole occupied hydrophobic pocket with the residues His480, glu302, phe221, Val373 and Val370. and are the same residues as the hydrophobic interactions for the three ligands L14, L13et L12.

letrozole   L14

L13   L12
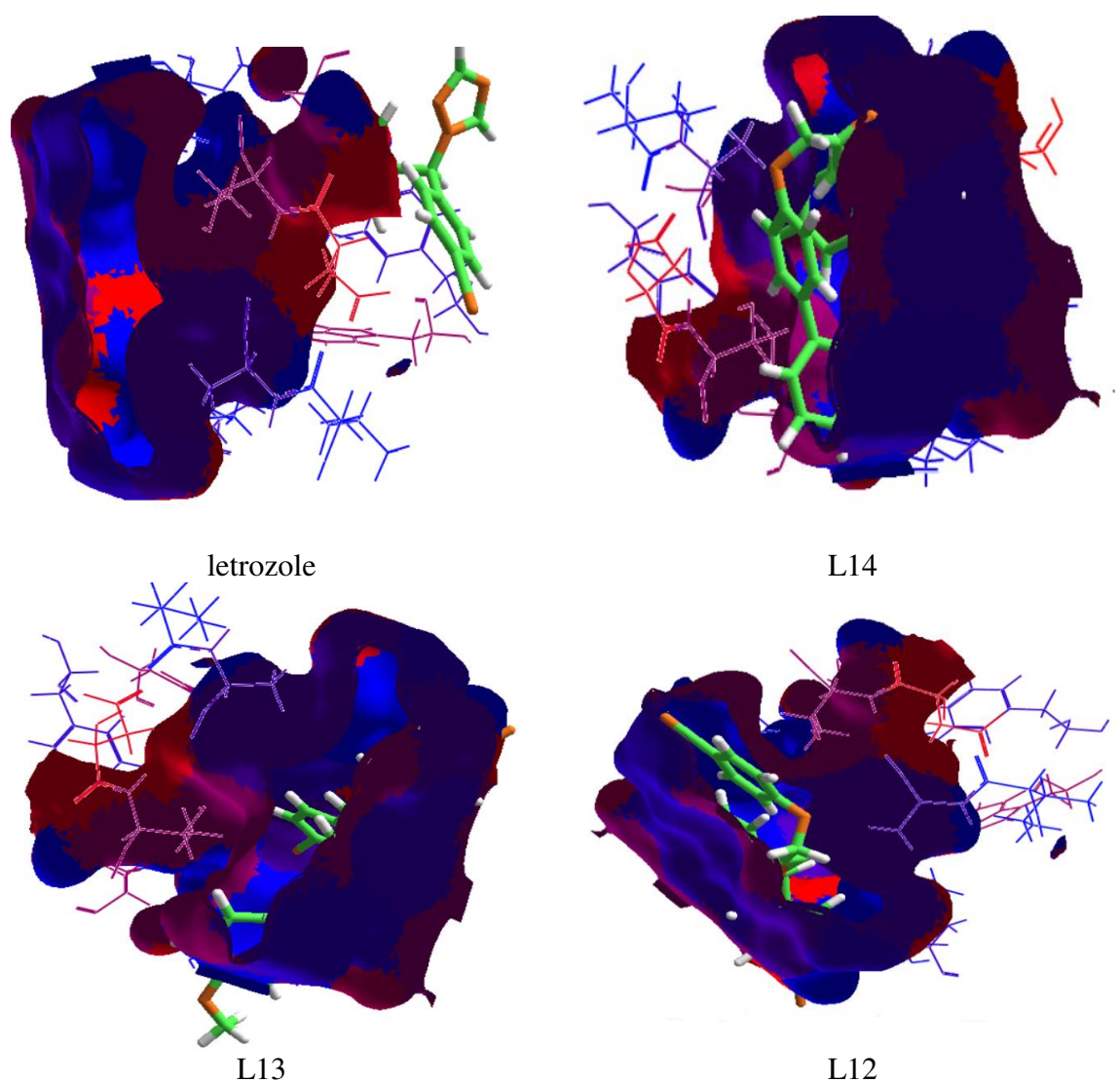
Figure.V.8 Hydrophobic bonding interactions between human placental aromatase cytochrome P450 and most active compounds and Letrozole.

letrozole   L14

L13   L12
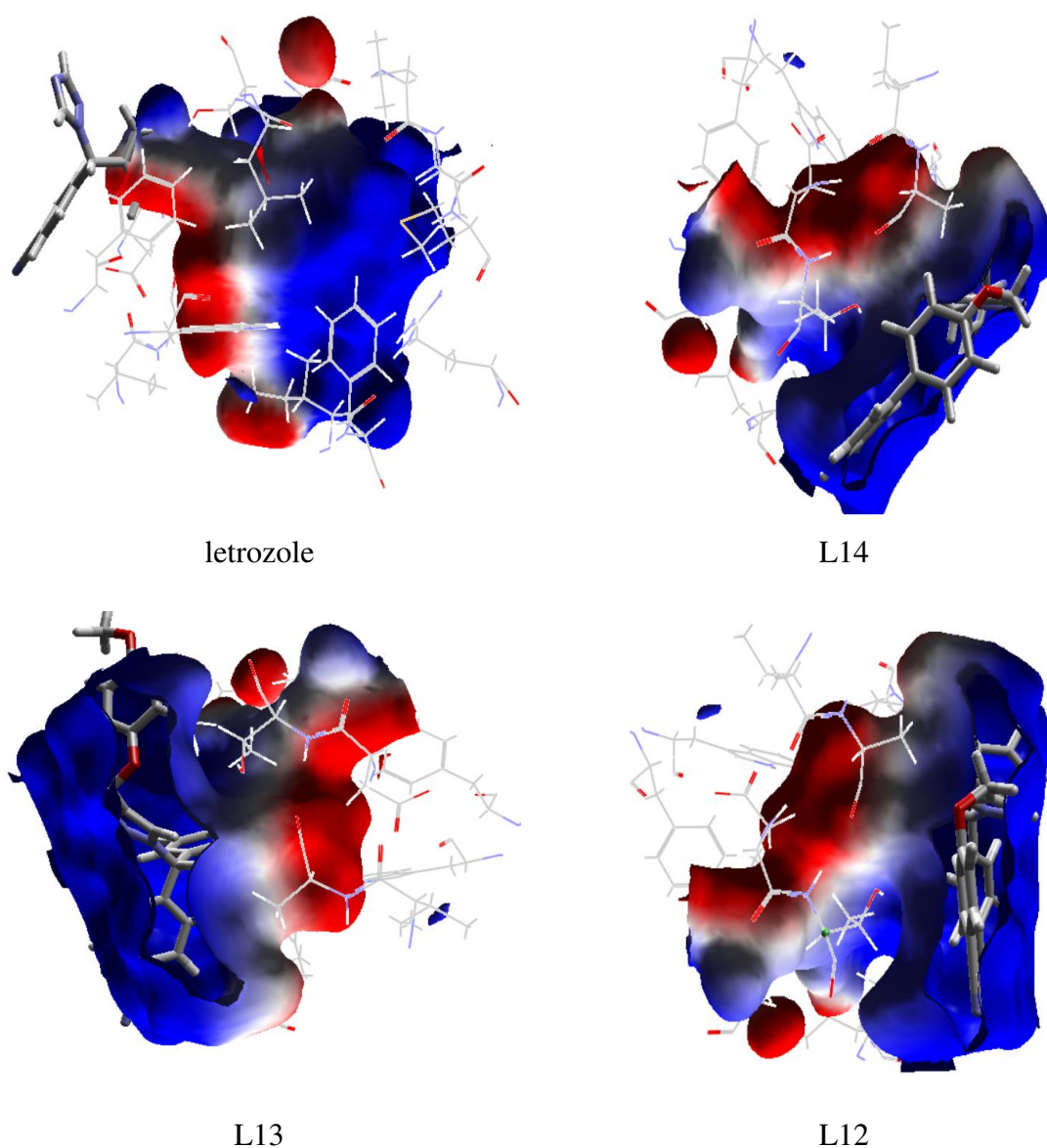
Figure.V.9 electrostatic bonding interactions between human placental aromatase cytochrome P450 and most active compounds and Letrozole

The residues comprising the catalytic cleft are Ile 305, Ala 306, Asp 309 and Thr 310 from the I-helix, Phe 221 and Trp 224 from the F-helix, Ile 133 and Phe 134 from the B-C loop, Val 370, Leu 372 and Val 373 from the K-helix–b3 loop, Met 374 from b3, and Leu 477 and Ser 478 from the b8–b9 loop). from the results it is found that all three ligands L14, L13 and L12 interact with the aromatase catalytic site residues.

The structure of most main content of 1,2,3- triazole was nonpolar, so this gave an advantage to hydrophobic for binding inside chains as receptor active site. The most active compounds have similar role binding with letrozole that has been proven as aromatase agent and occupied in cavity 1 of human placental aromatase cytochrome P450 toward

hydrophobic binding with aromatic ring, aliphatic chain and carbonyl group. Aromatic ring was important for binding.

3.2 Molinspiration Calculation

The CLogP (octanol / water partition co efficient) was calculated by the methodology developed by Molinspiration as a sum of fragment-based contributions and correlation factors. The molecular descriptors of best five compounds were tested to Lipinski's rule of five, interestingly all the ligands which we selected have molecular weight in the range of 416 – 467 (< 500). Low molecular weight drug molecules (< 500) are easily transported. Diffuse and absorbed as compared to heavy molecules[19].

Table V.3.In-silico prediction of ADME properties of compounds L14.L13 and L12.

Compound	Mi logP [c]	TPSA [f]	MW [d]	HBA [a]	HBD [b]	N ROT [e]
L14	6.22	87.54	467.54	6	0	7
L13	4.49	96.74	421.46	7	0	7
L12	4.18	111.13	416.44	7	0	6

a Number of hydrogen bond acceptors b Number of hydrogen bond donor's c Calculated octanol/water partition coefficient d Molecular weight. f Topological polar surface area e Number of rotatable bonds.

Molecular weight is an important aspect in therapeutic drug action, if it increases correspondingly, which in tum affects the drug action [20]. Number of hydrogen bond acceptors (O and N atoms) and number of hydrogen bond donors (NH and OH) in the tested compounds were found to within Lipinski limit range from 5-7 and less than 10 and 5 respectively.

Lipophilicity (log P value) and TPSA values are two important properties for the prediction of per oral bioavailability of drug molecules [21]. Permeability property of compounds were analyzed, the calculated log P value of L13 and L12 compounds was 4.49 and 4.18 (< 5). respectively. In contrast, the oral bioavailability of L14 is questionable where the mlogp was outranged. Which is the acceptable limit for the drugs to be able to penetrate through bio membranes.

Topological polar surface area TPSA was calculated as described by Veber et al[22]. O- and N-centere dpolar fragment were considered. TPSA has shown to be a very good descriptor characterizing drug absorption. Including intestinal absorption, bioavailability, Caco-2 permeability and BBB penetration. The highest degree of lipophilicity was found

with all the compounds, which are an indication for good lipid solubility that will help the drug to interact with the membranes. TPSA was calculated from the surface areas that are occupied by oxygen and nitrogen atoms and by hydrogen atoms attached to them. [23]. Thus, the TPSA is closely related to the hydrogen bonding potential of a compound. In our study, all ligands exhibited 86- 111 value of TPSA, indicates good bioavailability by oral route. Good bioavailability is more likely for compounds with ≤10 rotatable bonds and TPSA of ≤ 140 Å. As the number of rotatable bonds increases, the molecule becomes more flexible and more adaptable for efficient interaction with a particular binding pocket. Interestingly all the four compounds have 6-7 rotatable bonds and flexible.

3.3 ADMET properties

As derived from admet SAR server, reveal that L14 and L13 had better Human Intestinal Absorption (HIA) score than the control letrozole. Greater HIA denotes that the compound could be better absorbed from the intestinal tract upon oral administration. The results obtained for BBB penetrability greatly agreed with structures of the studied compounds. The compound has a less polar sesquiterpene lactone, was predicted to cross BBB, this is their order (L12, L13, L14 and letrozole). When it comes to predicting the efflux by P-glycoprotein P-gp all the compounds came out as a non-substrate and non-inhibitor of P-gp similar to our control molecule. AMES toxicity test is employed to know whether a compound is mutagenic or not. Similar to the control molecule, the l14 and l12 test ligands displayed negative AMES toxicity test, which means that the ligands are non-mutagenic. Carcinogenic profile also revealed that the ligands were non-carcinogenic similar to the control molecule.

The next important parameter is cytochrome P450 (CYP), which is known as isozymes group and it is involved in the metabolism of drugs, fatty acids, steroids, bile acids and carcinogens. Notwithstanding, some of the cytochrome P450 isoforms could be inhibited by one or more of the tested compounds.

Considering the cytochrome P450 (CYP) analysis, the major limitation of letrozole is its CYP450 3A4 substrate nature, lead to high drug–drug interaction and interruption in the and breast cancer.

Table V.4.*ADMET predictions using AdmetSAR*

Model	L14		L13		L12		letrozole	
	Result	Probability	Result	Probability	Result	Probability	Result	Probability
Absorption								
Blood-Brain Barrier	BBB+	0.9732	BBB+	0.9564	BBB+	0.9746	BBB+	0.9737
Human Intestinal Absorption	HIA+	1.0000	HIA+	1.0000	HIA+	0.9971	HIA+	0.9965
Caco-2 Permeability	Caco2+	0.5053	Caco2+	0.5369	Caco2+	0.5178	Caco2+	0.6347
Distribution								
Subcellular localization	Mitochondria	0.8808	Mitochondria	0.8460	Mitochondria	0.8475	Mitochondria	0.7175
Metabolism								
CYP450 2C9 Substrate	Non-substrate	0.7090	Non-substrate	0.6527	Non-substrate	0.7134	Non-substrate	0.7898
CYP450 2D6 Substrate	Non-substrate	0.8224	Non-substrate	0.8163	Non-substrate	0.8220	Non-substrate	0.9116
CYP450 3A4 Substrate	Non-substrate	0.5248	Substrate	0.5232	Non-substrate	0.5470	Non-substrate	0.6843
CYP450 1A2 Inhibitor	Inhibitor	0.6730	Inhibitor	0.5711	Inhibitor	0.6088	Non-inhibitor	0.8374
CYP450 2C9 Inhibitor	Inhibitor	0.6980	Non-inhibitor	0.5246	Inhibitor	0.5946	Non-inhibitor	0.9071
CYP450 2D6 Inhibitor	Non-inhibitor	0.8416	Non-inhibitor	0.8758	Non-inhibitor	0.8549	Non-inhibitor	0.9230
CYP450 2C19 Inhibitor	Inhibitor	0.7239	Non-inhibitor	0.5319	Inhibitor	0.6457	Inhibitor	0.8993
CYP450 3A4 Inhibitor	Inhibitor	0.5000	Non-inhibitor	0.5256	Non-inhibitor	0.5834	Inhibitor	0.6451
Toxicity								
Human Ether-a-go-go-Related Gene Inhibition	Weak inhibitor	0.6477	Weak inhibitor	0.5000	Weak inhibitor	0.5267	Weak inhibitor	0.9261
AMES Toxicity	Non AMES toxic	0.5678	AMES toxic	0.5240	Non AMES toxic	0.5270	Non AMES toxic	0.6371
Carcinogens	Non-carcinogens	0.8990	Non-carcinogens	0.9274	Non-carcinogens	0.9169	Non-carcinogens	0.8926
Fish Toxicity	High FHMT	0.9872	High FHMT	0.8696	High FHMT	0.9623	High FHMT	0.9605

4. CONCLUSION

The formation of a stable complex depends on the binding of the ligand in the active site. Our results, show that there are several interactions between ligands L14, L13 and L12 with the residues of the active site, which means the formation of complexes Stable, and thereafter, better binding of these ligands at the active site.

The highest degree of lipophilicity was found with all the compounds, which are an indication for good lipid solubility that will help the drug to interact with the membranes. The analysis of molecular docking gives place at the prospective identification of ligands. The molecules L14, L13 and L12 have the potential to inhibit the activity of 3EQM. The molecules L14, L13 and L12 demonstrate the better result in in-silico analysis. Moreover L13 has shown good hydrogen bonding which was not observed in our prototype Letrozole. All the designed ligands have shown no trace of carcinogenic.

Hence it has been predicted that all our designed ligands (Especially L12) can possibly act as new leads for the treatment of estrogen dependent diseases like endometriosis and breast cancer.

5. REFERENCES

1. Geisler, J., King, N., Dowsett, M., Ottestad, L., Lundgren, S., Walton, P., … Lønning, P. E. (1996). Influence of anastrozole (Arimidex), a selective, non-steroidal aromatase inhibitor, on in vivo aromatisation and plasma oestrogen levels in postmenopausal women with breast cancer. *British Journal of Cancer*, *74*(8), 1286-1291.

2. Fabian, C. J., & Kimler, B. F. (2005). Selective estrogen-receptor modulators for primary prevention of breast cancer. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *23*(8), 1644-1655.

3. Dowsett, M., Stein, R. C., Mehta, A., & Coombes, R. C. (1990). Potency and selectivity of the non-steroidal aromatase inhibitor cgs 16949a in postmenopausal breast cancer patients. *Clinical Endocrinology*, *32*(5), 623-634.

4. Kendall, A., Folkerd, E. J., & Dowsett, M. (2007). Influences on circulating oestrogens in postmenopausal women: relationship with breast cancer. *The Journal of Steroid Biochemistry and Molecular Biology*, *103*(2), 99-109.

5. Goss, P. E., & Strasser, K. (2001). Aromatase inhibitors in the treatment and prevention of breast cancer. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *19*(3), 881-894.

6. Okada, M., Yoden, T., Kawaminami, E., Shimada, Y., Kudoh, M., Isomura, Y., … Fujikura, T. (1996). Studies on Aromatase Inhibitors. I. Synthesis and Biological Evaluation of 4-Amino-4H-1, 2, 4-triazole Derivatives. Chemical and Pharmaceutical Bulletin, 44(10), 1871-1879.

7. Neves, M. A. C., Dinis, T. C. P., Colombo, G., & Sá e Melo, M. L. (2009). Fast Three-Dimensional Pharmacophore Virtual Screening of New Potent Non-Steroid Aromatase Inhibitors. *Journal of Medicinal Chemistry*, *52*(1), 143-150.

8. Doiron, J., Soultan, A. H., Richard, R., Touré, M. M., Picot, N., Richard, R., … Touaibia, M. (2011). Synthesis and structure-activity relationship of 1- and 2-substituted-1,2,3-triazole letrozole-based analogues as aromatase inhibitors. *European Journal of Medicinal Chemistry*, *46*(9), 4010-4024.

9. Daoud, I., Melkemi, N., Salah, T., & Ghalem, S. (2018). Combined QSAR, molecular docking and molecular dynamics study on new Acetylcholinesterase and Butyrylcholinesterase inhibitors. *Computational Biology and Chemistry*, *74*, 304-326.

10. Ghosh, D., Griswold, J., Erman, M., & Pangborn, W. (2009). Structural basis for androgen specificity and oestrogen synthesis in human aromatase. *Nature*, *457*(7226), 219-223.

11. Nantasenamat, C., Worachartcheewan, A., Prachayasittikul, S., Isarankura-Na-Ayudhya, C., & Prachayasittikul, V. (2013). QSAR modeling of aromatase inhibitory activity of 1-substituted 1,2,3-triazole analogs of letrozole. *European Journal of Medicinal Chemistry*, *69*, 99-114.

12. V.R. Avupati, P. N. Kurre, S. Rupa Bagadi, M. K.mar Muthyala and R. P. Yejella (2010) Denovo Based Ligand generation and Docking studies of PPARδ Agonists. Correlations between Predicted Biological activity vs. Biopharmaceutical Descriptors. 10, 74-86.

13. R. Storn, and K. Price, *Tech-report, International Computer Science Institute*, Berkley, **(1995).**

14. Lipinski C. A, Lombardo F, Dominy BW ,Feeny P J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug . Delivery. 1997. Rev,23, 3-25

15. Veber D F, Johnson S R, Cheng H Y, Smith B R, Ward K W, Kopple K D. Molecular properties that influence J. Med. Chem, 2002.45, 2615-2623

16. Cheng, F., Li, W., Zhou, Y., Shen, J., Wu, Z., Liu, G., … Tang, Y. (2012). admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. *Journal of Chemical Information and Modeling*, *52*(11), 3099-3105.

17. Neves, M. A. C., Dinis, T. C. P., Colombo, G., & Sá e Melo, M. L. (2009). Fast three-dimensional pharmacophore virtual screening of new potent non-steroid aromatase inhibitors. *Journal of Medicinal Chemistry*, *52*(1), 143-150.

18. Imberty, A., Gautier, C., Lescar, J., Pérez, S., Wyns, L., & Loris, R. (2000). An Unusual Carbohydrate Binding Site Revealed by the Structures of Two Maackia amurensis Lectins Complexed with Sialic Acid-containing Oligosaccharides. *Journal of Biological Chemistry*, *275*(23), 17541-17548.

19. Alberts, B., & Fulton, K. R. (2005). Election to the National Academy of Sciences: Pathways to membership. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(21), 7405-7406.

20. He, R., Chen, Y., Chen, Y., Ougolkov, A. V., Zhang, J.-S., Savoy, D. N., … Kozikowski, A. P. (2010). Synthesis and Biological Evaluation of Triazol-4-ylphenyl-Bearing Histone Deacetylase Inhibitors as Anticancer Agents. *Journal of Medicinal Chemistry*, *53*(3), 1347-1356.

21. V. Ochieng, P. J., Sumaryada, T., & Okun, D. (2017). Molecular docking and pharmacokinetic prediction of herbal derivatives as maltase-glucoamylase inhibitor. *Asian Journal of Pharmaceutical and Clinical Research*, *10*(9), 392-398.

22. Kelder, J., Grootenhuis, P. D., Bayada, D. M., Delbressine, L. P., & Ploemen, J. P. (1999). Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs. *Pharmaceutical Research*, *16*(10), 1514-1519

23. Ertl, P., Rohde, B., & Selzer, P. (2000). Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. *Journal of Medicinal Chemistry*, *43*(20), 3714-3717.

24. http://molegro.com/mvd-product.php

GENERAL CONCLUSION

GENERAL CONCLUSION

In this work, we applied the methods of Computational Approaches for Drug Design and Discovery, this study involved:

- In silico studies on structure activity relationships of substituted 1,2,3-triazole as aromatase inhibitor.
- QSAR model for predicting the aromatase inhibition activity of 1,2,3-triazole derivatives.
- Molecular docking studies and in silico ADMET of new substituted 1,2,3-triazole derivatives for anti-breast cancer activity.

We applied many methods of computational chemistry in this study. Quantum mechanics methods were used in the study of chemical reactivity of triazole and their derivatives, with methods: PM3, Ab initio/ (HF / 6-311 $^{++}$ G (d, p)) and DFT (B3LYP / 6-311 $^{++}$ G (d, p)), whose purpose is to determine the structural, electronic and energetic parameters associated with the molecules studied.

The efficiency of these methods used was confirmed by the comparison of the structural parameters between the results obtained by Ab initio and DFT theoretical methods. The nature of the substituent type (donor, acceptor) influences the core electronic and energetic parameters. Indeed, this study allows us to predict the of triazole derivatives.

The qualitative study of the structure-properties relationship was carried out on the 22 compounds. The molecules used in this study have pharmacological properties. The nature of the groups that bind to the basic nucleus of the molecules studied affects their physicochemical properties and consequently their pharmacological properties.

Molecular properties such as membrane permeability and oral bioavailability are usually associated with some basic molecular descriptors, such as log P (partition coefficient), molecular weight (MW), and the acceptors and donor for hydrogen bonding in a molecule.

In the pursuit of robust aromatase inhibitors, 1.2.3-triazole derivatives were employed in quantitative structure activity relationship (QSAR) study using multiple linear regression (MLR). The results demonstrated good predictive ability for the MLR model. After dividing the dataset into training and test set. The models were statistically robust internally ($R^2 = 0.982$) and the model predictability was tested by several parameters, including the external criteria (R^2pred = 0.851. CCC = 0.946). Insights gained from the present study are anticipated to provide pertinent information contributing to the origins of aromatase inhibitory activity and therefore aid in our on-going quest for aromatase inhibitors with robust properties.

As well as our interest was focused on the study of the interactions between the3EQM enzyme and the triazole derivatives, to better understand the mechanism of inhibition of these enzymes.

First, we tested molecular docking performances by Mol Dock Score and non-binding distances (Hydrogen and Steric).

The discussion was based essentially on these two parameters to explain the formation of the complexes (Enzyme-substrate), and subsequently the ligand binding at the active 3EQM site.

The study showed that the three ligands L12, L13 and L14 are the best inhibitors of 3eqm, these ligands possess the lowest energies compared to other seven with respect to the ligand of reference letrezole.

The application of Lipinski rules on the studied three ligand (L12.L13 and L14) shows that all these compounds, theoretically, will not have problems with oral bioavailability.

The last part of this work allowed us to learn about the possible pharmacokinetic properties of absorption, digestion, metabolism and excretion of the three ligands.

The ADME prediction also showed very encouraging results. Hence, these compounds may be tested in vitro as future aromatase inhibitors.

Abstract

Breast cancer is the most common type of female cancer. One class of hormonal therapy for breast cancer drugs -non steroidal aromatase inhibitors- are triazole analogues. In this work a fundamental and original research was made on the molecule of triazole heterocyclic, whose the aim is to predict the reactivity and biological activity studied of the compound. It is based on different computational and approaches used in computer aided -drug-design. *(SPR, QSAR, molecular docking, ADMET).*

A study of structure – property relationships (SPR) for 1,2,3 triazole derivatives has been carried.

A linear quantitative structure activity relationship model is obtained using Multiple Linear Regression (MLR) analysis as applied to a series of triazole derivatives with inhibitory activity of the aromatase. The accuracy of the proposed MLR model is illustrated using the following evaluation techniques*: cross validation, and external test.*

Docking process, the interaction and binding of ligands – protein were done and visualized using software Molegro Virtual Docking.

Molinspiration and ADMETSAR web servers used to calculate ADMET and physicochemical properties of the target compounds respectively.

The results are reported and discussed in the present investigation. A close agreement with experimental results was found which improves the affinity of the present work.

Key Word : 1,2,3-triazole, aromatase inhibitory, density functional theory, QSAR, MLR, ADMET, docking molecular

Résumé

Le cancer du sein est le type de cancer le plus répandu chez les femmes. Une classe d'hormonothérapie pour les médicaments anticancéreux - les inhibiteurs non stéroïdiens de l'aromatase - sont les analogues du triazole. Dans ce travail, une recherche fondamentale et originale a été faite sur la molécule de triazole hétérocyclique, dont le but est de prédire la réactivité et l'activité biologique des composés étudiés. Il repose sur des différentes approches informatiques et des calculs utilisés dans la conception de médicaments assistée par ordinateur. (*SPR, QSAR, amarrage moléculaire, ADMET*).

Une étude des relations structure-propriété (SPR) pour les dérivés du 1,2,3-triazole a été réalisée.

Un modèle quantitatif linéaire de la relation structure-activité a été obtenu à l'aide de l'analyse de la régression linéaire multiple (RLM) appliquée à une série de dérivés de triazole ayant une activité inhibitrice de l'aromatase. La précision du modèle proposé (RLM) est illustrée à l'aide des techniques d'évaluation suivantes : validation croisée et le test externe.

Le processus de docking, l'interaction et la liaison des ligands - protéines ont été réalisés et visualisés à l'aide du logiciel Molegro Virtual Docking.

Les serveurs Web Molinspiration et ADMETSAR sont utilisés pour calculer les propriétés ADMET et les propriétés physicochimiques des composés cibles, respectivement.

Les résultats sont rapportés et discutés dans la présente enquête. Un accord étroit avec les résultats expérimentaux a été trouvé, ce qui améliore l'affinité du travail actuel.

Mot clé: 1,2,3-triazole, inhibiteur de l'aromatase, théorie de la densité fonctionnelle, QSAR, MLR, ADMET, docking moléculair.

.