Université Mohamed Khider de Biskra
Faculté des sciences exactes et des sciences de la nature et de la vie
Département des Sciences de la matière

# MÉMOIRE DE MASTER

**Domaine :** Sciences de la Matière
**Filière :** Chimie
**Spécialité :** chimie pharmaceutique
**Réf. :** /

---

Présenté et soutenu par :
**Seridji Assila**
Le : 19/06/2023

# DESIGN AND SELECTION OF NEW HETEROCYCLIC MOLECULES OF MEDICAL INTEREST BY IN SILICO METHODS

**Conception et sélection de nouvelles molécules hétérocycliques d'intérêt médical par des méthodes in silico**

---

**Jury :**

| | | | |
|---|---|---|---|
| Mazri Radhia | M.C.B | Université de Biskra | Président |
| Belaidi Salah | Professeur | Université de Biskra | Rapporteur |
| Ouassaf Mebarka | M.C.A | Université de Biskra | Examinateur |

Année universitaire : 2022/2023

# ACKNOWLEDGMENTS

*First, and foremost, I would like to thank God Almighty for givingme the strength, knowledge, ability and opportunity to undertake this research study and to persevere and complete it satisfactorily. Without his blessings, this achievement would not have been possible.*

*I would like to express my great thanks to my supervisor, Professor* ***Belaidi Salah****, for his advice and guidance, which contributed to the completion of this work.*

*I would also like to thank the committee members, Mrs. President Dr.* ***Mazri Radhia ,*** *and Mrs. Examiner Dr.* ***Ouassaf Mebarka,*** *at the University of Biskra for agreeing to examine and judge the work.*

*I would like to thank all the members of group of computational and pharmaceutical chemistry,* ***LMCE*** *Laboratory at Biskra University, with whom I had the opportunity to work. They provided me a useful feedback and insightful comments on my work, especially to doctoral student* ***Chennai Yassmine*** *,and doctoral student* ***Khelfa Nedjla****.*

*At the end Thanks to all who participated (from near or far) to this work.*

# DEDICATION

*This thesis is profoundly dedicated to:*

*To my beloved parents whom raised me with a love of science and supported me in all my pursuits,*

*they are the reason for my success.*

*To my sisters Sara and Hiba, and all my family.*

*To my friends*

*To all those who are dear to me.*

# TABLE OF CONTENTS

## Chapter II: Computational approaches for drug design and discovery



## 2nd PART: RESULTS AND DISCUSSION

## Chapter III: Structural study, Electronics, MESP on the basic core using several molecular and quantum mechanical calculation methods



## Chapter IV: Qualitative study of the structure-activity relationship of a series of 1H-1, 2, 3-triazole derivatives

**2D, 3D**
2 Dimensional and 3 Dimensional

**A**

| | |
|---|---|
| ADME | Absorption,Distribution,Metabolism, and Excretion |
| AIDS | Acquired Immune Deficiency Syndrome |
| AMBER | Assisted Model Building with Energy Refinement |
| ANN | Artificial Neural Network |
| ATP | Adenosine Triphosphate |

**B**

| | |
|---|---|
| B3LYP | Becke, 3-parameter, Lee-Yang-Parr |

**C**

| | |
|---|---|
| CADD | Computer-Assisted Drug Design |
| CAT | Catalase |

**D**

| | |
|---|---|
| DFT | Density Functional Theory |
| DNPH | Dinitrophenylhydrazone |
| DNA | Deoxyribonucleic Acid |
| DPPH | 1,1-diphenyl-2-picrylhydrazyl |

**F**

| | |
|---|---|
| FR | Free Radical |

**H**

| | |
|---|---|
| HAT | Hydrogen Atom Transfer |
| HBD | Hydrogen-Bond Donors |
| HBA | Hydrogen-Bond Acceptors |
| HE | Hydration Energy |
| HF | Hartree-Fock |
| $H_2O_2$ | Hydrogen peroxide |
| HOMO | Highest Occupied Molecular Orbital |

**I**

| | |
|---|---|
| IC50 | Half Maximal Inhibitory Concentration |

**L**

| | |
|---|---|
| LE | Ligand Efficiency |
| LipE | Ligand lipophilicity efficiency |
| LOO | leave-one-out cross-validation |
| LogP | partition coefficient octanol/water |
| LUMO | Lowest Unoccupied Molecular Orbital |

**M**

| | |
|---|---|
| MLR | Multiple Linear Regression |
| MM | Molecular Mechanics |
| MP2 | Møller-Plesset level 2 |
| MPO | Multi-Parameter Optimization |
| MW | Molecular Weight |
| MESP | Molecular electrostatic surface potential |

**N**

| | |
|---|---|
| NH | number of heavy atoms |
| NRB | Number of Rotatable Bonds |

**O**

| | |
|---|---|
| $O_2^{\bullet-}$ | superoxide anion |

**P**

| | |
|---|---|
| PE | Predictive Error |
| PM3 | Parameterized Model number 3 |
| POL | Polarizability |
| PRESSE | Predicted Residual Sum of Squares |
| PSA | Polar Surface Area |

**Q**

| | |
|---|---|
| QM | Quantum mechanics |
| QSAR | Quantitative Structure–Activity Relationship |
| QSPR | Quantitative Structure–Property Relationship |

**R**

| | |
|---|---|
| $R^{\bullet}$ | Alkyl Radicals |
| RNS | Reactive Nitrogen Species |
| $RO^{\bullet}$ | Alkoxyles Radical |
| $ROO^{\bullet}$ | Peroxles Radical |
| ROS | Reactive Oxygen Species |

**S**

| | |
|---|---|
| SAG | Surface Area Grid |
| SAR | Structure–Activity Relationships |
| SE | standard error |
| SET | Single Electron Transfer |
| SOD | Superoxide Dismutase |
| SSY | Sum of The Squares of The Response Value |

**T**

| | |
|---|---|
| TPSA | Topological polar surface area |

**V**

| | |
|---|---|
| VDW | Van der Waals |

# General Introduction

The antioxidant activity remains a very dynamic axis of multidisciplinary scientific research. In 1956, Harman hypothesized that free radicals generated in the body were responsible for damaging many cellular components such as lipids, proteins, and nucleic acids…causing early signs of aging. Since then, the link between the spread of free radicals commonly known as Reactive Oxygen Species (ROS) and certain less serious pathologies continues to be confirmed, including cardiovascular diseases, inflammatory diseases , and some cancers, and others.

The human body has its own defense mechanism against the generation and propagation of ROS. These are endogenous enzyme systems such as glutathione peroxidase, superoxide dismutase, catalases and others . Nevertheless, an excessive generation of these predatory free radical species can cause an imbalance leading to oxidative stress. For detoxification, the body needs certain compounds known for their antioxidant activity provided mainly by food.(1)

Triazole nuclei, one of the most important and well-known heterocyclic compounds, are a common and indispensable feature of a wide variety of natural products and pharmaceuticals. Triazole nuclei are present as core structural components in many drug classes such as: Antibacterial, anti-inflammatory, analgesic, antiepileptic, antiviral, antineoplastic, antihypertensive, antimalarial, antidepressant, antihistamine, antioxidant etc...(2)

The structure-activity relationship (SAR) is a study that seeks to understand how changes in molecular structure result in changes in molecular properties, some of which are related to changes in biological potency. The SAR study follows rigorous rules and steps to end up with valid and reliable models allowing to synthesize a very large number of molecules simultaneously and to test their actions on therapeutic targets and give very attractive results.

QSAR (Quantitative Structure-Activity Relationship) modeling is a method that attempts to identify a quantitative relationship between the activity measured experimentally and theoretical or empirical descriptors related to the molecular structure. It could be implemented in laboratories and used in the pharmaceutical industry. The objective of QSAR modeling is to find accurate, applicable and robust models in order to find a relationship between structure and activity for the purpose of prediction but also interpretation.(3)

In our work, we report a computational study of 1H-1,2,3-triazole and a series of its derivatives; the main objective of this work is to apply the methods provided by computational chemistry to determine the structural and electronic properties of 1H-1,2,3-triazole; finally correlations between the chemical structure and the biological activity of a series of bioactive molecules to predict the expected biological activities in relation to their antioxidant activity.

The manuscript of this work is presented in five chapters divided into two parts; after a general introduction:

- The first part concerns a bibliographical synthesis, comprises two chapters containing respectively:
  - **Chapter 1:** Bibliographical research on stress oxidative and antioxidants activity.
  - **Chapter 2:** Bibliographical research on computational approaches for drug design and discovery .
- As for the second part, in which we will present and discuss the results of our calculations obtained for our study, it consists of three chapters containing, respectively:
  - **Chapter 3:** Structural study, Electronics, MESP on the basic core 1H-1,2,3-triazole using several molecular and quantum mechanical calculation methods.

In this chapter we made a comparison of the results obtained by the two calculation methods (MP2, DFT) on the molecule of 1H-1,2,3-triazole and experimental data.

  - **Chapter 4:** Qualitative study of the structure-activity relationship of a series of 1H-1,2,3-triazole derivatives and the application of selection methods (MPO, druglikeness, Golden triangle).

This chapter is devoted to the study of the structure activity/property relationship and the drug likeness properties of a bioactive series of 1H-1,2,3-triazole derivatives.

  - **Chapter 5:** Quantitative study of the QSAR properties of a series of 1H-1,2,3-triazole derivatives and application of chemometric methods.

Finally the last chapter includes a quantitative study which aimed to describe the relationships between the physicochemical properties and the biological activity of a series of bioactive derivatives of 1H-1,2,3-triazole by the use of a QSAR Model and the description of the statistical method used MLR and ANN , for the development of a statistical model of the biological activity.

At the end of this manuscript, we end with a general conclusion, which sums up our work.

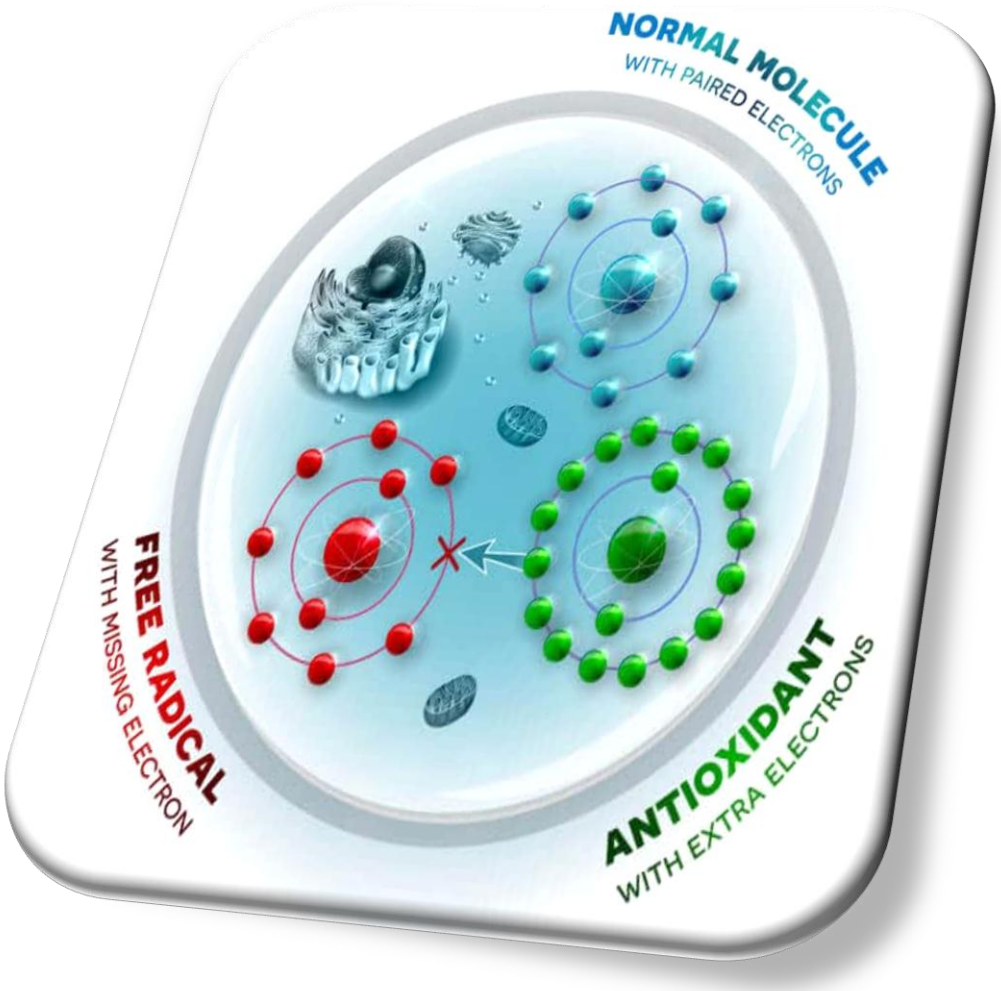
## References

(1). Ben Sghier FZ, Aid A. Conception de nouvelles molécules à activité anti-oxydante par des études 3D QSAR [memoir master]: University of mohamed khider biskra;2021.

(2). Kharb R, Sharma PC, Yar MS. Pharmacological significance of triazole scaffold. Journal of enzyme inhibition and medicinal chemistry. 2011;26(1):1-21.

(3). Belouerghi S, Azri S. Etude qualitative et quantitative des relations structure-activité d'une série de composés phénoliques [memoir master]: University of mohamed khider biskra; 2020.

# *Chapter I: General information on stress oxidative and antioxidants activity*



## I.1. Introduction:

Oxidative stress in biological systems is a complex process characterized by an imbalance between free radical (FR) production and the body's ability to eliminate these reactive species through the use of endogenous and exogenous antioxidants. In metabolic processes, a variety of reactions occur with reactive oxygen species (ROS) as promoters, such as hydrogen peroxide ($H_2O_2$) and superoxide radical anion ($O_2^{\bullet -}$), among others. A the presence an excess of ROS in the biological system can cause various pathologies, from cardiovascular diseases to promoting cancer. Biological systems have antioxidant mechanisms to control the enzymatic and non-enzymatic damage that allows ROS to be inactivated. Endogenous antioxidants are enzymes such as superoxide dismutase (SOD), catalase (CAT), glutathione peroxidase, or non-enzymatic compounds such as bilirubin and albumin.

When the body is exposed to high levels of ROS, the endogenous antioxidant system is disrupted and therefore does not ensure full protection of the body. To compensate for this lack of antioxidants, the body can fall back on exogenous antioxidants, which are supplied to it through food, dietary supplements or drugs. The most important exogenous antioxidants include phenolic compounds, carotenoids and vitamin C as well as some minerals such as selenium and zinc.

The reaction mechanism of antioxidant compounds is closely related to the reactivity and chemical structure of FR and the environment in which these active substances exist. Therefore, it is important to characterize ROS and, to a lesser extent, reactive nitrogen species (RNS) including precursors and free radicals.

There are many in vitro methods in the literature for assessing the effectiveness of antioxidant compounds in various matrices (plant extracts, serum, etc.) using hydrophilic, lipophilic, and amphiphilic media (emulsions). In vitro methods can be divided into two broad categories: 1/ single electron transfer reactions (SET) and 2/ hydrogen atom transfer reactions (HAT). These methods are widely used due to their speed and sensitivity.(1)

## I.2. Oxidative stress:

In general, oxidative stress is defined as the result of an imbalance in the balance between oxidative species and defense systems (antioxidants), resulting in the appearance of often irreversible damage to the cell.

Under physiological conditions, oxygen produces reactive species of oxygen (ROS, for reactive oxygen species) particularly toxic for cellular integrity, this is particularly the case in the mitochondria of a cell. Many ROS are free radicals, with oxidizing properties that bring them to react with their chemical environments, including a series of importance organic substrates (lipids, proteins, DNA, sugars, etc.). At the molecular level, these ROS can also play the role of secondary messengers by activating different factors or indirectly genes involved in the development of various pathologies.

It is important to note that oxidative stress can also be the consequence of the effect of environmental factors. Indeed, modern life confronts our body with factors causing an overproduction of ROS such as pollution, absorption of alcohol or medication, sun exposure or smoking. This leads either to a weakening of our antioxidant defenses or to the direct synthesis of ROS, which can cause cellular damage. Oxidative stress can develop during cardiovascular surgery, organ transplantation or respiratory distress but also in the event of an intense exercise poorly mastered. In parallel with these attacks, the situation worsens due to a less healthy and less balanced diet than before.

Determine an individual's oxidative stress status currently becomes a priority subject in terms of disease prevention. Indeed, many studies indicate that there is an association between the alteration of antioxidant defense systems and the development of more than 200 different physiopathology such as atherosclerosis or cancer, inflammatory diseases, diabetes and aging. (2)

### I.2.1. Reactive species:

Reactive species (which can be radical or not) are divided into two main categories:on the one hand ROS and on the other hand the reactive species of nitrogen (RNS).

Free radicals are a significant proportion of ROS. A free radical is defined as a chemical species, (e.g., atom or molecule) with one or more single electron (s) on its external layer of valence giving it great instability and therefore great reactivity. The genesis of these free radicals inside the organism is a physiological process in response to an exogenous factor such as aggression by microorganisms, heavy metals, ionizing radiation, ultra-violet rays or even cigarette smoke.

### I.2.1.1. Reactive oxygen species:

The ROS are mainly produced inside 2 cellular sites: on the one hand mitochondria and on the other hand the plasma membrane.(3)

❖ **Superoxide anion $O_2^{\bullet-}$:**

The superoxide anion is generated by different enzymatic systems, notably oxidases (E.G., NADPH oxidase in the lipid and cytochrome oxidase membrane in the mitochondrial respiratory channel). This reaction is made thanks to the transfer of an electronic oxygen -to -enezymatic cofactor electron.

$$O_2 + \acute{e} \xrightarrow{oxidase} O_2^{\bullet-}$$

Superoxide anion is often one designated as one of the species allowing the production of many other reactive oxygen species.

❖ **Hydrogen peroxide $H_2O_2$:**

The double prototonation of the superoxide anion $O_2^{\bullet-}$ can occur under favorable conditions in an acid milieu. This reaction is catalyzed by superoxide dismutase (SOD) leading to the formation of hydrogen peroxide $H_2O_2$.

$$2O_2^{\bullet-} + 2H^+ \xrightarrow{SOD} H_2O_2 + O_2$$

$H_2O_2$ Hydrogen peroxide is not a free radical, and therefore has greater stability compared to free radicals. However, it is considered a reactive oxygen species because it intervenes as a synthetic intermediary of other free radicals. Its physico-chemical properties allow it to cross the lipid membrane, it therefore plays an important role as a vector and diffuser of oxidizing stress.

❖ **The hydroxyl radical $OH^{\bullet}$:**

The hydroxyl radical is very toxic, it is formed by the very slow reaction.

$$H_2O_2 + H^+ \rightarrow OH^{\bullet} + H_2O$$

The hydroxyl radical can also come from the reduction of hydrogen peroxide by a transitional metal such as the ferrous ion. This reaction is called the Fenton reaction.

$$H_2O_2 + Fe^{2+} \rightarrow OH^{\bullet} + OH + Fe^{3+}$$

It is, and by far, the most reactive and toxic ROS, its lifespan being extremely short in the order of $10^{-9}s$. The hydroxyl radical reacts in a non-specific way with its environment

(such as DNA or proteins), it therefore intervenes as an initiator of lipid peroxidation resulting as the degradation of the lipid membrane. (4)

- **Alkyl radicals $R^{\bullet}$ , Alkoxyles $RO^{\bullet}$ and Peroxles $ROO^{\bullet}$:**

The oxidation of the cellular membrane by the hydroxyl radical leads to the formation of $ROO^{\bullet}$ peroxle radicals. The degradation of the latter according to the reaction of Fenton leads to the formation of highly reactive alkoxle radicals.

$$ROOH + Fe^{2+} \rightarrow RO^{\bullet} + OH + Fe^{3+}$$

### I .2.1.2. Reactive nitrogen species:

Reactive nitrogen species come from the reaction of ROS with nitric oxide $NO^{\bullet}$. These reactions are catalyzed by NO synthase, of which there are 3 types: neuronal, endothelial or inducible.

Under physiological conditions, nitrogen oxide is weakly reactive. It is the action of the superoxide anion on the nitrogen oxide which leads to the formation of reagents having a high reactivity. (5)

## I.2.2. Consequences of oxidative stress:

The main danger from free radicals comes from the damage they can cause when they react with important cellular components, such as DNA (6) , lipids (peroxidation), proteins (7) etc.

This oxidation causes damage throughout the body, accelerating aging (cardiovascular and neurodegenerative diseases, cancer, diabetes.) (8) and the degradation of cells and tissues. (9)

### I .2.2.1. DNA oxidative damage:

DNA is a molecule very sensitive to the attack by oxygen radicals, five main classes of mediated oxidative damage can be generated. The radical attack can be direct and lead to the oxidation of the bases, generating a large number of modified bases: Oxo guanine, 8 nitro guanine, formamidopyrimidine, 8 Oxo adenine, formimidouracile, 5 hydroxy cytosine, 5 hydroxy methyl uracile, thymine diol, oxazolone. But oxidative stress can also attack the bond between the base and deoxyribose, creating an abasic site, or attacking sugar itself, creating a single strand chain. (10) Indirect damage can result from the attack of lipids in

peroxidation generates mutagenic aldehydes, forming added on the basics of MDA-Guanine or Ethenoderivative type DNA.

The radical attack of proteins which are very numerous to come into contact with the DNA to protect it (histones) or to read it (enzymes and factors of replication or transcription), leads to bridging of proteins or adducts on bases of the lysinoguanine type.(11)

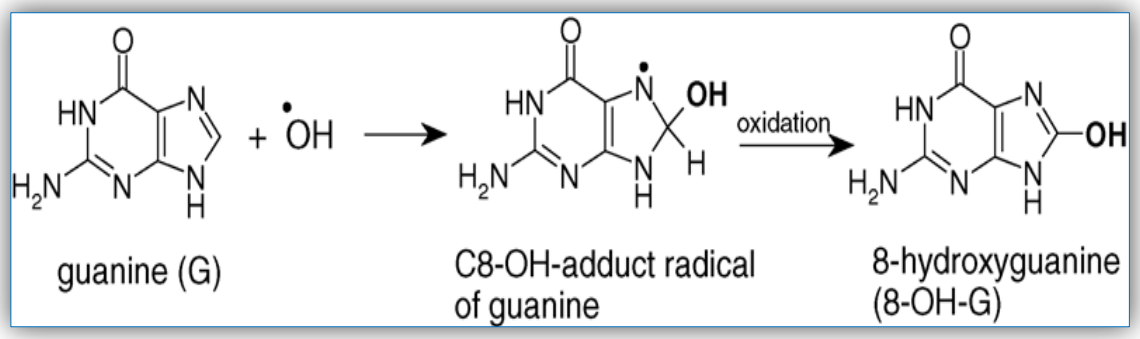



**Figure I.1. Reaction of guanine with hydroxyl radical.** (12)

**I .2.2.2. Lipid peroxidation:**

Lipids and mainly their AGPS are the privileged target of the attack by the hydroxyl radical capable of tearing a hydrogen from the carbon located between two double links.(10) This reaction is called lipid peroxidation. The Radical formed undergoes an internal rearrangement due to the displacement of the double connection closest to the single electron. (13)

Lipid peroxidation causes an increasing increase in the permeability of cell membranes inducing irreversible alteration of the functional properties of the cell, which can go as far as complete lysis.(10)

**I .2.2.3. Protein oxidation:**

The changes in primary, secondary and tertiary structures of proteins by free radicals are the basis of the formation of carbonyl protein derivatives via several mechanisms including the fragmentation and oxidation of amino acids. In the presence of dinitrophenylhydrazone (DNPH). The most sensitive to their action are aromatic amino acids such as tryptophan and tyrosine on which the radical is $OH^{\bullet}$ added, modifying the conformation of the protein. On the containers' amino acids atom of sulfur such as cysteine and methionine, oxidation by free radicals leads to the formation of disulfide bridges. (14)

## I .2.3. Oxidative stress diseases:

### I .2.3.1. Type 2 diabetes:

Type 2 diabetes or non -insulin dependent diabetes is a disease characterized by too high a blood glucose level. Its origin is metabolic, it comes from the progressive and insidious modification of the carbohydrate metabolism. Thus glucose is less well absorbed in the adipose and muscular tissues causing an increase in blood glucose. Its incidence increases significantly with age, clinical signs are generally observed after 40 years and the disease is diagnosed at the average age of 65 years. It is important to note that this diabetes must be differentiated from that of type 1 which results from an absence of active insulin.

There is a genetic component to type 2 diabetes but it cannot in any way explain the occurrence of the pathology on its own. This is correlated with age and the interaction between the genome and living conditions. Its incidence increases with age due to changes in cells and organs. In addition, physical inactivity, hypercholesterolemia and overweight are risk factors for the occurrence of this pathology.(15)

### I .2.3.2. Inflammation:

Inflammation is characterized by a response from living tissue to an assault.This phenomenon involves the immunity process which can be natural (phagocytosis for example) or specific (cellular or humoral). Inflammation is not synonymous with infection but an infection can be the cause of an inflammatory phenomenon.The different factors inducing inflammation are:

- Physical aggression (e.g., cold, heat).

- A chemical aggression (e.g., acid, base).

- An infection due to bacteria or viruses.

- A primary or secondary immune reaction (which is the consequence of reintroduction in the body of an antigen).

- Tissue necrosis.

Inflammation will aim to allow the elimination of the aggressor and cellular debris, as well as tissue repair. (16)

### I .2.3.3. Atherosclerosis:

High blood cholesterol helps the formation of the plates internal artery, if these plates are large they cut blood circulation, but most of the damage occurs when blood clots stand out and travel to the heart or the brain, which can trigger a heart attack or a stroke. Oxidative stress plays a critical role in plaque formation and with inherent vascular inflammation, can be a strong indicator of atherosclerosis.(17) Inflamed cells produce free radicals involved in cell degradation. In case of vascular inflammation, the endothelial cells of the vessels can be activated to produce oxidants.

### I. 2.3.4. Cancer:

Cancers involve several types of cells or tissues. The cancer cell, causing masses that gradually invade the territory of other organs. This invasion causes numerous damage and frequently leads to the death of the individual. The tumor cells have an abnormally increased production of reactive oxygen species. Human DNA is subject to the action of free radicals, 20 alterations of DNA per day which leads to the creation of mutant proteins., which can constitute an immortal cellular line recognized by the immune system which will fight by activating the defense mechanisms.

Consequently, the rate of free radicals will increase nearby and will be likely to mutate the surrounding cells again.(14)

## I.3. Antioxidants:

### I.3.1. Definition:

An antioxidant is by definition a more or less complex chemical species reducing oxidative stress within the organism. An antioxidant can therefore: (i) prevent the synthesis of free radicals by inhibiting the initiation of the reaction channels described above or (ii) directly deactivate the ROS. Antioxidants can be classified according to their modes of action: enzymatic systems, oxidant enzyme inhibitors, metal chelators and free radical traps.

The body has endogenous systems dedicated to this protective action. However, this line of defense is easily saturated. Many exogenous antioxidants are also present in the diet providing significant support in the antioxidant control. We find them in the fruits (apples, pears, red fruits ...), vegetables (broccoli, onion ...), drinks (coffee, tea, wine ...) as well as in spices, cocoa or cereals. These antioxidants are best known for their ability to react directly to react with free radicals by “neutralizing” them by reduction reaction. (5)

### I.3.2. Enzymatic antioxidant systems:

Among antioxidant defences, enzymes such as superoxide dismutase, glutathione peroxidase and catalase, remove free radicals and are located in subcellar organelles or in the cystol of eukaryotic cells.(18)

#### I.3.2.1. Catalase:

Catalase is a common enzyme present in almost all living organisms, which are exposed to oxygen, where it works to quickly catalyze the decomposition of hydrogen peroxide into gaseous oxygen and less reactive water molecules, by the following reaction:

$$2H_2O_2 \rightarrow O_2 + 2H_2O \quad (14)$$

#### I.3.2.2. Glutathione peroxidase:

Glutathione peroxidase (GSH-PX) is a selenium enzyme with the property of being able to catalyze the reduction of hydroxyperoxides. It is found in extracellular liquids as well as in cells, within cytosol and mitochondria. (19) It consists of 4 subunits each containing an atom of selenium. There are 5 isoforms from this enzyme varying depending on their location in the organism.

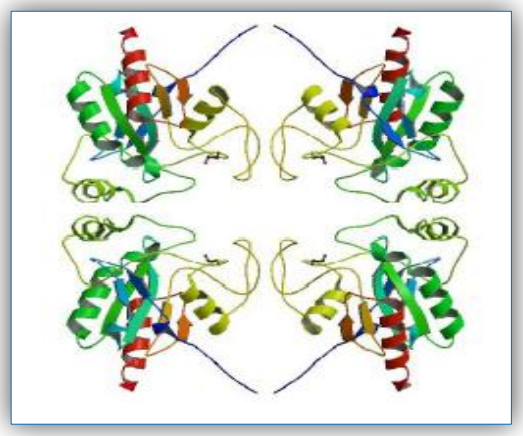


**Figure I.2. Three -dimensional structure of the glutathione peroxidase.** (20)

Its main activity is to allow the oxidation of glutathione by dimerization reaction with the formation of a disulphide bridge. This reaction leads to the generation of $H_2$ which can reduce the surrounding species.

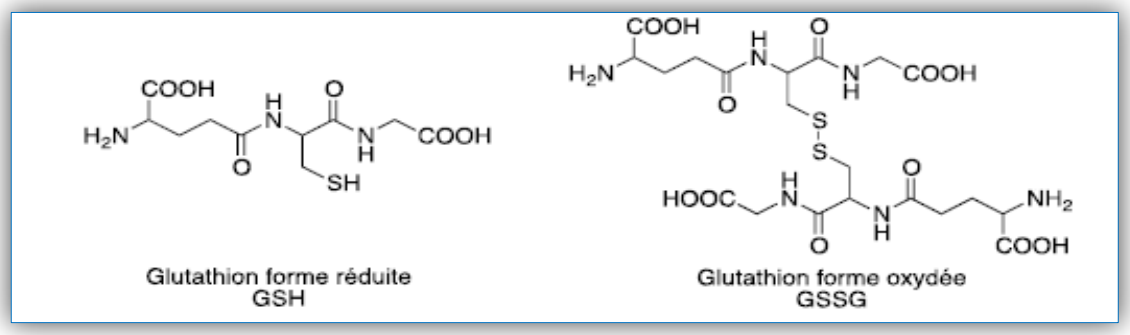



**Figure I.3. Glutathione dimerization reaction.**

### I.3.2.3. The dismutase superoxide:

The Dismutase Superoxide (SOD) is a metal protein with an activity enzymatic allowing it to catalyze the dismutation of the $O^{2-}$ anion. (21) There are several SOD, they differ in the metals (s) present in its structure which will (will) allow the enzyme-ligand binding. It is important to note that there are sods with two types of metals that can be copper, zinc, manganese and/or iron.

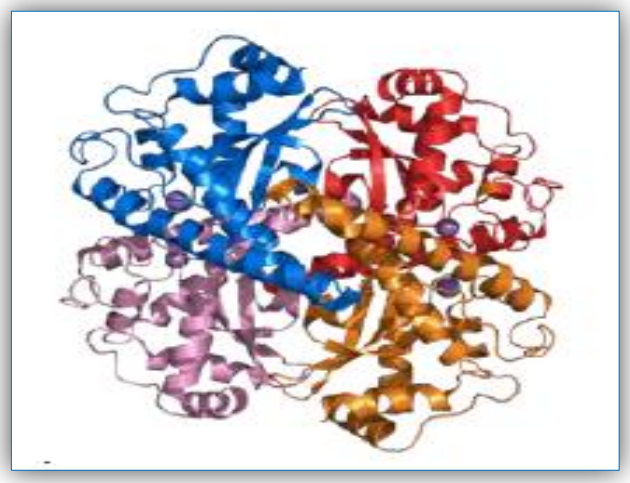


**Figure I.4. Three -dimensional structure of the dismutase superoxide.** (22)

In humans, there are three different SOD classes, catalyzing all the same reaction: The SOD copper and zinc that is found in cytosol and in the level of extracellular liquids, SOD iron and SOD in manganese, in the mitochondria.

### I.3.3. Non-enzymatic:

### I.3.3.1. Vitamins:

- **Vitamin E:**

Vitamin E is a liposoluble vitamin present in large quantities in vegetable oils (E.G., palm, olive and sunflower oil). Vitamin E is actually made up of eight different chemical compounds grouped into two sub-assemblies as a function of the presence of substitution groups and double bonds. So we find the sub-assemblies of tocopherols (α-, β-, γ-, or Δ-tocopherol) and tocotrienols (α-, β-, γ-, or Δ-Tocoprienol).

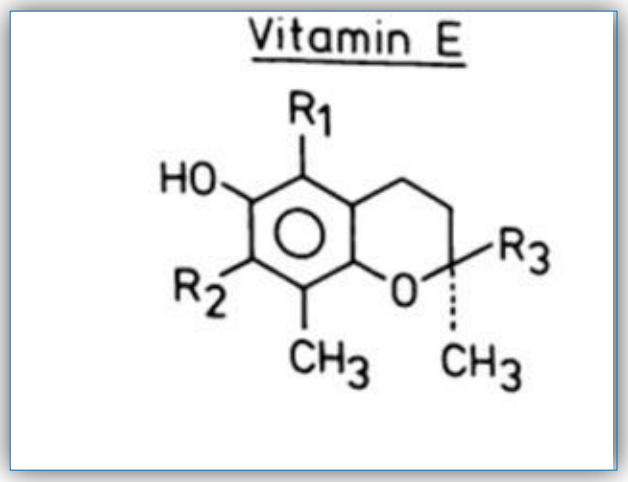



**Figure I.5. Chemical structure of Vitamin E.**

It will act as antioxidant against ROS (in parallel with vitamin C and glutathion) and more particularly in the inhibition of lipid peroxidation. The carbon chain increases the lipophilic character of the molecule and thus facilitates penetration into lipid bicouches, and allows intracellular direct action.

- **Vitamin C:**

Vitamin C, or ascorbic acid, is a hydrophilic molecule found in many fruits such as oranges, lemons and strawberries.(23) Vitamin C is a water-soluble vitamin, sensitive to heat, ultraviolet and oxygen. It is synthesized by plants and most animals, except in certain mammals such as humans. It is an antioxidant molecule capable of reacting directly with all ROS thus reducing lipid peroxidation and protein and DNA damage.

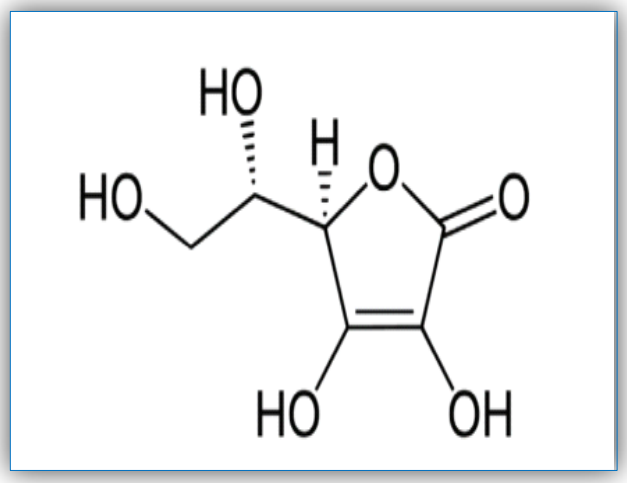



**Figure I .6.Chemical structures of Vitamin C.**

❖ **Vitamin A:**

Vitamin A is a liposoluble vitamin that is found in large quantities in the body especially at the level of the liver, its main storage place. There are two groups, namely retinoids (retinol, trétinoine, etc.) and provitamins A (mainly α- and β- carotenes).

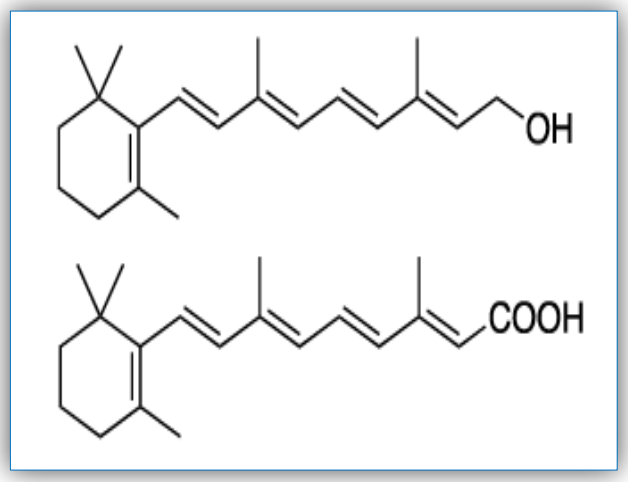



**Figure I.7. Chemical structure of some retinoid.** (24)

### I.3.3.2. Oligo-elements:

❖ **Selenium:**

Selenium is an oligoelement found in all organs of the human body but it is mainly located in the liver, kidneys, blood, brain, heart muscle and testicles.(14) Selenium is constituting selenoproteins, the main intracellular antioxidant, the peroxidase glutathionase.

It is found in particular in pork, beef and fish. He also plays the role of detoxification of heavy metals such as cadmium.

- **The copper:**

Copper is one of the essential cofactors of SOD given its ease of passing from the reduced state to the oxidized state. (25) We find copper especially in the liver, oysters and dark chocolate.

However, it also plays an important role in the initiation of reactions producing ROS by its transitional metal properties. (26) A significant copper concentration may be the revealing of oxidative stress. During the aging process, serum copper concentration is led to increase.(27)

- **Zinc:**

It is a cofactor of SOD capable of binding to many molecules, he intervenes in their stability and their activity. It intervenes in the synthesis of metalloproteins and sulfur rich protein which neutralize radicals. The Zn then partially inhibit the training reactions of oxygen species induced by iron or copper as such the analysis of the copper / zinc blood ratio can give interesting indications on the oxidant stress of individual.(28)

#### I.3.3.3. Polyphenols:

Plant polyphenols include a wide variety of compounds comprising, among other things, flavonoids, anthocyans and tannins. These are ubiquitous compounds that are found in plants. They attract attention since a few years because of their antioxidant properties. Indeed, they are capable of trapping free radicals, inhibiting lipid peroxidation by reducing hydroxyl, superoxides and peroxleal radicals. They are also able to trap metallic ions because they have chelators properties.(29)

### I.3.4. Methods of quantification of antioxidant activity:

Several methods are available to measure antioxidant activity of food and biological systems. They can be classified into two groups according to two mechanisms: either by the transfer of hydrogen atom, or by the transfer of a simple electron.

The first group techniques are used to assess lipid peroxidation using a lipid or lipoprotein substrate. The quantification of this property is expressed by measuring the degree of inhibition of oxidation.

Then, the methods of the second group are those which intervene in the measurement of the ability of trapping free radicals. They feature the scavenging of hydrogen peroxide ($H_2O_2$), hypochlorous acid (HOCl), hydroxyl $OH^{\bullet}$, superoxide anion ($O_2^{\bullet -}$), peroxyl ($ROO^{\bullet}$) and l nitric oxide ($NO^{\bullet}$).Among these techniques, we cite:

- DPPH test (It is based on the reduction of the free radical DPPH. It is a stable radical centered on Nitrogen. Its reduction by the H atom donor antioxidant, in particular phenolic compounds, leads to the formation of DPPH- H (colorless compound) and the radical AH).
- TEAC test (based on the ability of the antioxidant to inhibit the radical cation ABTS.+ compared to a reference antioxidant, trolox with a structure similar to vitamin E. the ABTS.+ with a blue-green color in contact with an H. donor leads to the colorless ABTS+ cation ).
- ORAC test (Oxygen radical absorbance capacity).
- Redox potential: The redox potential is an electrochemical measure of antioxidant power expressed in volts. It is a technique used in the evaluation of the antioxidant power of phenol. In the presence of free radicals (oxidizing agents) phenolic compounds are converted into phenoxy radicals.(30)

## References

(1). Santos-Sánchez NF, Salas-Coronado R, Villanueva-Cañongo C, Hernández-Carlos B, Shalaby E. Antioxidants. Chapter: Antioxidant Compounds and Their Antioxidant Mechanism. 2019:1-28.

(2). Hybertson BM, Gao B, Bose SK, McCord JM. Oxidative stress in health and disease: the therapeutic potential of Nrf2 activation. Molecular aspects of medicine. 2011;32(4-6):234-46.

(3). Kehrer JP. Free radicals as mediators of tissue injury and disease. Critical reviews in toxicology. 1993;23(1):21-48.

(4). Gutteridge JM, Halliwell B. Invited review free radicals in disease processes: a compilation of cause and consequence. Free radical research communications. 1993;19(3):141-58.

(5). Valko M, Rhodes C, Moncol J, Izakovic M, Mazur M. Free radicals, metals and antioxidants in oxidative stress-induced cancer. Chemico-biological interactions. 2006;160(1):1-40.

(6). Hadj Salem J. Extraction, identification, caractérisation des activités biologiques de flavonoïdes de Nitraria retusa et synthèse de dérivés acyles de ces molécules par voie enzymatique: Vandoeuvre-les-Nancy, INPL; 2009.

(7). Jacob L. L'insuffisance rénale aiguë: Springer Science & Business Media; 2007.

(8). Pincemail J, Defraigne J. Le CoEnzyme Q10 ou ubiquinone: un antioxydant particulier. Vaisseaux, Cœur, Poumons. 2003;8:55-60.

(9). Bonnet C, Alamigeon F, Micheels P. Guide complet des soins esthétiques: du coté de ma vie. Edition Eyrolles. 2010:14.

(10). Favier A. Le stress oxydant: Intérêt conceptuel et expérimental dans la compréhension des mécanismes des maladies et potentiel thérapeutique Review. l'actualitechimique, nouvembre, pp: 108-115. GHESTEM A, SEGUIN E, PARIS M, ORECCHIONI AM. 2001.

(11). Morris CJ, Earl JR, Trenam CW, Blake DR. Reactive oxygen species and iron—a dangerous partnership in inflammation. The international journal of biochemistry & cell biology. 1995;27(2):109-22.

(12). Dizdaroglu M, Jaruga P, Birincioglu M, Rodriguez H. Free radical-induced damage to DNA: mechanisms and measurement. Free Radical Biology and Medicine. 2002;32(11):1102-15.

(13). Fulbert J. Les radicaux libres en biologie clinique: origine, role pathogene et moyens de defense. Pathologie Biologie. 1992;40:66-77.

(14). Sara M, Iméne M, Benhammad NE. Étude in vitro de l'activité antioxydante de quelques souches des bactéries lactiques: Université-Jijel-; 2021.

(15). Bonnefont-Rousselot D. Glucose and reactive oxygen species. Current Opinion in Clinical Nutrition & Metabolic Care. 2002;5(5):561-8.

(16). Pasquier C. Stress oxydatif et inflammation. Revue française des laboratoires. 1995;1995(276):87-92.

(17). Bhattacharyya A, Chattopadhyay R, Mitra S, Crowe SE. Oxidative stress: an essential factor in the pathogenesis of gastrointestinal mucosal diseases. Physiological reviews. 2014;94(2):329-54.

(18). Fridovich I. Superoxide radical and superoxide dismutases. Annual review of biochemistry. 1995;64(1):97-112.

(19). Richard M, Belleville F, Chalas J, Ceballos-Picot I, Vitoux D, Boyer M, et al., editors. Les glutathion peroxydases: intérêt de leur dosage en biologie clinique. Annales de biologie clinique; 1997.

(20). Frei B, Stocker R, Ames BN. Antioxidant defenses and lipid peroxidation in human blood plasma. Proceedings of the National Academy of Sciences. 1988;85(24):9748-52.

(21). Peltier A-P, Russo-Marie F, Polla B. L'inflammation: John Libbey Eurotext; 1998.

(22). McCord JM, Fridovich I. Superoxide dismutase: the first twenty years (1968–1988). Free Radical Biology and Medicine. 1988;5(5-6):363-9.

(23). Fabre G, Bayach I, Berka K, Paloncýová M, Starok M, Rossi C, et al. Synergism of antioxidant action of vitamins E, C and quercetin is related to formation of molecular associations in biomembranes. Chemical Communications. 2015;51(36):7713-6.

(24). Higdon J. Antioxidant Vitamins and Health: Cardiovascular Disease, Cancer, Cataracts, and Aging by Claude Fernand Bourgeois, 2003, 306 pages, hardcover, 72. HNB Publishing, New York Am J Clin Nutr.1:239.

(25). Jomova K, Valko M. Advances in metal-induced oxidative stress and human disease. Toxicology. 2011;283(2-3):65-87.

(26). Laliberté J, Labbe S. The molecular bases for copper uptake and distribution: lessons from yeast. Medecine Sciences: M/S. 2008;24(3):277-83.

(27). Del Corso L, Pastine F, Protti M, Romanelli A, Moruzzo D, Ruocco L, et al. Blood zinc, copper and magnesium in aging. A study in healthy home-living elderly. Panminerva Medica. 2000;42(4):273-7.

(28). Rostan EF, DeBuys HV, Madey DL, Pinnell SR. Evidence supporting zinc as an important antioxidant for skin. International journal of dermatology. 2002;41(9):606-11.

(29). Delattre J, Beaudeux J-L, Bonnefont-Rousselot D. Radicaux libres et stress oxydant: aspects biologiques et pathologiques: Editions Tec & Doc; 2005.

(30). Belouerghi S, Azri S. Etude qualitative et quantitative des relations structure-activité d'une série de composés phénoliques [memoir master]: University of mohamed khider biskra; 2020.

# *Chapter II: Computational approaches for drug design and discovery*

## II.1. Introduction:

Drug discovery plays an important role in the growth of any pharmaceutical company and society, as newer and safer drugs are launched in the market with the sole objective of improving the therapeutic value and safety of drugs. The pharmaceutical industry has consistently shown that it can discover and develop innovative medicines for a wide range of diseases.(1)

Drug research, as it is called today, began when chemistry had reached the peak of its career, allowing chemical principles and theories to be applied to problems outside the scope of chemistry, and when pharmacology became an independent scientific discipline on its own.(2)

Despite the diversity of research and development approaches in pharmaceutical companies, turnover rates are unacceptably high. One of the factors contributing to the high attrition rate is that the drug has unacceptable absorption, distribution, metabolism, elimination and toxicity (ADMET) side effects and development must therefore be discontinued. This factor accounts for approximately 50% of all costly drug development failures,(3) and it is generally accepted that these issues should be addressed as early as possible in the drug discovery process.(4, 5) Clearly, pitfalls in the current drug discovery process require an unconventional approach that not only reduces development time, but also reduces associated costs.(6)

## II.2. Computer-aided drug design basic principles:

CADD refers to the application of computer science to the discovery, design and optimization of biologically active compounds. Some of the most commonly used techniques are briefly described.(7)

### II.2.1. Molecular mechanics:

MM is often applied to large systems to calculate molecular structures and relative potential energies of molecular conformations or atomic arrangements.(8, 9) Electrons in the studied system were not explicitly considered, but each atom - specifically the nucleus and associated electrons- was considered as a single particle. The exclusion of electrons in MM is justified by the Born-Oppenheimer approximation,(10) which states that electron and nuclear motions can be separated and considered separately.The energy difference between

conformations is important in such calculations, not the absolute value of the potential energy.

MM can be simply viewed as a ball and spring model with classical forces between atoms and molecules. (11) These forces are explained by potential energy functions incorporating structural features such as length, bond angle, and torsion angle. The potential energy function is equipped with parameters designed to reproduce the experimental properties.(9) The MM or rather the total potential energy of a molecule is described as the sum of the bond stretch energy($\mathbf{E_{str}}$) , the bond angle bending energy($\mathbf{E_{bend}}$), the torsion energy($\mathbf{E_{tor}}$) and energy of interactions between non-bonded atoms ($\mathbf{E_{nb}}$) . The energetic contributions of the latter constitute the van der Waals ($\boldsymbol{E_{vdw}}$) and electrostatic ($\mathbf{E_{elec}}$) interactions:(12)

$$\mathbf{E_{tot}} = \mathbf{E_{str}} + \mathbf{E_{bend}} + \mathbf{E_{tor}} + \mathbf{E_{vdw}} + \mathbf{E_{elec}}$$

$$\boldsymbol{E_{tot}} = \sum_{\boldsymbol{bonds}} \boldsymbol{K_r}\left(\boldsymbol{r} - \boldsymbol{r_{eq}}\right)^{\mathbf{2}} + \sum_{\boldsymbol{angles}} \boldsymbol{K_r}\left(\boldsymbol{\theta} - \boldsymbol{\theta_{eq}}\right)^{\mathbf{2}} + \sum_{\boldsymbol{dihedrals}} \frac{\boldsymbol{v\,n}}{\mathbf{2}}[\mathbf{1} - \ \cos(\boldsymbol{n\emptyset} - \boldsymbol{r})] + \sum_{\boldsymbol{i<j}} \left[\frac{\boldsymbol{Aij}}{\boldsymbol{rij^{12}}} + \frac{\boldsymbol{Bij}}{\boldsymbol{rij^{6}}} + \frac{\boldsymbol{qiqj}}{\boldsymbol{\sum rij}}\right]$$

- **Force fields in molecular mechanics :**

We call the mathematical model that represents the potential energy of a molecule in molecular mechanics a force field. It is important to note that force fields are a purely empirical approach. Designates both the mathematical equation (potential energy function) and the experimental parameters that compose it. There are different force fields in molecular mechanics:

- ✓ **MM2/MM3/MM4:**

MM2 is the first force field developed by Allinger et al. It was initially designed for simple molecules (alkanes, alkenes, non-conjugated alkynes, amines, etc.), but its improved versions MM3 (1989) and MM4 (1996) allow it to process increasingly complex organic molecules.

- ✓ **CHARM (Bio+):**

Developed by Karplus et al, for the calculation of biomolecules. Its design is similar to that of AMBER. Although initially this force field was intended for amino acids and proteins, now it is about other bimolecular ones.

- ✓ **AMBER:**

AMBER (Assisted Model Building with Energy Refinement), was written by Kollman the field is established for proteins and nucleic acids (UCSF, 1994). It has been used for polymers and for other small molecules.(11)

## II.2.2. Quantum mechanics:

The QM method treats molecules as collections of nuclei and electrons without any reference to "chemical bonding". QM is important for understanding the behavior of systems at the atomic level. QM methods apply the QM laws to approximate the wave function and solve the Schrödinger equation.

The solution to the Schrödinger equation is in terms of the motions of the electrons, which in turn lead directly to molecular structure and energy, among other observables, as well as information about bonding. However, the Schrödinger equation cannot actually be solved for any system other than one electron (the hydrogen atom), and approximations are necessary. According to QM, an electron attached to an atom cannot possess any arbitrary energy or occupy any position in space. These characteristics can be determined by solving the time-independent Schrödinger equation:(12)

$$\mathbf{H = T + V}$$

Where *H* is the Hamiltonian operator (sum of kinetic energy), *T* the potential energy, and *V* the operator. *H* can also be defined as:

$$\boldsymbol{H = \left[-\frac{h^2}{8\pi^2}\sum_i \frac{1}{m\,j} \div \left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\right] + \sum_i \sum_{<j} \left(\frac{ei\;ej}{r_{ij}}\right)}$$

These methods are techniques for solving the Schrödinger equation of systems with many electrons. QM methods include Ab initio (13) density functional theory (DFT) (14, 15) and semi empirical calculations.(16, 17)

They use data fitted to experimental results in order to simplify the calculations.(11) The semi-empirical methods contain a default minimum basis (STO-3G) on the other hand

the pure quantum methods use different bases and correlations according to the type of calculation. The HyperChem software provides these semi-empirical (CNDO, INDO, MINDO3, PM3, RM1...) and quantum (Ab-initio and DFT) methods.

❖ **Density functional theory :**

Density functional theory (DFT) is currently the most successful quantum mechanical modeling method for computing the electronic structure (mainly the ground state) of many-body systems, especially atoms, molecules and condensed phases. In chemistry, DFT is used to predict various molecular properties such as B. Molecular structure, vibrational frequency, atomization and ionization energies, electrical and magnetic properties, reaction pathways, etc. Modern DFT calculations are based on two theorems of Hohenberg and Kohn. This proves that the electron energy of the ground state molecule can be completely determined by the electron density ρ(r). The electron density ρ(r) can be defined as follows, where r is the electron's spatial variable and s is the electron's spin variable. (12)

$$\mathbf{P(r) = N \Sigma\, s_1 \ldots \Sigma\, s_N \int dr_2 \ldots \int r_N |\Psi(r_1, s_1, r_2, s_2 \ldots\ldots r_N, s_N)|^2 \int P(r) dr}$$

## II.2.3. Drug-likenes rules for drug discovery:

The fastest way to assess the drug-like properties of a compound is to use "Rules". Rules are a set of guidelines about the structural properties of compounds that are more readily absorbed after oral administration. The attribute values associated with the rules can be calculated quickly by examining the structure or calculated using widely available software.

❖ **Lipinski Rules (Oral drug properties) :**

Lipinski's Rule of Five, also known as Pfizer's Rule of Five or simply Rule of Five (RO5), is a rule of thumb for assessing drug similarity or determining whether a compound with a particular pharmacological or biological activity has properties that make a drug similar to what it might be in humans Orally active drug. This rule was developed in 1997 by Christopher A. Lipinski.(18) The rules are very influential in this regard.(19)

The Lipinski rule states that orally active drugs generally only violate the following criteria:

1/ No more than 5 hydrogen bond donors (nitrogen or oxygen atoms with one or more hydrogen atoms).

2/ No more than 10 hydrogen bond acceptors (nitrogen or oxygen atoms).

3/ The molecular weight is less than 500 Daltons.

4/ The octanol-water partition coefficient log P is not greater than 5.

Note that all numbers are multiples of five, which is where the rule's name comes from.(20)

- **Veber Rules :**

Additional rules were introduced by Veber et al.They investigated structural features that increase oral bioavailability in rats. They concluded that molecular flexibility, polar surface area (PSA) and number of hydrogen bonds are important determinants of oral bioavailability.

Steering thresholds can be calculated manually or by software.PSA is calculated using software and is closely related to hydrogen bonding.(21) 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).(22)

## II.2.4. Quantitative structure-activity relationships (QSAR):

### II.2.4.1. Introduction:

Quantitative structure-activity relationship modeling (QSAR) is the construction of models for predicting biological activity based on structure and molecular information in chemical libraries. Widely used in drug discovery and development, the QSAR concept has been widely used to relate molecular information to biological activity and other physicochemical properties, hence the name Quantitative Structure-Property Relationships (QSPR). Was attached. QSAR is a widely accepted predictive and diagnostic technique for discovering relationships between chemical structure and biological activity. QSAR emerged and was developed to meet the needs and desires of experienced chemists for predicting biological responses.(12) It goes into the practice of most aspects of agricultural chemistry, medicinal chemistry, and finally chemistry.(23)

A QSAR is the final result of a computational process that begins with an adequate description of the molecular structure and ends with some conclusions, assumptions, and predictions about the behavior of the molecule in the environment, physicochemical, and biological systems under analysis.(24) The result of a QSAR calculation is a set of

mathematical formulas relating chemical structure to biological activity. Multivariate QSAR analysis uses all molecular descriptors from different representations of the molecule (1D, 2D, and 3D representations) to compute the model and search for the best descriptor valid for the properties analyzed increase. The general mathematical form of QSAR is given by the following equation:

**Biological activity = *f* (physiochemical properties and/or structural properties)**

QSAR models are not only used for the prediction of properties but are also helpful in selection of alternative mechanism of action, determination of useful structural characteristics, projecting new design methodologies and help in proposing new hypotheses for future research work.

### II.2.4.2. 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.

### II.2.4.3. QSAR properties that characterize the molecular structure:

- **Molecular Weight (MW) :**

Molecular weight descriptor has been used as a descriptor in systems such as transport studies where diffusion is the mode of operation. It is an important variable in QSAR studies pertaining to cross resistance of various drugs in multi-drug resistant cell lines.(25)

Molecular weight is correlated with the size of the molecule.(26) High molecular weight compounds are likely to show high toxicity as promiscuity of compounds is also likely to increase(27) Additionally, the systemic clearance of a compound is inversely proportional to the molecular weight(28)

- **Hydration Energy (HE):**

Hydration energy is very important in the selectivity filter of ion channels where the drug is almost entirely strip from the hydration water. (29)

Indeed, in the biological environments the polar molecules are surrounded by water molecules. They are established hydrogen bonds between a water molecule and these molecules. The donor sites of the proton interact with the oxygen atom of water and the acceptor sites of the proton interact with the hydrogen atom. The first corresponds to the complex with the strongest hydrogen bond. These hydrated molecules are dehydrated at least partially before and at the time of their interaction. These interactions of weak energy, which we observe in particular between messengers and receivers, are generally reversible.(25)

❖ **Partition Coefficient (Log P) :**

The transport, the passage through the membranes and the pharmacological activity of a molecule can be conditioned by its partition between a lipid phase and an aqueous phase, that is to say its hydrophilic character. This can be quantified by the Octanol-Water partition coefficient, noted (log P), which measures the differential solubility of a solute in these two immiscible solvents. (30) It is an important measure for the identification of drug similarity, according to Lipinski's rule, orally delivered drugs should have log P values greater than or equal to -2 and less than or equal to 5.(31) It is defined by the following formula:

$$\textbf{Log P} = \frac{[\textbf{drug concentration}]_{\textbf{octanol}}}{[\textbf{drug concentration}]_{\textbf{water}}}$$

Allows to estimate the bioavailability of a molecule:

- 0 < Log P < 3: Optimal biological activity (permeability, solubility).
- Log P < 0: Compounds that are too hydrophilic (poor permeability of the lipid bilayer).
- Log P > 3: Compounds that are too lipophilic (poor aqueous solubility) .(31)

❖ **Polarizability (Pol):**

Polarizability refers to a phenomenon caused by the moment of the electric charges of the atom. A molecule placed in an electric field E undergoes a deformation and acquires an induced electric dipole m:oment proportional to the field E. . Atomic polarizability increases with the size of atoms.

$$\mathbf{P(e) = \varepsilon o\ \alpha\ E}$$

Where: P: dipole that is created.

α: polarizability.

E: electric field.

❖ **Molecular Volume :**

Molecular volume is a function of MM and structure and takes into account all accessible conformations available to the molecule under physiological conditions. This actually refers to the rotating bonds and the number of rings in the molecule.

The volume is defined by the relation: $\mathbf{V = \frac{MM}{d}}$

Where: MM: the molecular weight.

d: density.(31)

❖ **Molecular Refractivity (MR) :**

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.(32)

This parameter is a measure of the volume occupied by an atom or group of atoms. The molar refractivity is a constitutive-additive property that is calculated by Lorenz-Lorentz formula:

$$\boldsymbol{MR = \frac{n^2-1}{n^2+2} \times \frac{Mw}{p}}$$

where n is the refraction index, Mw is the molecular weight, and ρ is the density. The $\frac{\mathbf{n^2-1}}{\mathbf{n^2+2}}$ term provides a correction factor by defining how easily the substituent can be poralized, whereas the Mw/ρ term defines a volume. Molar refractivity is related to the lipophilicity, volume, and steric of the molecules. Moreover, it has been correlated with the London dispersive force that acts in the drug-receptor interaction.(33)

**II.2.4.4. 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:(34) Principal component analysis (PCA), Cluster analysis (CA), Simple liner regression (SLR), multiple liner regression (MLR). Stepwise multiple liner regression (MLR ), Principle component regression (PCR) ,Continuum Regression (CR),Partial least squares (PLS) ,Genetic function approximation (GFA) ,Genetic partial least squares (GPLS) ,Logistic regression (LR) , K-Nearest Neighbors classification (KNN) ,Neural Network(NN) , Discriminant analysis (DA) , Decision Trees (DT), Canonical Correlation (CC).(35)

- **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:

$$\boldsymbol{Y = b_0 + b_1 X_1 + b_2 X_2 + \cdots \ldots + b_m X_m + e}$$

MLR analysis computes the regression coefficients bi by minimizing the residuals, which quantify the deviation between the data (Y) and the model (Y'), as in simple linear regression.(36)

- **Artificial neural networks :**

Artificial neural network (ANN) models are nonlinear models that can be used to predict biological activity in large molecular datasets**.** In contrast to classical statistical methods such as regression analysis or partial least squares analysis, artificial neural networks are capable of investigating complex nonlinear relationships. Therefore, neural networks are well suited for drug design and QSAR **.**They are used to simulate a wide variety of nonlinear complex systems from the fields of pharmaceutical, technological, psychological, and medicinal chemistry. For example, artificial neural networks have been successfully used to predict and synthesize new organic compounds.(37)

**II.2.4.5. Testing the overall significance of the regression:**

❖ **Correlation coefficient R (and coefficient of determination $R^2$):**

It is the most widespread statistical indicator is the correlation coefficient, which evaluates the part of the variance of the activity / the target property explained by the model.

$$\mathbf{R} = \sqrt{1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2}}$$

Where: R: is the correlation coefficient.

$y_i$ , $\hat{y}_i$ :are , respectively, the observed and calculated values of the dependent variable.

$\bar{y}_i$: is the mean value of observed values.

These coefficients are not affected by the unit of measurement chosen and reflect a good correlation between the target activity and the initial activity if $R^2$ is close to 1 (ideal case).

❖ **The adjusted coefficient of determination $R^2_{adj}$:**

$$\mathbf{R^2adj} = \frac{\mathbf{R^2(n-1)-p}}{\mathbf{n-p-1}}$$

Where: n is the number of dependent variables (the molecules); p is the number of independent variables (descriptors); $R^2$ is the coefficient of determination.(38)

❖ **Standard Deviation (S):**

Standard deviation (S) is a statistical measure of the spread or uncertainty around the mean. It is defined by the equation:(39)

$$\mathbf{S} = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n - p - 1}}$$

In particular, the smaller the standard deviation, the better the correlation.

❖ **Fischer Statistic (F):**

The Fisher F index is also used to measure the level of statistical significance of the model, that is to say the quality of the choice of descriptors constituting the model.

$$\mathbf{F} = \frac{\sum(\hat{y}\mathbf{i} - \bar{\mathbf{y}}_\mathbf{i})^2}{\sum(\mathbf{y_i} - \hat{\mathbf{y}}_\mathbf{i})^2} \frac{(\mathbf{n} - \mathbf{p} - \mathbf{1})}{\mathbf{p}}$$

yi, ŷi: are, respectively, the observed and calculated values of the dependent variable.

ȳi: is the mean value of predicted values.

n is the number of dependent variables (the molecules).

p is the number of independent variables (descriptors).(40)

❖ **Quality Factor (Q):**

Quality factor is calculated by equation:

$$\mathbf{Q} = \frac{\mathbf{R}}{\mathbf{S}}$$

Where r is variance and S is standard deviation. Over fitting and chance correlation, due to excess number of descriptors, can be detected by Q value. Positive value for this QSAR model suggests its high predictive power and lack of over fitting.(39)

**II.2.4.6. Validation of QSAR Models:**

Additionally, after obtaining the model equations, it is important to assess the stability and goodness of fit of the model before using it to predict biological activity. Validity aims to demonstrate the reliability and importance of the method for a particular purpose. Therefore, validation of the QSAR model must be performed. There are two validation methods for QSAR models: internal and external validation techniques to determine the confidence and strength of the model. In general, QSAR modeling involves a systematic, multi-step process. These include creating datasets, selecting and generating molecular descriptors, deriving mathematical or statistical models, training and validating models using training datasets, and testing models on test datasets. (12)

**a) Internal validation parameters:**

Whatever the ultimate objective of the QSAR model, it must be validated before it can be interpreted or used for predictive purposes. There are different solutions to ensure the validity of a model. The LOO (leave-one-out) cross-validation is a process that allows to test the predictive precision of a model. This method is based on the calculation of some statistical parameters such as: PRESS, SSY, Spress, $R^2_{cv}$ et $R^2_{adj}$, PE.(40)

**Table II.1. Statistical parameters for cross-validation.**

| Statistic | Definition | Formula |
|---|---|---|
| **PRESS** | predicted residual sum of squares | $\mathbf{PRESS} = \sum (\mathbf{Y_{obs}} - \mathbf{Y_{cal}})^2$ |
| **TSS** | Total Sum of Squares | $\mathbf{TSS} = \sum (Y_{obs} - \overline{Y})^2$ |
| $\mathbf{R^2_{adj}}$ | The square of the correlation Adjusted | $\mathbf{R^2_{adj}} = 1 - (1 - \mathbf{R^2}) \left( \frac{\mathbf{n-1}}{\mathbf{n-p-1}} \right)$ |
| $\mathbf{R^2_{cv}}$ | The square of the correlation Coefficient | $\mathbf{R^2_{cv}} = 1 - \frac{\mathbf{PRESS}}{\mathbf{TSS}}$ |
| $\mathbf{S_{PRESS}}$ | Standard prediction error validation | $\mathbf{S_{PRESS}} = \sqrt{\frac{\mathbf{PRESS}}{\mathbf{n}}}$ |
| **PE** | Prediction error | $\mathbf{PE} = \frac{\mathbf{0.6745(1 - R^2)}}{\sqrt{\mathbf{n}}}$ |

The PRESS statistic (predicted residual sum of squares) seems to be the most important parameter for a good estimation of the actual predictive error of the models. Its small value indicates that the model predicts better than chance and can be considered statistically significant.(11)

**b) External validation:**

This method consists in predicting the activity of a series of molecules generally called test series which are not in the development series of the model, this validation is characterized by the parameter $R^2_{test}$. Recently, several studies have shown the inadequacy of the $R^2$, $R^2_{CV}$ parameters to verify the predictive power of QSAR models. Therefore, other parameters should be checked for this purpose. These parameters are known as "external validation criteria" or often referred to as "Tropsha criteria".(41)

**Table II.2. Statistical parameters for external validation.**

| Statistic | Definition | Formula |
|---|---|---|
| $R^2_{pred}$ | Predictive residual sum of squares (external validation) | $R^2_{pred} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2}$ |
| $R^{\circ 2}$ | The squared correlation coefficient predicted versus observed | $R^{\circ 2} = 1 - \frac{\sum_{i=1}^{ntest}(\hat{y}_i - y_i{}^{r^\circ})^2}{\sum_{i=1}^{ntest}(\hat{y}_i - \bar{\hat{y}})^2}$, $y_i{}^{r^\circ} = K' \hat{y}_i$ |
| $R'^{\circ 2}$ | The squared correlation Coefficient versus predicted activities | $R'^{\circ 2} = 1 - \frac{\sum_{i=1}^{ntest}(y_i - \hat{y}_i{}^{r^\circ})^2}{\sum_{i=1}^{ntest}(y_i - \bar{y})^2}$, $\hat{y}_i{}^{r^\circ} = K' y_i$ |
| K | The slopes of regression | $K = \frac{\sum_{i=1}^{ntest} y_i \hat{y}_i}{\sum_{i=1}^{ntest} \hat{y}_i{}^2}$ |
| K' | The slopes of regression | $K' = \frac{\sum_{i=1}^{ntest} y_i \hat{y}_i}{\sum_{i=1}^{ntest} y_i^2}$ |

### II.2.4.7. 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 $R^2$ for external test set, $R^2 > 0.6$
- $\frac{(R^2 - R_0^2)}{R^2} < 0.1$ and $0.85 \leq K \leq 1.15$ (for test set).

- $\frac{(R^2-R_0'^2)}{R^2} < 0.1$ and $0.85 \leq K' \leq 1.15$ (for test set).
- $|R^2 - R_0'^2| < 0.3$ (for test set).

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. (42)

### II.2.4.8. Applications of the QSAR study:

Some QSAR studies seem to be academic only, and these models have a large number of applications, such as:

- Optimized pharmacological activity.
- Rational design of surfactants, fragrances, dyes, fine chemicals and many other products.
- Identify hazardous compounds or predict inventory of existing compounds at an early stage of product development.
- Predict toxicity and side effects of new compounds.
- Prediction of toxicity of environmental species.
- Predict various physicochemical properties of molecules.
- Predict the fate of molecules released into the environment.(11)

## II.3. Conclusion:

This memory work was carried out within the computational and pharmaceutical chemistry team.

The study of the electronic and structural properties of triazole and its derivatives was carried out by molecular modeling (molecular mechanics, DFT , and QSAR), using HyperChem (8.0.7) and Gaussian (09) software in a PC and in an (HP Micro-processeur Intel® Xeon® CPU E5-2620 0 @ 8GO DE RAM) computing station.

Qualitative study of polar area (PSA), log D and number of rotary bonds using Marvin Sketch software. And For quantitative study XLSTAT and JMP software was used.

## References

(1). Ratti E, Trist D. Continuing evolution of the drug discovery process in the pharmaceutical industry. Pure and applied chemistry. 2001;73(1):67-75.

(2). Drews J. Drug discovery: a historical perspective. science. 2000;287(5460):1960-4.

(3). Bharath E, Manjula S, Vijaychand A. In silico drug design tool for overcoming the innovation deficit in the drug discovery process. Int J Pharm Pharm Sci. 2011;3(2):8-12.

(4). Alomar MJ. Factors affecting the development of adverse drug reactions. Saudi pharmaceutical journal. 2014;22(2):83-94.

(5). Jochum C, Gasteiger J. Canonical numbering and constitutional symmetry. Journal of Chemical Information and Computer Sciences. 1977;17(2):113-7.

(6). Baldi A. Computational approaches for drug design and discovery: An overview. Systematic reviews in Pharmacy. 2010;1(1):99.

(7). Hagler A, Lifson S, Dauber P. 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. 1979;101(18):5122-30.

(8). Rogers DW. Computational Chemistry using the PC: John Wiley & Sons; 2003.

(9). Lewars EG. Introduction to Quantum Mechanics in Computational Chemistry. Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics. 2011:85-173.

(10). Hotokka M. Calculation of vibrational frequencies by molecular mechanics. Handbook of Vibrational Spectroscopy. 2006.

(11). Aouisset D. In silico analysis of pyrazole derivatives applied to drug design [Memoir of master]: University Mohammed Khider Biskra; 2022.

(12). Khamouli S. Study of the structure, QSAR properties and molecular docking of some amino-pyrimidine derivatives as Novel Mycobacterium tuberculosis inhibitors for drug design [Doctorat ]: University Mohammed Khider Biskra; 2019.

(13). Perdew JP, Ruzsinszky A. Fourteen easy lessons in density functional theory. International Journal of Quantum Chemistry. 2010;110(15):2801-7.

(14). Burke K. Perspective on density functional theory. The Journal of chemical physics. 2012;136(15):150901.

(15). Hehre WJ. A guide to molecular mechanics and quantum chemical calculations: Wavefunction Irvine, CA; 2003.

(16). Dewar MJ, Zoebisch EG, Healy EF, Stewart JJ. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. Journal of the American Chemical Society. 1985;107(13):3902-9.

(17). Comba P, Hambley TW, Martin B. Molecular modeling of inorganic compounds: John Wiley & Sons; 2009.

(18). Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced drug delivery reviews. 2012;64:4-17.

(19). Kerns E, Di L, Properties D-l. Concepts, Structure Design and Methods: from ADME to Toxicity Optimization. Academic press San Diego, CA, USA:; 2008.

(20). Lipinski CA. Lead-and drug-like compounds: the rule-of-five revolution. Drug discovery today: Technologies. 2004;1(4):337-41.

(21). Veber DF, Johnson SR, Cheng H-Y, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. Journal of medicinal chemistry. 2002;45(12):2615-23.

(22). Satyanarayanajois SD. Drug design and discovery: methods and protocols: Humana Press; 2011.

(23). Eriksson L, Jaworska J, Worth AP, Cronin MT, McDowell RM, Gramatica P. Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. Environmental health perspectives. 2003;111(10):1361-75.

(24). Golbraikh A, Shen M, Xiao Z, Xiao Y-D, Lee K-H, Tropsha A. Rational selection of training and test sets for the development of validated QSAR models. Journal of computer-aided molecular design. 2003;17:241-53.

(25). Kerassa A, Belaidi S, Harkati D, Lanez T, Prasad O, Sinha L. Investigations on Molecular Structure, Electronic Properties, NLO Properties and Comparison of Drug-

Likeness of Triazolothiadiazole Derivatives by Quantum Methods and QSAR Analysis. Reviews in Theoretical Science. 2016;4(1):85-96. https://doi.org/10.1166/rits.2016.1050.

(26). Perkins R, Fang H, Tong W, Welsh WJ. Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology. Environmental Toxicology and Chemistry: An International Journal. 2003;22(8):1666-79.

(27). Hansch C, Leo A, Taft R. A survey of Hammett substituent constants and resonance and field parameters. Chemical reviews. 1991;91(2):165-95.

(28). Lantican BP, Muir RM. Isolation and properties of the enzyme system forming indoleacetic acid. Plant Physiology. 1967;42(8):1158.

(29). Lalitha P, Sivakamasundari S. Calculation of molecular lipophilicity and drug likeness for few heterocycles. Oriental Journal of Chemistry. 2010;26(1):135.

(30). Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced drug delivery reviews. 1997;23(1-3):3-25.

(31). Louafi f. Etude de la relation structure-activité dans la molécule du furane et ses dérivés [Memoire of Master ]: University of mohammed khider biskra; 2013.

(32). Pankaj P, Kapupara CR, Matholiya A. Dedakiya and RD Tusharbindu. Inter Int bull Drug res. 2011;1(1):1-10.

(33). Wang J, Xie X-Q, Hou T, Xu X. Fast approaches for molecular polarizability calculations. The Journal of Physical Chemistry A. 2007;111(20):4443-8.

(34). Shao J. Linear model selection by cross-validation. Journal of the American statistical Association. 1993;88(422):486-94.

(35). Skrehota O. Quantitative structure-property relationship modeling algorithms, challenges and IT solutions: Masaryk University Faculty of Informatics Brno, Czechoslovakia; 2010.

(36). Wu W, Zhang C, Lin W, Chen Q, Guo X, Qian Y, et al. Quantitative structure-property relationship (QSPR) modeling of drug-loaded polymeric micelles via genetic function approximation. PloS one. 2015;10(3):e0119575.

(37). Boudergua S, Alloui M, Belaidi S, Al Mogren MM, Ibrahim UAE, Hochlaf M. QSAR modeling and drug-likeness screening for antioxidant activity of benzofuran derivatives. Journal of Molecular Structure. 2019;1189:307-14.

(38). Belouerghi S, Azri S. Etude qualitative et quantitative des relations structure-activité d'une série de composés phénoliques [memoir master]: University of mohamed khider biskra; 2020.

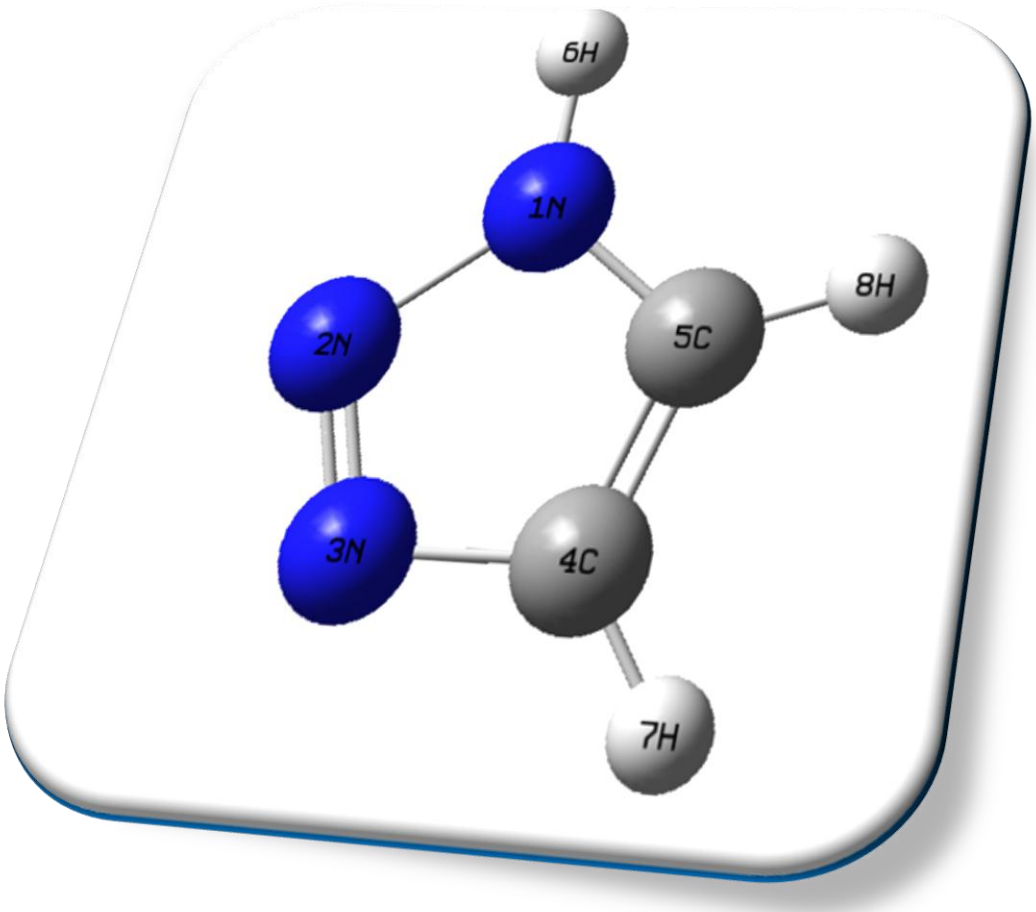
(39). Verma RP, Hansch C. QSAR modeling of taxane analogues against colon cancer. European journal of medicinal chemistry. 2010;45(4):1470-7.

(40). Tropsha A, Gramatica P, Gombar VK. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. QSAR & Combinatorial Science. 2003;22(1):69-77.

(41). Guebbabi K, Kedidi N. Etude 2D-QSAR DE Dérivés de benzophénone en tant qu'agents antipaludique [memoir master]: university of mohammed khider biskra; 2021.

(42). Ouassaf M, Belaidi S, Shtaiwi A, Chtita S. Quantitative Structure Activity Relationship (QSAR) Investigations and Molecular Docking Analysis of Plasmodium Protein Farnesyltransferase Inhibitors as Potent Antimalarial Agents. Jordan Journal of Pharmaceutical Sciences. 2022;15(3):315-40. https://doi.org/10.35516/jjps.v15i3.407.

# *Chapter III: Structural study, Electronics, MESP on the basic core using several molecular and quantum mechanical calculation methods*



## III.1. Introduction:

Triazole nuclei, one of the most important and well-known heterocyclic compounds, are a common and indispensable feature of a wide variety of natural products and pharmaceuticals. Triazole nuclei are present as core structural components in many drug classes such as: Antibacterial, anti-inflammatory, analgesic, antiepileptic, antiviral, antineoplastic, antihypertensive, antimalarial, antidepressant, antihistamine, antioxidant …etc. (1)

Molecular modeling involves the use of theoretical calculation methods (molecular mechanics, quantum mechanics and semi-empirical) to determine the geometry of a molecule and evaluate the associated physico-chemical properties.

There are many methods of theoretical chemistry aimed at determining the physical or chemical properties of isolated molecules, quantum chemical methods that allow to accurately determine the electronic properties of molecules, (2) and play an important role in obtaining molecular geometries and predicting various properties, (3) on the other hand molecular mechanics methods which are based on empirical parameters which make it possible in particular to determine the structural parameters.

Two molecular orbitals, so-called frontier orbitals, play a special role here:

HOMO (Highest Occupied Molecular Orbital) explains the nature of the electron donor (nucleophile) of the molecule. The higher the energy of this OM, the easier it is for the molecule to donate electrons. The LUMO (lowest unoccupied orbital of the molecule) explains the electrophilic (electrophilic) nature of the molecule. The lower the energy of this MO, the easier it is for the molecule to accept electrons.(2)

Thus, over time it has been found that electron delocalization between HOMO and LUMO in general becomes the main factor that determines the ease of chemical reaction and the pathway of stereo selectivity, independent of intra- and intermolecular processes.(4)

We are interested in this chapter in a structural study of the basic core 1,2,3-triazole.

## III.2. Study of the structural and electronic properties of the basic core of 1H-1,2,3-triazole:

### ❖ Triazole:

Triazole, also known as pyrrodiazole is one of the classes of organic heterocyclic compounds containing a five-membered diunsaturated ring structure composed of three nitrogen atoms and two carbon atoms at non-adjacent positions. The simplest form of the triazole family is triazole itself. Triazole is a white to pale yellow crystalline solid with a weak, characteristic odour.(1) 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. (5)

We have studied in detail the structural and electronic parameters of the privileged conformation of the basic core of 1H-1, 2,3-triazole (figure III.1).

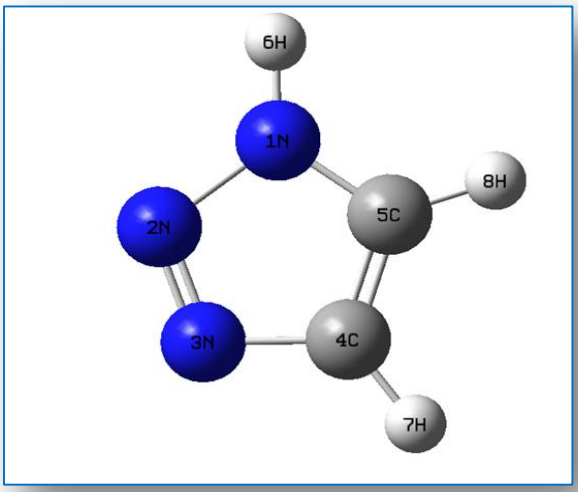



**Figure III.1. 3D conformation of 1H-1,2,3-triazole.**
**(Gauss View 5.0.8)**

### ❖ Material and methods:

The optimization process (the overall minimum energy condition) was applied using the molecular mechanics calculation method, force field MM + using the Polak-Ribière algorithm has as criterion a root of the mean square of the gradient equal to 0.1

kcal.mol$^{-1}$.The obtained structure was re-optimized using the PM3 method by the semi-empirical method.

Then in the next step, a parallel study was made using the Logiciel Gaussian09, we performed geometry optimization calculations of 1,2,3-triazole.We used the method MP2 and DFT. using the three bases 6-31G+ and 6-31G(d,p) and 6-31G++(d,p) . This investigation will determine the best as well as the appropriate basis for these systems.

To determine these structural parameters and properties, calculations have been carried out which target the following characteristics:

• The distances between the linked atoms. **Table III.1**

• The valence angles formed by three atoms linked. **Table III.2**

• The charges of each atom by the DFT method and the MP2 method (Mulliken's Charges). **Table III.3**

**Table III.1. Calculated bond length values of 1H-1,2,3-triazole.**

| Parametres | EXP [8] | DFT/B3LYP | | | MP2 | | |
|---|---|---|---|---|---|---|---|
| | | 6-31G+ | 6-31G(d,p) | 6-31G++(d,p) | 6-31G+ | 6-31G(d,p) | 6-31G++(d,p) |
| **N1-N2** | 1.355 | 1.386 | 1.352 | 1.351 | 1.405 | 1.352 | 1.353 |
| **N2-N3** | 1.309 | 1.328 | 1.303 | 1.303 | 1.381 | 1.329 | 1.331 |
| **N3-C4** | 1.370 | 1.384 | 1.365 | 1.366 | 1.395 | 1.361 | 1.362 |
| **C4-C5** | 1.378 | 1.383 | 1.375 | 1.377 | 1.397 | 1.380 | 1.383 |
| **C5-N1** | 1.356 | 1.366 | 1.356 | 1.357 | 1.378 | 1.355 | 1.356 |

**Table III.2. Calculated values of the valence angle of 1H-1,2,3-triazole.**

| Parametres | EXP [9] | DFT/B3LYP | | | MP2 | | |
|---|---|---|---|---|---|---|---|
| | | 6-31G+ | 6-31G(d,p) | 6-31G++(d,p) | 6-31G+ | 6-31G(d,p) | 6-31G++(d,p) |
| **N1-N2-N3** | 108.2 | 106.17 | 106.83 | 106.92 | 105.36 | 106.26 | 106.28 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **N2-N3-C4** | 108.2 | 109.28 | 109.23 | 109.27 | 108.37 | 108.71 | 108.76 |
| **N3-C4-C5** | 109.9 | 108.92 | 108.83 | 108.70 | 109.92 | 109.48 | 109.37 |
| **C4-C5-N1** | 104.4 | 104.36 | 103.48 | 103.53 | 104.40 | 103.22 | 103.31 |
| **C5-N1-N2** | 110.2 | 111.24 | 111.61 | 111.55 | 111.93 | 112.31 | 112.25 |

**Table III.3. The net atomic charges of 1H-1,2,3-triazole.**

| **Atomes** | **DFT/B3LYP** | | | **MP2** | | |
|---|---|---|---|---|---|---|
| | **6-31G+** | **6-31G (d,p)** | **6-31G++ (d,p)** | **6-31G+** | **6-31G (d,p)** | **6-31G++ (d,p)** |
| **N1** | -0.372 | -0.309 | -0.142 | -0.510 | -0.414 | -0.251 |
| **N2** | -0.048 | -0.078 | -0.026 | -0.007 | -0.057 | 0.005 |
| **N3** | -0.199 | -0.271 | -0.201 | -0.189 | -0.301 | -0.226 |
| **C4** | -0.218 | 0.041 | -0.094 | -0.281 | 0.005 | -0.109 |
| **C5** | -0.036 | 0.082 | -0.179 | -0.064 | 0.053 | -0.136 |

- **Interpretation of the results:**

The theoretical results obtained by the different methods of semi-empirical and quantum calculation (MP2, DFT) with different bases to note a good correlation between the calculated and experimental values for the geometric parameters.

- **For the atomic distances:**

According to Table 1 we see that the variation of MP2 [0.000-0.050] between the results obtained by the three bases and the experimental values. On the other, hand that the variation in the DFT method [0.000-0.031] between the results obtained by the three bases and the experimental values.

However, the variation between the values calculated by three bases of the DFT method values 6-31G+ [0.005-0.031] and 6-31G(d,p) [0.000-0.006] and 6-31G++(d,p) [0.001-0.006] compared to the experimental.

❖ **For the valence angles:**

According to Table 2 the difference between the experimental values and the values calculated by the three bases of the MP2 method [0.0˚-2.84˚], are very large compared to the values of the three DFT bases [0.04˚-2.03˚].

In addition, the difference between the experimental values and the values of two separate bases of the DFT method are: 6-31G+ [0.04˚-2.03˚] and 6-31G (d,p) [0.92˚-1.41˚] and 6-31G++ (d,p) [0.87 ˚-1.35˚].

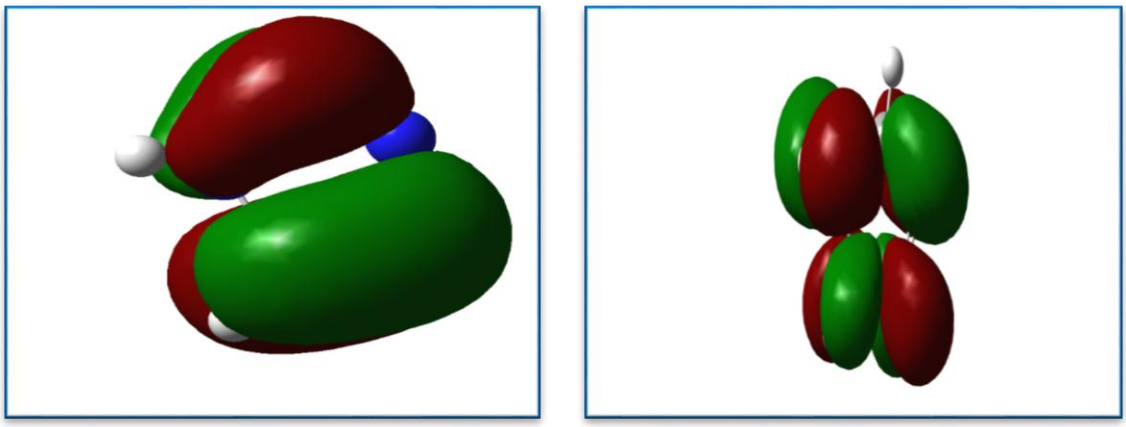
After having compared the results obtained by the two methods MP2 and DFT by different bases, we see that the values are close, but the closest to the experimental values are the results of the base 6-31G++ (d,p) of the DFT method.

❖ **For the net atomic charges:**

According to Table III.1 the atoms N1, N2, N3 have negative charges which favor the electrophilic attack and the atom C4 and C5 have positive charges which favor the nucleophilic attack.

The efficiency of DFT/B3LYP method with 6-311++G (d,p) basis set may be scrutinized by comparison with the results obtained by MP2 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 1H-1,2,3-triazole derivatives in other parts of this work.

The important aspect of the frontier electron theory is the emphasis on the highest occupied and lowest vacant molecular orbitals (HOMO and LUMO), instead of thinking about the total electron density in a nucleophile, we should be thinking to the location of the HOMO orbital because the electrons in this orbital are freer to participate in the reaction. Similarly, frontier orbital theory predicts that a site where the lowest unoccupied orbital is located is a good electrophilic site.(2)



**HOMO** **LUMO**

**Figure III.2. Frontiers orbitals HOMO and LUMO of the basic core 1 H 1,2,3-triazole.**

We can see that the molecular homo orbital focuses mainly on the N1, N2, C4 and C5 of the cycle 1H-1,2,3-triazole with a certain relocation along the N1-N2 and C4-C5 links while that the LUMO molecular orbital is spread over the entire molecule. This distribution of Frontiers orbitals have shown the existence of the relocation of π electrons in the conjugated system of the molecule.

## III.3. Molecular electrostatic potential:

The molecular electrostatic potential is an established tool for explaining the reaction behavior of various chemical systems in electrophilic and nucleophilic reactions. Studies of biological recognition processes and hydrogen bonding were performed to predict reactive sites for electrophilic and nucleophilic attack of the molecules under study.(5)

MEP is a useful property for studying reactivity since an approaching electrophile is attracted to negative regions (where the electron scattering effect is dominant).(4) The values of the electrostatic potential are represented by different colors. The positive, negative, and neutral regions of the electrostatic potential of the molecules are represented by color gradients. In general, the potential increases in the order of red < orange < yellow < green < Blue. The red color indicates the maximum negative region and the blue color represents the maximum positive region. (6, 7)

The MEP was calculated at the DFT optimized geometry (Figure III.3).

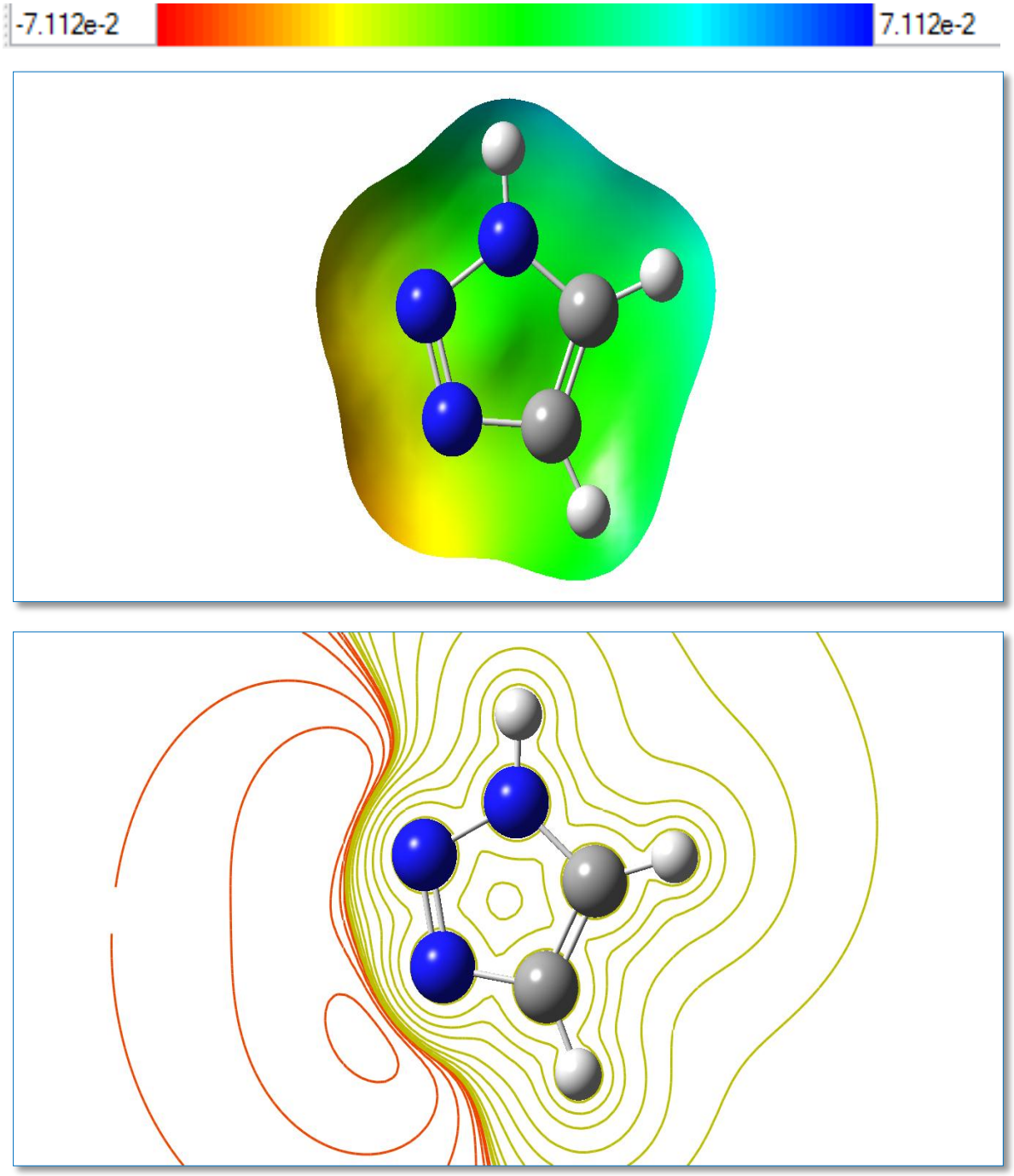



**Figure III.3. 3D MESP surface map and 2D MESP contour map for 1H 1,2,3- triazole (Gauss view 5.0.8)**

The MESP surface map (Figure III.3) for 1H-1,2,3-triazole shows that the maximum negative region is localized (orange to yellow) over N2, N3 atoms which accepts electrophilic attack and the maximum positive region is localized on NH group (bleu), indicating a possible site for nucleophilic attack.

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

## III.4. Conclusion:

In this chapter, we carried out a structural and electronic comparison with the different quantum theoretical methods (MP2 , DFT) by different bases. In addition, it was noticed after the comparison, that there is a similarity between the obtained calculation results and the experimental results. However, The (DFT) with base 6-31G++ (d, p) method is the most suitable method for doing calculations on the 1H-1, 2, 3-triazole nucleus.

# References

(1). Kharb R, Sharma PC, Yar MS. Pharmacological significance of triazole scaffold. Journal of enzyme inhibition and medicinal chemistry. 2011;26(1):1-21.

(2). Aouisset D. In silico analysis of pyrazole derivatives applied to drug design [Memoir of master]: University Mohammed Khider Biskra;2022.

(3). Davidson ER. Quantum theory of matter: introduction. Chemical Reviews. 1991;91(5):649.

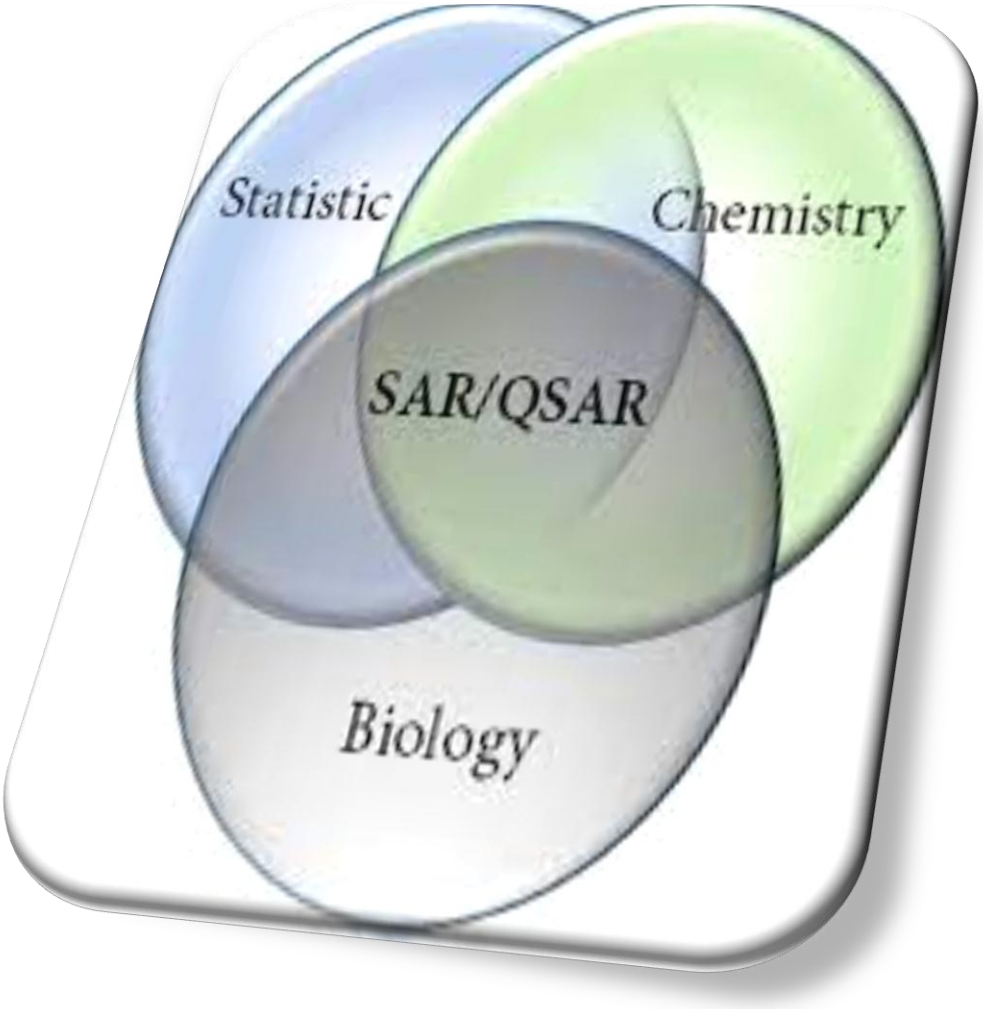
(4). Düğdü E, Ünver Y, Ünlüer D, Tanak H, Sancak K, Köysal Y, et al. Synthesis, structural characterization and comparison of experimental and theoretical results by DFT level of molecular structure of 4-(4-methoxyphenethyl)-3, 5-dimethyl-4H-1, 2, 4-triazole. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy. 2013;108:329-37.

(5). Ouassaf M. Contribution à la découverte de médicaments par une étude computationnelle de plusieurs séries de molécules hétérocycliques [Doctorat] : University Mohamed Khider-Biskra ; 2019.

(6). Chandrasekar S, Balachandran V, Evans H-S, Latha A. 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. 2015;143:136-46.

(7). Almi Z, Belaidi S, Segueni L. Structural Exploration and Quantitative Structure-Activity Relationships Properties for 1.2. 5-Oxadiazole Derivatives. Reviews in Theoretical Science. 2015;3(3):264-72.

# *Chapter IV: Qualitative study of the structure-activity relationship of a series of 1 H-1, 2, 3-triazole derivatives*



## IV.1. Introduction:

QSAR, also called quantitative structure-property relationship, is the process by which chemical structures with specific effects, such as biological activity or chemical reactivity, are correlated.

Studying the relationship between the structure and biological activity of certain molecules concede for the establishment of correlations between structural parameters and properties of molecules.(1)

So a quantitative structure-activity or property relationship (QSAR or QSPR) is a methodology consisting in highlighting a quantitative relationship between a macroscopic magnitude (biological activity, physico-chemical property) and the molecular structure of the compounds studied.(2)

Drug similarity is a qualitative concept used in drug design, which is estimated from the molecular structure even before the substance is synthesized and tested. The calculation of the medicinal property can give us a better hypothesis of the biological activity of certain molecules. The theoretical calculation of certain properties of a molecule can fulfill the essential parameters to demonstrate a certain biological activity.(3)

## IV.2. Study of the QSAR properties of the series of 1H-1,2,3-triazole derivatives:

### IV.2.1. Chemical structures of 1H-1,2,3-triazole derivatives:

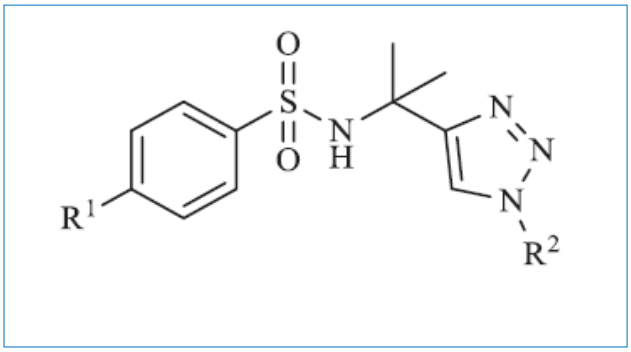
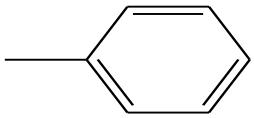
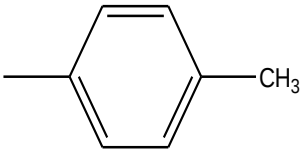
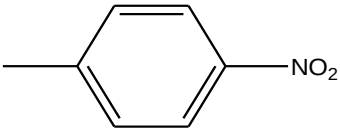
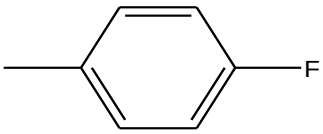
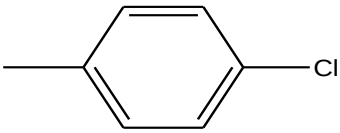
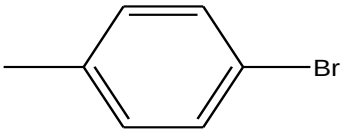
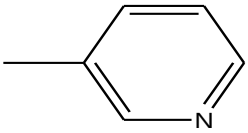
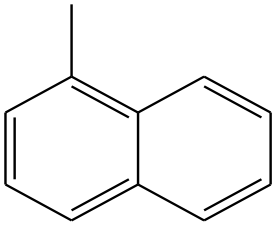
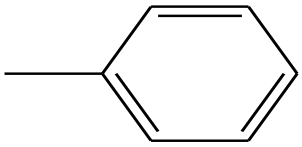
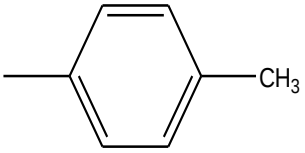
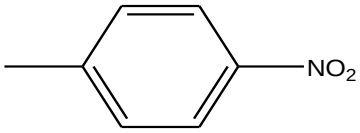
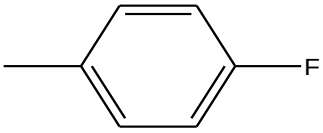
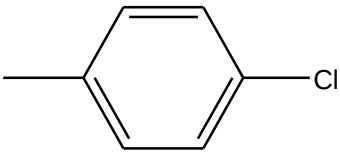
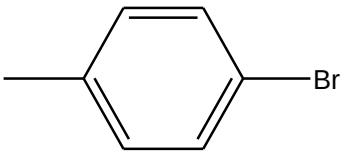
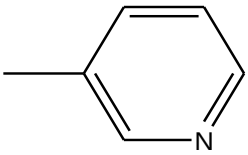
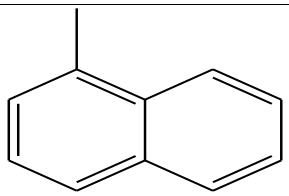
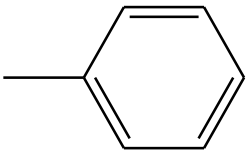


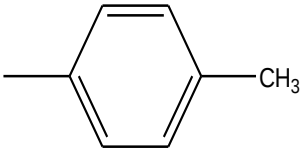
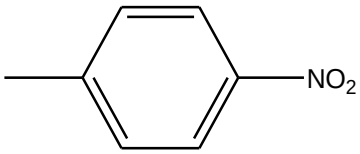
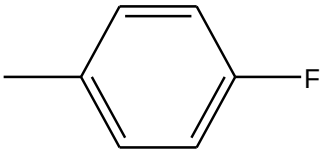
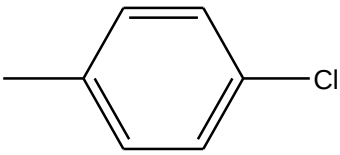
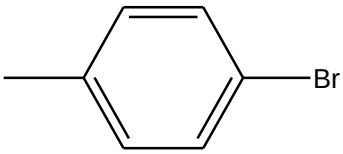
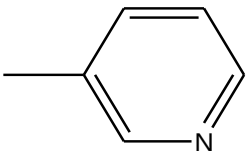
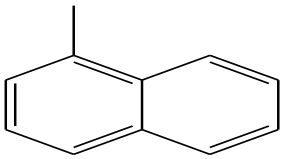
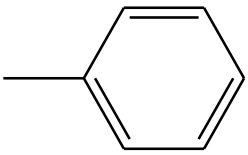

**Figure IV.1. General structure of 1H-1,2,3-triazole derivatives.**

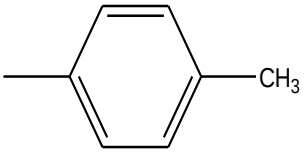
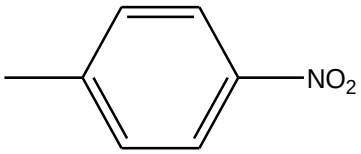
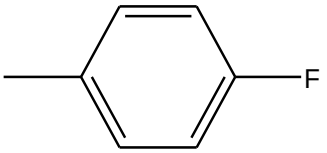
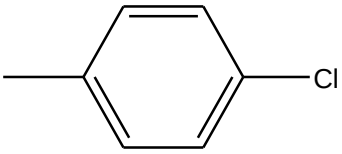
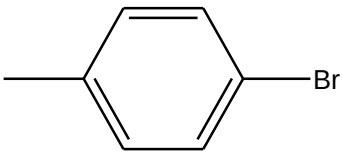
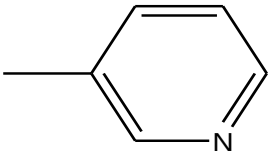
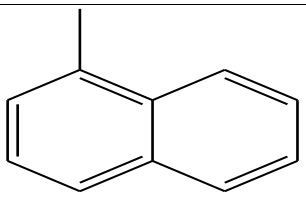
An important objective of this study was to assess the physico chemical domain of derivatives of 1H-1.2.3-triazole reported in the literature to a biological activity. We have selected a bioactive series of derivatives of 1H-1.2.3-triazole. The derivatives of the twenty two of its bioactive molecules are represented in Table IV.1.

**Table IV.1. Chemical structures of 1H-1,2,3-triazole derivatives.**(4) (chemdraw)

| Compounds | $R^1$ | $R^2$ |
|---|---|---|
| **01** | —H |  |
| **02** | —H |  |
| **03** | —H |  |
| **04** | —H |  |
| **05** | —H |  |
| **06** | —H |  |
| **07** | —H |  |
| **08** | —H |  |

| | | |
|---|---|---|
| **09** | —F |  |
| **10** | —F |  |
| **11** | —F |  |
| **12** | —F |  |
| **13** | —F |  |
| **14** | —F |  |
| **15** | —F |  |
| **16** | —F |  |
| **17** | —Cl |  |

| | | |
|---|---|---|
| **18** | —Cl |  |
| **19** | —Cl |  |
| **20** | —Cl |  |
| **21** | —Cl |  |
| **22** | —Cl |  |
| **23** | —Cl |  |
| **24** | —Cl |  |
| **25** | —Br |  |

| | | |
|---|---|---|
| **26** | —Br |  |
| **27** | —Br |  |
| **28** | —Br |  |
| **29** | —Br |  |
| **30** | —Br |  |
| **31** | —Br |  |
| **32** | —Br |  |

### IV.2.2. Study of the physico-chemical properties of 1H-1,2,3-triazole derivatives:

In this part of work, we studied seven physical and chemical properties of thirty two of 1H-1,2,3-triazole derivatives in relation to their activity, was performed by QSAR methods, using HyperChem software (8.0.7).

The properties concerned are surface area (SAG), molar volume (V), energy of hydration (HE), octanol/water partition coefficient (logP), molar refractivity (MR), polarizability (Pol) and molecular weight (MW).

The results obtained in the following table:

**Table IV.2. QSAR parameters of 1H-1, 2,3-triazole derivatives.**

| Compound | Surface area ($Å^3$) | Hydratation Energy (Kcal/mol) | Volume ($Å^2$) | Polarizability ($Å^3$) | Refractivity ($Å^3$) | molecular weight (amu) | LogP |
|---|---|---|---|---|---|---|---|
| **01** | 543.21 | -10.50 | 941.37 | 34.20 | 107.12 | **342.42** | 0.43 |
| **02** | 569.80 | -9.99 | 998.56 | 36.04 | 111.41 | 356.44 | 0.58 |
| **03** | 572.51 | -16.06 | 1006.35 | 36.04 | 112.34 | 387.14 | -4.25 |
| **04** | 517.30 | -10.03 | 894.02 | 34.12 | 107.25 | 360.41 | -0.17 |
| **05** | 533.56 | -9.99 | 927.15 | 36.13 | 111.84 | 376.86 | 0.20 |
| **06** | 544.34 | -9.99 | 947.00 | 36.83 | 114.66 | 421.31 | 0.48 |
| **07** | **504.20** | -11.73 | **871.94** | 33.49 | **103.55** | 343.40 | -1.04 |
| **08** | 568.20 | -10.77 | 1044.23 | 40.38 | 125.32 | 392.47 | 0.50 |
| **09** | 552.62 | -10.18 | 954.61 | 34.11 | 107.25 | 360.40 | -0.17 |
| **10** | 562.48 | **-9.04** | 999.41 | 35.95 | 111.53 | 374.43 | -0.02 |
| **11** | 599.87 | **-16.48** | 1024.43 | 35.95 | 112.47 | 405.40 | **-4.85** |
| **12** | 548.23 | -9.85 | 958.37 | 34.02 | 107.38 | 378.40 | -0.77 |
| **13** | 570.28 | -9.87 | 994.19 | 36.04 | 111.97 | 394.85 | -0.40 |
| **14** | 579.59 | -9.80 | 1010.82 | 36.74 | 114.79 | 439.30 | -0.12 |
| **15** | 546.17 | -11.94 | 942.07 | **33.40** | 103.76 | 361.39 | -1.52 |
| **16** | 591.44 | -11.03 | 1071.08 | 40.29 | 125.45 | 410.47 | -0.10 |
| **17** | 573.69 | -11.29 | 992.22 | 36.13 | 111.84 | 376.86 | 0.20 |
| **18** | 575.72 | -9.45 | 1032.67 | 37.97 | 116.12 | 390.89 | 0.36 |
| **19** | 588.35 | -15.60 | 1045.75 | 37.97 | 117.05 | 421.86 | -4.47 |
| **20** | 571.34 | -10.34 | 1005.65 | 36.04 | 111.97 | 394.85 | -0.40 |
| **21** | 575.36 | -10.17 | 1025.11 | 38.06 | 116.56 | 411.31 | -0.02 |
| **22** | 583.40 | -10.21 | 1043.86 | 38.76 | 119.37 | 455.76 | 0.26 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **23** | 537.82 | -11.48 | 962.65 | 35.42 | 108.35 | 377.85 | -1.14 |
| **24** | 605.08 | -11.08 | 1099.17 | 42.13 | 130.04 | 426.92 | 0.28 |
| **25** | 558.58 | -10.16 | 993.62 | 36.83 | 114.66 | 421.31 | 0.48 |
| **26** | 605.65 | -9.22 | 1054.23 | 38.66 | 118.94 | 435.34 | **0.63** |
| **27** | 598.09 | -14.97 | 1059.18 | 38.67 | 119.87 | 466.31 | -4.2 |
| **28** | 561.23 | -9.81 | 998.75 | 36.74 | 114.79 | 439.30 | -0.12 |
| **29** | 586.98 | -9.78 | 1041.41 | 38.76 | 119.37 | 455.76 | 0.26 |
| **30** | 604.04 | -9.70 | 1066.27 | 39.45 | 122.19 | 500.21 | 0.53 |
| **31** | 556.31 | -11.64 | 983.46 | 36.12 | 111.09 | 422.30 | -0.99 |
| **32** | **621.44** | -10.89 | **1108.99** | **43.01** | **132.86** | **471.37** | 0.55 |

Small value: **green** Big value: **red**

- **Results interpretation:**

According to the results obtained, we notice that the values of the polarizability are generally proportional to the values of the surfaces and the volumes. (Direct Correlation Relationship).

It is also observed that most of the values of the polarizability and of the molar refractivity increase relatively with the size and the molecular weight of the 1H-1,2,3-triazoles studied.

This result is in agreement with the Lorentz-Lorenz formula, which gives a relationship between the polarizability, the molar refractivity and the molecular size. (5)

Therefore, according to relation shows that the molar refractivity and the polarizability increase with the volume and the molecular mass.

For example, compounds 16, 24 and 32 carry bulky substituents have large values of polarizability (40.29$Å^3$), (42.13$Å^3$), (43.01$Å^3$) with high values of molar refractivity (125.45 $Å^3$),(130.04$Å^3$),(132.86$Å^3$) respectively.

The hydration energy in absolute values, the largest are that of compounds 03 (16.48Kcal/mol), 11 (16.06Kcal/mol) and the small values are that of compounds 10 (9.04Kcal/mol) and 26 (9.22Kcal/mol).In fact, in biological environments, Hydration energy is the energy that is released when water molecules attach to ions. These are hydrogen bonds established between them.

For all the compounds of the series have the same number of donor sites in the proton that equal to 1 but have deferential numbers of acceptable sites. In our structure of

Compounds studied nitrogen plays two acceptor roles and donors simultaneously. For example, the compound 11 which has one proton donor sites and eight proton acceptor sites.

Lipophilic is a property that has a major effect on solubility, absorption, Distribution, metabolism and excretion properties, as well as pharmacological activity. Hansch and Leo explained that highly lipophilic molecules will be distributed in the lipid interior of the membranes and will be retained there.(6) For a good bioavailability oral, and optimal the log P must be (0 <log p <3). For log p too high, the drug has a low solubility and for a log P too low; the drug has difficulty penetrating Lipid membranes. (7)

According to the obtained results, they belong to the interval [-4.85 to 0.63]. Compounds 11, 19 presents the low division coefficient [-4.85,-4.47] respectively. When the division coefficient is rather weak, it therefore has a better gastric tollerance. The compounds 2 and 26 which respectively have values greater than 0.63, 0.58 have the ability to depend on plasma proteins.

## IV.3. Drug like and multi-parameter optimization (MPO):

A successful, effective, and safe drug must exhibit balanced properties, including efficacy for its intended target, adequate absorption, distribution, metabolism, and excretion (ADME) properties, and an acceptable safety profile. Achieving this balance between often conflicting requirements is a major challenge in drug discovery. Methods for simultaneously optimizing many factors in a design are often described under the term "multiparameter optimization" (MPO).

In this review article, we will describe how MPO can be used to efficiently design and select high-quality compounds and describe the range of methods that have been used in drug discovery, including; simple "rules of thumb" such as Lipinski's rule; and the Veber and Ghose et al. On the other hand, metric methods aim to combine potency with other parameters into a single metric that can be monitored during optimization, i.e. the older methods of Ligand Efficiency (LE) and Lipophilic Ligand Efficiency (LipE),(8) then for modern someone methods ,find golden triangle.(9)

### IV.3.1. Representation of “drug-like” calculations based on Lipinski:

A major contributor in the field of the characterization of “drug-like” compounds is Lipinski with the “rule of 5”.(10) This rule is the most used for the identification of “drug-

like" compounds**.** (11) According to this rule, compounds that do not pass at least two of the following criteria are very likely to have absorption or permeability problems:

- Molecular mass ≤ 500 Da
- log P ≤ 5
- binding acceptors H ≤ 10
- H bond donors ≤ 5

The "rule of 5" was developed from orally administrable compounds that successfully passed phase II clinical trials. It is therefore not a method for distinguishing compounds that are potentially drugs from those that are not, but rather a method for identifying compounds with low absorption or low permeability**.**

**Table IV.3. Lipinski parameters of 1H-1, 2, 3-triazole derivatives.**

| Compound | HBA | HBD | MW | Log P | Number of Violations |
|---|---|---|---|---|---|
| **01** | 4 | 1 | 342.42 | 0.43 | 0 |
| **02** | 4 | 1 | 356.44 | 0.58 | 0 |
| **03** | 5 | 1 | 387.14 | -4.25 | 0 |
| **04** | 5 | 1 | 360.41 | -0.17 | 0 |
| **05** | 5 | 1 | 376.86 | 0.20 | 0 |
| **06** | 4 | 1 | 421.31 | 0.48 | 0 |
| **07** | 4 | 1 | 343.40 | -1.04 | 0 |
| **08** | 4 | 1 | 392.47 | 0.50 | 0 |
| **09** | 5 | 1 | 360.40 | -0.17 | 0 |
| **10** | 5 | 1 | 374.43 | -0.02 | 0 |
| **11** | 6 | 1 | 405.40 | -4.85 | 0 |
| **12** | 6 | 1 | 378.40 | -0.77 | 0 |
| **13** | 6 | 1 | 394.85 | -0.40 | 0 |
| **14** | 5 | 1 | 439.30 | -0.12 | 0 |
| **15** | 5 | 1 | 361.39 | -1.52 | 0 |
| **16** | 5 | 1 | 410.47 | -0.10 | 0 |
| **17** | 5 | 1 | 376.86 | 0.20 | 0 |
| **18** | 5 | 1 | 390.89 | 0.36 | 0 |
| **19** | 6 | 1 | 421.86 | -4.47 | 0 |
| **20** | 6 | 1 | 394.85 | -0.40 | 0 |
| **21** | 6 | 1 | 411.31 | -0.02 | 0 |
| **22** | 5 | 1 | 455.76 | 0.26 | 0 |

| | | | | | |
|---|---|---|---|---|---|
| **23** | 5 | 1 | 377.85 | -1.14 | 0 |
| **24** | 5 | 1 | 426.92 | 0.28 | 0 |
| **25** | 4 | 1 | 421.31 | 0.48 | 0 |
| **26** | 4 | 1 | 435.34 | 0.63 | 0 |
| **27** | 5 | 1 | 466.31 | -4.2 | 0 |
| **28** | 5 | 1 | 439.30 | -0.12 | 0 |
| **29** | 5 | 1 | 455.76 | 0.26 | 0 |
| **30** | 4 | 1 | 500.21 | 0.53 | 0 |
| **31** | 4 | 1 | 422.30 | -0.99 | 0 |
| **32** | 4 | 1 | 471.37 | 0.55 | 0 |

- **Results interpretation:**

The Lipinski rule is the most used for compounds (Drug-like). We remember that this rule is intended to identify compounds posing absorption and permeability problems.(12)

According to the results obtained in the table above, it is observed that all the logP number values of the 1H-1,2,3-triazole derivatives studied are less than 5 for lipophilic, ranging from-4.85 to 0.63 so these compounds are better solubilized in aqueous and lipid solutions. On the one hand, a negative value for logP indicates that the too hydrophilic compound. It therefore has good solubility in water better gastric tolerance and effective elimination by kidneys. On the other hand, a positive value for the logP indicates that the compound is too lipophilic .Thus, it has a good permeability through the biological membrane, a better bond of plasma proteins, elimination by metabolism, but a low solubility and gastric tolerance.

We have for example the compounds (3, 19, and 11) take the values of logP respectively (-4.23, -4.47,-4.85) this indicate that these compounds have good solubility and better gastric tolerance. We also take the compound 26 which has a maximum logP (0.63) so it has better permeability through the biological membrane.

All the compounds of the series comply with rules 3 and 4, this means that the hydrogen acceptor numbers less than 10 (O, F, Br, Cl) and the number of hydrogen donors less than 5 (OH, NH) .results in increased hydration energy and low permeability .Because the HBAs which are of a large number leads to a low permeability through a bilayer membrane. The smaller number leads to better permeability.(13)

And for weight All compounds in the series have molar masses below 500 Da except 30 (500.21), and all compounds are likely soluble and easily cross cell membranes.

Based on these criteria it can be said that the compounds of the series are acceptable to be administered orally.(14)

Therefore, for the number of violation we notice that all the compounds have a null violation for the Lipinski rule.

All these results present all the compounds meet the Lipinski rules (rules of five), and suggest that these compounds theoretically would have no problems with oral bioavailability.

### IV.3.2. Veber's rules:

Two other criteria introduced by Veber PSA (Polar Surface area) and number of links that can rotate ≤ 10, are often used in addition to the "rule of 5". These limits were established from measurements of the oral bioavailability of drug candidates.(15)

There are two defined descriptors for ideal oral bioavailability:

- Rotating links are less than 10.
- Polar area is less than 140 $Å^2$.

PSA is used to calculate percent absorbance (% ABS) according to the following equation: % ABS = 109 ± 0.345 x PSA.(16)

**Table IV.4. Veber's rules for 1H- 1, 2, 3-triazole derivatives.**

| Compound | NRB | PSA ($Å^2$) | ABS% |
|---|---|---|---|
| **01** | 5 | 85.26 | 79.58 |
| **02** | 5 | 85.26 | 79.58 |
| **03** | 6 | 131.08 | 63.78 |
| **04** | 5 | 85.26 | 79.58 |
| **05** | 5 | 85.26 | 79.58 |
| **06** | 5 | 85.26 | 79.58 |
| **07** | 5 | 98.15 | 75.14 |
| **08** | 5 | 85.26 | 79.58 |
| **09** | 5 | 85.26 | 79.58 |
| **10** | 5 | 85.26 | 79.58 |
| **11** | 6 | 131.08 | 63.78 |
| **12** | 5 | 85.26 | 79.58 |
| **13** | 5 | 85.26 | 79.58 |
| **14** | 5 | 85.26 | 79.58 |

| | | | |
|---|---|---|---|
| **15** | 5 | 98.15 | 75.14 |
| **16** | 5 | 85.26 | 79.58 |
| **17** | 5 | 85.26 | 79.58 |
| **18** | 5 | 85.26 | 79.58 |
| **19** | 6 | 131.08 | 63.78 |
| **20** | 5 | 85.26 | 79.58 |
| **21** | 5 | 85.26 | 79.58 |
| **22** | 5 | 85.26 | 79.58 |
| **23** | 5 | 98.15 | 75.14 |
| **24** | 5 | 85.26 | 79.58 |
| **25** | 5 | 85.26 | 79.58 |
| **26** | 5 | 85.26 | 79.58 |
| **27** | 6 | 131.08 | 63.78 |
| **28** | 5 | 85.26 | 79.58 |
| **29** | 5 | 85.26 | 79.58 |
| **30** | 5 | 85.26 | 79.58 |
| **31** | 5 | 98.15 | 75.14 |
| **32** | 5 | 85.26 | 79.58 |

**Note:** Using MarvinSketch 17.1.9 software to calculate NRB and PSA.

- **Results interpretation:**

According to the results obtained in the table above, it is observed that all the NRB number values of the 1H- 1, 2, 3-triazole derivatives studied are less than 10. The low number of rotary bonds in the compounds studied indicates that these ligands upon binding to a protein change their conformation only slightly.

For the PSA results, we note that all the compounds of the series studied have values below 140 Å², which shows the good prediction of oral bioavailability and transport through biological membranes.

For the percent absorption (%ABS) values, all compounds can be assured to have a large %ABS ranging from 63.78 % to 79.58 %, indicating that these compounds should have good cell membrane permeability.

Therefore, we conclude that all the studied compounds respect Veber's rule.

### IV.3.3. Efficiency of ligand "LE":

Ligand Efficiency It is a Metric methods originated from the observation that smaller compounds tend to have better physico-chemical properties and good ADMEs than larger compounds.

The smallest compound tends to have the best physicochemical properties and good ADME with respect to ligand efficacy**,** on the other hand larger compounds have weaker physicochemical and ADME properties.(1)

The calculation of this parameter is performed as follows:

**pIC50 = - log (IC50)** **LE = 1.4 pIC50 / NH**

**pIC50**: biological activity.

**NH**: is the number of heavy atoms.

**Table IV.5. Ligand efficacy of 1H- 1,2,3-triazole derivatives.**

| Compounds | pIC50(4) | NH | LE |
|---|---|---|---|
| **01** | 2.161 | 24 | 0.126 |
| **02** | 2.303 | 25 | 0.129 |
| **03** | 2.431 | 27 | 0.126 |
| **04** | 2.378 | 25 | 0.133 |
| **05** | 2.367 | 25 | 0.132 |
| **06** | 2.270 | 25 | 0.127 |
| **07** | 2.247 | 24 | 0.131 |
| **08** | 2.365 | 28 | 0.118 |
| **09** | 2.613 | 25 | 0.146 |
| **10** | 2.526 | 26 | 0.136 |
| **11** | 2.616 | 28 | 0.131 |
| **12** | 2.620 | 26 | 0.141 |
| **13** | 2.717 | 26 | 0.146 |
| **14** | 2.578 | 26 | 0.139 |
| **15** | 2.445 | 25 | 0.137 |
| **16** | 2.549 | 29 | 0.123 |
| **17** | 2.327 | 25 | 0.130 |
| **18** | 2.140 | 26 | 0.115 |
| **19** | 2.447 | 28 | 0.122 |
| **20** | 2.196 | 26 | 0.118 |
| **21** | 2.230 | 26 | 0.120 |
| **22** | 2.266 | 26 | 0.122 |
| **23** | 2.453 | 25 | 0.137 |
| **24** | 2.400 | 29 | 0.116 |
| **25** | 2.323 | 25 | 0.130 |
| **26** | 2.294 | 26 | 0.123 |
| **27** | 2.385 | 28 | 0.119 |
| **28** | 2.236 | 26 | 0.120 |
| **29** | 2.185 | 26 | 0.118 |

| | | | |
|---|---|---|---|
| **30** | 2.350 | 26 | 0.126 |
| **31** | 2.401 | 25 | 0.134 |
| **32** | 2.348 | 29 | 0.113 |

- **Results interpretation:**

From the results obtained in Table IV.5, we observe that LE decreases with the increase in the number of heavy atoms, this means that to obtain a high and significant LE, compounds with weak heavy atoms and this can be explained by the correlation that exists between the size of the compound and their physico-chemical properties.

It has been suggested that a contributing factor to the typically observed for LE in larger ligands may be due to less favorable binding entropies for larger and more flexible ligands [36]. And that to see clearly in the table IV.6 as for example for the compounds (4, 5, 7, 9) having the small values of heavy atoms take respectively the efficiency (0.133, 0.132, 0.131, 0.3297, 0.146), while the two compounds (24 and 32) that have the largest value of heavy atom 29 corresponding to the small values of ligand efficiency 0.116 and 0.113 respectively, that may have bad physicochemical properties and ADME. On the other hand, the efficiency varied proportionally with the pIC50 values.

### IV.3.4. Ligand Lipophilicity Efficiency "LLE":

Is a parameter that combines both activity and lipophilicity, It is defined as a measure of the efficiency of a ligand to bind to a given target.

On the other hand, we will study the lipophilic efficiency (LipE) of the ligand, since it is well established that the lipophilicity of a molecule plays a crucial role in determining its ability to be a drug candidate.(17) To maximize potency while keeping lipophilicity as low as possible, due to the association between high lipophilicity and several issues including low solubility, membrane permeability, metabolic stability, etc.(18)

For an LLE of 5 to 7 or even more, it can be said that these optimized compounds should be more selective. (19)

Leeson and Springthorpe (17) define Lipophilic Efficiency (LipE) as follows:

$$\mathbf{LLE = LipE = pIC50 - log\ P}$$

**Table IV.6. Ligand Lipophilicity Efficiency of 1H- 1, 2, 3-triazole derivatives.**

| Compounds | pIC50(4) | LogP | LLE |
|---|---|---|---|
| **01** | 2.161 | 0.43 | 1.731 |
| **02** | 2.303 | 0.58 | 1.723 |
| **03** | 2.431 | -4.25 | 6.681 |
| **04** | 2.378 | -0.17 | 2.548 |
| **05** | 2.367 | 0.20 | 2.167 |
| **06** | 2.270 | 0.48 | 1.790 |
| **07** | 2.247 | -1.04 | 3.287 |
| **08** | 2.365 | 0.50 | 1.865 |
| **09** | 2.613 | -0.17 | 2.783 |
| **10** | 2.526 | -0.02 | 2.546 |
| **11** | 2.616 | -4.85 | 7.466 |
| **12** | 2.620 | -0.77 | 3.390 |
| **13** | 2.717 | -0.40 | 3.117 |
| **14** | 2.578 | -0.12 | 2.698 |
| **15** | 2.445 | -1.52 | 3.965 |
| **16** | 2.549 | -0.10 | 2.649 |
| **17** | 2.327 | 0.20 | 2.127 |
| **18** | 2.140 | 0.36 | 1.780 |
| **19** | 2.447 | -4.47 | 6.917 |
| **20** | 2.196 | -0.40 | 2.596 |
| **21** | 2.230 | -0.02 | 2.250 |
| **22** | 2.266 | 0.26 | 2.006 |
| **23** | 2.453 | -1.14 | 3.593 |
| **24** | 2.400 | 0.28 | 2.120 |
| **25** | 2.323 | 0.48 | 1.843 |
| **26** | 2.294 | 0.63 | 1.664 |
| **27** | 2.385 | -4.20 | 6.585 |
| **28** | 2.236 | -0.12 | 2.356 |
| **29** | 2.185 | 0.26 | 1.925 |
| **30** | 2.350 | 0.53 | 1.820 |
| **31** | 2.401 | -0.99 | 3.391 |
| **32** | 2.348 | 0.55 | 1.798 |

- **Results interpretation:**

According to the lipophilic ligand efficiency obtained in Table IV.6 it can be seen that the LLE change values between 1.664 and 7.466. We takes for example the compounds (3, 19, 27) having values of LLE respectively (60681, 6.917, 6.585) which belong to the range 5-7 this indicates that these compounds have been successfully optimized. On the other

hand, the values of the lipophilic efficiency of the ligand for the compounds (1,7,14) are respectively (1.731,3.287,2.698) it is noted that none of the compounds reaches an LLE greater than 5.In these cases ,the affinity gain is accompanied by an increase in lipophilicity . In this respect, the optimization was not as optimal as in the first example.

**IV.3.5. Golden Triangle:**

The Golden Triangle is a visualization tool advanced from in vitro permeability, in vitro clearance, and computational data to advice medicinal chemists conception metabolically stable, permeable, and potent drug candidates.

Classify compounds as permeable and stable and plot molecular weight (MW) versus octanol: buffer (pH 7.4) partition coefficient (log D) or estimated octanol: buffer (pH 7.4) partition coefficient (log D) the relationship graph shows useful trends.(9)

Analysis of at least two orthogonal trends, such as permeability and clearance, can be extremely effective in balancing and optimizing multiple properties. In addition, molecular weight and logD affect potency-efficiency calculations, allowing potency, clearance and permeability to be optimized simultaneously.(9)

For a successful golden triangle construction, we move the design properties to an area with a baseline of log $D_{7.4}$=-2.0 to log $D_{7.4}$= 5, axes of Mass weight (MW) is from 200 and a peak in the symmetric triangle between Log D7.4 = 1 and 2 and MW = 450 (peak coordinates: Log $D_{7.4}$ =1.5 and MW = 450); these limits give a triangular shape, called golden triangle. These trends lead to a shaped area known as the Golden Triangle and molecules in this area that: are low clearances and permeable should follow the golden triangle rule.

Log D is a measure of logP at a given pH for a compound of a certain pKa. This parameter thus takes into account the concentration of ionized molecules.(1)

**Table IV.7. Distribution coefficients of 1H-1, 2,3-triazoles.**

| Compounds | MW | Log D |
|---|---|---|
| **01** | 342.42 | 2.91 |
| **02** | 356.44 | 3.38 |
| **03** | 387.14 | 2.87 |
| **04** | 360.41 | 3.05 |
| **05** | 376.86 | 3.43 |
| **06** | 421.31 | 3.70 |

| | | |
|---|---|---|
| **07** | 343.40 | 1.60 |
| **08** | 392.47 | 3.91 |
| **09** | 360.40 | 3.05 |
| **10** | 374.43 | 3.52 |
| **11** | 405.40 | 3.00 |
| **12** | 378.40 | 3.19 |
| **13** | 394.85 | 3.57 |
| **14** | 439.30 | 3.84 |
| **15** | 361.39 | 1.74 |
| **16** | 410.47 | 4.05 |
| **17** | 376.86 | 3.43 |
| **18** | 390.89 | 3.90 |
| **19** | 421.86 | 3.38 |
| **20** | 394.85 | 3.57 |
| **21** | 411.31 | 3.95 |
| **22** | 455.76 | 4.22 |
| **23** | 377.85 | 2.12 |
| **24** | 426.92 | 4.43 |
| **25** | 421.31 | 3.70 |
| **26** | 435.34 | 4.17 |
| **27** | 466.31 | 3.66 |
| **28** | 439.30 | 3.84 |
| **29** | 455.76 | 4.22 |
| **30** | 500.21 | 4.50 |
| **31** | 422.30 | 2.39 |
| **32** | 471.37 | 4.71 |

**Note:** Using MarvinSketch 17.1.9 software to calculate LogD.

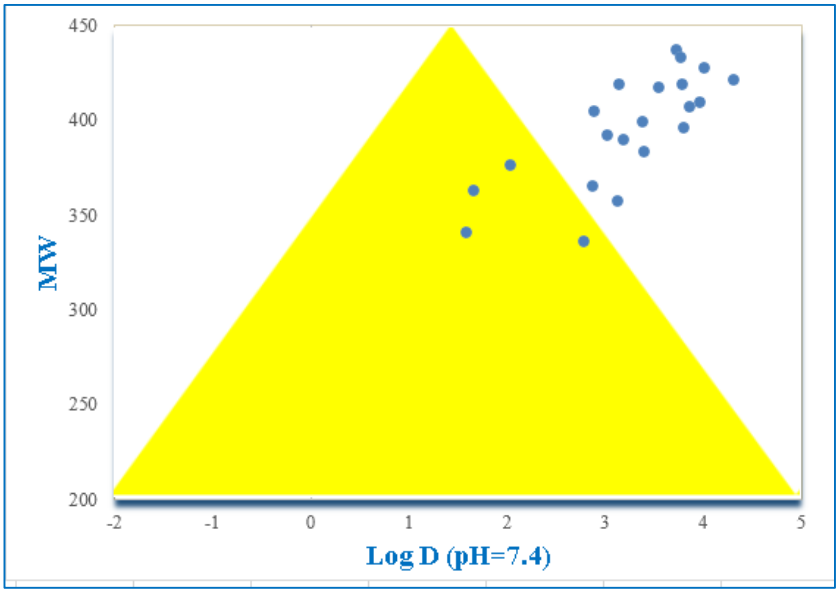



**Figure .IV. 2. The Golden Triangle.**

**Results interpretation:**

Compounds that reside within the golden triangle are more likely to be both metabolically stable and possess good membrane permeability than those that lie outside.

However, in our case, the golden triangle helps us to sort out because it shows that the majority of the compounds studied are outside the triangle except for compounds 1, 7, 15 and 23.

These compounds (1, 7, 15 and 23) are located inside the golden triangle, so that these derivatives have a good permeability of the active ingredient and an efficient hepatic clearance.

## IV.4. Calculation of Padel descriptors :

The PaDEL descriptor was used to calculate a pool of descriptors of the optimized molecules of 1H-1,2,3-triazole derivatives.

**nAtom:** Number of atoms.(20)

**Zagreb:** The Zagreb indices is topological descriptors treat the structure of the compound as a graph, with atoms as vertices and covalent bonds as edges. (21) Various researchers use the Zagreb indices in their QSPR and QSAR studies, they are also included in a number of programs used for the routine computation of topological indices. (22)

The original Zagreb indices are defined as follows:

$$\mathbf{M_1} = \sum_{\mathbf{vertices}} a_i^2$$

$$\mathbf{M_2} = \sum_{\mathbf{edges}} \mathbf{a_i}\ \mathbf{a_j}$$

The first zagreb indice $\mathbf{M_1}$: defined as the sum of the squared vertex degrees $\mathbf{a_i}$ ,over all graph vertices i.

The second zagreb indice $\mathbf{M_2}$: defined as the sum of the vertex degrees $\mathbf{a_i}\ \mathbf{a_j}$ , taken over all edges ij. (23)

**Table IV.8. Padel descriptors.**

| Compounds | Zagreb | nAtom |
|---|---|---|
| **01** | 128 | 42 |
| **02** | 134 | 45 |
| **03** | 144 | 44 |
| **04** | 134 | 42 |
| **05** | 134 | 42 |
| **06** | 134 | 42 |
| **07** | 128 | 41 |
| **08** | 154 | 48 |
| **09** | 134 | 42 |
| **10** | 140 | 45 |
| **11** | 150 | 44 |
| **12** | 140 | 42 |
| **13** | 140 | 42 |
| **14** | 140 | 42 |
| **15** | 134 | 41 |
| **16** | 160 | 48 |
| **17** | 134 | 42 |
| **18** | 140 | 45 |
| **19** | 150 | 44 |
| **20** | 140 | 42 |
| **21** | 140 | 42 |
| **22** | 140 | 42 |
| **23** | 134 | 41 |
| **24** | 160 | 48 |
| **25** | 134 | 42 |
| **26** | 140 | 45 |
| **27** | 150 | 44 |
| **28** | 140 | 42 |
| **29** | 140 | 42 |
| **30** | 140 | 42 |
| **31** | 134 | 41 |
| **32** | 160 | 48 |

## IV.5. Conclusion:

In this chapter, a qualitative study of the structure-activity relationship of a series of 1H-1,2,3-triazole derivatives has been presented. Also offers a structural comparison between the thirty compounds studied. She also provides an in-depth discussion of drug likeness.

Polarizability values are generally proportional to volume and area values and to refractivity. Compound 32 takes the important values for polarizability 43.01 $Å^3$ and refractivity 132.86 $Å^3$ .

The hydration energy in absolute value, the largest being that of compound 11 (16.48 kcal/mol) .It is therefore the best distribution in the tissues.

In this study, all compounds have logP values less than 5 are soluble in aqueous solution and therefore capable of expecting the surface of the membranes and have a gastric tolerance.

By comparison, of the calculation results and the criteria of the Lipinski rule, it was noticed that all the compounds of the series confirm these criteria. This shows that the compounds studied theoretically do not present any oral bioavailability problem.

Another study based on Veber's rules shows that the compounds of the series studied are all in correlation with this rule, that is to say that the number of rotary bonds less than 10 to all the compounds, and the PSA values are all less than 140 $Å^2$ . Also the most compounds having the greatest values of %ABS (79.58 %) so they present a better membrane permeability.

For the study of LE we find that compound 9 has the greatest value which is 0.146, which allows it good physico-chemical and ADME properties.

For the study of LLE we find that the compounds (3, 19, 27) having values of LLE respectively (60681, 6.917, 6.585) which belong to the range 5-7 this indicates that these compounds have been successfully optimized.

The Compounds (1, 7, 15 and 23) lie inside the golden triangle, so these derivatives have a good permeability of the active ingredient and an efficient hepatic clearance.

## References

(1). Aouisset D. In silico analysis of pyrazole derivatives applied to drug design [Memoir of master]: University Mohammed Khider Biskra; 2022.

(2). Dermeche K, Tchouar N, Belaidi S, Salah T. Qualitative Structure-Activity Relationships and 2D-QSAR Modeling of TNF-α Inhibition by Thalidomide Derivatives. Journal of Bionanoscience. 2015;9(5):395-400.

(3). Kerassa A, Belaidi S, Harkati D, Lanez T, Prasad O, Sinha L. Investigations on Molecular Structure, Electronic Properties, NLO Properties and Comparison of Drug-Likeness of Triazolothiadiazole Derivatives by Quantum Methods and QSAR Analysis. Reviews in Theoretical Science. 2016;4(1):85-96. https://doi.org/10.1166/rits.2016.1050.

(4). Kaushik C, Chahal M, Luxmi R, Kumar D, Kumar A, Kumar M, et al. Synthesis, characterization and biological activities of sulfonamide tagged 1, 2, 3-triazoles. Synthetic Communications. 2020;50(22):3443-61.

(5). Kragh H. The Lorenz-Lorentz formula: Origin and early history. Substantia. 2018;2(2):7-18.

(6). Ghose AK, Crippen GM. 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. 1987;27(1):21-35.

(7). Burgot G, Bona M, Christen M, Burgot J. First relationships between detoxication properties of some 1, 2-dithiole-3-thiones and their log P. International journal of pharmaceutics. 1996;129(1-2):295-9.

(8). D. Segall M. Multi-parameter optimization: identifying high quality compounds with a balance of properties. Current pharmaceutical design. 2012;18(9):1292.

(9). Johnson TW, Dress KR, Edwards M. Using the Golden Triangle to optimize clearance and oral absorption. Bioorganic & medicinal chemistry letters. 2009;19(19):5560-4.

(10). Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced drug delivery reviews. 1997;23(1-3):3-25.

(11). Lipinski CA. Lead-and drug-like compounds: the rule-of-five revolution. Drug discovery today: Technologies. 2004;1(4):337-41.

(12). Aurélien M. Création et utilisation de chimiothèques optimisées pour la recherche «in silico» de nouveaux composés bioactifs [Thesis of Doctorat]: University of Rleans; 2006.

(13). Lipinski CA. F.. Lombardo, BW Dominy, PJ Feeney. Adv Drug Delivery Rev. 1997;23:3-30.

(14). Qaneinasab S, Bayat Z. Prediction of Henry's law constant of benzimidazole derivatives using quantum chemical calculation. Journal of Chemical and Pharmaceutical Research. 2011;3(4):561-4.

(15). Veber DF, Johnson SR, Cheng H-Y, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. Journal of medicinal chemistry. 2002;45(12):2615-23.

(16). Remko M, Swart M, Bickelhaupt FM. Theoretical study of structure, pKa, lipophilicity, solubility, absorption, and polar surface area of some centrally acting antihypertensives. Bioorganic & medicinal chemistry. 2006;14(6):1715-28.

(17). Leeson PD, Springthorpe B. The influence of drug-like concepts on decision-making in medicinal chemistry. Nature reviews Drug discovery. 2007;6(11):881-90.

(18). Edwards MP, Price DA. Role of physicochemical properties and ligand lipophilicity efficiency in addressing drug safety risks. Annual Reports in Medicinal Chemistry. 45: Elsevier; 2010. p. 380-91.

(19). Pliska V, Testa B, van de Waterbeemd H, Mannhold R, Kubinyi H, Timmerman H. Lipophilicity in drug action and toxicology: VCH Weinheim; 1996.

(20). Ouassaf M, Belaidi S, Shtaiwi A, Chtita S. Quantitative Structure Activity Relationship (QSAR) Investigations and Molecular Docking Analysis of Plasmodium Protein Farnesyltransferase Inhibitors as Potent Antimalarial Agents. Jordan Journal of Pharmaceutical Sciences. 2022;15(3):315-40.

(21). Khamouli S. Study of the structure, QSAR properties and molecular docking of some amino-pyrimidine derivatives as Novel Mycobacterium tuberculosis inhibitors for drug design [Doctorat ]: University Mohammed Khider Biskra; 2019.

(22). Miličević A, Nikolić S. On variable Zagreb indices. Croatica chemica acta. 2004;77(1-2):97-101.

(23). Bonche D, Trinajstič N. Overall molecular descriptors. 3. Overall Zagreb indices. SAR and QSAR in Environmental Research. 2001;12(1-2):213-36.

# *Chapter V: Quantitative structure-activity relationship (QSAR) of a series of 1H-1,2,3-triazole derivatives and application of chemometric methods*

## V.1. Introduction :

Quantitative structure-activity relationship (QSAR) analysis is based on the general principle of medicinal chemistry that the biological activity of a ligand or compound is related to its molecular structure or properties, and structurally similar molecules may have similar biological activities.(1) Such molecular structural information is encoded in molecular descriptors and a QSAR model defines mathematical relationships between descriptors and biological activities of known ligands to predict unknown ligands activities. QSAR methods have been applied in several scientific studies including chemistry, biology, and toxicology and drug discovery to predict and classify biological activities of virtual or newly synthesized compounds.(2-5)

Multiple linear regression (MLR) is a mathematical tool that quantifies the relationship between a dependent variable and one or more independent variables, it was used to develop QSAR models, and all the variables that have been included in the model are significant.(6)

There are a large number of studies suggesting that nonlinear methods can offer significant predictive performance in modeling. The application of nonlinear methods have shown considerable interest in QSAR study. The nonlinear method of the artificial neural network (ANN) takes into account all the nonlinear relationships between the physico-chemical descriptors and the corresponding biological activity. Indeed, one of the most notable attributes of ANN is its ability to make predictions about new data with a degree of accuracy similar to that obtained with the training base. The descriptors appearing in the QSAR model are intended for the virtual screening of new molecules that have potent and enhanced activity.(7)

In this work, we are interested in the study of the physicochemical properties of 1H-1,2,3-triazole and its derivatives which make it possible to predict the physicochemical parameters which are influential on the biological activity, in order to predict the biological activity of new molecules, it is for this reason the choice of a series of thirty two derivatives of 1H-1,2,3-triazole with different physico-chemical descriptors has been made,using multiple linear regression (MLR) as well as artificial neural network (ANN) methods .

## V.2. Quantitative study of the structure-activity relationship:

In this part of this work we will establish a quantitative structure-activity relationship between the antioxidant activity of a series derived from 1H-1,2,3-triazole and their structural descriptors, therefore we propose a quantitative model, and we try to interpret the activity of these molecules based.

For this purpose, based on the results obtained in the previous chapter, of preoptimization molecular structures and their minimizations, we have highlighted all possible molecular descriptors. These are the different physical properties and chemicals, known as physico-chemical descriptors as well as the parameter de Lipinski, Veber, by the use of the module " QSAR properties of the hyperchem 8.07 and Marvinsketch 17.1.9 and other descriptors using Padel were used as independent variables and have been correlated with biological activities of 1H-1,2,3-trazole derivatives to generate QSAR models. For that purpose, multiple linear regression MLR, and artificial neural networks (ANNs) are used.

**Table V.1. Values of molecular descriptors.**

| Compounds | V ($Å^2$) | SA ($Å^3$) | EH (kcal/mol) | POL ($Å^3$) | MW (amu) | REF ($Å^3$) | logP | HBA |
|---|---|---|---|---|---|---|---|---|
| **01** | 941.37 | 543.21 | -10.50 | 34.20 | 342.42 | 107.12 | 0.43 | 4 |
| **02** | 1006.35 | 572.51 | -16.06 | 36.04 | 387.14 | 112.34 | -4.25 | 5 |
| **03** | 1044.23 | 568.20 | -10.77 | 40.38 | 392.47 | 125.32 | 0.50 | 4 |
| **04** | 954.61 | 552.62 | -10.18 | 34.11 | 360.40 | 107.25 | -0.17 | 5 |
| **05** | 999.41 | 562.48 | -9.04 | 35.95 | 374.43 | 111.53 | -0.02 | 5 |
| **06** | 958.37 | 548.23 | -9.85 | 34.02 | 378.40 | 107.38 | -0.77 | 6 |
| **07** | 994.19 | 570.28 | -9.87 | 36.04 | 394.85 | 111.97 | -0.40 | 6 |
| **08** | 1010.82 | 579.59 | -9.80 | 36.74 | 439.30 | 114.79 | -0.12 | 5 |
| **09** | 942.07 | 546.17 | -11.94 | 33.40 | 361.39 | 103.76 | -1.52 | 5 |
| **10** | 992.22 | 573.69 | -11.29 | 36.13 | 376.86 | 111.84 | 0.20 | 5 |
| **11** | 1032.67 | 575.72 | -9.45 | 37.97 | 390.89 | 116.12 | 0.36 | 5 |
| **12** | 1045.75 | 588.35 | -15.60 | 37.97 | 421.86 | 117.05 | -4.47 | 6 |
| **13** | 1025.11 | 575.36 | -10.17 | 38.06 | 411.31 | 116.56 | -0.02 | 6 |
| **14** | 1099.17 | 605.08 | -11.08 | 42.13 | 426.92 | 130.04 | 0.28 | 5 |
| **15** | 993.62 | 558.58 | -10.16 | 36.83 | 421.31 | 114.66 | 0.48 | 4 |
| **16** | 1054.23 | 605.65 | -9.22 | 38.66 | 435.34 | 118.94 | 0.63 | 4 |
| **17** | 1059.18 | 598.09 | -14.97 | 38.67 | 466.31 | 119.87 | -4.2 | 5 |
| **18** | 1041.41 | 586.98 | -9.78 | 38.76 | 455.76 | 119.37 | 0.26 | 5 |
| **19** | 1108.99 | 621.44 | -10.89 | 43.01 | 471.37 | 132.86 | 0.55 | 4 |
| **20t** | 998.56 | 569.80 | -9.99 | 36.04 | 356.44 | 111.41 | 0.58 | 4 |
| **21t** | 894.02 | 517.30 | -10.03 | 34.12 | 360.41 | 107.25 | -0.17 | 5 |
| **22t** | 947.00 | 544.34 | -9.99 | 36.83 | 421.31 | 114.66 | 0.48 | 4 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **23t** | 1024.43 | 599.87 | -16.48 | 35.95 | 405.40 | 112.47 | -4.85 | 6 |
| **24t** | 1071.08 | 591.44 | -11.03 | 40.29 | 410.47 | 125.45 | -0.10 | 5 |
| **25t** | 1066.27 | 604.04 | -9.70 | 39.45 | 500.21 | 122.19 | 0.53 | 4 |
| **26t** | 983.46 | 556.31 | -11.64 | 36.12 | 422.30 | 111.09 | -0.99 | 4 |

**Table V.1. Continued**

| Compounds | HBD | nAtom | NH | NRB | PSA ($Å^2$) | LogD | Zagreb |
|---|---|---|---|---|---|---|---|
| **01** | 1 | 42 | 24 | 5 | 85.26 | 2.91 | 128 |
| **02** | 1 | 44 | 27 | 6 | 131.08 | 2.87 | 144 |
| **03** | 1 | 48 | 28 | 5 | 85.26 | 3.91 | 154 |
| **04** | 1 | 42 | 25 | 5 | 85.26 | 3.05 | 134 |
| **05** | 1 | 45 | 26 | 5 | 85.26 | 3.52 | 140 |
| **06** | 1 | 42 | 26 | 5 | 85.26 | 3.19 | 140 |
| **07** | 1 | 42 | 26 | 5 | 85.26 | 3.57 | 140 |
| **08** | 1 | 42 | 26 | 5 | 85.26 | 3.84 | 140 |
| **09** | 1 | 41 | 25 | 5 | 98.15 | 1.74 | 134 |
| **10** | 1 | 42 | 25 | 5 | 85.26 | 3.43 | 134 |
| **11** | 1 | 45 | 26 | 5 | 85.26 | 3.90 | 140 |
| **12** | 1 | 44 | 28 | 6 | 131.08 | 3.38 | 150 |
| **13** | 1 | 42 | 26 | 5 | 85.26 | 3.95 | 140 |
| **14** | 1 | 48 | 29 | 5 | 85.26 | 4.43 | 160 |
| **15** | 1 | 42 | 25 | 5 | 85.26 | 3.7 | 134 |
| **16** | 1 | 45 | 26 | 5 | 85.26 | 4.17 | 140 |
| **17** | 1 | 44 | 28 | 6 | 131.08 | 3.66 | 150 |
| **18** | 1 | 42 | 26 | 5 | 85.26 | 4.22 | 140 |
| **19** | 1 | 48 | 29 | 5 | 85.26 | 4.71 | 160 |
| **20t** | 1 | 45 | 25 | 5 | 85.26 | 3.38 | 134 |
| **21t** | 1 | 42 | 25 | 5 | 85.26 | 3.05 | 134 |
| **22t** | 1 | 42 | 25 | 5 | 85.26 | 3.7 | 134 |
| **23t** | 1 | 44 | 28 | 6 | 131.08 | 3 | 150 |
| **24t** | 1 | 48 | 29 | 5 | 85.26 | 4.05 | 160 |
| **25t** | 1 | 42 | 26 | 5 | 85.26 | 4.5 | 140 |
| **26t** | 1 | 41 | 25 | 5 | 98.15 | 2.39 | 134 |

**t: test set**

## V.2.1. Multiple Linear Regression (MLR):

### V.2.1.1. QSAR model development:

Despite being the oldest, MLR still remains one of the most popular approaches to build QSAR models.This is due to its simple practicaluse, ease of interpretation and transparency. Indeed, the key algorithm is available and accurate predictions can be provided.(8) The values of the calculated descriptors are those listed in (Table V.1) . Data were randomly divided into two groups: a training set (internal validation) and a testing set (external validation).

In this step, we tried to develop the QSAR model, in other words, from the relationships identified in isolation, it was a question of using methods of analysis allowing to group the different parameters into a single relationship to explain correlations between physico-chemical and pIC50 biological activities of 1H-1,2,3-triazole derivatives, this analysis was performed using the statistical software XLSTAT.

The selection of a set of appropriate descriptors from a large number of them requires a method that is able to discriminate between parameters.

Among the various QSAR equations, the best QSAR models were selected based on various statistical parameters such as:

- Correlation coefficient R that measures the degree of line association between two variables .It varies in value from 0 to 1.
- The square of Correlation coefficient ($R^2 > 0.6$) which is the relative measure of the goodness of fit.
- The standard error of the estimate representing the absolute measure of the goodness of fit.
- The Fischer value (F) (F the Fisher ratio), reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High F-test values indicate that the model is statistically significant.(9)

A correlation matrix between parameters was performed on all descriptors. Nevertheless, the analysis revealed five independent descriptors for the development of the model.

The correlation between the biological activity and the descriptors, expressed by the following relationships:

**pIC50 = 2.240+0.006*SA-0.214*POL+0.101*Zagreb-0.398*NH+0.170*logD**

N=19 ; R=0.93 ; $R^2 = 0.865$; SE=0.083 ; F=11. 25 ; Q=10.85 ; $N_{test} = 7$ ; $R^2_{test} = 0.8$

Where: **N:** is the number of compounds (training set), **R:** is the correlation coefficient, $\mathbf{R^2}$: is coefficient of determination, **F:** is Fisher's statistic, **SE:** is the standard error of estimate and **Q:** is the goodness of fit or adaptation, $\mathbf{N_{test}}$**:** is the number of compounds ( test set), $\mathbf{R^2}$: is coefficient of determination of test set.

**SA:** surface area.

**POL:** polarizability.

**NH:** number of heavy atoms.

**Log D:** octanol/water partition coefficient (pH=7.4)

**Zagreb:** Zagreb indices.

- **Results and discussion:**

The values of the variance fraction can vary between 0 and 1. The QSAR model must consider an $R^2 > 0.6$ for it to be valid.

The values R = 0.93 and $R^2 = 0.865$ allowed us to firmly indicate the correlation between the different descriptors used such as: SA, NH, POL, logD and Zagreb and the biological activity studied (antioxidant).

Low value of standard error of estimate ($< 0.3$) indicates the accuracy of the statistical fit. All the values of the t-statistic are significant which confirms the significance of each descriptor.

The calculated F value for the generated QSAR model exceeds the tabulated F value by large margin as desired for a meaningful regression. Furthermore, the calculated F Value also determines a confidence limit superior to 95% for this model. The positive value of quality factor (Q) for QSAR model suggests its high predictive power and lack of over fitting, low standard deviation of the model demonstrates accuracy of the model.

❖ **Quantification of descriptors :**

RML statistical analysis determines and quantifies the correlations between the descriptors and the target variable. It also indicates the relative contribution of each descriptor to the overall explanation of the activity.

In the model equation, we notice that SA, Zagreb, logD with positive coefficients suggest that biological activity increases with increasing values of these descriptors. On the one hand, the negative coefficients of POL and NH, suggest the opposite.

The correlation matrix for the pIC50 biological activity and the descriptors selected to build the QSAR-2D model are presented in the following table:

**Table V.2. Model correlation matrix.**

| | pIC50 | SA | POL | Zagreb | NH | LogD |
|---|---|---|---|---|---|---|
| **pIC50** | 1.000 | | | | | |
| **SA** | -0.274 | 1.000 | | | | |
| **POL** | -0.322 | 0.878 | 1.000 | | | |
| **Zagreb** | 0.050 | 0.746 | 0.865 | 1.000 | | |
| **NH** | 0.063 | 0.751 | 0.834 | 0.992 | 1.000 | |
| **LogD** | -0.273 | 0.767 | 0.840 | 0.586 | 0.546 | 1.000 |

### V.2.1.2. 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.

The predictive powers of the equations were validated by leave–one–out (LOO) cross–validation method , cross-validation is a practical and reliable method for testing the significance of a model. (10)

In order to verify the predictive capacities of our selected MLR model, we resorted to its validation by the use of the "Leave-One-Out" technique (LOO-technique). The model developed was validated by calculating the following statistical parameters: **PRESS** (sum of squares of predicted residual), **SSY** (sum of squares of response value); $\mathbf{R^2_{cv}}$(global predictive ability), $\mathbf{R^2_{adj}}$, **PE** (the predictive error of the correlation coefficient) and $\mathbf{S_{PRESS}}$ (prediction uncertainty).(9)

**Table V. 3. Cross-validation parameters.**

| Model | PRESS | SSY | PRESS/SSY | Spress | r2cv | r2adj | 6PE |
|---|---|---|---|---|---|---|---|
| **1** | **0.064** | **0.473** | **0.135** | **0.058** | **0.865** | **0.813** | **0.125** |

PRESS is an important cross-validation parameter because it is a good approximation of the actual predictive error of the model. Its value being lower than SSY indicates that this model predicts better than chance and can be considered statistically significant. According to the results presented in (table V.3) this value is equal to 0.064, the model is statistically significant.

In our case, for the proposed model Press < SSY indicating that it has good predictive power and is better than chance .For a good model, Press/SSY should be less than 0.4 and its value in our model is 0.135.

The indication of model performance is obtained from $\mathbf{R^2_{cv}}$ (the global prediction capability). The high value of $\mathbf{R^2_{cv}}$ and $\mathbf{R^2_{adj}}$ are essential criteria for the best qualification of the QSAR model. Our result of these two values for this QSAR model was 0.865 and 0.813 respectively.

$S_{PRESS}$ (prediction uncertainty) is a good parameter to use to decide the uncertainty of the prediction .The lower the value of this parameter, the better the predictive ability of the model. In our case, this parameter has a small value of 0.058, which is why the prediction ability is the best for this model.

The prediction error of the correlation coefficient (PE) is another parameter used to determine the predictive power of the proposed model. We calculated the value of 6PE from the proposed model and it is presented in Table V.3. For this model the condition R > 6PE is satisfied and therefore they can be considered as having a good predictive power.

**Table V. 4. Experimental, predicted and residual values of (log (1/IC50)) of bioactive series by MLR and ANN.**

| Compounds | MLR | | | ANN | |
|---|---|---|---|---|---|
| | **Exp(pIC50)** (10) | **Pred (pIC50)** | **Residu** | **Pred(PIC50)** | **Residu** |
| **01** | 2.161 | 2.241 | -0.080 | 2.167 | -0.006 |
| **02** | 2.431 | 2.448 | -0.017 | 2.434 | -0.003 |
| **03** | 2.365 | 2.281 | 0.084 | 2.363 | 0.002 |
| **04** | 2.613 | 2.552 | 0.061 | 2.606 | 0.007 |
| **05** | 2.526 | 2.508 | 0.018 | 2.553 | -0.027 |
| **06** | 2.620 | 2.774 | -0.154 | 2.622 | -0.002 |
| **07** | 2.717 | 2.650 | 0.067 | 2.671 | 0.046 |
| **08** | 2.578 | 2.503 | 0.075 | 2.584 | -0.006 |
| **09** | 2.445 | 2.439 | 0.006 | 2.445 | 0.000 |
| **10** | 2.327 | 2.318 | 0.009 | 2.329 | -0.002 |
| **11** | 2.140 | 2.225 | -0.085 | 2.195 | -0.055 |
| **12** | 2.447 | 2.430 | 0.017 | 2.440 | 0.007 |
| **13** | 2.230 | 2.212 | 0.018 | 2.179 | 0.051 |
| **14** | 2.400 | 2.437 | -0.037 | 2.395 | 0.005 |
| **15** | 2.323 | 2.332 | -0.009 | 2.323 | 0.000 |
| **16** | 2.294 | 2.314 | -0.020 | 2.294 | 0.000 |
| **17** | 2.385 | 2.390 | -0.005 | 2.398 | -0.013 |
| **18** | 2.185 | 2.182 | 0.003 | 2.185 | 0.000 |
| **19** | 2.348 | 2.400 | -0.052 | 2.348 | 0.000 |
| **20t** | 2.303 | 2.310 | -0.007 | 2.307 | -0.004 |
| **21t** | 2.378 | 2.325 | 0.053 | 2.229 | 0.149 |

| | | | | | |
|---|---|---|---|---|---|
| **22t** | 2.270 | 2.099 | 0.171 | 2.070 | 0.200 |
| **23t** | 2.616 | 2.590 | 0.026 | 2.503 | 0.113 |
| **24t** | 2.549 | 2.679 | -0.130 | 2.503 | 0.046 |
| **25t** | 2.350 | 2.200 | 0.150 | 2.228 | 0.122 |
| **26t** | 2.401 | 2.215 | 0.186 | 2.224 | 0.177 |

The figure below (Figure III.1) presents the correlation between predicted and experimental values for the studied biological activity of 1H-1,2,3-triazole derivatives. We can observe that the predicted pIC50 values are in an acceptable agreement and regular distribution with experimental ones with correlation coefficient $R^2$ for the training set ($R^2$ = 0.865) and test set ($R^2$ = 0.80) indicate the significant correlation between different independent variables with antioxidant activity.

It indicates that this model can be successfully applied to predict the activity studied.

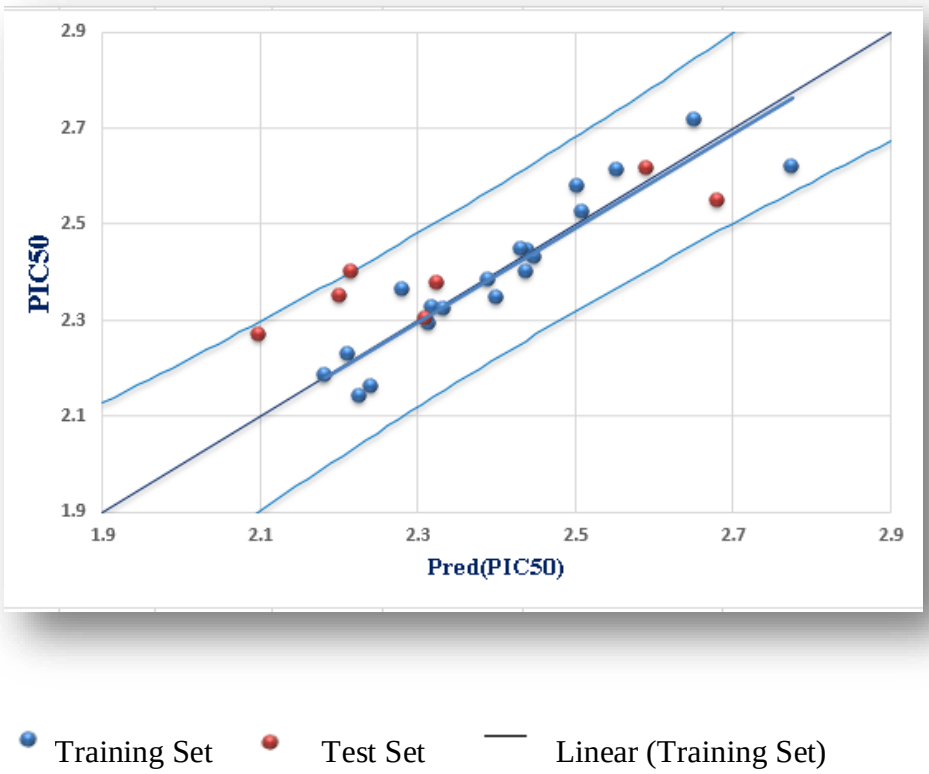



**Figure .V.1. Correlation of experimental and predicted pIC50 as derived using MLR.**

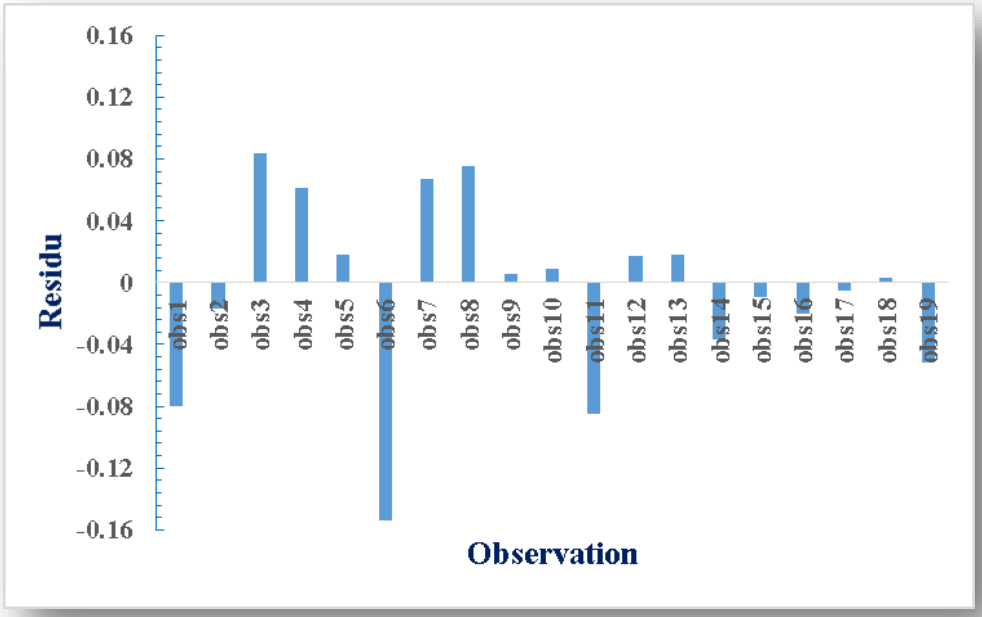



**Figure .V.2. Correlation between experimental and residual biological activity by MLR**

To investigate the presence of systematic error in the development of the QSAR models, the residual values of the biological activity were plotted against the experimental values, as shown in Figure .V.2. The distribution of residuals on both sides of the x-axis indicates the absence of the systemic error.

It indicates that this model can be successfully applied to predict biological activity for all triazoles used in QSAR model development.

## V.2.2. Artificial neural networks :

After confirming the robustness of a developed MLR-QSAR model with good statistical qualities, we used the five selected molecular descriptors in this model as input layers to build a model artificial neural network (ANN), by using a JMP 8.0.2 software.

Artificial neural networks (ANNs) models are non-linear models useful to predict the biological activity of large data sets of molecules . In contrast to classical statistical methods such as regression analysis or partial least squares analysis, ANNs enable the investigation of complex and nonlinear relationships. Neural networks are therefore ideally suited for use in drug design and QSAR. They are applied for simulating various non-linear complex systems of pharmaceutical, engineering, psychology and medicinal chemistry domains . For

instance, ANN was successfully used for the prediction and synthesis of new organic chemical compounds.(11)

In this work, ANN contained five inputs corresponding to the five descriptors selected from the correlation matrix, five hidden neurons, and one output neuron which is pIC50 (Figure .V.3). The number of artificial neurons in the hidden layer was adjusted experimentally, five neurons in the hidden layer permitted to attain the best correlation between experimental and predicted data.

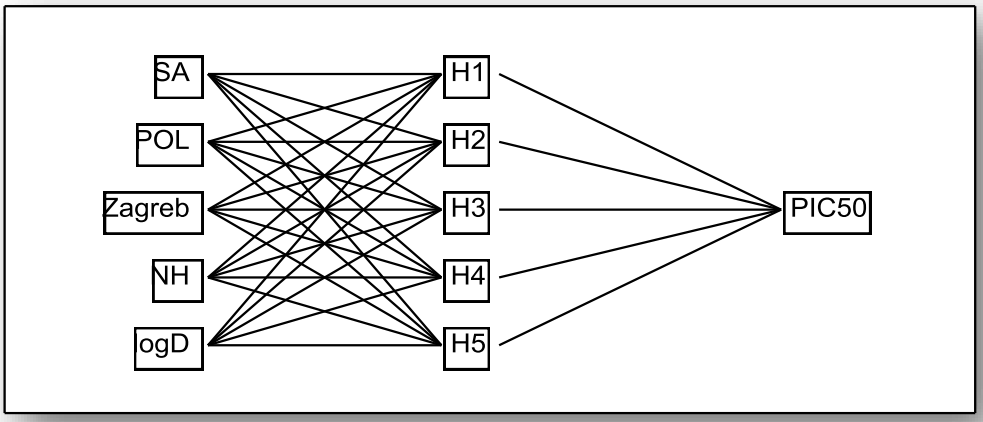



**Figure .V.3. Structure of ANN.**

Then, a good correlation of experimental and predicted PIC50 by ANN is found. This is shown in Figure .V.4. This network is characterized by a relatively high values of $R^2$ and a low value of SEE indicating that the proposed model is reliable and well adjusted, and a high value of F. $R^2 = 0.981$ and SEE=0.022, F=52.34. Furthermore,the robustness of the model was further confirmed by the significant value of the test data set $(R^2 = 0.802)$. We can conclude that the ANN model with (5-5-1) architecture is able to establish a satisfactory relationship between the five descriptors and the antioxidant activity.

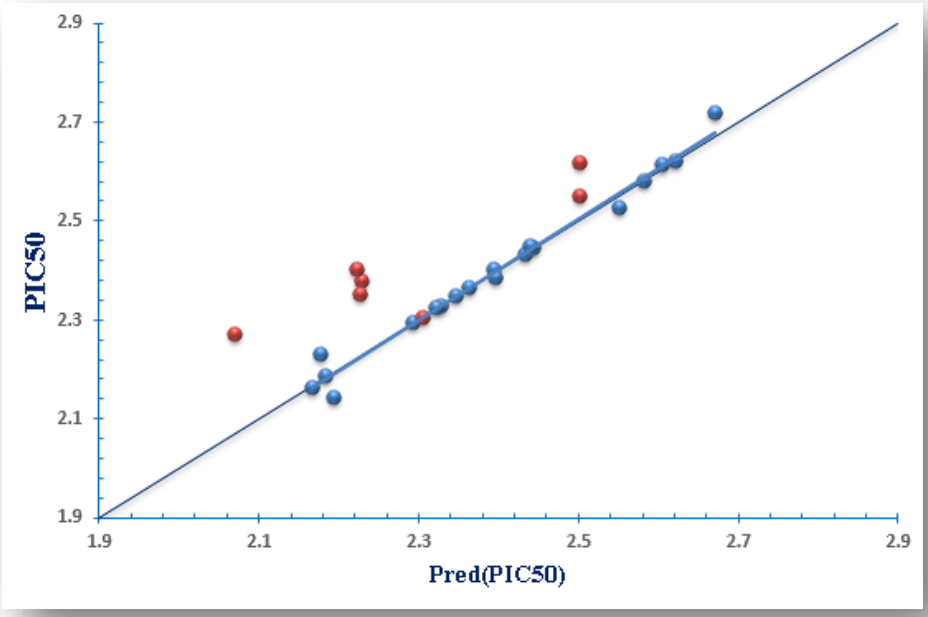



**Figure .V.4. Correlation of experimental and predicted pIC50 as calculated by ANN.**

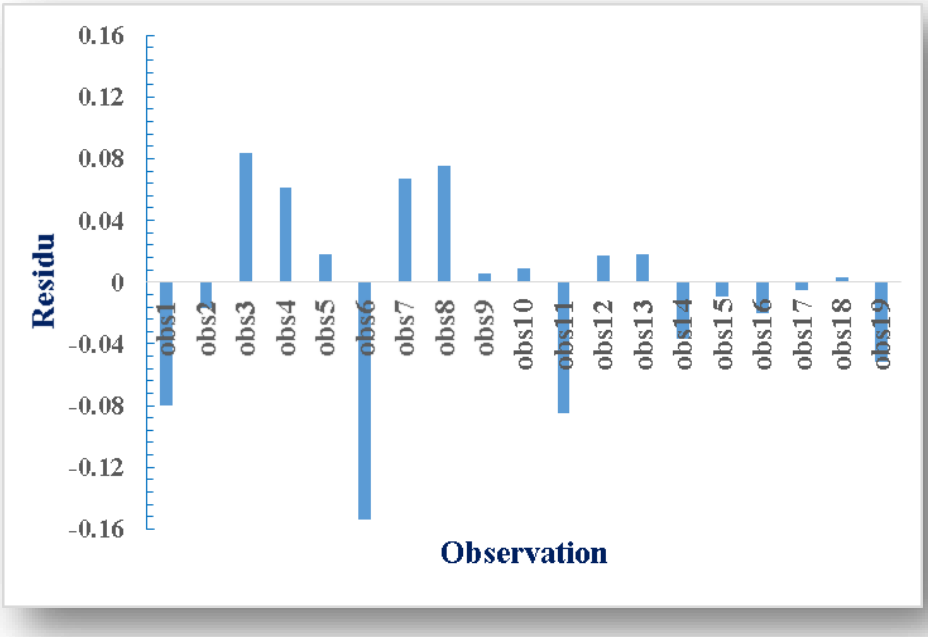


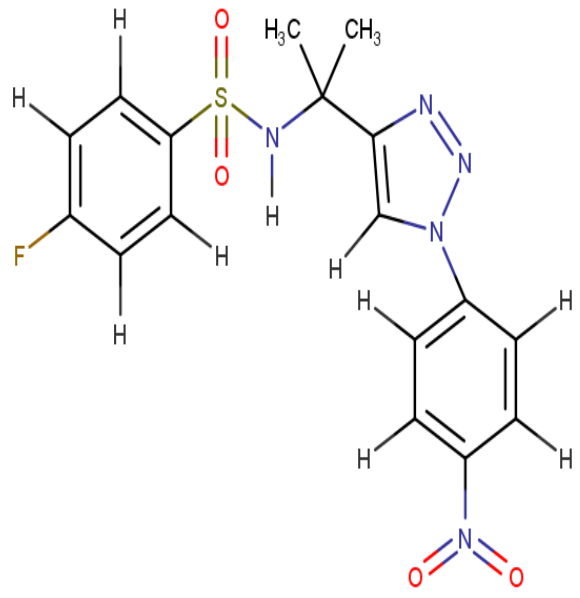
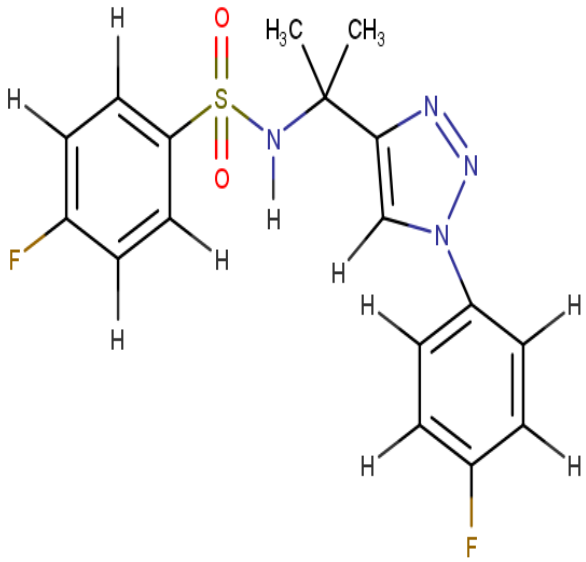

**Figure .V.5. Correlation between experimental and residual biological activity by ANN**

## V.3. Identification of leads:

At the end of this study, it is necessary to know after the application of our various filters, the most relevant compounds (lead) for the pathology to be treated.(12)

A "lead" compound is a molecule that can become a drug in the future, so it must interact with the target of the pathogen to inhibit, kill or modify its basic activity. For a molecule to have a strong chance of being a drug candidate, it must have ADME-T properties suitable for use in human or animal medicine.

After comparing the results,we observe that the compound 06 and 23 respect most of rules. Therefore the compounds 06 and 23 were identified as promising molecules (Leads), because they passed most of the tests successfully. However, it must be confirmed by additional in vivo and vitro studies. These two compounds present both a high biological activity and optimal chemical and biological properties, resulting in these compounds having the potential to be successful and effective future drugs.

| | |
|---|---|
|  |  |
| Compound 06 | Compound 23 |

## V.4. Conclusion:

The present study of QSAR has made it possible to determine the quantitative relationship between the physico-chemical properties of the compounds with their activities has been carried out.

MLR regression analysis was used to develop the model and predict biological activity from molecular descriptors belonging to the 1H-1,2,3-triazole derivative series. Our developed QSAR model is based on the following descriptors: SA, nHeavyAtom, POL, logD and Zagreb.

The QSAR model indicates that these descriptors have significant relationships with the observed bioactivity. We observed a high similarity between the values experimental values and the predicted values of the activity, which indicates the excellent quality of the QSAR model.

After that an artificial neural network (ANN) was used as a nonlinear method to create a QSAR model and interpret the nonlinear relationship between the antioxidant activity of 1H-1,2,3-triazole derivatives and their structural descriptors selected by the MLR method.The results confirm that the ANN nonlinear model has good stability and good predictive ability compared to that of the linear model (MLR). Nevertheless,both models remain satisfactory and exhibit a high predictive power.

By comparison of the different selection methods applied to the compounds studied: two lead molecules of the series were found which have high ADME properties.

## References

(1). Esposito E. AJ Hopfinger et JD Madura,«. Methods for Applying the Quantitative Structure-Activity Relationship Paradigm» Methods in Molecular Biology. 2004;275:131-213.

(2). Dermeche K., Tchouar N., Belaidi S., Salah T., Qualitative Structure-Activity Relationships and 2D-QSAR Modeling of TNF-α Inhibition by Thalidomide Derivatives, Journal of Bionanoscience, 2015,9, 395-400.

(3). Hansen C, Telzer BR, Zhang L. Comparative QSAR in toxicology: examples from teratology and cancer chemotherapy of aniline mustards. Critical reviews in toxicology. 1995;25(1):67-89.

(4). Perkins R, Fang H, Tong W, Welsh WJ. Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology. Environmental Toxicology and Chemistry: An International Journal. 2003;22(8):1666-79.

(5). Belaidi S, Salah T , Melkemi N, Sinha L, Prasad O, Structure activity relationship and quantitative structure-activity relationships modeling of antitrypanosomal activities of alkyldiamine cryptolepine derivatives, Journal of Computational and Theoretical Nanoscience, 2015, 12 (9), 2421-2427.

(6). Belaidi S., Mazri R., Belaidi H., Lanez T., Bouzidi D., Electronic structure and physico-chemical property relationship for thiazole derivatives, Asian Journal of Chemistry, 2013, 25, 9241.

(7). Kerassa, A. , Belaidi, S., Harkati, D., Lanez, T., Prasad, O., and Sinha, L., Investigations on Molecular Structure, Electronic Properties, NLO Properties and Comparison of Drug-Likeness of Triazolothiadiazole Derivatives by Quantum Methods and QSAR Analysis, Reviews in Theoretical Science,2016,4 ,85. https://doi.org/10.1166/rits.2016.1050.

(8). Soualmia F, Belaidi S, Tchouar N, Lanez T, Boudergua S. QSAR Studies and Structure Property/Activity Relationships Applied in Pyrazine Derivatives as Antiproliferative Agents Against the BGC823. Acta Chimica Slovenica. 2021;68(4):882-95.

(9). Aouisset D. In silico analysis of pyrazole derivatives applied to drug design [Memoir of master]: University Mohammed Khider Biskra; 2022.

(10). Kaushik C, Chahal M, Luxmi R, Kumar D, Kumar A, Kumar M, et al. Synthesis, characterization and biological activities of sulfonamide tagged 1, 2, 3-triazoles. Synthetic Communications. 2020;50(22):3443-61.

(11). Ouassaf M, Belaidi S, Lotfy K, Daoud I, Belaidi H. Molecular docking studies and ADMET properties of new 1.2.3 triazole derivatives for anti-breast cancer activity. Journal of Bionanoscience. 2018;12:26-36. https://doi.org/10.1166/jbns.2018.1505.

(12). Boudergua S, Alloui M, Belaidi S, Al Mogren MM, Ibrahim UAE, Hochlaf M. QSAR modeling and drug-likeness screening for antioxidant activity of benzofuran derivatives. Journal of Molecular Structure. 2019;1189:307-14.

(13). Corinne H. Search for high affinity inhibitors of DNA polymerase beta by virtual screening and NMR: University Toulouse III-Paul Sabatier; 2008.

# General Conclusion

In this work, we used computational chemistry to study the antioxidant activity on the series of triazole derivatives which present good drug candidates for antioxidant treatment.

In this work, we applied molecular modeling methods on molecules based on the nucleus of 1H-1,2,3-triazole. Thus, our study mainly focused on the following topics:

In this work, we used computational chemistry to study the biological antioxidant activity on the series of triazole derivatives which present good drug candidates for antioxidant treatment.

In this work, we applied molecular modeling methods on molecules based on1H-1,2,3-triazole.So our study was mainly focused on the following topics:

- A conformational analysis on the basic core of 1H-1,2,3-triazole.
- A qualitative study on the structure-properties/activity relationships of a bioactive series of 1H-1,2,3-triazole derivatives.
- A quantitative study of structure-activity relationships on a series of twenty six 1H-1,2,3-triazole derivatives, of which a QSAR model was developed with successful statistical tests.

We first focused our attention on the study of the structural and electronic properties of the basic nucleus: 1H-1,2,3-triazole. The results demonstrate that the structural and electronic comparison of the nucleus 1H-1,2,3-triazole presents similar results using different computational methods: quantum MP2 and DFT methods.The efficiency of the method (DFT) has been demonstrated to do the calculations and to further our study of the basic core of 1H-1,2,3-triazole.

After the basic molecule study, we have studied the structure-activity relationship (SAR) and drug likeness proprieties of a bioactive series of 1H-1,2,3-triazole derivatives have shown that structural units involved in the biological activity. In our case, the Lipinski and Veber rules are validated. Therefore, theoretically, all compounds will not have a problem with oral bioavailability.

Next to, we have applied Quantitative structure activity relationship (QSAR) analysis to a series of 26 of 1H-1,2,3-triazole derivatives using a combination of various physicochemical and quantum descriptors, Using multiple linear regression (MLR) as well as artificial neural network (ANN) methods.

Firstly, A multiple linear regression (MLR) procedure was used to model the relationships between molecular descriptors and the chemotherapeutic activity of the 1H-1,2,3-triazole derivatives.Our results suggest the best QSAR model with the following descriptors: SA, NH, POL, logD and Zagreb. Good agreement between experimental and predicted activity values, obtained in the validation procedure, indicated the good quality of the derived QSAR model. The best QSAR model developed show a good predictive correlation coefficient $R^2 = 0.865$ and external predictive ability of prediction $R^2_{test} = 0.80$ was developed by MLR. The proposed model has good robustness and predictability when verified by internal and external validation.

After confirming the robustness of a developed MLR-QSAR model with good statistical qualities, we used the five selected molecular descriptors in this model as input layers to build a model artificial neural network (ANN) .

An artificial neural network (ANN) was used as a nonlinear method to create a QSAR model and interpret the nonlinear relationship between the antioxidant activity of triazole derivatives and their structural descriptors selected by the MLR method.The best neural network model has a [5-5-1] architecture with a high coefficient of determination ($R^2 = 0.981$) for training set and ($R^2 = 0.802$) for the test set. The results confirm that the ANN nonlinear model has good stability and good predictive ability compared to that of the linear model (MLR). Nevertheless,both models remain satisfactory and exhibit a high predictive power.

Finally, by evaluation of the results obtained, two compounds were identified as series leads with high "ADME" properties.

## ABSTRACT

The purpose of this work carried out during this memory concerns a fundamental and novel research which aims to discuss and predict the activity of a series of thirty two derivatives of 1H-1,2,3-triazole which can be antioxidants.

In the first step of this research work, different theoretical calculation methods (**MP2 and DFT/B3LYP)** were used to study the basic core and validate a method to calculate the structures of the series of 1H-1,2,3-triazoles.We found that the **DFT/B3LYP** method was the most appropriate, so this method was used to determine the structural, electronic parameters associated with the molecules studied. Then, a qualitative structure-activity relationship **(SAR)** study was also performed for this bioactive series of 1H-1,2,3-triazole derivatives using different **MPO** methods.

Finally, a quantitative study was also carried out to predict the biological activity of the studied compounds and its derivatives by suggesting the best **QSAR** model by means of the method of multiple linear regression **(MLR)** and the method of artificial neural networks **(ANN)**, which are used to quantify the relationships between molecular descriptors and antioxidant activity.



***Keywords*** **: 1H-1,2,3-triazole , DFT , SAR , MPO , QSAR, MLR, ANN.**



## RÉSUMÉ

Le but du ce travail effectué au cours de ce mémoire concerne une recherche fondamentale et originale qui vise à discuter et à prédire l'activité d'une série de trente deux dérivés de 1H-1,2,3-triazole qui peuvent être des antioxidants.

Dans le cadre de premier temps de ces travaux de recherche, des différentes méthodes de calcul théorique (**MP2** et **DFT/B3LYP)** ont été utilisé pour etudier le noyau de base et valider une methode pour calculer les structures de la série des 1H-1,2,3-triazoles . Nous avons trouvé que la méthode **DFT/B3LYP** était la plus appropriée, donc cette méthode a été utilisée pour déterminer les paramètres structuraux, électroniques associés aux molécules étudiées . Ensuite, une étude qualitative de la relation structure-activité **(SAR)** a été effectuée également pour cette série bioactive de dérivés de 1H-1,2,3-triazole en utilisant des différentes méthodes **MPO**.

Enfin, une étude quantitative a été effectuée également pour prédire l'activité biologique des composés étudiés et ses dérivés en suggérant le meilleurs modèle **QSAR** au moyen de La méthode de la régression linéaire multiple **(MLR)** et la méthode de réseaux de neurones artificiels **(ANN)**,qui sont utilisées pour quantifier les relations entre les descripteurs moléculaires et l'activité anti-oxidante.



***Mots clés*** **: 1H-1,2,3-triazole, DFT, SAR , MPO , QSAR, MLR, ANN.**



## الملخص

الهدف من العمل الذي تم إجراؤه خلال هذه الأطروحة يتعلق ببحث أساسي وجديد يهدف إلى مناقشة وتوقع نشاط سلسلة من اثنين وثلاثين مشتقًا من1، 2، 3 تريازول والتي يمكن أن تكون مضادات للأكسدة.

كخطوة أولى في هذا البحث، تم استخدام طرق حساب نظرية مختلفة **(MP2 و DFT / B3LYP)** لدراسة النواة الأساسية والتحقق من صحة طريقة لحساب هياكل سلسلة 1، 2، 3 تريازول وجدنا أن طريقة **DFT / B3LYP** هي الأنسب ،لذلك تم استخدام هذه الطريقة لتحديد المعلمات الهيكلية، الإلكترونية المرتبطة بالجزيئات المدروسة. بعد ذلك، تم إجراء دراسة العلاقة النوعية بين البنية والنشاط **(SAR)** لهذه السلسلة النشطة بيولوجيًا لمشتقات1، 2، 3 تريازول باستخدام طرق **MPO** المختلفة.

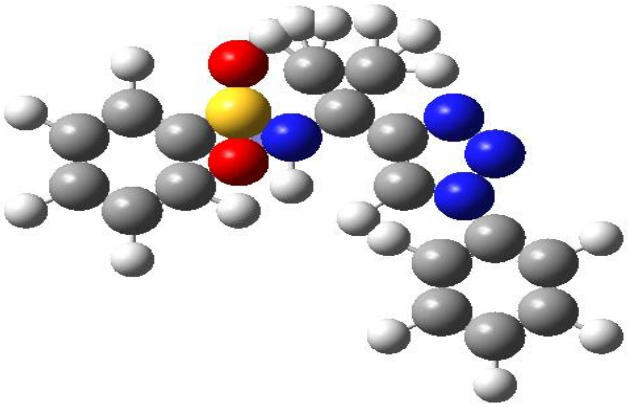
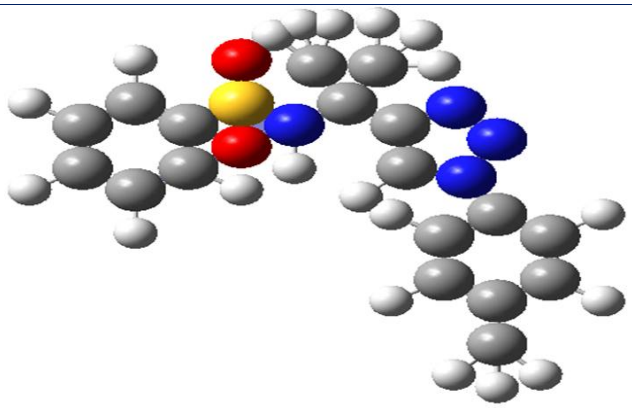
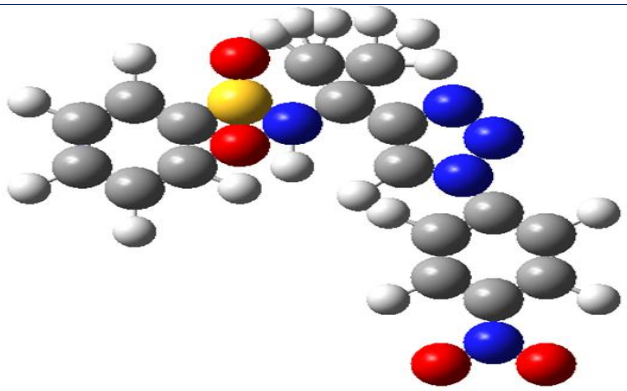
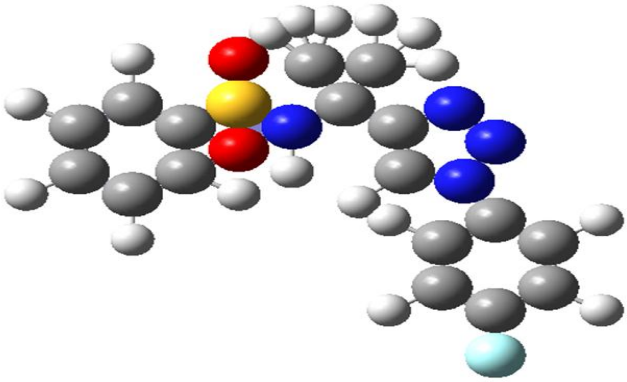
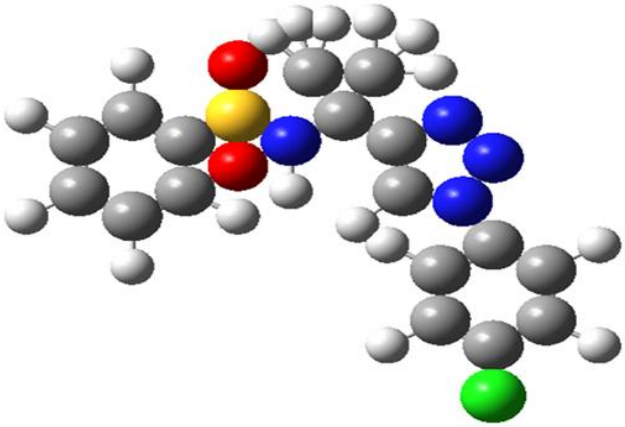
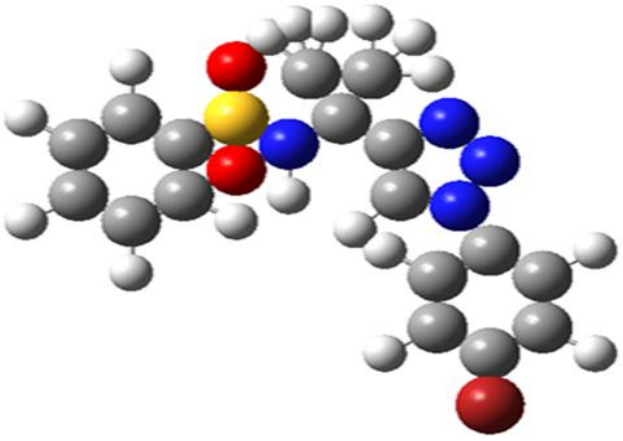
وفي الأخير، تم القيام بإجراء دراسة كمية أيضًا للتنبؤ بالنشاط البيولوجي للمركبات المدروسة ومشتقاتها من خلال اقتراح أفضل نموذج **QSAR** عن طريق طريقة الانحدار الخطي المتعدد **(MLR)** وطريقة الشبكات العصبية الاصطناعية **(ANN)**، والتي تُستخدم لتحديد العلاقات بين الواصفات الجزيئية والنشاط المضاد للأكسدة.

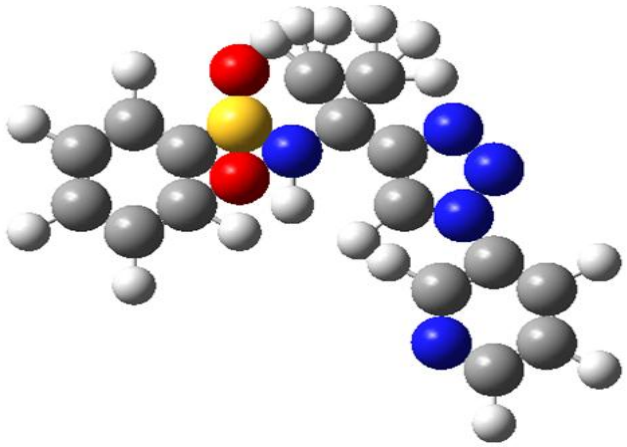
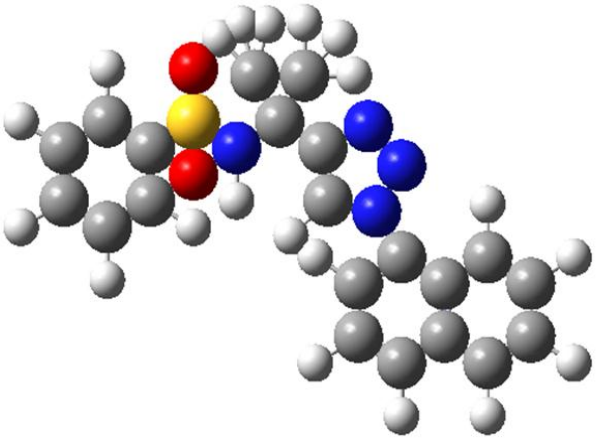
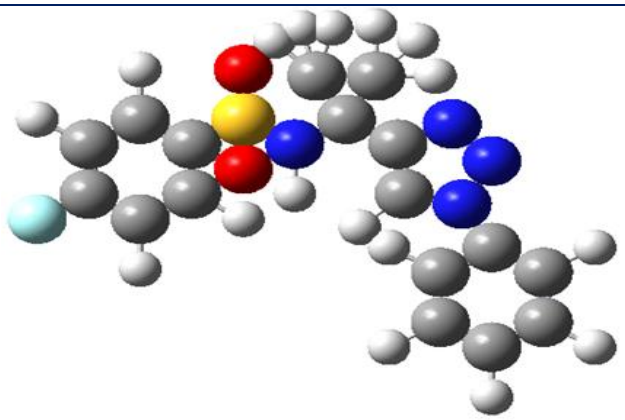
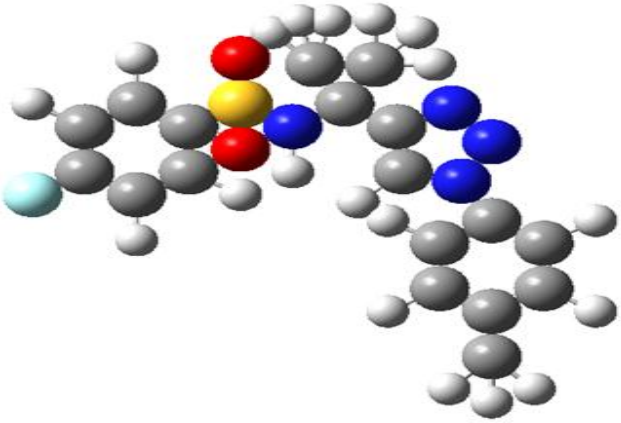
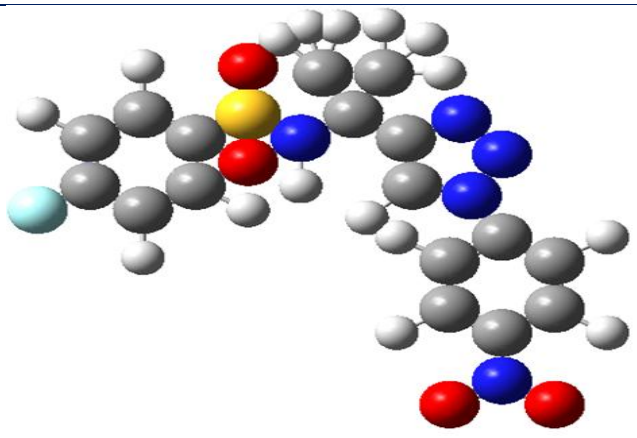
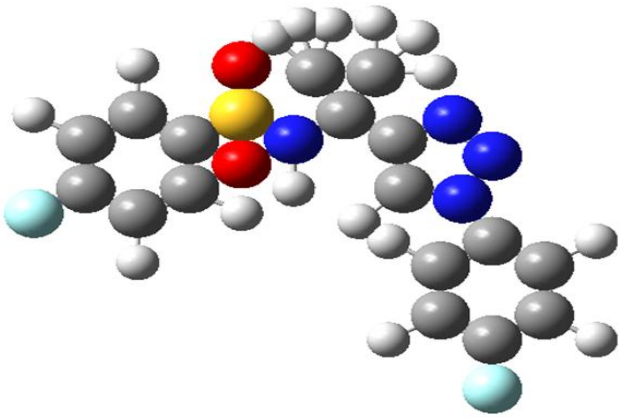


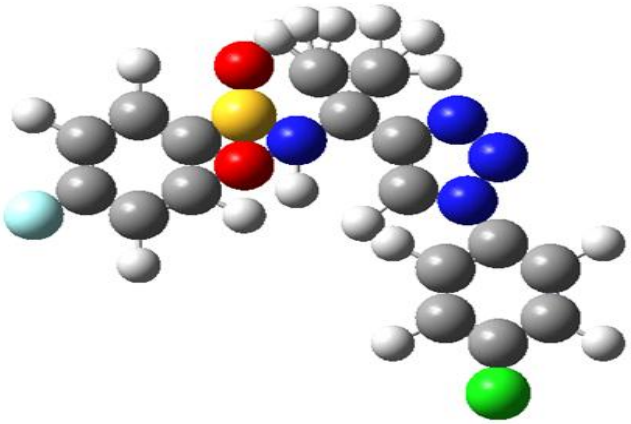
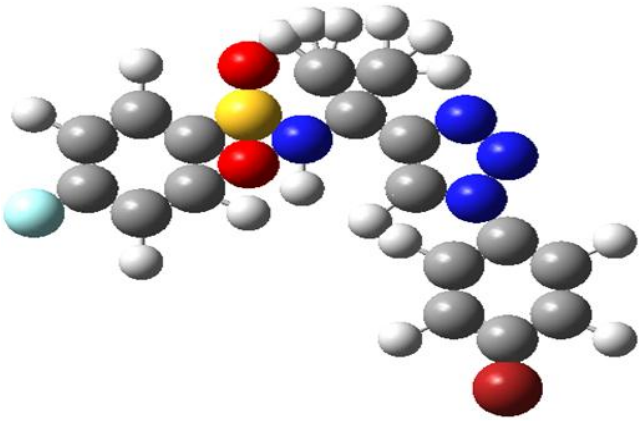
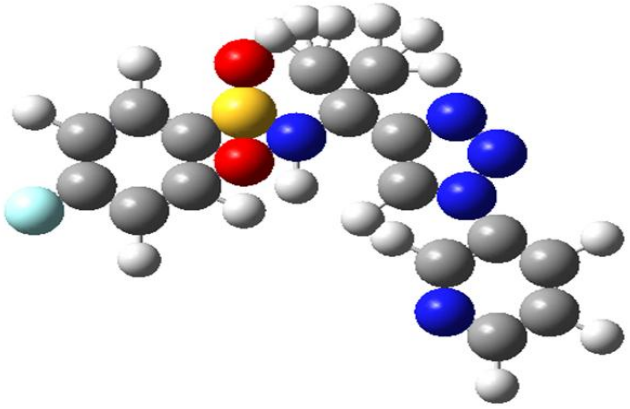
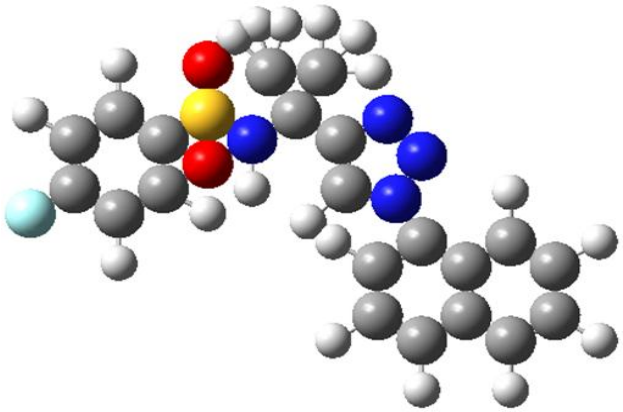
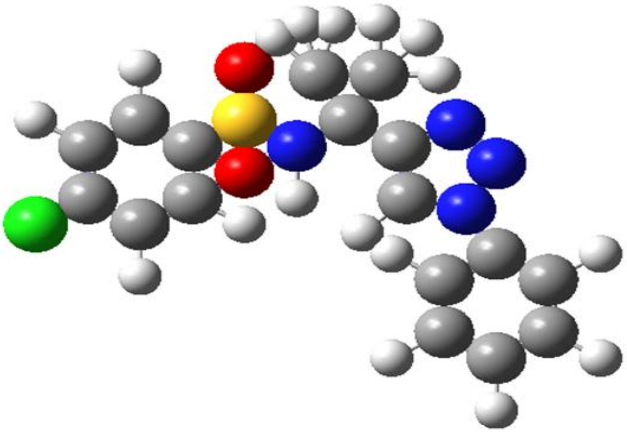
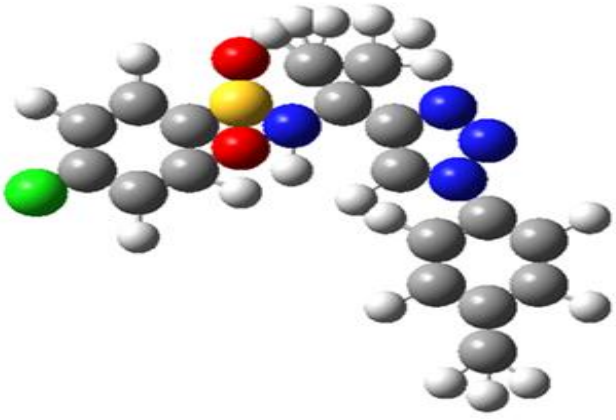
***الكلمات الأساسية*****:1، 2، 3 تريازول, ANN, MLR, QSAR, MPO, SAR, DFT.**

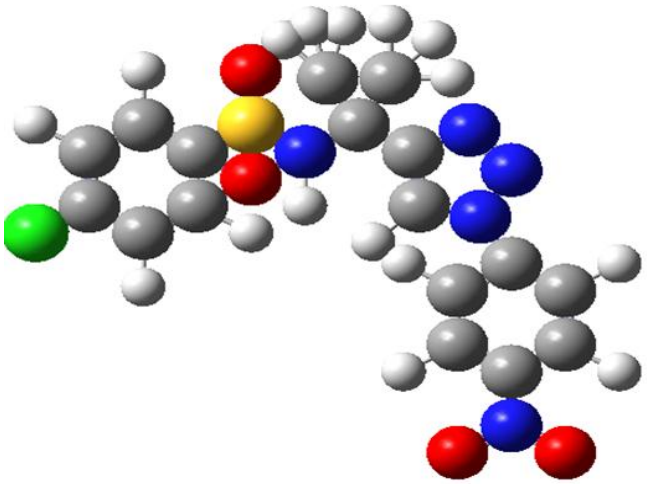
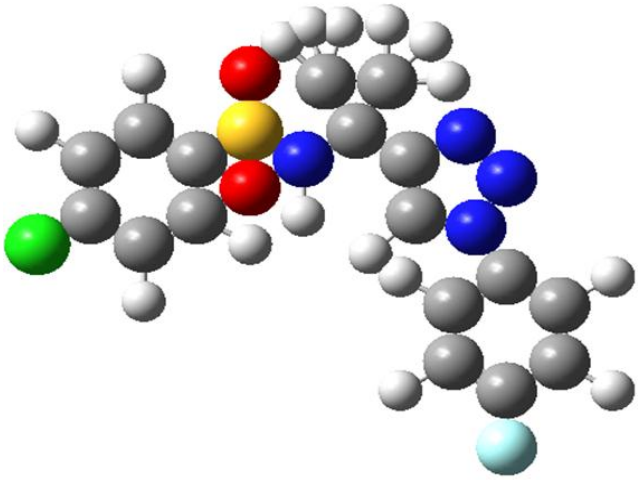
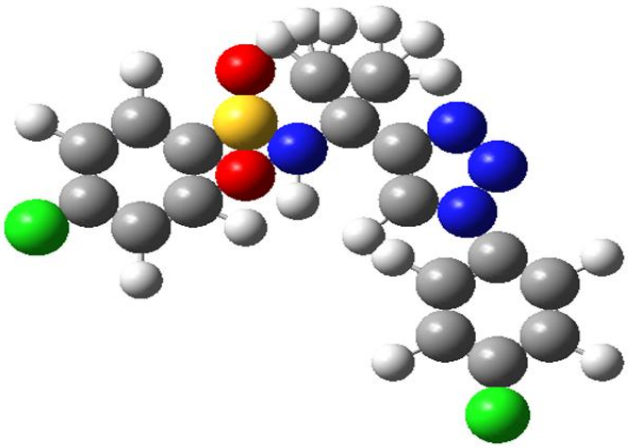
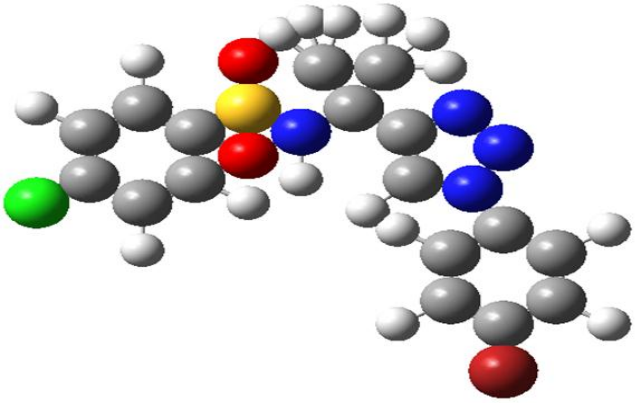
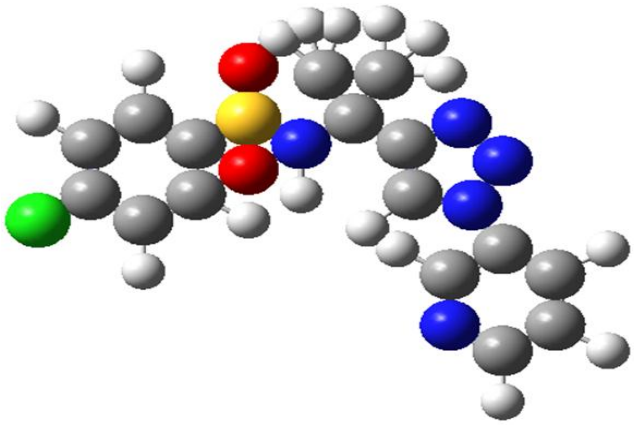
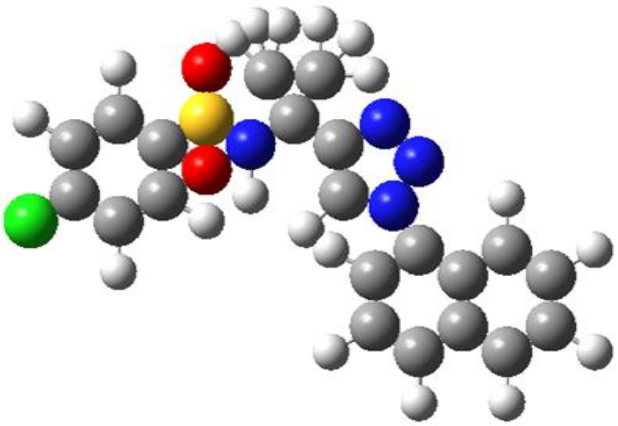
# Appendix

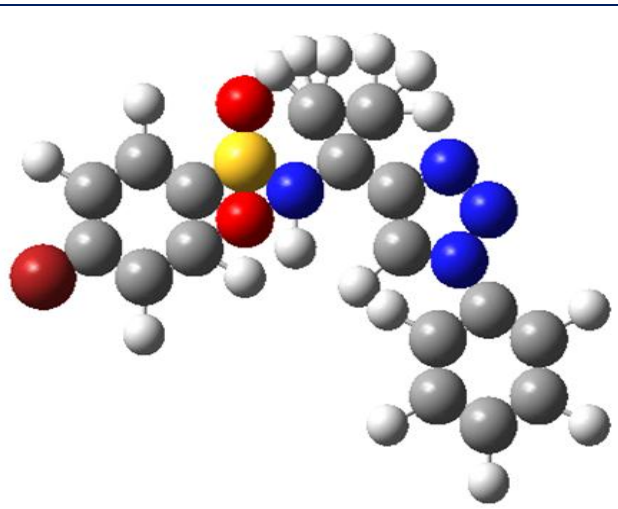
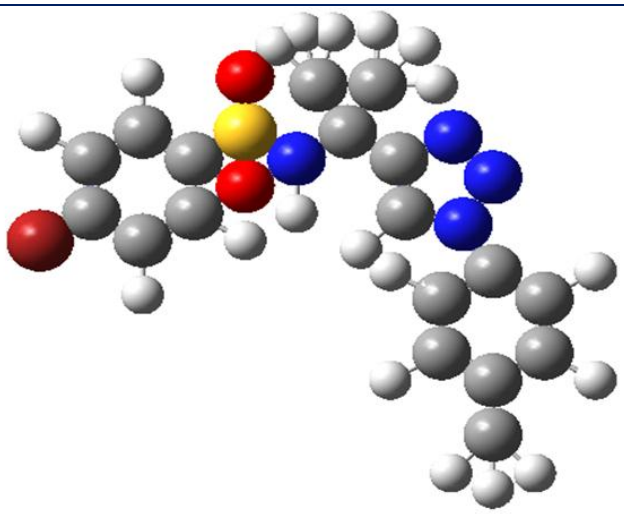
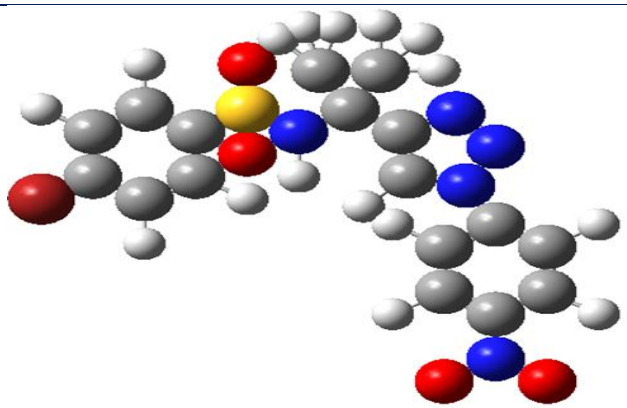
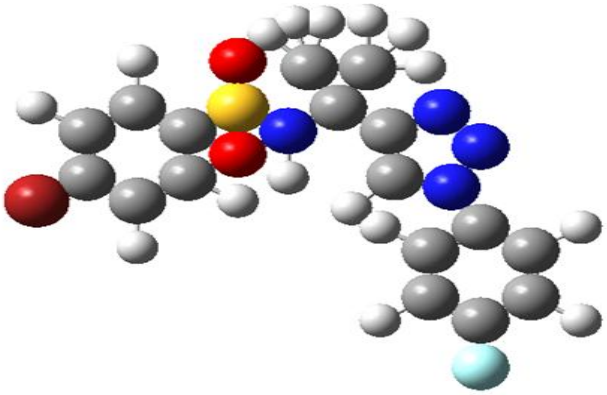
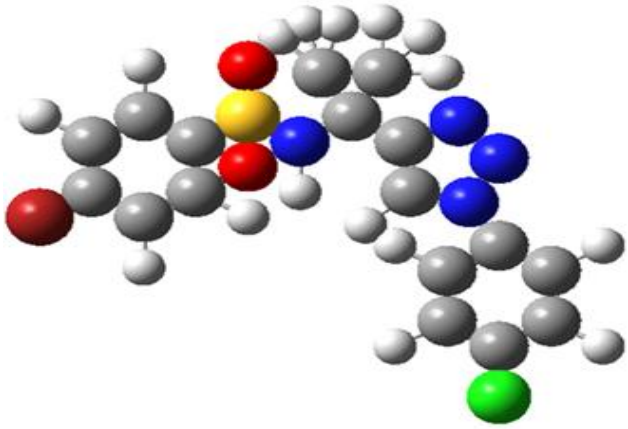
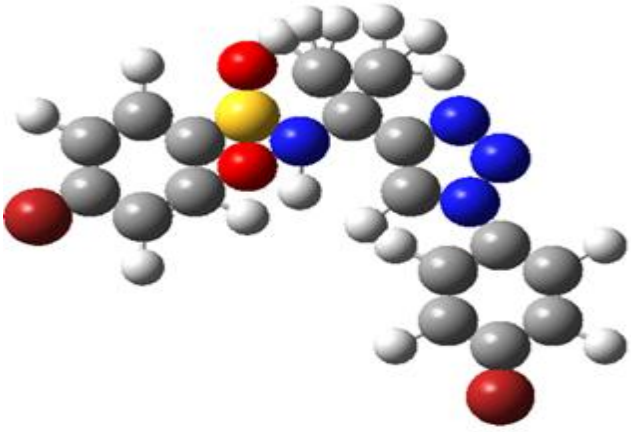
**3D PRESENTATION OF THE STRUCTURES OF 1H-1,2,3-TRIAZOLE DERIVATIVES**

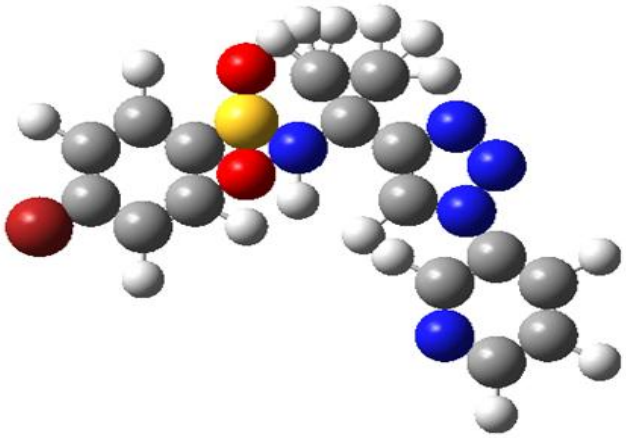
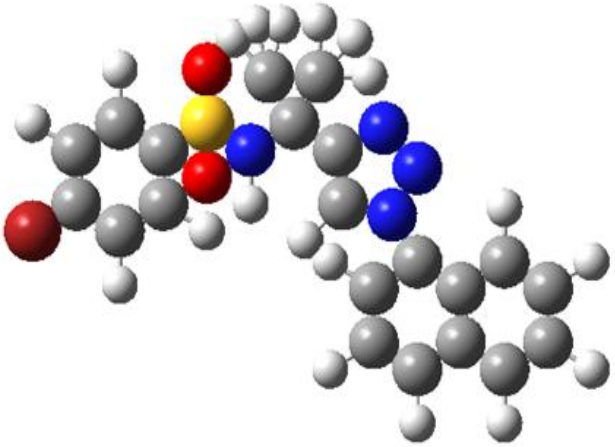
| | |
|---|---|
|  |  |
| **Compound 01** | **Compound 02** |
|  |  |
| **Compound 03** | **Compound 04** |
|  |  |
| **Compound 05** | **Compound 06** |

| | |
|---|---|
|  |  |
| **Compound 07** | **Compound 08** |
|  |  |
| **Compound 09** | **Compound 10** |
|  |  |
| **Compound 11** | **Compound 12** |

| | |
|---|---|
|  |  |
| **Compound 13** | **Compound 14** |
|  |  |
| **Compound 15** | **Compound 16** |
|  |  |
| **Compound 17** | **Compound 18** |

| | |
|---|---|
|  |  |
| **Compound 19** | **Compound 20** |
|  |  |
| **Compound 21** | **Compound 22** |
|  |  |
| **Compound 23** | **Compound 24** |

| | |
|---|---|
|  |  |
| **Compound 25** | **Compound 26** |
|  |  |
| **Compound 27** | **Compound 28** |
|  |  |
| **Compound 29** | **Compound 30** |

| | |
|---|---|
|  |  |
| **Compound 31** | **Compound 32** |