MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE MOHAMED KHIDER BISKRA

FACULTE DES SCIENCES EXACTES
ET
DES SCIENCES DE LA NATURE ET DE LA VIE

Département des Sciences de la matière

THESE

Présentée par

Toufik SALAH

En vue de l'obtention du diplôme de :

Doctorat en chimie

Option :
Chimie Moléculaire

Intitulée:

Etude par la chimie computationnelle des corrélations 2D-QSAR et 3D-QSAR de quelques composés bioactifs.

Soutenue le : 02 Fev. 2017

Devant la commission d'examen :

M. Mahmoud OMARI	Prof.	Univ. Biskra	Président
M. Salah BELAIDI	Prof.	Univ. Biskra	Directeur de thèse
M. Noureddine GHERRAF	Prof.	Univ. Oum el Bouaghi	Examinateur
M. Nadjib MELKEMI	MC/A	Univ. Biskra	Examinateur

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

MOHAMED KHIDER BISKRA UNIVERSITY

FACULTY OF EXACT SCIENCES AND SCIENCES OF NATURE AND LIFE

Matter Sciences Department

THESIS

Presented by

Toufik SALAH

In order to obtain the diploma of:

PhD in Chemistry

Option :
Molecular chemistry

Entitled:

2D-QSAR & 3D-QSAR modeling of some antitrypanosomal compounds based on computational chemistry

Defended on : 02 Feb. 2017

In front of the thesis committee members:

Mr. Mahmoud OMARI	Prof.	Biskra Univ.	President
Mr. Salah BELAIDI	Prof.	Biskra Univ.	Supervisor
Mr. Noureddine GHERRAF	Prof.	Oum el Bouaghi Univ.	Examiner
Mr. Nadjib MELKEMI	MC/A	Biskra Univ.	Examiner

To my dear parents

To my fiancee Yulia

To my sister Ourida
and her husband
Zineddine

To all those who are dear to me

Acknowledgements

In the Name of Allah, the Beneficent, the Merciful

First praise is to Allah, the Almighty, on whom ultimately we depend for sustenance and guidance. I thank him for giving me strength and ability to complete this study.

Second, my sincere appreciation goes to my supervisor Prof Salah BELAIDI, whose guidance, careful reading and constructive comments were valuable. His timely and efficient contribution helped me and made me who I am today.

I want to express my deep gratitude to my Coach or Dear-Friend Dr. Nadjib MELKEMI for the trust, the insightful discussion, offering valuable advice, for his support during the whole period of my research.

I wish to extend my sincere thanks to Prof Mahmoud Omari, Professor at the University of Biskra, for accepting to chair the committee of my dissertation.

I also would like to thank the committee members, Dr. Nadjib Melkemi and Prof. Noureddine GHERRAF, for agreeing to examine and judge the work.

I also wish to thank the group of computational and pharmaceutical chemistry, LMCE Laboratory at Biskra University, its leadership and the staff for providing me with an academic base, which has enabled me to take

up this study. Therefore, I would like to thank all the people who contributed in some way to the work described in this thesis.

Finally, and most importantly, I would like to thank my fiancee Yulia. Her support, encouragement, quiet patience and unwavering love were undeniably the bedrock upon which the past three years of my life have been built in big happiness. Her tolerance of my occasional vulgar moods is a testament in itself of her unyielding devotion and love. I thank my parents, Ahcene and Rokia, for their faith in me and allowing me to be as ambitious as I wanted. It was under their watchful eye that I gained so much drive and an ability to tackle challenges head on. Also, I thank my beloved sister Ourida and her husband Zineddine who were responding all my requests with big love, unending encouragement and support.

Table of Contents

CHAPTER 2: Materials and Methods

CHAPTER 3: Results and Discussion

LIST OF TABLES

LIST OF FIGURES & SCHEMES

LIST OF ABBREVIATIONS

2D: Two dimensional.

3D: Three dimensional.

Asn: Asparagine.

Asp: Aspartic.

B3LyP: Becke, 3-parameter, Lee-Yang-Parr.

Babesia SPP: Babesia species.

CADD: Computer Aided Drug Design.

CCDD: Cambridge Crystallographic Data Center.

CSD: Cambridge Structural Database.

Cys: Cysteine.

d: Density.

D: Debye.

DFT: Density Functional Theory.

DNA: DeoxyriboNucleic Acid.

Gln: Glutamine.

Glu: Glutamic.

Gly: Glycine.

GTO: Orbital Type Gaussian.

HBA: Hydrogen Bond Acceptor.

HBD: Hydrogen Bond Donor.

HE: Hydration Energy.

His: Histidine.

HOMO: Highest Occupied Molecular Orbital.

HSAB: Hard and Soft Acids and Bases.

IBM SPSS: International Business Machines, Statistical Package for the Social Science.

IUPAC: International Union of Pure and Applied Chemists.

LBDD: Ligand Based Drug Design.

LCAO: Linear Combination of Atomic Orbitals.

Leu : Leucine.

LipE : Lipophilic Efficiency.

LUMO: Lowest Unoccupied Molecular Orbital.

MD: Molecular Docking.

MESP: Molecular ElectroStatic Potential.

Met: Methionine.

MLR: Multiple Linear Regression.

MM: Molecular Mechanics.

MPO: Multiple Parameters Optimization.

MR: Molar Refractivity.

MW: Molecular Weight.

NADPH: Nicotinamide Adenine Dinucleotide PHosphate.

NBO: Natural Bond Orbital.

NMR: Nuclear Magnetic Resonance.

NRB: Number of Rotatable Bonds.

Pdb: Protein Data Bank.

PE: Predictive Error.

Pol: Polarizability.

PRESS: Predicted Residual Sum of Squares.

PSA: Polar Surface Area.

QSAR: Quantitative Structure Activity Relationship.

RCSB: Research Collaboratory for Structural Bioinformatics.

RMSD: Root Mean Square Deviation.

SAG: Surface Area Grid.

SAR: Structure Activity Relationship.

SBDD: Structure Based Drug Design.

Ser: Serine.

SPR: Structure Property Relationship.

STO: Orbital Type Slater.

T. b. Gambiense : Trypanosoma Brucei Gambiense.

T. b. Rhodesiense : Trypanosoma Brucei Rhodesiense.

*T. Cruz*i: *Trypanosoma Cruzi.*

Trp: Tryptophan.

amu: Atomic Mass Unit.

USD: United States Dollar.

VFF: Valence Force Field.

WHO: World Health Organization.

GENERAL INTRODUCTION

Drug discovery and developing a new medicine is a long, complex, costly and highly risky process that has few peers in the commercial world. This is why computer-aided drug design (CADD) approaches are being widely used in the pharmaceutical industry to accelerate the process that helps in decreasing the time and cost of discovery researches. The cost benefit of using computational tools in the lead optimization phase of drug development is substantial [1].

It is beneficial to apply computational tools in hit-to-lead optimizations to cover a wider chemical space while reducing the enormous number of compounds that must be synthesized and tested in vitro. Therefore, the research-based pharmaceutical industry has increasingly employed modern medicinal chemistry methods, such as: structure-based analysis (Geometric and electronic optimizations), Structure-activity/property-relationships (QSAR modeling), Ligand-based screening (Pharmacophore based screening), Docking poses and energy profiles for hit analogs (Molecular docking analysis), prediction of favorable affinity or optimize drug metabolism and pharmacokinetics (ADMET), etc. [2-6].

Nowadays, many researches have been carried out for the development of anti-inflammatory, antioxidant, anticancer and antiparasitic drugs, where we find antitrypanosomal compounds that have attracted the attention of several researchers. Seeing that currently available drugs have insufficient efficiency and excessive toxicity. In addition, for the diseases caused by this family of parasites, which called Trypanosomatidae, there is no vaccine [7].

Among the potential diseases caused by trypanosomes, are Sleeping sickness and Chagas disease, they are known as vector-borne parasitic diseases [8]. Current therapeutic options for Trypanosoma infections are not different from those available more than 20 years ago and are far from ideal, limited by toxicity and the emergence of drug-resistant parasites [9, 10].

Thus, the development of new and effective drugs as well as, the discovery of new molecular targets for new therapies is an urgent need to treat these vector-borne diseases. In the face of, many researches have been carried out for the development of trypanocidal drugs; nevertheless, no drug could achieve a parasitological cure. In opposite, some of them presented serious toxic side effects. Although, some compounds were found to be important candidates as inhibitors against the parasitic flagellate *T. Cruzain and T. Rhodsain* [11-13].

Following our interest in this field, we applied several computational methods on the most effective trypanocidal candidates in order to, study structures and properties of these ligands (the work is published in: Reviews in Theoretical Science, Volume 04, Number 03, pp. 355-364 (9) ; 2015), evaluate their biological response (the work is published in: Journal of

Computational and Theoretical Nanoscience, Volume 09, Number 12, pp. 2421-2427 (6) ; 2015) and better understand Trypanosomes inhibition mode (the work is published in: Journal of Theoretical and Computational Chemistry, Volume 03, Number 15, pp. 1-17 (17); 2016).

This work consists of three main chapters; the content of each part is briefly described below:

- Chapter 1 : *Trypanocidal Drug Resistance Vs Drug Discovery Activities*

Here, we present general informations about the use of computational chemistry in drug discovery area, and then we discuss all the possibilities of Trypanosomes infections in parasites world, followed by the description of each compound that will be used as a candidate in our investigation.

- Chapter 2 : *Materials and Methods*

This part contains definitions, theories, approaches for each method (i.e. Molecular Modeling, QSAR modeling, Molecular Docking, etc.) and material (i.e. Gaussian, HyperChem, Molegro Virtual Docker, LigandScout, etc.) that will be used in our researches.

- Chapter 3 : *Results and Discussion*

This chapter consists in discussing all the important results based on different approaches.

At first, we aim to study the molecular geometry, electronic properties and substitution effects of 1,3,4-thiadiazole nucleus using ab-initio and DFT methods, Afterward, we extend a dataset of 1,3,4-thiadiazole derivatives to identify compounds with higher potency and with better balance of properties respectively to their antitrypanosomal activity against *Trypanosoma Cruzi*.

Then, we develop the relationship between physicochemical properties of Cryptolepine derivatives with respect to their antitrypanosomal activities against *Trypanosoma Cruzain* and *Trypanosoma Rhodesain* using QSAR modeling.

Finally, we attempt to determine the active center sites of trypanocidal compounds using the Conceptual DFT and to identify the most favorable ligand conformation among the trypanocidal compounds under study when bound to the target *T. Cruzain* using the Molecular Docking analysis.

Chapter 1

TRYPANOCIDAL DRUG RESISTANCE

VS

DRUG DISCOVERY ACTIVITIES

The resistance towards available drugs is rapidly becoming a major worldwide problem. The need to design new compounds to deal with this resistance has become one of the most important areas of research nowadays.

1.1 Computational Chemistry for Drug Discovery

1.1.1 Historical Overview

Until the second half of the nineteenth century, the natural world was the principal source of chemical compounds capable of exerting therapeutic activity. Before then, plants, animals and minerals were manipulated by medicine men, shamans, witch doctors and regular physicians alike in what was usually a futile attempt to cure disease [14].

As an interdisciplinary endeavor with an industrial base, drug research is not much older than a century.

Drug research, as we know it today, began its career when chemistry had reached a degree of maturity that allowed its principles and methods to be applied to problems outside of chemistry itself and when pharmacology had become a well-defined scientific discipline in its own right [15]. By 1870, some of the essential foundations of chemical theory had been laid. Avogadro's atomic hypothesis had been confirmed and a periodic table of elements established. Chemistry had developed a theory that allowed it to organize the elements according to their atomic weight and valence. There was also a theory of acids and bases. In 1865, August Kekule formulated his pioneering theory on the structure of aromatic organic molecules [16, 17].

This benzene theory gave a decisive impulse to research on coal-tar derivatives, particularly dyes. In turn, the evolution of dye chemistry had a profound influence on medicine. The selective affinity of dyes for biological tissues led Paul Ehrlich, a medical student in the laboratory of the anatomist Wilhelm Waldeyer (between 1872 and 1874) at the University of Strasbourg, to postulate the existence of "chemoreceptors".

Ehrlich later argued that certain chemoreceptors on parasites, microorganisms, and cancer cells would be different from analogous structures in host tissues, and that these differences could be exploited therapeutically. It was the birth of chemotherapy, a particular type of drug therapy, that in the course of the 20th century led to unprecedented therapeutic triumphs [18].

Analytical chemistry, in particular the isolation and purification of the active ingredients of medicinal plants, also demonstrated its value for medicine in the 19th century. In 1815, F. W. Serturner isolated morphine from opium extract [19]. Papaverin was isolated in 1848, but its

antispasmodic properties were not discovered until 1917 [20]. As active ingredients from plants became available, many pharmacies addressed the problem of providing standardized preparations of these often still impure drugs. Coal-tar, an abundant by-product of the industrialization, contained many of the aromatic or aliphatic building blocks that became the toolkit of medicinal chemistry from its beginnings to the present [21].

Finally, pharmacology, which had its roots in the physiological experiments of François Magendie and Claude Bernard, claimed its place among the medical disciplines. Under the leadership of Oswald Schmiedeberg, the institute of pharmacology at the University of Strasbourg laid many of the intellectual and experimental foundations of pharmacology between 1871 and 1918 [22]. However, none of the institutions that had supported these seminal efforts-pharmacies, university laboratories, or the chemical companies producing dyes-represented suitable platforms for the newly emerging drug research that was driven by chemistry but increasingly controlled by pharmacology and by clinical sciences.

New institutions to support interdisciplinary drug research and development had to be created. They either grew out of pharmacies or were founded as pharmaceutical divisions in chemical or dye companies. A new way of finding, characterizing, and developing medicines led to the formation of a new industry [23, 24].

1.1.2 Drug Discovery Process

The process of drug development is time-consuming and costintensive. Several years are required for lead identification, optimization, and *in-vitro* and *in-vivo* testing before the first clinical trials are started. Preapproval costs of a new drug exceed 800 million USD [25]. It is well known that about 90% of drug candidates fail in the first phase of clinical trials [26, 27].

The reality is that the use of computers and computational methods permeates all aspects of drug discovery today. Those who are most proficient with the computational tools have the advantage for delivering new drug candidates more quickly and at lower cost than their competitors.

‘Is there really a case where a drug that’s on the market was designed by a computer?’ the phrasing of the question suggests misunderstanding and oversimplification of the drug discovery process.

First, it is the rare case today when an unmodified natural product like taxol becomes a drug. It is also inconceivable that a human with or without computational tools could propose

a single chemical structure that ends up as a drug; there are far too many hurdles and subtleties along the way.

Most drugs now arise through discovery programs that begin with identification of a biomolecular target of potential therapeutic value through biological studies including. A multidisciplinary project team is then assembled with the goal of finding clinical candidates, i.e., druglike compounds that are ready for human clinical trials, which typically selectively bind to the molecular target and interfere either with its activity as a receptor or enzyme. Molecular libraries are then screened, and the resulting leads are optimized in a cycle that features design, synthesis and assaying of numerous analogs [28].

1.1.3 Computational Tools in Drug Discovery

Computational chemistry uses physics-based algorithms and computers to simulate chemical events and calculate chemical properties of atoms and molecules. In drug design and discovery, diverse computational chemistry approaches are used to calculate and predict events, such as the drug binding to its target and the chemical properties for designing potential new drugs [28].

Computational methods are nowadays, routinely used to accelerate the long and costly drug discovery process. Typically, once the drug discovery target is selected, drug discovery activities (Fig. 1) are divided into those for:

A- The hit identification phase, in which the aim is the identification of chemical compounds with a promising activity toward the target.

B- The lead generation phase, in which hit compounds are improved in potency against the target.

C- The lead optimization phase, in which lead compounds are optimized, generating druglike molecules ultimately able to exert their beneficial pharmacological effect in patients.

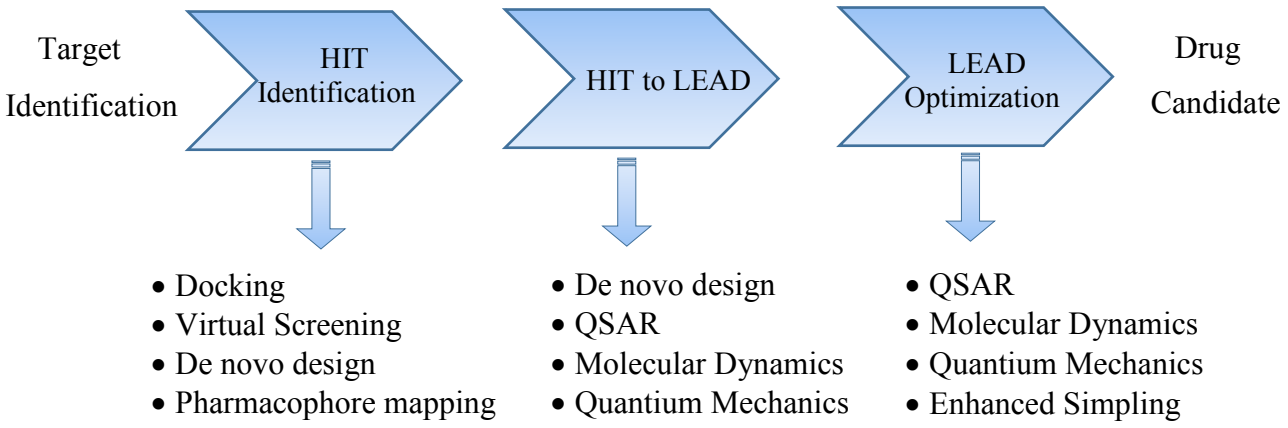


Fig. 1: Drug Discovery pipeline.

Computations can help in all these drug discovery activities, from drug target identification, commonly a receptor or an enzyme, to the design and optimization of a new drug-like compound. While computational methods for target identification rely mainly on computational sciences such as bioinformatics and computational genomics, different computational approaches are used, once the target has been identified and the search for small molecule inhibitors has commenced, starting from the hit identification phase, moving toward the lead generation and optimization phases (Fig. 1). At that point, routinely applied computational chemistry approaches include methods for structure-based drug design (SBDD), when structural data of the target protein are available, and ligand-based drug design (LBDD), when structural information of the target is missing or not fully reliable. Overall, these methods facilitate the identification of promising chemical scaffolds that interfere favorably with the target's function, producing a positive pharmacological effect.

Experimental biochemical and pharmacological data on new compounds, such as their *in-vitro* inhibitory potency and *in-vivo* efficacy, can be used to check the computational predictions while also forming the basis upon which better models can be constructed, leading to the design of superior compounds [28, 29].

The impact of computational chemistry on drug discovery has been intensified in the last few decades by the rapid development of faster architectures and better algorithms for time-affordable high-level computations.

1.1.4 Limitations

Every chemist wants to use computer to help limit failures in the lab. Computational chemistry is most useful when it is able to assist experimentalist by limiting. However, computational methods, they all have their strengths and weaknesses such as [30]:

A- Larger systems are less accurate then smaller systems.

B- It only can simulate real conditions only if programmed.

C- Need to be experimentally tested for accuracy.

D- Does not look at the whole picture.

1.1.5 Medicinal Chemists Today

Medicinal chemists today are facing a serious challenge because of the increased cost and enormous amount of time taken to discover a new drug, and also because of fierce competition amongst different drug companies.

Pharma Industry development became one of the important targets for different Governments in the last three decades, especially with the tremendous achievements of the multinational companies in this area of industry.

Challenges of this industry in third world have been increased, the discovery of a new active marketed molecule costs millions of USD, which increase the problem we are facing. Many worldwide companies have been merged to increase their ability in new drug discovery area. Most of them on the other hand adopted new technologies to discover and develop new drug entities, among these are computational and modeling techniques, which help in decreasing the time and cost of discovery researches, that's why importance of Computer Aided Drug Design and Molecular Modeling is increasing nowadays [31].

1.2 Human Trypanosomiasis Diseases

Human Trypanosomiasis is a vector-borne parasitic disease caused by intra-macrophage protozoa that belong to the genus Trypanosoma [8]. Current therapeutic options for Trypanosoma infections are not different from those available 20 years ago and are far from ideal, limited by toxicity and the emergence of drug-resistant parasites [9, 10]. Thus, the development of new and effective drugs as well as, the discovery of new molecular targets for new therapies is an urgent need to treat these vector-borne diseases.

1.2.1 Historical Overview

Human African Trypanosomiasis, also called as Sleeping sickness; it is an old disease. It was known to the slave traders, who rejected Africans with the characteristic swollen cervical glands, because they knew that these people would die untimely deaths [32]. There have been three particularly severe epidemics during the twentieth century. The first was from 1896 until 1906 in Uganda and the Congo basin, the second during the 1920s, and the third began in the 1970s and continues until the present time. Intensive systematic screening by mobile teams, of many millions of people per year at risk, halted the epidemic of the 1920s. The illness was practically eliminated by 1960. Such active population screening was not continued, at least partly because the disease had nearly disappeared from Africa. Not surprisingly, with the breakdown of the control system, the disease has re-emerged again as a major health problem [33].

Human American Trypanosomiasis, also known as Chagas disease, at first observed in 1909 in the intestine of Triatomine bugs at Minas Gerais province in Brazil by Carlos Chagas. Trypanosome was transmitted to monkeys and then sought in vertebrates; it is first in cat's blood that was found then in humans.

The work done at this time through the motivation and mobilization of the medical personnel has led to excellent results. The disease was controlled in its main foci and each patient could benefit from a treatment. Political, ethnic and economic problems arose after the independence of most countries that contributed to the re-emergence of the disease in the last thirty years. Nowadays, fight and control strategies have progressed but unfortunately are always difficult to set up and maintain because of local conditions and the lack of resources [33-35].

1.2.2 Human African Trypanosomiasis

1.2.2.1 Definition

Human African Trypanosomiasis, is a deadly parasitic disease endemic to sub-Saharan Africa [35]. It is caused by infection with protozoan parasites belonging to the genus Trypanosoma. They are transmitted to humans by tsetse fly bites (Glossina genus) which have acquired their infection from human beings or from animals harbouring human pathogenic parasites.

Tsetse flies are found just in sub-Saharan Africa though only certain species transmit the disease. For reasons that are so far unexplained, in many regions where tsetse flies are found, sleeping sickness is not.

Rural populations living in regions where transmission occurs and which depend on agriculture, fishing, animal husbandry or hunting are the most exposed to the tsetse fly and therefore to the disease.

The disease develops in areas ranging from a single village to an entire region. Within an infected area, the intensity of the disease can vary from one village to the next [36-39].

1.2.2.2 Distribution

There are four main endemic areas from which countries affected by Trypanosomiasis are classified according to the levels of prevalence of the disease. In each country, the spatial distribution of the reported cases is very heterogeneous and appears as foci and even micro-foci.

Countries where Trypanosomiasis is epidemic have a high prevalence and high transmission level. These are Angola, the Democratic Republic of Congo, Uganda and Sudan.

Countries with a high endemicity have a relatively high prevalence and a steady increase. This is observed in Cameroon, Congo, Côte d'Ivoire, Central African Republic, Guinea, Mozambique, Tanzania and Chad.

Countries with low endemicity are Benin, Burkina Faso, Gabon, Ghana, Equatorial Guinea, Kenya, Mali, Nigeria, Togo and Zambia.

Some countries have an unknown epidemiological status to this day, such as Burundi, Botswana, Ethiopia, Liberia, Namibia, Rwanda, Senegal and Sierra Leone (Fig. 2) [40].

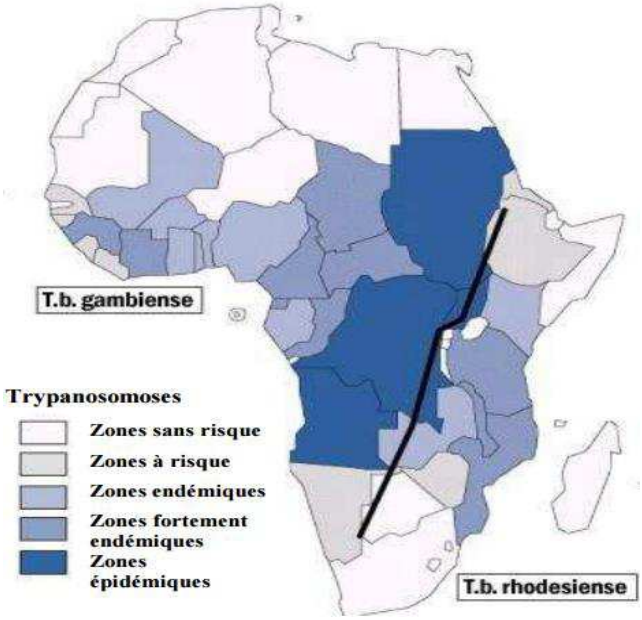


Fig. 2: Geographical distribution of Trypanosomiasis *T. b. Gambiense and T. b. Rhodesiense* in sub-Saharan Africa, 2015 (WHO).

1.2.2.3 Signs and Symptoms

In the first stage, the Trypanosomes multiply in subcutaneous tissues, blood and lymph. This is also called haemo-lymphatic stage, which entails bouts of fever, headaches, joint pains and itching.

In the second stage the parasites cross the blood-brain barrier to infect the central nervous system. This is known as the neurological or meningo-encephalic stage. In general this is when more obvious signs and symptoms of the disease appear: changes of behaviour, confusion, sensory disturbances and poor coordination.

Disturbance of the sleep cycle, which gives the disease its name, is an important feature. Without treatment, sleeping sickness is considered fatal although cases of healthy carriers have been reported.

1.2.2.4 Transmission

The disease is mostly transmitted through the bite of an infected tsetse fly but there are other ways in which people are infected:

- Mother-to-child infection: the Trypanosome can cross the placenta and infect the fetus.

- Mechanical transmission through other blood-sucking insects is possible, however, it is difficult to assess its epidemiological impact.
- Accidental infections have occurred in laboratories due to pricks with contaminated needles.
- Transmission of the parasite through sexual contact has been also documented [7, 40].

1.2.2.5 Forms of Human African Trypanosomiasis

Human African Trypanosomiasis takes two forms, depending on the parasite involved:

- Trypanosoma brucei gambiense, which is found in 24 countries in west and central Africa. This form currently accounts for over 98% of reported cases of sleeping sickness and causes a chronic infection. A person can be infected for months or even years without major signs or symptoms of the disease. When more evident symptoms emerge, the patient is often already in an advanced disease stage where the central nervous system is affected.
- *Trypanosoma brucei Rhodesiense*, which is found in 13 countries in eastern and southern Africa. Nowadays, this form represents under 2% of reported cases and causes an acute infection. First signs and symptoms are observed a few months or weeks after infection. The disease develops rapidly and invades the central nervous system. Only Uganda presents both forms of the disease, but in separate zones.

Another form of Trypanosomiasis occurs mainly in Latin America. It is known as American Trypanosomiasis or Chagas disease. The causal organism belongs to a different Trypanosoma subgenus and is transmitted by a different vector [7, 41].

1.2.3 American Human Trypanosomiasis

1.2.3.1 Definition

Chagas disease, is a potentially life-threatening illness caused by the protozoan parasite *Trypanosoma Cruzi (T. Cruzi)*. Chagas disease is found mainly in endemic areas of 21 Latin American countries, where it is mostly vector-borne transmitted to humans by contact with faeces or urine of triatomine bugs, known as 'kissing bugs', among many other names, depending on the geographical area.

About 10 million people worldwide are estimated to be infected with *Trypansosoma Cruzi*, mainly in Latin America where Chagas disease is in the endemic state. Over 25 million people are exposed to the disease. In 2008, Chagas disease have killed more than 10,000 people.

The cost of treatment for Chagas disease remains substantial. In Colombia alone, the annual cost of medical care for all patients with the disease was estimated to be about 267 million USD in 2008. Spraying insecticide to control vectors would cost nearly 5 million USD annually [42].

1.2.3.2 Distribution

Chagas disease occurs mainly in the continental part of Latin America, with the exception that the disease has not occurred in the Caribbean isles. In the past decades, however, it has been increasingly detected in the United States of America, Canada, and many European and some Western Pacific countries (Fig. 3). This is due mainly to population mobility between Latin America and the rest of the world [34, 43].

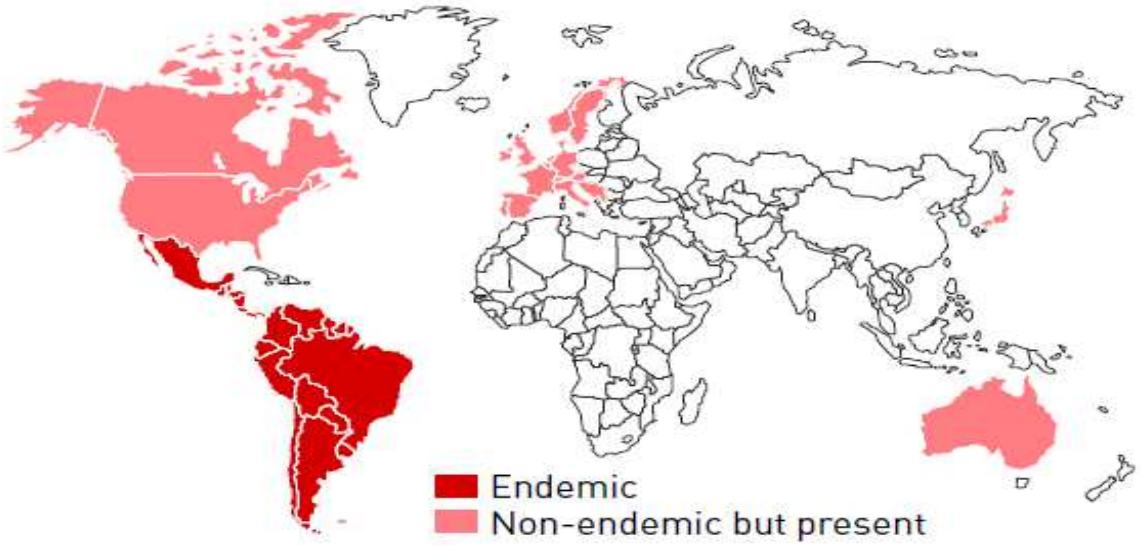


Fig. 3: Geographical distribution of *Trypanosoma Cruzi* overall the world, 2015 (WHO).

1.2.3.3 Signs and Symptoms

Chagas disease presents itself in two phases. The initial, acute phase lasts for about two months after infection. During the acute phase, a high number of parasites circulate in the blood but in most cases symptoms are absent or mild. In less than 50% of people bitten by a triatomine bug, characteristic first visible signs can be a skin lesion or a purplish swelling of the lids of one eye. Additionally they can present fever, headache, enlarged lymph glands, pallor, muscle pain, difficulty in breathing, swelling, and abdominal or chest pain.

During the chronic phase, the parasites are hidden mainly in the heart and digestive muscles. Up to 30% of patients suffer from cardiac disorders and up to 10% suffer from digestive (typically enlargement of the oesophagus or colon), neurological or mixed alterations. In later years, the infection can lead to sudden death or heart failure caused by progressive destruction of the heart muscle and its nervous system [34, 44].

1.2.3.4 Transmission

In Latin America, *T. Cruzi* parasites are mainly transmitted by contact with faeces/urine of infected blood-sucking triatomine bugs. These bugs, vectors that carry the parasites, typically live in the wall or roof cracks of poorly-constructed homes in rural or suburban areas. Normally they hide during the day and become active at night when they feed on human blood. They usually bite an exposed area of skin such as the face, and the bug defecates close to the bite.

The parasites enter the body when the person instinctively smears the bug faeces or urine into the bite, the eyes, the mouth, or into any skin break.

T. Cruzi can also be transmitted by [38, 45]:

- Consumption of food contaminated with *T. Cruzi* through, for example, contact with infected triatomine bug faeces or urine.
- Blood transfusion from infected donors.
- Passage from an infected mother to her newborn during pregnancy or childbirth.
- Organ transplants using organs from infected donors.
- Laboratory accidents.

1.3 Parasites

"You had no right to be born; for you make no use of life. Instead of living for, in, and with yourself, as a reasonable being ought, you seek only to fasten your feebleness on some other person's strength"

1.3.1 Historical Overview

During our relatively short history on Earth, humans have acquired an amazing number of parasites, about 300 species of helminth worms and over 70 species of protozoa. Many of these are rare and accidental parasites, but we still harbor about 90 relatively common species, of which a small proportion cause some of the most important diseases in the world, inevitably, these are the ones that have received the most attention [46].

Since most of these parasitic diseases occur mainly in the tropics, the field of parasitology has tended to overlap with that of tropical medicine, and thus the histories of these two fields are intertwined. There is, however, much more to the history of human parasitology than this, and our understanding of parasites and parasitic infections cannot be separated from our knowledge of the history of the human race.

Human evolution and parasitic infections have run hand in hand, and thanks to the spinoffs from the Human Genome Project, we now know much more about the origins of the human race than ever before. Sometime, about 150,000 years ago, Homo sapiens emerged in eastern Africa and spread throughout the world, possibly in several waves, until 15,000 years ago at the end of the Ice Age humans had migrated to and inhabited virtually the whole of the face of the Earth, bringing some parasites with them and collecting others on the way. In particular, the spread and present distribution of many parasites throughout the world has largely been the result of human activities [47-49].

In recent years parasites have become a powerful model system for the study ecology and evolution, with practical applications in disease prevention. They plague billions of people, kill millions annually, and inflict debilitating injuries such as blindness and disfiguration on additional millions. World Health Organization estimates that one person in every four harbors parasitic worms [50].

1.3.2 Definitions

Parasitism is a very common way of life, and probably the prevalent means of obtaining food among organisms. Parasitism involves intimate association of two living organisms of different kinds. One called "host" provides accommodation and food for another, which called "parasite" [51].

A parasite is an organism living in or on another living organism, obtaining from it part or all of its organic nutriment, commonly exhibiting some degree of adaptive structural modification, and causing some degree of real damage to its host [52]. Therefore, we can say, the parasite lives at the expense of the host, but there are several types of host-parasite relationship and parasite behavior.

There are different kinds of parasites which can be listed as [53]:

- Ectoparasite – a parasitic organism that lives on the outer surface of its host, e.g. lice, ticks, mites etc.
- Endoparasites – parasites that live inside the body of their host, e.g. Entamoeba histolytica.
- Obligate Parasite - This parasite is completely dependent on the host during a segment or all of its life cycle, e.g. Plasmodium spp.
- Facultative parasite – an organism that exhibits both parasitic and non-parasitic modes of living and hence does not absolutely depend on the parasitic way of life, but is capable of adapting to it if placed on a host. E.g. Naegleria fowleri
- Accidental parasite – when a parasite attacks an unnatural host and survives. E.g. Hymenolepis diminuta (rat tapeworm).
- Erratic parasite - is one that wanders in to an organ in which it is not usually found. E.g. Entamoeba histolytica in the liver or lung of humans.

Also, there are different kinds of hosts as listed below:

- Definitive host – a host that harbors a parasite in the adult stage or where the parasite undergoes a sexual method of reproduction.
- Intermediate host - harbors the larval stages of the parasite or an asexual cycle of development takes place. In some cases, larval development is completed in two different intermediate hosts, referred to as first and second intermediate hosts.

- Paratenic host – a host that serves as a temporary refuge and vehicle for reaching an obligatory host, usually the definitive host, i.e. it is not necessary for the completion of the parasites life cycle.
- Reservoir host – a host that makes the parasite available for the transmission to another host and is usually not affected by the infection.
- Natural host – a host that is naturally infected with certain species of parasites.
- Accidental host – a host that is under normal circumstances not infected with the parasite.

1.3.3 Distribution of Parasites in the Animal World

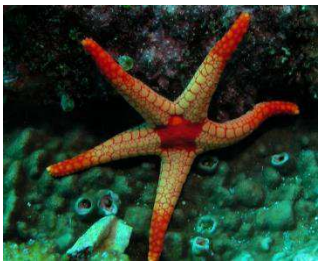
Parasites are found throughout the living world. Parasites, themselves, belong either to the world of mushrooms (Parasitism by fungiis is not mentioned here) or to the animal world.

The parasites distributed into several groups. Whereas, some groups are composed almost exclusively of parasites. Although, most have both parasitic and free species. Thus, they can be found as vertebrates, which have very little parasitic species, or we can say that they exist only among fish (crustaceans).

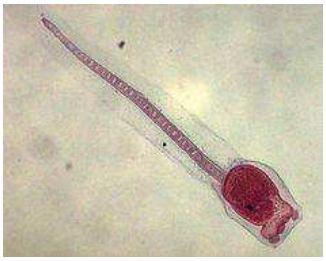
Parasites distribution groups are described as below [54, 55]:

A. Groups do not contain parasite, such as:

Sponges



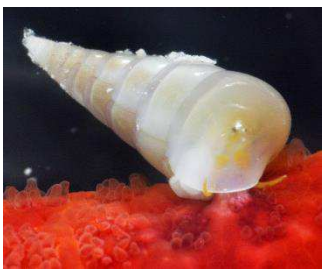
Echinoderms



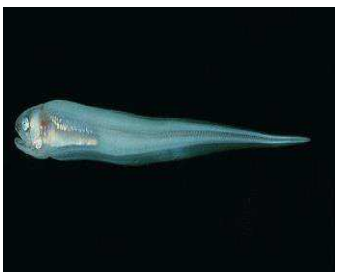
Urochordates

B. Groups restrict number of parasites, such as:

Coelenterata



Molluscs

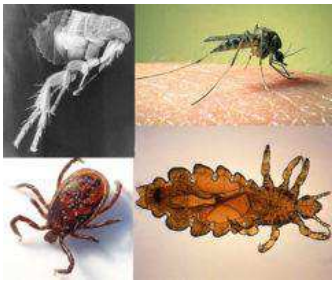


Vertebrates

C. Groups of numerous parasites, such as:

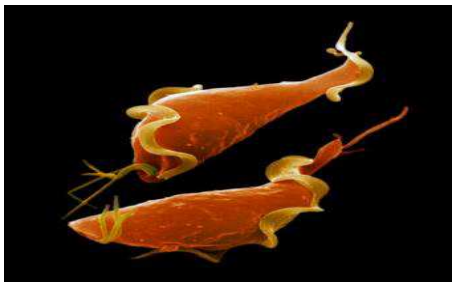
Crustaceans

Mites



Insects

D. Groups of excessively numerous parasites, such as:



Protozoa

Helminths

1.3.4 Classification of Human Parasites

Parasites form part of the animal kingdom which comprises about 800,000 identified species. The classification of parasites is controversial as there is no universally accepted system [53].

From biological and morphological sides, parasites are classified into four groups [53]:

A. Protozoa (unicellular endowed with motion): according to cases, they are moving through charm sprites (Rhizopods), flagella, undulating membrane or eyelashes. They appear as asexual or sexual potential, movable or encysted and intra- or extracellular.

B. Helminth or worm (Multicellular with differentiated tissues). They are known as Adults of both sexes in the form larval, embryonic or ovular.

C. Fungi, they constitute a full reign, which are microscopic fungi identified as isolated spores or consolidated or free filaments or tissue based.

D. Arthropods, molluscs, pararthropodes or annelids, they are metazoans, multicellular with differentiated tissues; they can be in adult forms males and females.

The Eco-classification of parasites is made of a sequence of steps as bellow:

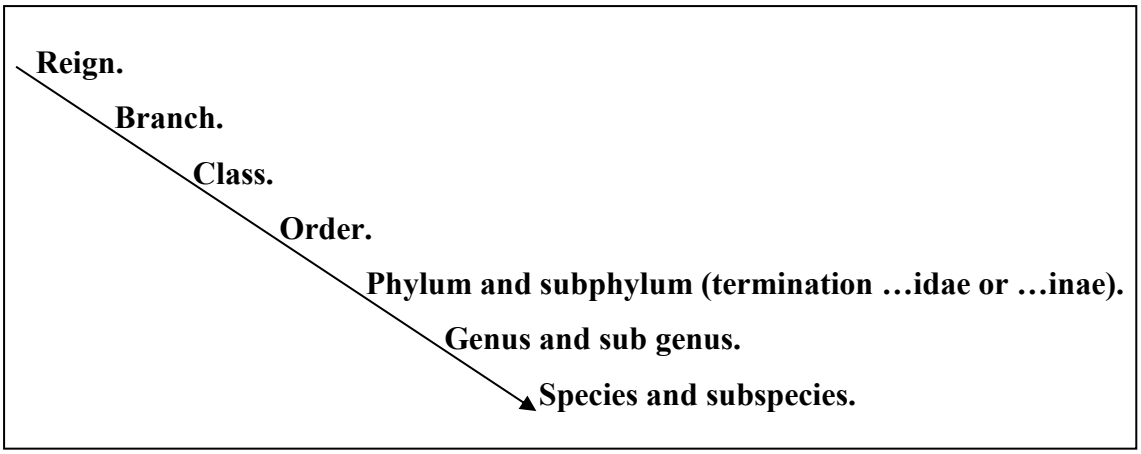


Fig. 4: Genus and species of parasites resulting from a sequence of steps.

1.3.5 Human Vector-Borne Infections

Parasites can cause disease in humans. Some parasitic diseases are easily treated and some are not. The burden of these diseases often rests on communities in the tropics and subtropics, but parasitic infections also affect people in developed countries [56].

Table. 1: Human vector-borne infections.

Infection (Disease)	Causative Agent	Vector (Common Name)
Protozoal		
Malaria	*Plasmodium* spp.	Mosquitoes
Leishmaniasis	*Leishmania* spp.	Sandflies
Chagas disease	*Trypanosoma Cruzi*	Triatomid bugs
East African Trypanosomiasis	*Trypanosoma brucei Rhodesiense*	Tsetse flies
West African Trypanosomiasis	*Trypanosoma brucei gambiense*	Tsetse flies
Babesiosis	*Babesia* spp.	Ticks
Helminthic		
Filariasis	*Wuchereria bancrofti*	Mosquitoes
Filariasis	*Brugia malayi*	Mosquitoes
Filariasis	*Dirofilaria* spp.	Mosquitoes
Filariasis	*Mansonella perstans*	Biting midges
Filariasis	*Mansonella streptocerca*	Biting midges
Filariasis	*Mansonella ozzardi*	Biting midges
Onchocerciasis	*Onchocerca volvulus*	Black flies
Loiasis	*Loa loa*	Deer flies
Dog tapeworm infection	*Dipylidium caninum*	Dog lice and fleas, human fleas
Rat tapeworm infection	*Hymenolepis diminuta*	Rat fleas, beetles, grain beetles
Dwarf tapeworm	*Hymenolepis nana*	Grain beetles (rare)

1.3.6 Protozoa Parasites

"Protozoa" means "first animal", the simplest form of animal life

1.3.6.1 What is a Protozoa

Protozoa are unicellular eucaryotic microorganisms that lack cell walls, can grow in marine habitat, soil, fresh water, symbiotic or parasites in other organisms. Protozoa depends on nutrition, temperature, pH and some depends on sunlight [57].

1.3.6.2 Protozoa Characteristics [57]

Protozoa parasite has special characteristics as fellow:

- Eucaryotic unicellular: 1 – 150 µm.
- No cell walls.
- Mostly motile, with flagella or cilia, or amoeboid.
- Chemoheterotrophs.
- Similar to animal, only unicellular.
- Feed by ingesting particulate matter (phagocytosis) and engulfing liquid or dissolved nutrition (pinocytosis).
- Most protozoa are parasites and have two forms: Trophozoite and Cysteine.
- Protozoa are distinguished from prokaryotes by their eukaryotic nature, from algae by their lack of chlorophyll, from fungi by their lack of cell walls, and from slime molds by their inability to fruiting bodies.

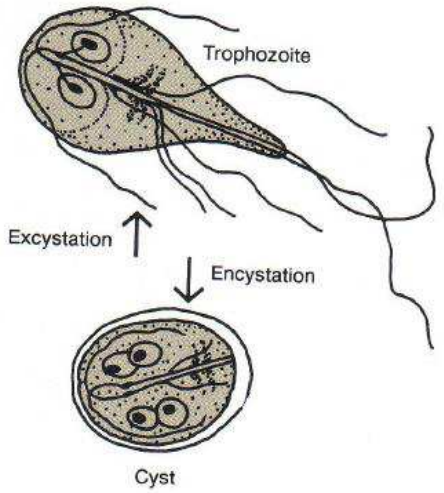


Fig. 5: Protozoa forms.

1.3.6.3 Classification of Protozoa

Old classification, based on mortality [57]:

- Sarcodina (amoeboid): Entamoeba histolytica.
- Mastigophores (with flagella): Trypanosoma brucei.
- ciliates (with cilia) : Balantidium coli.
- Sporozoa (no mature form): Plasmodium, Toxoplasma.

New classification (started since 1986), based on cellular structure by electron microscope [56, 57]:

- Phylum: Sarcomastigophora: Trypanosoma.
    - Sub-Phylum Mastigophora.
    - Sub-phylum Opalinata.
    - Sub-phylum Sarcodina.
- Phylum: Labyrinthomorpha : Labyrinthula.
- Phylum: Apicomplexa: Toxoplasma.
- Phylum: Myxozoa: Ceratomyxa.
- Phylum: Microspora: Encephalitozoon.
- Phylum: Ascetospora: Marteilia.
- Phylum: Ciliophora: Balantidium.

Another classification has been created based on the morphological criteria, whereas, Protozoa are represented by seven groups [58]:

- Flagellated protozoa.
- Amoeboid protozoa.
- Sporozoites.
- Microsporidia.
- Haplosporidium.
- Myxosporean.
- Protozoa ciliates.

1.3.7 Flagellate Protozoa (Trypanosomatidae) blood and Tissues Parasites

Tissue flagellated protozoan belong to the phylum Kinetoplastida and to the Trypanosomatid order. They have one to four flagella. The kinetoplast that contains DNA is a self-reproducible organelle bound to mitochondria (Fig. 6).

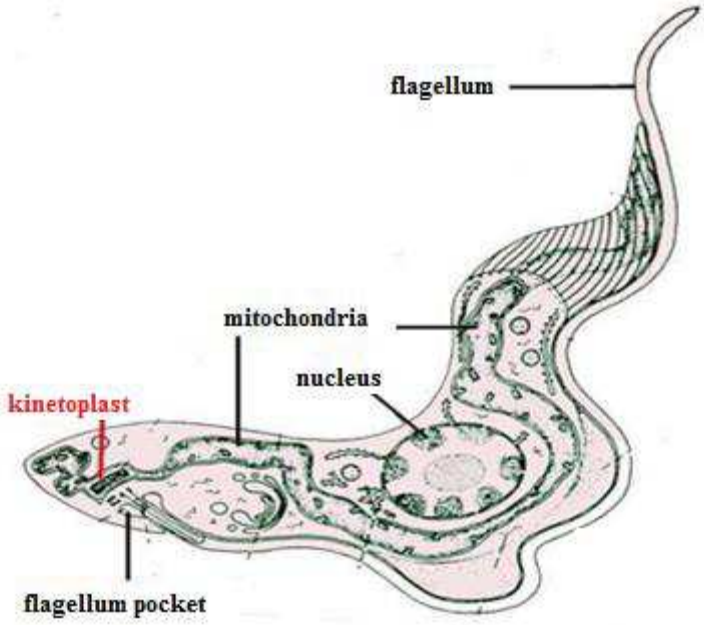


Fig. 6: Trypanosoma structure as trypomastigote.

The family of the Trypanosomatidae contains four genus: Leptomonas, Crithidia, Trypanosomes and Leishmania. The genus Trypanosoma is divided into subgenus based on the behavior of vertebrate and invertebrate hosts such as Trypanozoon, Duttonella, Nannomonas, Pycnomonas, Tejeraia, Schizotrypanum, Herpetosoma, Megatrypanum, Endotrypanum.

Species of the genus Trypanosomawhich infect humans in Africa (T. Brucei) belong to the subgenus Trypanozoon and in America (*T.Cruzi*) to subgenus Schizotrypanum [7]

1.3.8 Responsible Enzymes for Human African Trypanosomiasis and Chagas Diseases

1.3.8.1 Trypanosoma Brucei

Most Trypanosomes cannot infect humans as they die in human serum, but two mutants of *Trypanosoma Brucei* have evolved the ability to survive in human serum. These subspecies are *Rhodesian* and *Gambian*, which are the causative agents of Human African Trypanosomiasis which called 'sleeping sickness', these parasites are transmitted by vector, tsetse fly (Fig. 7), most of which belong to the species Glossina palpalis [59]. These parasites

are a great source of human suffering and economic loss in a large area of the African continent [60].

The treatment of Trypanosomiasis at *T.B. Rhodesian* is harder than the infection by *T.B.gambiens*. In addition, it is passing to the nervous stage more faster [61].

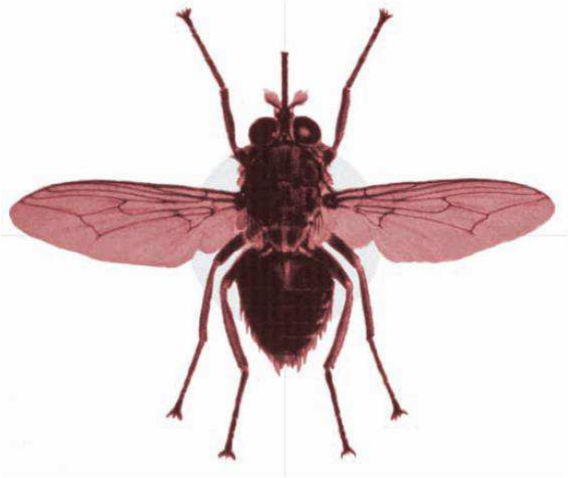


Fig. 7: Tsetse fly.

1.3.8.2 Trypanosoma Cruzi

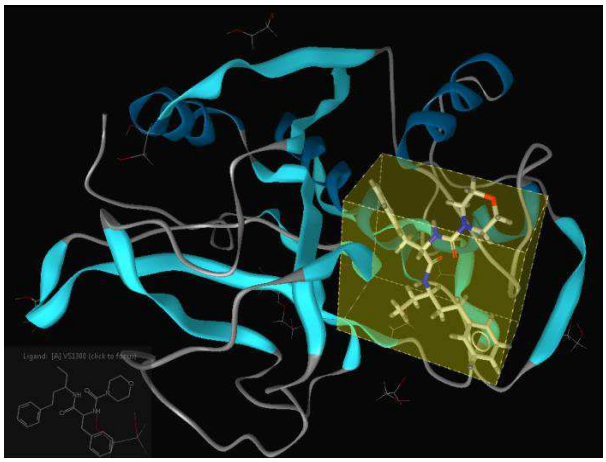
It is the causative agent of Human American Trypanosomiasis, which called 'Chagas disease'; this parasite is transmitted to humans by blood-sucking insects (Fig. 8) of the family Reduviidae (Triatominae) [62, 63].

Trypanosoma Cruzi, is an unbeaten adversary. Immunological or chemotherapeutic protection is not available and it is becoming, like other higher Trypanosomatidae, a favorite for studies intended to elucidate how mitochondrial DNA joins in regulating relations between life in the invertebrate vector and in the mammalian host. Thanks largely to its easy cultivation and apparently indefinite retention of infectivity for vector and mammalian hosts: it is much more amenable to experimental manipulation than the salivarian Trypanosomes, especially those of the T. Brucei group. It shares with Leishmania amenability to analysis of adapt ion to intracellularity. *T. Cruzi* under conditions not well defined, can form in culture infective metacyclic trypomastigotes, a morphogenetic process far more difficult to effect with salivarian Trypanosomes. This joining of urgent practical and timely biological motives makes *T. Cruzi* despite the hazards attending its use, an ever more popular object even among workers not specifically concerned with prophylaxis and therapy [64].

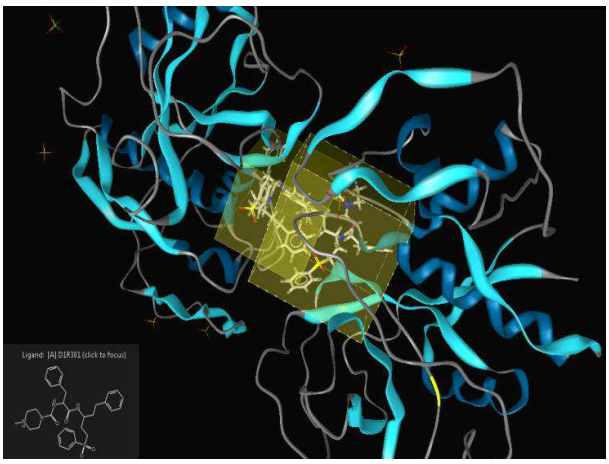
Fig. 8: Reduviid bug.

1.3.9 T. Brucei and T. Cruzi Structures

They are papain-like from the family of cysteine proteases (Cathepsin L-like enzyme), they are capable of degrading the extracellular compounds, *T.Rhodesain* and *T.Cruzi* (Fig. 9) contain the triad-catalytic active site (Cys \ His \ ASN).



Rhodesain bound to vinyl sulfone derived inhibitor



Cruzi bound to vinyl sulfone derived inhibitor

Fig. 9: Crystal structures of *T.Rhodesain* (2P86) and *T.Cruzi* (2OZ2).

1.4 Heterocyclic Compounds

The IUPAC Gold Book describes heterocyclic compounds as:

Cyclic compounds having as ring members atoms of at least two different elements, e.g. quinoline, 1,2-thiazole, bicyclo[3.3.1]tetrasiloxane [65].

Rings are considered as "heterocycles" only if they contain at least one atom selected from halogen, N, O, S, Se or Te as a ring member. Heterocyclic rings may be present as distinct entities or condensed, either with carbocycles or among themselves.

1.4.1 Historical Overview

Heterocycles are common structural units in marketed drugs and in medicinal chemistry targets in the drug discovery process. Over 80% of top small molecule drugs by US retail sales in 2010 contain at least one heterocyclic fragment in their structures. In fact, heterocyclic moieties are present in the structures of all top 10 brand name small molecule drugs (Fig. 10).

The one reason behind such high prevalence of oxygen, sulfur, and especially nitrogen-containing rings in drug molecules is obvious. The research process that leads to identification of an effective therapeutic treatment is largely based on mimicking nature by "fooling" it in a very subtle way. Because heterocycles are the core elements of a wide range of natural products such as nucleic acids, amino acids, carbohydrates, vitamins, and alkaloids, medicinal chemistry efforts often evolve around simulating such structural motifs. However, heterocycles play a much bigger role in the modern repertoire of medicinal chemists. Some of the drug properties that can be modulated by a strategic inclusion of heterocyclic moiety into the molecule include: 1) Potency and selectivity through bioisosteric replacements, 2) Lipophilicity, 3) Polarity, and 4) Aqueous solubility [65].

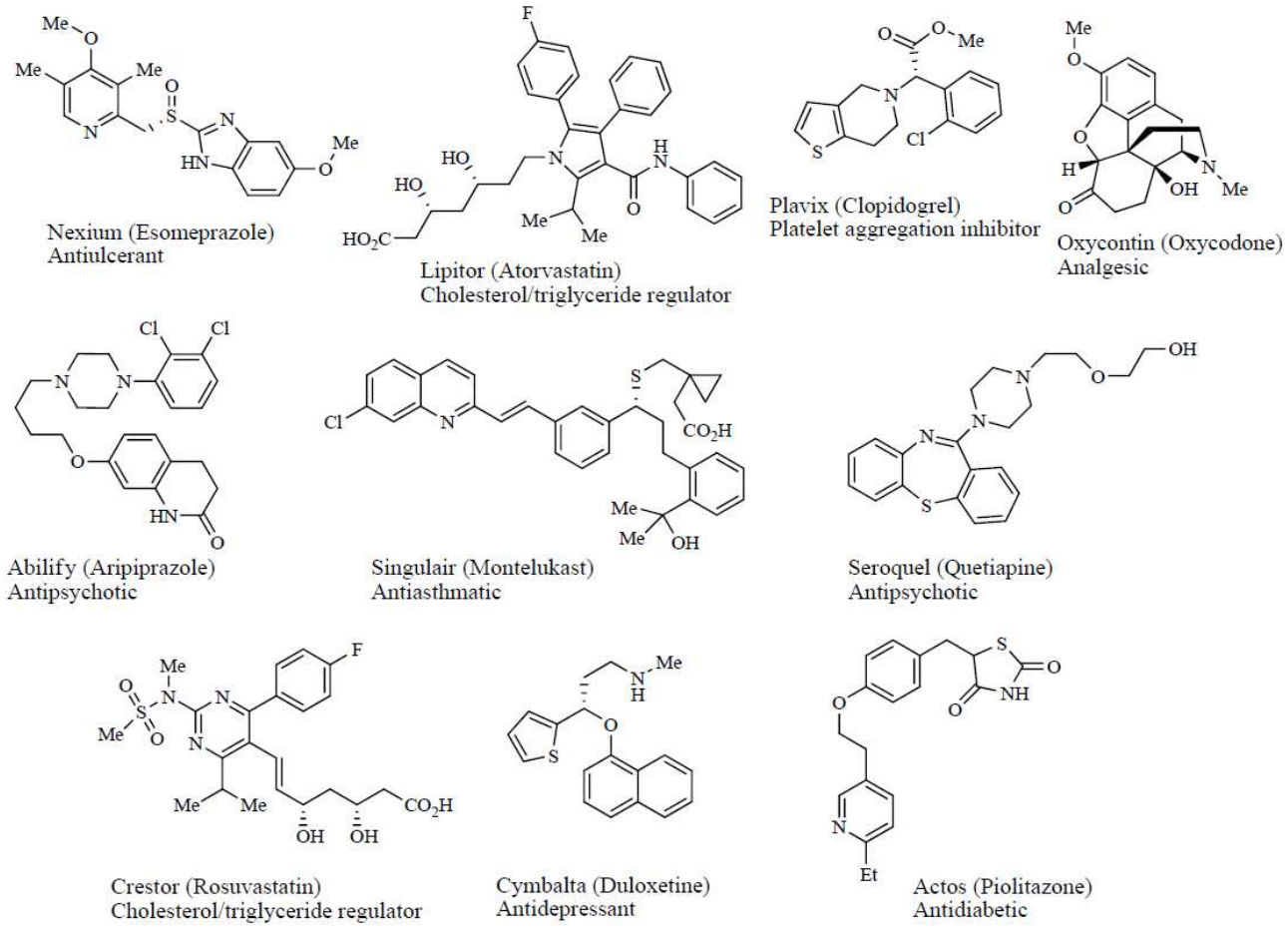


Fig. 10: Top 10 brand name small molecule drugs by retail sales in 2010.

1.4.2 Cryptolepine

Cryptolepine, a naturally occurring indoloquinoline alkaloid used as an antimalarial drug in Central and Western Africa, this molecule presents a moderate antitrypanosomal activity [11].

1.4.2.1 Generality

Cryptolepine (5-methyl-10*H*-indolo[3,2-b]quinoline) (Fig. 11-A) is the major alkaloid from the roots of the west African climbing shrub Cryptolepis sanguinolenta, an herbal drug used in traditional medicine for the treatment of malaria [66, 67]. Cryptolepine and its hydrochloride (Fig. 11-B), display potent *in-vitro* antiplasmodial activity [68] but also present cytotoxic properties that preclude their clinical use [69]. The cytotoxicity can be ascribed to the ability of cryptolepine to intercalate into DNA and inhibit topoisomerase II as well as DNA synthesis [70, 71].

However, studies on the possible mode of antimalarial action suggest that cryptolepine is able to inhibit hemozoin formation (i.e., the heme detoxification process), which is also the mechanism reported for 4-aminoquinolines. A basic amino side-chain is a major chemical feature required for 4-aminoquinoline accumulation in the parasite. Based on this structural requirement, Joao Lavrado et al. designed a novel group of cryptolepine analogues, (Fig. 11-C), that incorporate an alkyldiamino side-chain at C11, aiming to increase accumulation in the parasite food vacuole [12].

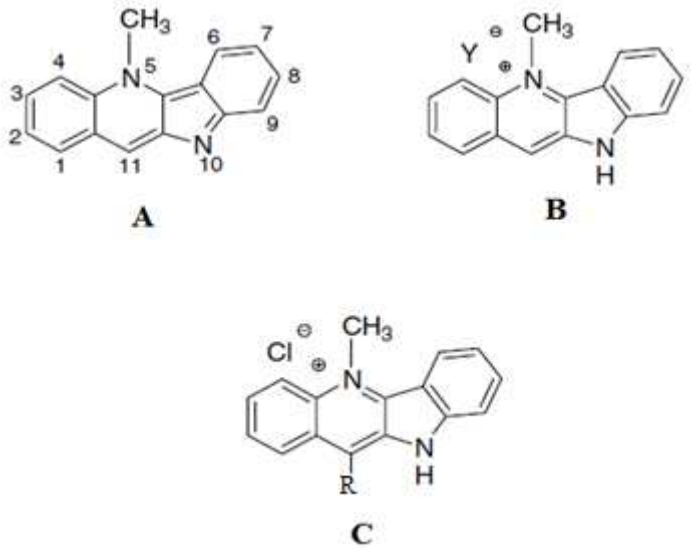


Fig. 11: Structures of : 'A' Cryptolepine, 'B' Cryptolepine hydrochloride (Y = Cl) and 'C' Cryptolepine derivatives with basic side-chains at C-11.

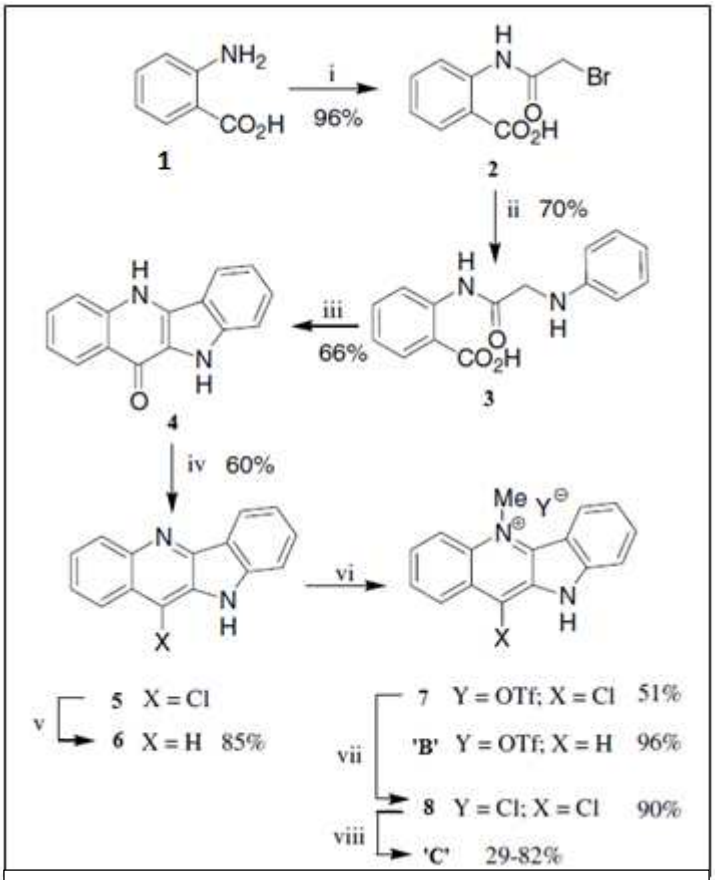
1.4.2.2 Cryptolepine Derivatives

Cryptolepine hydrochloride, 'B' (Y = Cl), and its C11 alkyldiamine derivatives, 'C', were synthesized via 11-chloroquindoline intermediate, 5, according to the route depicted in Scheme 1. Anthranilic acid, 1, was treated with bromoacetyl bromide to afford the bromoacetyl derivative 2, which was treated with aniline to give compound 3. Acid-catalyzed cyclization of 3 with polyphosphoric acid (PPA) gave the quindolone 4, which, by reaction with $POCl_3$, gave the corresponding 11-chloroquindoline 5. Hydrogenation of 5 with 10% Pd-C at 60 psi provided the quindoline 6. Reaction of 5 and 6 with methyl triflate formed the corresponding triflate 7 and cryptolepine triflate ('B', Y = OTf).

Initial attempts to prepare compounds 'C' involved the reaction of triflate 7 with alkyldiamines, but the final product contained always triflate and chloride anions.

Alternatively, compound 7 was first treated with sodium carbonate and then with HCl to give the key intermediate 11-chlorocryptolepine hydrochloride, 8.

Finally, cryptolepine derivatives containing an alkyldiamine side-chain at C11, 'C', were obtained by reacting 8 with the appropriate amine, in 29–82% yield. Most of the alkyldiamines used in this final step were purchased and used without further purification.



Scheme 1: Synthesis of cryptolepine derivatives. Reagents and conditions: (i) Bromoacetyl bromide, DMF/Dioxane, rt; (ii) aniline, DMF, 120 °C; (iii) PPA, 130 °C; (iv) $POCl_3$, 120 °C; (v) H2 Pd-C 10%, NaOAc, AcOH, 60 psi; (vi) 5 or 6, methyl triflate, anhydrous toluene, rt; (vii) (a) 5% Na_2CO_3; (b) HCl–Et_2O; (viii) RNH_2, AcOEt, reflux.

1.4.3 1,3,4-thiadiazol

1,3,4-thiadiazole, a heterocyclic molecule that holds a special place among the major pharmaceutical products; where, many studies have been carried out on this one rather than other isomers; 1,3,4-thiadiazole and its derivatives have a considerable attention to enhance the antibacterial and antiparasitic activity, particularly against Trypanosomes [11].

1.4.3.1 Generality

Thiadiazole is a 5- membered ring system containing hydrogen-binding domain, sulfur atom, and two-electron donor nitrogen system that exhibit a wide variety of biological activity. They occur in four isomeric forms in the nature viz. 1,2,3-thiadiazole; 1,2,5-thiadiazole; 1,2,4-thiadiazole; and 1,3,4-thiadiazole (Fig. 12).

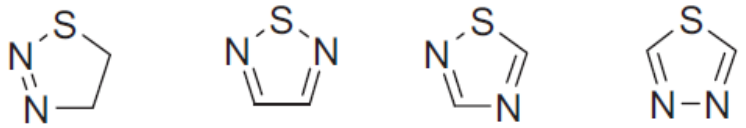


Fig. 12: Thiadiazole structures.

Thiadiazole nucleus is present as a core structural component in an array of drug categories such as antimicrobial, anti-inflammatory, analgesic, antiepileptic, antiviral, antineoplastic, and antitubercular agents. The broad and potent activity of thiadiazole and its derivatives has established them as pharmacologically significant scaffolds [72].

1.4.3.2 1,3,4-thiadiazol Derivatives

A glance at the standard reference work shows that many studies have been carried out on the 1,3,4-thiadiazole rather than other isomers. The 1,3,4-thiadiazole nucleus is one of the most important and well-known heterocyclic nuclei, which is a common and integral feature of a variety of natural products and medicinal agents. 1,3,4-thiadiazole and its derivatives have a considerable attention to enhance the antibacterial and antiparasitic activity particularly against Trypanosomes [73].

Considering this panorama and trying to circumvent this undesirable profile, several 1,3,4-thiadiazole-2-arylhydrazone (Fig. 13) derivatives have been designed and synthesized as attractive antichagasic drug candidates. Due to the hypothesis that the introduction of a radical scavenger subunit linked to its heterocyclic scaffold could modulate the production of toxic nitro anion radical species, they could potentially avoid mutagenic properties [74, 75].

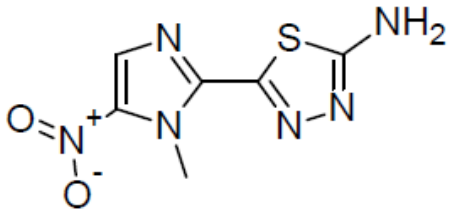


Fig. 13: 1,3,4-thiadiazole-2-arylhydrazone structure.

1.4.4 Nifurtimox

Nifurtimox, a cheap, orally administered drug not yet fully validated for use in these vector borne diseases. Whereas, Nifurtimox presents serious toxic side effects, there are also doubts as to whether this drug is capable of achieving parasitological cure [11].

1.4.4.1 Generality

The Nitrofuran drug, Nifurtimox (Fig. 14), used against Chagas disease, contains a nitro group that is central to activity. It also has been used in trials, with limited success, against arsenical refractory T. Brucei in West Africa [75].

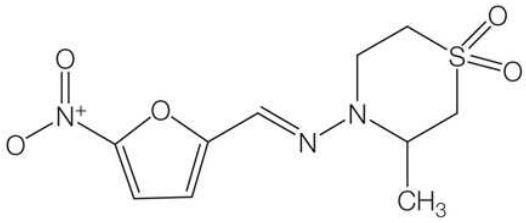


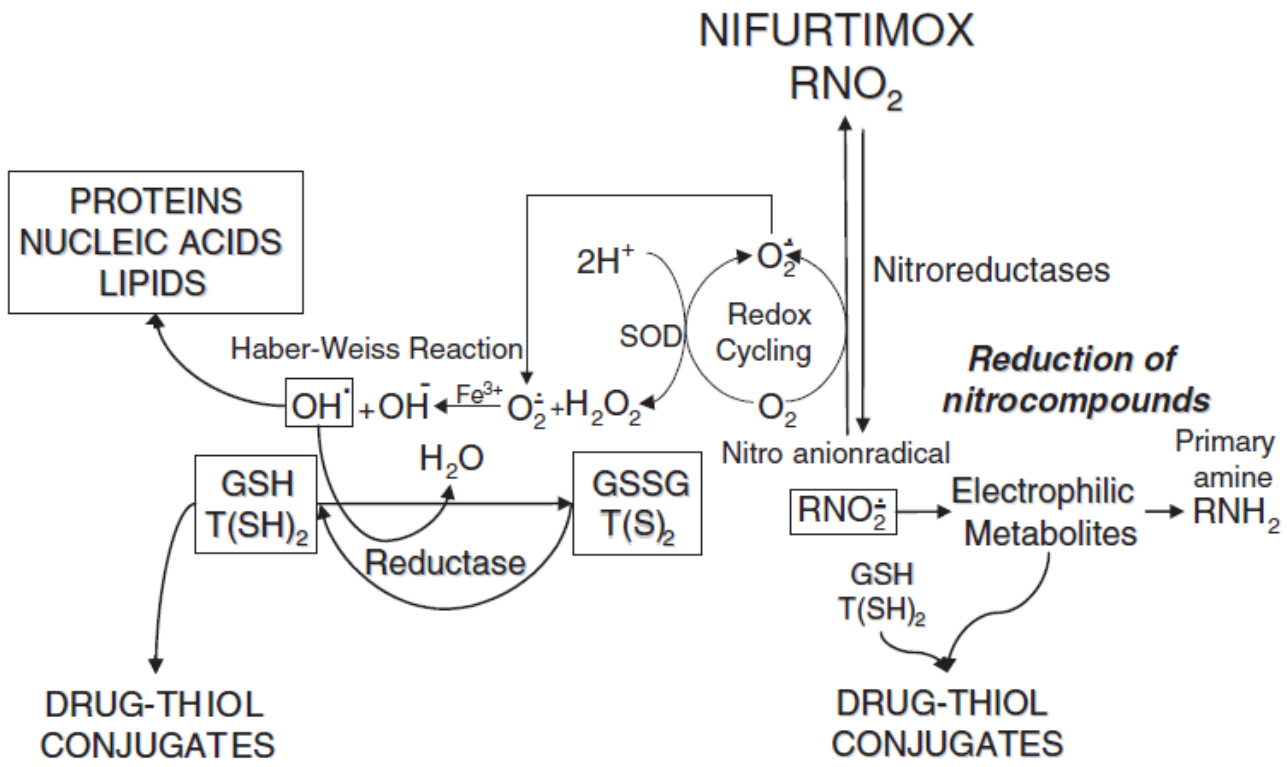
Fig. 14: Nifurtimox structure.

Uptake of Nifurtimox into *T. Cruzi* has been reported to occur via passive diffusion across the plasma membrane [76]. Studies have not been extended yet to T. Brucei, but it is also likely to enter these cells via passive diffusion. The one-electron reduction of the nitro group generates a potent free radical, which, in turn, generates reduced oxygen metabolites (such as superoxide, hydroxyl free radical, and hydrogen peroxide) believed to cause the death of the parasite [77]. However Nifurtimox is not yet fully validated for use in these vector borne diseases, therefore, new drugs are urgently needed.

1.4.4.2 Mechanism of Action of Nifurtimox

Nifurtimox acts through the formation of free radicals and/or electrophilic metabolites (Scheme 3). The nitro group is reduced to an amino group by the action of Nitroreductases, with the formation of various free radical intermediates and electrophilic metabolites. This process begins with a reaction catalyzed by NADPH-cytochrome P-450 reductase, which acts on the nitro group of R-NO_2-type molecule, producing an intermediary nitro anion radical (R-NO_2^-). This radical undergoes redox recycling with molecular oxygen, which partially reduces it and regenerates the drug. As seen in Scheme 2, the superoxide anion (O_2^-) undergoes superoxide dismutase-catalyzed dismutation to afford O_2 and H_2O_2 [78, 79].

The superoxide anion (O_2^-) and hydrogen peroxide (H_2O_2), in the presence of Fe^{3+}, form the hydroxyl free radical (Haber-Weiss reaction). These free radicals, mainly OH^U, bind to lipids, proteins, and DNA, damaging them [80].



Scheme 2: Role of Glutathione and Trypanothione in the action and metabolism of the antichagasic drug Nifurtimox. The nitro group of Nifurtimox is reduced to free radicals or electrophilic metabolites by *T. Cruzi* cytochrome P450-related nitroreductases. The Nifurtimox-derived free radicals may undergo redox cycling with oxygen and it is produced H_2O_2 by the further action of superoxide dismutase (SOD). The produced oxygen derived free radicals and electrophilic metabolites bind to intracellular macromolecules damaging them. In the parasite, Trypanothione ($T(SH)_2$) and glutathione (GSH) neutralize the Nifurtimox derived metabolites by conjugation producing drug-thiol conjugates that will be further metabolized to mercapturates in the mammal host. Free radicals are neutralized by oxidation of reduced GSH or $T(SH)_2$. Trypanothione reductase reduces oxidized Trypanothione ($T(S)_2$).

When Nifurtimox is added to *T. Cruzi* infected cells, a characteristic ESR spectrum corresponding to the nitro anion appears. Furthermore, the Nifurtimox concentration (10–20 μM) at which epimastigote culture is inhibited is similar to that required for maximum production of superoxide anion, and for the exit of hydrogen peroxide from the cell to begin [81, 82].

These and other experiments suggest that the intracellular reduction of Nifurtimox, generating the nitro radical, followed by redox cycling, and production of O_2^- and H_2O_2, is the main mechanism of action of Nifurtimox against *T. Cruzi* [83].

Finally, the family of the Trypanosomatidae continues to impose a burden of infectious diseases to the least equipped countries regarding sending new drugs to dispensaries. Regarding the diseases caused by this family of parasites (African Trypanosomiasis and Chagas disease) there is no vaccine, in addition, currently available drugs have insufficient efficiency and excessive toxicity [7].

Chapter 2

MATERIALS AND METHODS

2.1 Elements of the Theory

2.1.1 The Foundation of Theoretical Chemistry

The foundation of theoretical chemistry is the Schrödinger equation:

$$\hat{H}\,\psi = E\,\psi \quad \ldots\ldots\ldots\ldots\ldots (1)$$

Where $\hat{H}$ is the Hamiltonian operator, ψ is the wavefunction (eigenfunction for a given Hamiltonian) and E is the energy of the system. The wavefunction ψ describes the system and takes as variables the positions of electrons and nuclei in the system, leading to the following equation:

$$\hat{H}\,\psi_i\left(\overrightarrow{x_1},\ldots,\overrightarrow{x_N},\overrightarrow{R_1},\ldots,\overrightarrow{R_M}\right) = \hat{E}\,\psi_i\left(\overrightarrow{x_1},\ldots,\overrightarrow{x_N},\overrightarrow{R_1},\ldots,\overrightarrow{R_M}\right) \quad \ldots\ldots\ldots\ldots\ldots (2)$$

$\overrightarrow{x_N}$ describing the positions of the electrons N and $\overrightarrow{R_M}$ describing the positions of the nuclei M. A knowledge of ψ allows the properties of the system to be deduced.

The first step for solving the Schrödinger equation is to establish the form of the Hamiltonian operator.

$$H = T_e + T_n + V_{ne} + V_{ee} + V_{nn} \quad \ldots\ldots\ldots\ldots\ldots (3)$$

Where T_e and T_n are the kinetic energy terms for the electrons and the nuclei respectively; V_{ne} represents the attractive potential between electrons and nuclei, and V_{ee} and V_{nn} the inter-electronic and inter-nuclear repulsion potentials.

2.1.2 The Born-Oppenheimer Approximation

The wavefunctions ψ are functions of the position of both the nuclei and the electrons of the system. However, since a nucleus is much heavier than an electron (approximately 1900 times more), its movements compared to the electrons are negligible. In this case, they can be considered to be frozen and their kinetic energy set to zero but they still contribute to the potential energy of the system.

ψ is now only dependent on the kinetic energy of the electrons (T_e), the electron-nuclear attraction (V_{ne}) and the electron-electron repulsion (V_{ee}), so the Hamiltonian becomes:

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N}\frac{1}{r_{ij}} + V_{nn} \quad \ldots\ldots\ldots\ldots\ldots (4)$$

Where V_{nn} represents the nucleus-nucleus repulsion and is a constant. We see that this equation above is factorisable.

$$\hat{H}_{elec} = \sum_i\left(-\frac{1}{2}\nabla_i^2 - \sum_A\frac{Z_A}{r_{iA}} + \sum_{j>i}\frac{1}{r_{ij}}\right) \quad \ldots\ldots\ldots\ldots\ldots (5)$$

And the wavefunction depends now only on the electronic coordinates:

$$H_{elec}\,\psi_{i(elec)}\left(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_i,\vec{x}_j,\ldots,\vec{x}_N\right)=E_{elec}\,\psi_{i(elec)}\left(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_i,\vec{x}_j,\ldots,\vec{x}_N\right)$$

…… (6)

The electrons are also described by their spin quantum number. This spin can take two values, -1/2 or +1/2, which are defined by the alignment of the spin with respect to an arbitrary axis [84].

2.2 Molecular Modeling

2.2.1 Elements of Computational Chemistry

2.2.1.1 Chemical Drawing

A vast number of organic molecules are known. In order to distinguish one from another, chemists give them names. There are two kinds of names: trivial and systematic. Trivial names are often brand names (such as the amino acid phenylanine). Trivial names don't give any real clue as to the structure of a molecule, unless you are the recipient of divine inspiration. The IUPAC systematic name for phenylanine is 2-amino-3-phenyl-propionic acid. Any professional scientist with a training in chemistry would be able to translate the systematic name into Figure 15.

There are various conventions that we can follow when drawing chemical structures, but the conventions are well understood amongst professionals. First of all, we haven't shown the hydrogen atoms attached to the benzene ring (or indeed the carbon atoms within), and we have taken for granted that we understand that the normal valence of carbon is four. Everyone understands that hydrogens are present, and so we needn't clutter up an already complicated drawing.

In the Figure 15 the benzene ring as alternate single and double bonds, yet we understand that the C—C bonds in benzene are all the same. This may not be the case in the molecule shown; some of the bonds may well have more double bond character than others and so have different lengths, but once again it is a well-understood convention. Sometimes a benzene ring is given its own symbol Ph or Φ. Then again, we drawn the NH_2 and the OH groups as 'composites' rather than showing the individual O–H and N–H bonds, and so on. We followed to some extent the convention that all atoms are carbon atoms unless otherwise stated.

Much of this is personal preference, but the important point is that no one with a professional qualification in chemistry would mistake this drawing for another molecule. Equally, given the systematic name, no one could possibly write down an incorrect molecule.

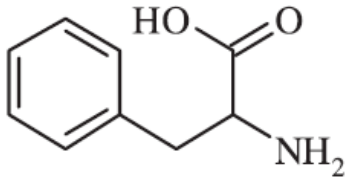


Fig. 15: 2D drawing of phenylanine.

The aim of this chapter is to show how chemistry is a well-structured science, with a vast literature. There are a number of important databases that contain information about syntheses, crystal structures, physical properties and so on. Many databases use a molecular structure drawing as the key to their information, rather than the systematic name. Structure drawing is therefore a key chemical skill.

2.2.1.2 Three Dimensional Effects

Chemical drawings are inherently two-dimensional objects; they give information about what is bonded to what. Occasionally, the lengths of the lines joining the atoms are scaled to reflect accurate bond lengths.

Molecules are three-dimensional entities, and that is where the fun and excitement of chemistry begins. In order to indicate three-dimensional effects on a piece of paper, we might draw the molecule CBrClFH (which is roughly tetrahedral) as in Figure 16. The top left-hand drawing is a two-dimensional one, with no great attempt to show the arrangement of atoms in space. The next three versions of the same molecule show the use of ‘up’, ‘down’ and ‘either’ arrows to show the relative dispositions of the bonds in space more explicitly. The bottom two drawings are two dimensional attempts at the three-dimensional structure of the molecule and its mirror image. Note that the molecule cannot be superimposed on its mirror image. The central carbon atom is a chiral center, and the two structures are called enantiomers.

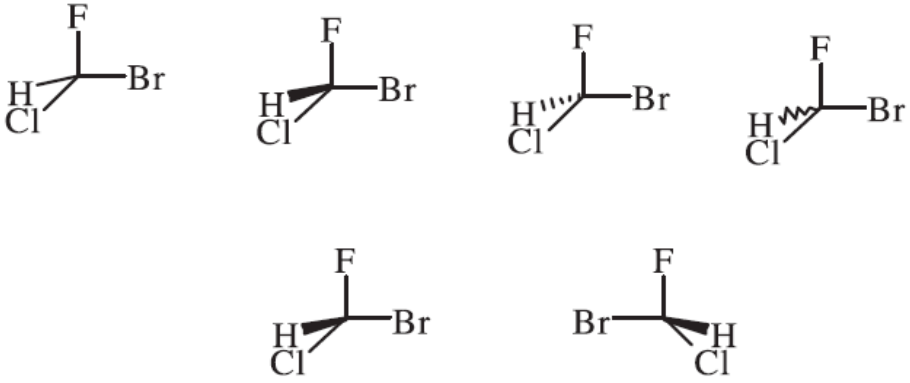


Fig. 16: 2D drawings.

2.2.1.3 Modeling

The title of this and many other texts includes the word ‘modelling’, which begs the question as to the meaning of the word ‘model’.

Model, a preliminary solid representation, generally small, or in plastic material, to be followed in construction: something to be copied: a pattern: an imitation of something on a smaller scale: a person or thing closely resembling another...

This definition captures the status of modelling in the 1970s, and Figure 17 shows a photograph of a plastic model of L-phenylanine. Such plastic models were fine in their day, but they took a considerable time to build, they tended to be unstable and, more importantly, they had to know the molecular structure before they could actually build one. Not only that, they gave no sense of temperature in that they didn’t vibrate or show the possibility of internal rotation about single bonds, and they referred only to isolated molecules [85].

Fig. 17: Plastic model of L-phenylanine.

As you know, a given molecule may well have many stable conformations. Plastic models gave clues as to which conformations were unlikely on the grounds of steric repulsions, but mostly they did not help to identify the ‘true’ molecular geometry.

We have come a long way since then. Computer simulation has long taken over from mechanical model building, which helps to predict the molecular geometry, to simulate the temperature, to allow for solvent effects, etc.

2.2.1.4 Molecular Structure Databases

Molecular geometries can be determined for gas-phase molecules by microwave spectroscopy and by electron diffraction. In the solid state, the field of structure determination is dominated by X-ray and neutron diffraction and many crystal structures are known. Also, nuclear magnetic resonance (NMR) also has a role to play, especially for proteins.

Over the years, a vast number of molecular structures have been determined and there are several well-known structural databases. One is the Cambridge Structural Database (CSD) (http://ccdc.cam.ac.uk=), which is supported by the Cambridge Crystallographic Data Centre (CCDC). The CCDC was established in 1965 to undertake the compilation of a computerized

database containing comprehensive data for organic and metal–organic compounds studied by X-ray and neutron diffraction. At the time of writing, there are some 272 000 structures in the database [86].

For each entry in the CSD, three types of information are stored. First, the bibliographic information: who reported the crystal structure, where they reported it and so on. Next comes the connectivity data; this is a list showing which atom is bonded to which in the molecule. Finally, the molecular geometry and the crystal structure. The molecular geometry consists of cartesian coordinates. The database can be easily reached through the Internet, but individual records can only be accessed on a fee-paying basis.

The Protein Data Bank (PDB) is the single worldwide repository for the processing and distribution of three-dimensional biological macromolecular structural data. It is operated by the Research Collaboratory for Structural Bioinformatics [87].

At the time of writing, there were 19 749 structures in the databank, relating to proteins, nucleic acids, protein–nucleic acid complexes and viruses. The databank is available free of charge to anyone who can navigate to their site http://www.rcsb.org/. Information can be retrieved from the main website. A four-character alphanumeric identifier, such as 1PCN, represents each structure. The PDB database can be searched using a number of techniques, all of which are described in detail at the homepage.

2.2.1.5 File Formats

The PDB (.pdb) file format is widely used to report and distribute molecular structure data. A typical .pdb file for phenylanine would start with bibliographic data, then move on to the cartesian coordinates (expressed in angstroms and relative to an arbitrary reference frame) and connectivity data as shown below (Fig.18). The only parts that need concern us are the atom numbers and symbols, the geometry and the connectivity.

```
HETATM    1      N     PHE    1     -0.177     1.144     0.013
HETATM    2      H     PHE    1      0.820     1.162    -0.078
HETATM    3      CA    PHE    1     -0.618     1.924     1.149
HETATM    4      HA    PHE    1     -1.742     1.814     1.211
HETATM    5      C     PHE    1     -0.290     3.407     0.988
HETATM    6      O     PHE    1      0.802     3.927     0.741
HETATM    7      CB    PHE    1      0.019     1.429     2.459
HETATM    8      1HB   PHE    1      0.025     0.302     2.442
HETATM    9      2HB   PHE    1      1.092     1.769     2.487
HETATM   10      CG    PHE    1     -0.656     1.857     3.714
HETATM   11      CD1   PHE    1     -0.068     1.448     4.923
HETATM   12      HD1   PHE    1      0.860     0.857     4.900
HETATM   13      CD2   PHE    1     -1.829     2.615     3.757
HETATM   14      HD2   PHE    1     -2.301     2.975     2.829
HETATM   15      CE1   PHE    1     -0.647     1.783     6.142
HETATM   16      HE1   PHE    1     -0.176     1.457     7.081
HETATM   17      CE2   PHE    1     -2.411     2.946     4.982
HETATM   18      HE2   PHE    1     -3.338     3.538     4.999
HETATM   19      CZ    PHE    1     -1.826     2.531     6.175
HETATM   20      HZ    PHE    1     -2.287     2.792     7.139
HETATM   21      O     PHE    1     -1.363     4.237     1.089
HETATM   22      H           22     -0.601     1.472    -0.831
HETATM   23      H           23     -1.077     5.160     0.993
CONECT    1      2     3    22
CONECT    2      1
CONECT    3      1     4     5     7
CONECT    4      3
CONECT    5      3     6    21
CONECT    6      5
CONECT    7      3    10     8     9
CONECT    8      7
CONECT    9      7
CONECT   10      7    11    13
CONECT   11     10    15    12
CONECT   12     11
CONECT   13     10    17    14
CONECT   14     13
CONECT   15     11    19    16
CONECT   16     15
CONECT   17     13    19    18
CONECT   18     17
CONECT   19     15    17    20
CONECT   20     19
CONECT   21      5    23
CONECT   22      1
CONECT   23     21
END
```

Fig. 18: Cartesian coordinates and connectivity data of Phenylanine.

Records 1 - 23 identify the atoms so for example, atom 1 is a Nitrogen with Cartesian coordinates x = - 0.177Å, y = 1.144 Å and z = 0.013 Å. The PHE identifies the aminoacid residue, of which there is just one in this simple case. The record CONECT 1 2 3 22 tells us that atom 1 (N) is joined to atoms 2, 3 and 22 [87].

2.2.1.6 Three Dimensional Display

As we progress through the time, there are several suitable packages with different ways to represent a molecular model; it is what referred as the Rendering. First, we have the line representation of phenylanine, Figure 19 line models are drawn using lines. Each representation has a number of options. In this case, we set the option to perceive aromaticity so the benzene ring is shown as an aromatic entity rather than single and double bonds. The molecule can be colored, rotated and otherwise manipulated.

There are other representations, based on space filling. For example, the CPK (Corey–Pauling–Koltun) rendering shown in Figure 19 refers to a popular set of atomic radii used for depicting space-filling pictures.

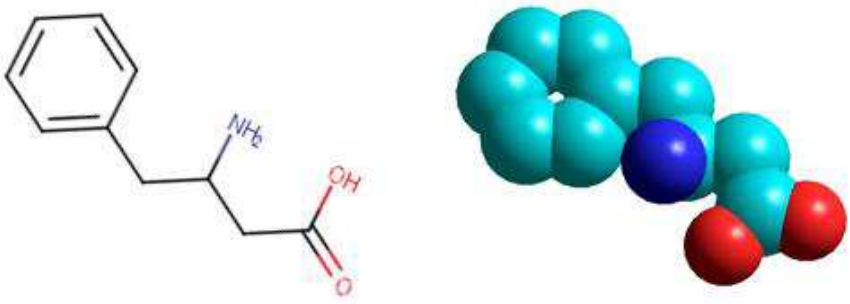


Fig. 19: Line and CPK space-filling rendering of Phenylanine.

2.2.1.7 Computer packages

Over the years, several chemical computer packages have appeared in the marketplace. They are all very professional and all perform much the same function. Which one you choose is a matter of personal preference; in the present work, we are going to use MarvinSketch 6.2.1 [88] and Hyperchem 8.08 [89] for our chemical drawing illustrations. Gaussian 09 [90] for chemical properties calculations, IBM SPSS [91] for Chemometrics utilities and Molegro Virtual Docker [92] as a tool for chemical interactions and illustrations.

2.2.2 Molecular mechanics

2.2.2.1 Introduction

Molecular modellers usually have a quite different objective; they want a force field that can be transferred from molecule to molecule, in order to predict (for example the geometry of a new molecule) by using data derived from other related molecules. They make use of the bond concept, and appeal to traditional chemists' ideas that a molecule comprises a sum of bonded atoms; a large molecule consists of the same features we know about in small molecules, but combined in different ways [85].

Molecular mechanics describes molecules in terms of "bonded atoms", which have been distorted from some idealized geometry due to non-bonded van der Waals and Coulombic interactions. The success of molecular mechanics models depends on a high degree of transferability of geometrical parameters from one molecule to another, as well as predictable dependence of the parameters on atomic hybridization.

For example, carbon-carbon single bond lengths generally fall in the small range from 1.45 to 1.55Å, and increase in length with increasing "p character" of the carbon hybrids. Thus, it is possible to provide a fairly accurate "guess" at molecular geometry in terms of bond lengths, bond angles and torsion angles, provided that the molecule has already been represented in terms of a particular valence structure. The majority of organic molecules fall into this category.

The molecular mechanics "energy" of a molecule is described in terms of a sum of contributions arising from distortions from "ideal" bond distances (stretch contributions), bond angles (bond contributions) and torsion angles (torsion contributions), together with contributions due to "non-bonded" (van der Waals and Coulombic) interactions. It is commonly referred to as a "strain energy", meaning that it reflects the "strain" inherent to a "real" molecule relative to some idealized form [93].

$$E^{strain} = \sum_A^{bonds} E_A^{stretch} + \sum_A^{bond\ angles} E_A^{bend} + \sum_A^{torsion\ angles} E_A^{torsion} + \sum_A^{non\ bonded\ atoms} \sum_B E_{AB}^{non-bonded} \quad \text{..... (7)}$$

2.2.2.2 Force Fields

Molecular mechanics models differ both in the number and specific nature of the terms which they incorporate, as well as in the details of their parameterization. Taken together, functional form and parameterization, constitute what is termed a force field.

We have been vague so far about which variables are the 'correct' ones to take. Chemists visualize molecules in terms of bond lengths, bond angles and dihedral angles, yet this information is also contained in the set of Cartesian coordinates for the constituent atoms. Both are therefore 'correct'; it is largely a matter of personal choice and professional training.

Spectroscopists have developed systematic simplifications to the force fields in order to make as many as possible of the small terms vanish. If the force field contains only 'chemical' terms such as bond lengths, bond angles and dihedral angles, then it is referred to as a valence force field (VFF). There are other types of force fields in the literature, intermediate between the VFF and the general force field [92].

With these principles in mind, many different force fields were build in professional molecular modelling programs such as : Dreiding, MM1, MM2, Amber, OPLS…etc.

2.2.2.3 Limitations of Molecular Mechanics

The primary advantage of molecular mechanics models, is their simplicity. Except for very small systems, computation cost is completely dominated by evaluation of non-bonded van der Waals and Coulombic terms, the number of which is given by the square of the number of atoms.

However, the magnitude of these terms falls off rapidly with increasing interatomic distance and, in practice, computation cost scales linearly with molecular size for sufficiently large molecules.

Molecular mechanics calculations may easily be performed on molecules comprising several thousand atoms. Additionally, molecular mechanics calculations are sufficiently rapid to permit extensive conformational searching on molecules containing upwards of 100-200 atoms.

The fact that molecular mechanics models are parameterized may also be seen as providing an advantage over quantum chemical models. It is possible, at least in principle, to construct molecular mechanics models which will accurately reproduce known experimental data, and hopefully will anticipate (unknown) data on closely-related systems.

There are important limitations of molecular mechanics models. First, they are limited to the description of equilibrium geometries and equilibrium conformations. Because the mechanics "strain energy" is specific to a given molecule (as a measure of how far this molecule deviates from its "ideal arrangement"). Two important exceptions are: calculations involving isomers with exactly the same bonding, e.g., comparison of cis and trans-2-butene, and conformational energy comparisons, where different conformers necessarily have exactly the same bonding.

In addition, molecular mechanics calculations reveal nothing about bonding or, more generally, about electron distributions in molecules. As will become evident later, information about electron distributions is key to modeling chemical reactivity and selectivity. There are, however, important situations where purely steric effects are responsible for trends in reactivity and selectivity, and here molecular mechanics would be expected to be of some value.

Finally, it needs to be noted that molecular mechanics is essentially an interpolation scheme, the success of which depends not only on good parameters, but also on systematics among related molecules. Molecular mechanics models would not be expected to be highly successful in describing the structures and conformations of "new" (unfamiliar) molecules outside the range of parameterization [93].

2.2.3 Quantum Mechanics

2.2.3.1 Introduction

There are two approaches to quantum mechanics. One is to follow the historical development of the theory from the first indications that the whole fabric of classical mechanics and electrodynamics should be held in doubt to the resolution of the problem in the work of Planck, Einstein, Heisenberg, Schrödinger, and Dirac. The other is to stand back at a point late in the development of the theory and to see its underlying theoretical structure [94].

Quantum mechanics can be thought of roughly as the study of physics on very small length scales, although there are also certain macroscopic systems it directly applies to. In quantum mechanics, particles have wavelike properties, and a particular wave equation, the Schrodinger equation, governs how these waves behave.

What are we looking for in quantum mechanics? we search to solve the equation of Schrödinger of an electron using Born-Oppenheimer approximations. We can use the Born-Oppenheimer approximation to construct an electronic Hamiltonian, which neglects the kinetic

energy term of the nuclei since the weight of a typical nucleus is thousands of times greater than that of an electron.

$$\widehat{H} = -\frac{\hbar}{2m_e}\nabla_i^2 + \sum_I \sum_{J<1} \frac{Z_I Z_J e^2}{r_{IJ}} - \sum_I \sum_j \frac{Z_I e^2}{r_{iI}} + \sum_J \sum_{i>j} \frac{e^2}{r_{ij}} \quad \ldots\ldots\ldots\ldots\ldots (8)$$

This Hamiltonian is used in the Schrödinger equation describing the motion of the electrons in the field of the fixed nuclei:

$$\widehat{H}_{ele}\,\Psi_{ele} = E_{eff}\,\Psi_{ele} \quad \ldots\ldots\ldots\ldots\ldots (9)$$

Solving this equation for the electronic wave function will produce the effective nuclear potential function E_{eff} that depends on the nuclear coordinates and describes the potential energy surface of the system.

For bond electronic problem, Ψ should satisfy two requirements: antisymmetricity and normalization. Ψ should change sign when two electrons of the molecule interchange and the integral of Ψ over all space should be equal to the number of electrons of the molecule [95].

2.2.3.2 HF and DFT Methods

Many aspects of molecular structure and dynamics can be modeled using classical methods in the form of molecular mechanics and dynamics. The classical force field is based on empirical results, averaged over a large number of molecules. Because of this extensive averaging, the results can be good for standard systems, but there are many important questions in chemistry that can not at all be addressed by means of this empirical approach. If one wants to know more than just structure or other properties that are derived only from the potential energy surface, in particular properties that depend directly on the electron density distribution, one has to resort to a more fundamental and general approach: quantum chemistry. The same holds for all non-standard cases for which molecular mechanics is simply not applicable [96].

In these methods, we use a set of basic functions from which we define all the molecular orbitals of the chemical system studied by LCAO approximation (linear combination of atomic orbitals).

The choice of basic orbital is essential for quantum chemistry calculations. There are two types of basic functions mainly used in electronic structure calculations: the Orbital Type Slater (STO) and those of Gaussian type (GTO).

The question here is, what basis to choose? Is it the biggest possible?

The answer is: all depend of your studied system.

With the increase in computing power, the minimum size of a calculation base is currently around 6-31G (d)

- Double zeta at least.
- With polarization at least on the heavy atoms.
- Diffuse functions in the case of an anion or a system performing hydrogen bonds or a system having free doublets.
- A good basis: 6-31G** (or 6-31++G**).
- A very good basis: 6-311 ++ G (2df, 2p).

2.2.3.3 Limitations of Quantum Mechanics

The advantages of Quantum Mechanics is that it takes in consideration the electronic systems (wave function of an electron for the hartree-fock methods HF and electronic density for the method of density functional theory DFT), also it allows to study the chemical reactivity and many physicochemical properties are accessible such as spectral, dipole moments, etc..

However, these methods drawback calculations could be sometimes very long, the size of the system studied is limited and there must be a very good calculation equipment.

2.2.4 Electronic Charges and their Properties

As far as we can tell, there are four fundamental types of interactions between physical objects. There is the *weak nuclear interaction* that governs the decay of beta particles, and the *strong nuclear interaction* that is responsible for binding together the particles in a nucleus. The familiar *gravitational interaction* holds the earth very firmly in its orbit round the sun, and finally we know that there is an *electromagnetic interaction* that is responsible for binding atomic electrons to nuclei and for holding atoms together when they combine to form molecules.

2.2.4.1 Point Charges

It turns out that there are two types of electric charge in nature, which we might choose to call type X and type Y. Experimental evidence shows the existence of an electrostatic force between electric charges, the force between two X-type charges is always repulsive, as is the force between two Y-type charges. The force between an X and a Y-type is always attractive, for this reason, the early experimenters decided to classify charges as positive or negative. Therefore, the best known fundamental particles responsible for these charges are electrons and protons.

The term ‘point charge’ is a mathematical abstraction; obviously, electrons and protons have a finite size. Therefore, a point charge is one whose dimensions are small compared with the distance between them.

The concept of a point charge may strike you as an odd one, but once we have established the magnitude of the force between two such charges, we can deduce the force between any arbitrary charge distributions on the grounds that they are composed of a large number of point charges [94].

2.2.4.2 Charge Distribution

In general, charge distribution is the distribution of electrons of an atom or molecule in atomic or molecular orbitals. For many purposes we may regard the structure of a molecule as known when we can specify the equilibrium configuration of its nuclei. It is this structure, or a fair approximation to it, which is deducible from, for example, an NMR experiment, microwave spectroscopy, or neutron diffraction, etc.

Many molecular properties can be interpreted only if we also know in some detail the distribution of the electrons. For example, to understand why a particular configuration of the nuclei yields and equilibrium geometry requires a fairly minute description of the electron density distribution in the molecule. Moreover, any prediction of the interactions of a molecule with its surroundings depends in great measure on a knowledge of the total charge distribution [97].

The understanding of the total charge distribution in molecules allows the determination of many different properties such as: The frontier molecular orbitals HOMO and LUMO, Dipole moment, Electrostatic potential, the preferential sites of nucleophilic or electrophilic attacks, etc.

2.2.5 Molecular Modeling Calculations

2.2.5.1 Molecular Geometry Optimization

Geometry optimization is usually the first step in calculation, it is the process of finding the coordinates that minimize the energy. Within the Born-Oppenheimer approximation, the geometry of a molecule at zero absolute temperature corresponds to the minimum of the total energy:

$$\{x_i, y_i, z_i\}_{T=0} = arg \min_{\{x_i, y_i, z_i\}} ([\Psi|\hat{H}|\Psi]) \quad \ldots\ldots\ldots\ldots\ldots (10)$$

We usually look for zero gradient (a gradient is the steepest ascent direction), but not all such points are minima [98].

2.2.5.2 Conformation Search

As discussed above, the geometry optimization methods start with an initial geometry and then change that geometry to find a lower-energy shape. This usually results in finding a local minimum of the energy as depicted in Figure 20. This local minimum corresponds to the conformer that is closest to the starting geometry. In order to find the most stable conformer (a global minimum of the energy), some type of algorithm must be used and then many different geometries tried to find the lowest-energy one.

Conformation search algorithms are an automated means for generating many different conformers and then comparing them based on their relative energies. Due to the immensely large number of possible conformers of a large molecule, it is desirable to do this with a minimum amount of CPU time. Quite often, all bond lengths are held fixed in the course of the search, which is a very reasonable approximation. Frequently, bond angles are held fixed also, which is a fairly reasonable approximation [99].

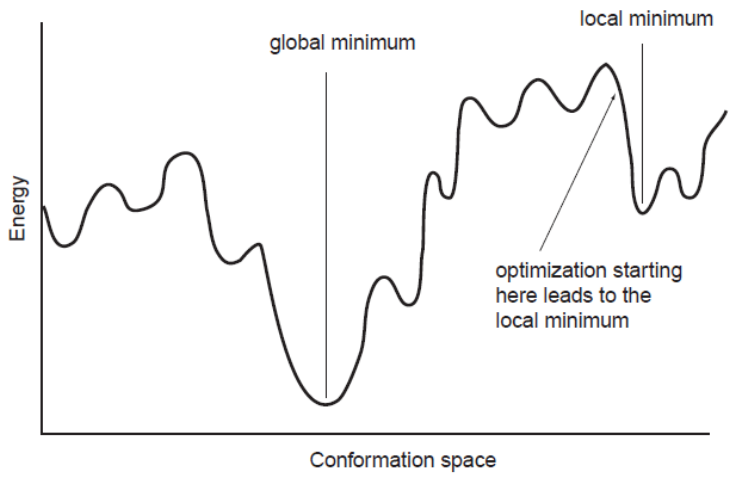


Fig. 20: Example of all possible conformers' energy of a simple molecule.

2.2.5.3 Geometric and Electronic Parameters

2.2.5.3.1 Geometric Parameters

A- Bond length

Is the distance between atomic centers involved in a chemical bond. The notion of bond length is defined differently in various experimental methods of determination of molecular geometry; this leads to small (usually 0.01 – 0.02 Å) differences in bond lengths obtained by different techniques. For example, in gas-phase electron-diffraction experiments, the bond length is the interatomic distance averaged over all occupied vibrational states at a given temperature. In an X-ray crystal structural method, the bond length is associated with the distance between the centroids of electron densities around the nuclei. In gas-phase microwave spectroscopy, the bond length is an effective interatomic distance derived from measurements on a number of isotopic molecules, etc [65].

B- Valence angle

The shape of a molecule is determined by its Valence angles, the angles made by the lines joining the nuclei of the atoms in the molecule. The bond angles of a molecule, together with the bond lengths, define the shape and size of the molecule. In Figure 21, we can see that there are six bond angles Cl-C-Cl in CCl_4, all of which have the same value. That bond angle, 109.5°, is characteristic of a tetrahedron. In addition, all four C-Cl bonds are of the same length (1.78 Å). Thus, the shape and size of CCl_4 are completely described by stating that the molecule is tetrahedral with C-Cl bonds of length 1.78 Å [65].

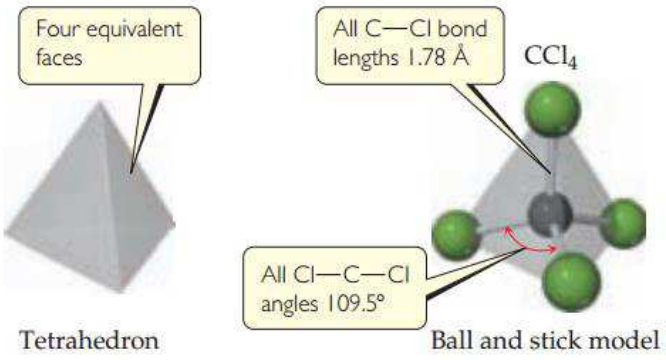


Fig. 21: Tetrahedral shape of CCl_4.

C- Dihedral Angle (Torsion angle)

In a chain of atoms A-B-C-D, the dihedral angle between the plane containing the atoms A,B,C and that containing B,C,D. In a Newman projection (Fig.22) the torsion angle is the angle (having an absolute value between 0° and 180°) between bonds to two specified (fiducial) groups, one from the atom nearer (proximal) to the observer and the other from the further (distal) atom. The torsion angle between groups A and D is then considered to be positive if the bond A-B is rotated in a clockwise direction through less than 180° in order that it may eclipse the bond C-D: a negative torsion angle requires rotation in the opposite sense [65].

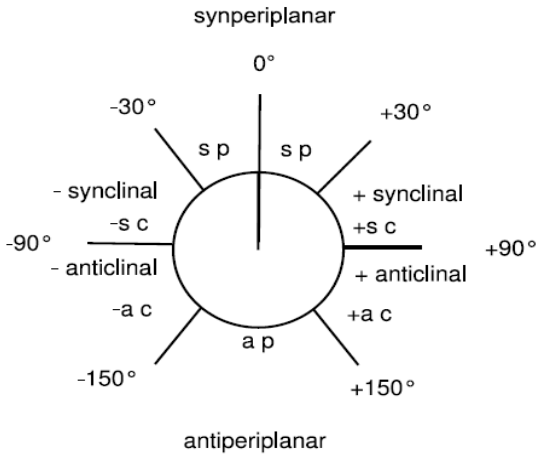


Fig. 22: Newman projection.

2.2.5.3.2 Electronic Parameters

A- MESP

The molecular electrostatic potential (MESP) surface which is a plot of electrostatic potential mapped onto the iso-electron density surface [100], the importance of the MESP lies in the fact that it simultaneously displays the molecular size and shape as well as positive, negative and neutral electrostatic potential regions in terms of the electrostatic surface, which explain the investigation of the molecular structure with its physiochemical property relationships [101, 102].

B- Atomic Charges

A main problem in comparing different point-charge models is the missing of clear criterion for the quality of the charges. This is probably the reason why so many charge models have been suggested such as:

- CHelp G: Produce charges fit to the electrostatic potential at points selected according to the CHelp G scheme.
- NBO: Requests a full Natural Bond Orbital analysis.
- Mulliken: retains only the density terms involving pairs of basis functions on different centers.
- HLY: Hu, Lu, and Yang charge fitting method.

Furthermore, different applications put different demands on the charges. For example, in molecular dynamics, molecules move, so the charges must be able to describe the electrostatics properly in all accessible points in the phase space, and they should also be invariant to changes in the internal coordinates of the molecule [103-105].

C- Heat of Formation

The heat of Formation is known as the change in enthalpy accompanying the formation of one mole of a compound from its elements in their natural and stable states, under standard conditions of one atmosphere at a given temperature [106].

This is equal to the energy required to ionize the valence electrons of the atoms involved. The heat of formation is defined as:

$$\Delta H_f = E_{elec} + E_{nuc} - E_{iso} + E_{atom} \quad \ldots\ldots\ldots\ldots\ldots (11)$$

Where E_{elec} is the the electronic energy, E_{nuc} is the nuclear-nuclear repulsion energy, E_{iso} is the energy required to strip all the valence electrons of all the atoms in the system and E_{atom} is the total heat of atomization of all the atoms in the system [106].

D- Molecular Orbitals (Homo-1, HOMO, LUMO, LUMO+1, DeltaE)

The most important orbitals in a molecule are the frontier molecular orbitals, called highest occupied molecular orbital (HOMO) which explains the ionization energy and lowest unoccupied molecular orbital (LUMO) which explains the electron affinity in a molecule.

These orbitals determine the way the molecule interacts with other species including the interactions between a ligand and its receptor [107]. Whereas, electrophilic and nucleophilic attack will most likely occur at atoms where the coefficients of the corresponding atomic orbitals in HOMO and LUMO, respectively, are large [108].

Sometimes, even, the first excited state is degenerate the terms HOMO and LUMO can not be give clear distinctions. Therefore, we can also use the second highest occupied and second

lowest unoccupied molecular orbitals (HOMO -1 and LUMO +1, respectively), whereas, these orbitals are lower in energy than HOMO and LUMO [11].

E- Dipole moment

Electric dipole consists of a pair of charges of equal magnitude and opposite signs (+ q and - q), separated by a distance (r). The dipole moment of an electric dipole is a vector directed from the negative to the positive charge. The magnitude of the dipole moment is measured in Debye (D) [108].

If the positive and negative charges in a molecule do not overlap, the molecule possesses a permanent dipole moment (μ) (polar molecule). Molecular dipole moment is usually calculated using the following formula:

$$\mu = \sum_{i=1}^{n} q_i \times r_i \quad \ldots\ldots\ldots\ldots\ldots (12)$$

Where q is the partial charge of atom i and r is the radius-vector of an atom i from the origin of the coordinate system (Centre of charge or Centre of the weight).

2.2.5.4 Physicochemical Properties

A- Surface Area Grid

The molecular surface (SAG) with a grid calculation (solvent-accessible or Van der Waals) much slower than the approximate-surface calculation, but it is more accurate for a given set of atomic radii. It is recommended that this method can be used as a benchmark for the approximate surface area calculation. The grid method used for SAG calculation is that described by Bodor et al [109].

B- Molecular Volume

The Molecular volume (V) is the volume occupied by a body; it is a specific volume (at normal temperature and pressure) and its calculation is very similar to that of the surface (Grid) it employs the same grid method described by Bodor et al.

The volume is defined by the following relationship [110]:

$$V = \frac{MW}{d} \quad \ldots\ldots\ldots\ldots\ldots (13)$$

Where MW is the molecular weight and d is the density.

C- Hydration Energy

Hydration is the formation of a solution involves the interaction of the solute with the solvent molecules, different liquids can be used as solvents, but water is the solvent the most

commonly used. When water is used as solvent, the dissolution process is called hydration [111]. The calculation of the hydration energy (HE) is based on the exposed surface which depends on the type of molecular groups atoms which can be hydrogen bond donors such as OH, NH, PH.. or acceptor groups such as O, N, P..

In the organism, polar molecules are not surrounded by water molecules, which allow them to form hydrogen bonds between them; obviously, the donor sites interact with the oxygen atom of the water and the acceptor sites with hydrogens [112]. The hydration energy (HE) is a key factor determining the stability of different molecular conformations [109].

D- Log P

Lipophilicity is an important property that has a significant impact on the solubility, absorption, distribution, metabolism and excretion of drugs. Hansch and Leo estimated that molecules that carry a high lipophilicity will be retained within the membrane lipids [113].

The best method to estimate the ability of a compound to dissolve in aqueous and in organic environements is the measure of the lipophilicity.

$$P = \frac{[Solute]_{Octanol}}{[Solute]_{water}} \quad \ldots\ldots\ldots\ldots\ldots (14)$$

The partition coefficient P, defined as the ratio of molar concentration of a chemical dissolved at equilibrium in octanol phase $[Solute]_{octanol}$ to its molar concentration in aqueous phase $[Solute]_{water}$ [114-116].

E- Molar Refractivity

The molar refractivity (MR) is a steric parameter that is dependent on the spatial array of the aromatic ring in the synthesized compounds [117]. The refractivity is a special case of the molecular volume, it is a refractivity reduced to a quantity of material, as well as the refractivity is the quality of the refractive power of a body, it is used in the radio-electricity, biology, etc. [118, 119].

In chemistry, the molecular refractivity is an important criterion to measure the steric factor, it is important in the case where the substitute has π electrons or lone pairs [113].

This method is parameterized for organic systems with a variety of heteroatoms. The molar refractivity is defined by the following relationship:

$$MR = V\,\frac{n^2-1}{n^2+2} \quad \ldots\ldots\ldots\ldots\ldots (15)$$

Where V is the molecular volume and n is the refraction index.

F- Polarizability

We call the polarizability (Pol) of a molecule, the ease with which the electron cloud is deformed by an external electric field. The polarizability of a molecule can improve its aqueous solubility. This feature plays a very important role in the modeling of many molecular properties [113].

Highly polarizable molecules can be expected to have strong attractions with other molecules. The polarizability is defined by the following relationship:

$$P(e) = \varepsilon_0 \, \alpha \, E \quad \ldots\ldots\ldots\ldots\ldots (16)$$

Where is P(e) is Polarizability coefficient, ε_o is the dielectric constant and α is the induced electric dipole moment.

G- Molecular Weight

The molecular weight (MW) of a system calculation is based on a general applicability method [89]. The molecular weight is an important criterion in the passive diffusion of drugs through the tightly packed aliphatic side chains of the bilayer membrane. It is related to the size of the molecule, as molecular size increases, a larger cavity must be formed in water in order to solubilize the compound [97].

H- Hydrogen bonding

Hydrogen bonding is another very important binding force, which called the electrostatic attraction between two polar groups that occurs when a hydrogen (H) atom covalently bound to a highly electronegative atom such as nitrogen (N) or oxygen (O) experiences the electrostatic field of another highly electronegative atom nearby.

One of the most common examples of hydrogen bonds are those formed in liquid alcohols. Most OH groups make a hydrogen bond to an Oxygen of an adjacent alcohol, thereby creating a network of hydrogen bonds.

In order for a hydrogen bond to occur there must be both a hydrogen donor and an acceptor present. The donor in a hydrogen bond (HBD) is the atom to which the hydrogen atom participating in the hydrogen bond is covalently bonded, and is usually a strongly electronegative atom such as N,O, or F.

The hydrogen acceptor (HBA) is the neighboring electronegative ion or molecule, and must possess a lone electron pair in order to form a hydrogen bond as shown in Figure 23.

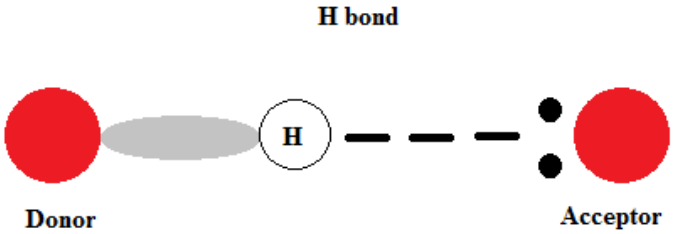


Fig. 23: Hydrogen bond.

I- Polar Surface Area

The polar surface area (PSA) defined as the sum of surfaces of polar atoms in a molecule. This parameter is easy to understand and, most importantly, provides good correlation with passive molecular transport through membranes, and so allows estimation of transport properties of drugs. It has been successfully applied for the prediction of intestinal absorption [120, 121].

J- Number of Rotatable Bond

The number of rotatable bonds (NRB) is linked to the microsomal clearance; compounds that have a large NRB (high flexibility) are most likely to adopt a conformation recognized by the cytochrome P450.

Bonds that are not rotatable are [122]:

- Non single bonds & Bonds connected to terminal triple bonds, including bonds to cyano groups.
- Ring bonds & Bonds connecting two aromatic rings with collectively three or more ortho substituents.
- Bonds to hydrogen and other monovalents atoms (halogens).
- Bond to terminal atoms, including CH_3, NH_2 and OH.
- The C-N bond between a carbonyl and amide nitrogen (also goes for the C-N bond in thioamides and the S-N bond in sulfonamides).

K- Ligand Efficiency

A recent trend has been to combine potency with other parameters into a single metric which may be monitored during optimization. One commonly used example is the ligand efficiency (LE), which is defined as :

$$LE = 1.4\, pIC_{50} \,/\, N_H \quad \ldots\ldots\ldots\ldots\ldots (17)$$

Where pIC_{50} = -log (IC_{50}), the IC_{50} is expressed in molar concentration and N_H is the number of heavy atoms [123].

The Ligand Efficiency (LE) is a useful tool used to penalize large compounds over small compounds with similar potency because larger compounds tend to have poorer physicochemical and ADME properties.

L- Lipophilic Efficiency

Lipophilic Efficiency used to maximize potency while maintaining as low as possible the lipophilicity [124, 125], due to the association between high lipophilicity and several issues including poor solubility, membrane permeation, metabolic stability, etc. along with an increased risk of non-specific interactions and toxicity [126].

The Lipophilic Efficiency (LipE) is defined as:

$$LipE = pIC_{50} - \log\, P \quad \ldots\ldots\ldots\ldots\ldots (18)$$

2.3 QSAR Modeling

Quantitative structure-activity relationship (QSAR) is among the most practical tool in computational chemistry which has done much to enhance our understanding of fundamental processes and phenomena in medicinal chemistry and drug design [91].

The fundamental idea of QSAR consists of the possibility of relationships between molecular structure or properties [127, 128] derived from molecular structure and a molecular response. QSAR can be regarded as a computer-derived rule that quantitatively describes the biological activity in terms of chemical descriptors; it has been frequently used to predict biological activities of new compounds [129-132].

2.3.1 Biological Data

Biological parameters are measurable and quantifiable (Specific Enzyme or Hormone concentration, phenotype distribution of specific genes in a population, etc.) that serve as indices for the physiological assessments related to health, disease risk, psychiatric disorders, environmental exposure and its effects, metabolic processes, drug addiction, etc.

Biological data are usually expressed on a logarithmic scale. Moreover, inverse logarithm of the activity (log 1 / C) are also used to obtain mathematical values higher when structures are biologically very effective.

Some examples of biochemical and biological data, used in QSAR analysis are listed in the table bellow [133]:

Table. 2: Types of Biological data utilized in QSAR analysis.

Source of Activity	Biological Parameters
1. Isolated receptors	
Rate constants	Log K_{cat} ; Log K_{uncat} ; Log K ;
Michaelis-Menten constants	Log $1/K_m$
Inhibition constants	Log $1/K_i$
Affinity data	pA_2 ; pA_1
2. Cellular systems	
Inhibition constants	Log $1/IC_{50}$
Cross resistance	Log CR
In-vitro biological data	Log 1/C
Mutagenicity states	Log TA_{98}
3. '*In-vivo*' systems	
Biocencentration factor	Log BCF
In-vivùo reaction rates	Log I (Induction)
Pharmacodynamic rates	Log T (Total clearance)

2.3.2 Descriptors

Theoretical descriptors derived from physical and physico-chemical theories show some natural overlap with experimental measurements. Several quantum chemical descriptors, surface areas, and volume descriptors are examples of such descriptors also having an experimental counterpart.

With respect to experimental measurements, the greatest recognized advantages of the theoretical descriptors are usually (but not always) in terms of cost, time, and availability.

Each molecular descriptor takes into account a small part of the whole chemical information contained into the real molecule and, as a consequence, the number of descriptors is continuously increasing with the increasing request of deeper investigations on chemical and biological systems. Different descriptors have independent methods or perspectives to view a molecule, taking into account the various features of chemical structure. Molecular descriptors have now become some of the most important variables used in molecular modeling, and, consequently, managed by statistics, chemometrics, and chemoinformatics [134].

There are many different descriptors and many different kinds of classification based on the effect (steric, hydrophobic, electronic, etc.) or based on its dimension that called Descriptors Block, whereas we find four types of molecular descriptors 0D, 1D, 2D and 3D [134].

- Molecular descriptors 0D, 1D: contains numbers of atoms, functional groups, molecules properties, and charges (eg. Molecular Weight, Molecular Refractivity, etc.).
- Molecular descriptors 2D: contains simple descriptors and other derivatives of algorithms applied to a topological representation (eg. Number of Cl atoms, presence of hydroxyl group, etc.).
- Molecular descriptors 3D: contains descriptors derived from a geometric representation also called geometric descriptors (eg. Molecular volume, Surface area, etc.).

2.3.3 Multiple Linear Regression

There is a wide range of multivariate methods, of which a list not-limited of most popular methods, includes the following: PCA, Factor analysis, PCR, PLS, MANOVA, Neural Networks, Genetic Algorithms, etc. Among these methods, we find the multiple linear regression methods (MLR) whereas; we consider the samples as points in a space defined by variables and the coordinates of an individual are given by its values for each of these variables [135,136].

MLR expresses a single dependent variable (Y) as a linear combination of multiple independent variables (X) :

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k \quad \ldots\ldots\ldots\ldots\ldots (19)$$

Whereas, $\beta_1 \ldots \beta_k$ are coefficients of the regression, and β_0 is a constant, the regression model can be built in a stepwise manner.

After having determined the QSAR model, several questions show up.

A- The relationship between Y and X_1, ... X_k is it significant?

The answer is to use the Sig test as fellow:

Test: H_0: $\beta_1 = \ldots = \beta_k = 0$ (Y = β_0 does not depend on X)

H_1: at least one $\beta_k \neq 0$ (Y depends on at least one X)

We need to reject the hypothesis H_0 by the test of significance, whereas the parameter Sig is the conditional probability that a relationship as strong as the one observed in the data would be present, if the null hypothesis were true. It is often called the p-value. Typically a value of less than 0.05 is considered significant [91, 137].

B- How to choose the best QSAR model?

The answer is to use some statistical parameters as fellow [138, 139]:

- F: is the ratio of two mean squares. When the F value is large and the significance level is small the null hypothesis can be rejected. In other words, F is a statistic parameter used for testing the null hypothesis that inclusion of an additional variable does not result in a significant increase in R^2.
- SE: (standard error of the estimate) is a measure of how much the value of a test statistic varies from sample to sample. It is the standard deviation of the sampling distribution of a statistic.
- R: is the correlation coefficient between the observed and predicted values of the dependent variable. It ranges in value from 0 to 1.
- Q: is the factor of the quality that suggests the power of prediction.

C- How to validate a QSAR model?

The answer is to use the cross-validation technique called Leave One Out (LOO) based on the following parameters [140-144]:

- PRESS: Predicted Residual Sum of Squares is the difference between an observed value and the value predicted by the model.

- SSY: is the sum of the squares of the distances of the observed values for a variable compared to the average of this variable, the sum of squares allows to measure the total change in a variable.
- SPRESS: the predictive ability of the models is evaluated by the root mean square error.
- R^2_{cv} : The change in the R^2 statistic that is produced by adding or deleting an independent variable. If the R^2 change associated with a variable is large, that means that the variable is a good predictor of the dependent variable.
- R^2_{adj}: the sample R^2 tends to optimistically estimate how well the models fits the population. The model usually does not fit the population as well as it fits the sample from which it is derived. Adjusted R^2 attempts to correct R^2 to more closely reflect the goodness of fit of the model in the population.
- PE: the prediction error of the correlation coefficient is used to determine the predictive power of the models proposed.

2.4 Conceptual DFT

A branch of DFT has been developed since the late 1970s and early 1980s, called "conceptual DFT" by its protagonist, R. G. Parr. [145]. It was developed based on the idea that the electron density is the fundamental quantification for describing atomic and molecular ground states.

Parr and co-workers, and later on a large community of chemically orientated theoreticians, were able to give sharp definitions for chemical concepts which were already known and had been in use for many years in various branches of chemistry (electronegativity being the most prominent example), thus affording their calculation and quantitative use.

This step initiated the formulation of a theory of chemical reactivity which has gained increasing attention in the literature in the past decade. The breakthrough in the dissemination of this approach was the publication in 1989 of Parr and Yang's Density Functional Theory of Atoms and Molecules [146].

In recent years, Conceptual Density Functional Theory offered a perspective for the interpretation/ prediction of experimental/theoretical reactivity [13, 147-150] data based on a series of response functions to perturbations in the number of electrons and/or external potential [1451-153].

To avoid any confusion, it should be noted that the term "conceptual DFT" does not imply that the other branches of DFT did not contribute to the development of concepts within DFT. "Conceptual DFT" concentrates on the extraction of chemically relevant concepts and

principles from DFT [154]. Thus, we list some concepts that have been largely used in our field.

1- Electronegativity and the Electronic Chemical Potential.
2- Global Hardness and Softness.
3- The Electronic Fukui Function, Local Softness, and Softness Kernel.
4- Local Hardness and Hardness Kernel.
5- The Molecular Shape FunctionsSimilarity.
6- The Nuclear Fukui Function and Its Derivatives.
7- Spin-Polarized Generalizations.
8- Solvent Effects.
9- Time Evolution of Reactivity Indices.

2.5 Molecular Docking

The research-based pharmaceutical industry has increasingly employed modern medicinal chemistry methods, including molecular modeling, this field has progressed hand-in-hand with advances in biomolecular spectroscopic methods such as X-ray crystallography and nuclear magnetic resonance (NMR), which have enabled striking progress in molecular and structural biology. These techniques have allowed the resolution of more than 100,000 three-dimensional protein structures, providing vital structural information about key macromolecular drug targets [155].

Efforts in storing, organizing and exploring such information have generated a growing demand for robust and sophisticated computational tools. Based on this perspective, the accurate integration of computational and experimental methods has provided the up-to-date understanding of the intricate aspects of intermolecular recognition [156].

2.5.1 Molecular Docking: State of the Art

Molecular Docking is a computational procedure that attempts to predict the noncovalent/covalent binding of a small molecule (Ligand) as a lead for further drug development with a macromolecule (Receptor) obtained from data banks or MD simulations, etc. [11].

Molecular docking methods generally involve a sampling algorithm for selecting best possible conformation and orientation of ligands within target protein binding site. Hundreds to thousand orientational and conformational configurations are generated then scored/ranked

by a scoring function. Scoring functions are another critical component of molecular docking methods and can be broadly categorized into force field based, empirical and knowledge based methods. Although molecular docking methods were pretty successful in rank-ordering compounds in virtual screening, however, their pose prediction performance varies depending upon the type of small molecule ligand and target protein.

Docking can also be used to try to predict bound conformation, binding energy and affinity [157]. Thus, the purpose of drug discovery is to derive drugs that more strongly bind to a given protein target than the natural substrate [11].

2.5.2 Types of Molecular Docking

There are two major types of molecular docking, *Rigid Docking* consists of obtaining the preferred conformation of a protein-ligand system whereas every single molecule maintains a fixed internal geometry. In this case, the relaxation of the internal geometry of each entity, interacting in the complex, is not taken into account.

However, it is quite conceivable that the protein and ligand structures can be modified during the process of molecular docking in order to optimize the best interaction between the two entities, it is what called the Flexible Docking [157].

2.5.3 Ligand Conformational Search

There are three type of approaches to search through ligand conformational space [158]:

A- Exhaustive:

Performing an exhaustive search through the whole conformational space quickly becomes infeasible due to the exponential growth with respect to the number of rotational bonds (i.e., the ligand flexibility). Thus, most algorithm use a branch-and-bound approach to prune the conformational space. The geometric and chemical properties of the binding site are limiting factors that allow to reject many conformations [159, 160].

B- Stochastic:

The two dominant approaches in this category are *Tabu search* and *genetic algorithms*. In tabu search, small random perturbations are applied to the current conformation followed by their ranking according to some fitness function that in many cases only evaluates geometric constraints (to increase performance). Rejected conformations are marked (tabu) to avoid re-exploring them. In the

genetic approach, a population of conformations is evaluated at each iteration and a set of random perturbations ("mutations") and breeding rules are applied to the most fitted conformations in order to improve the fitness of the next generation [161, 162].

C- Simulation:

Simulation methods are generally based on molecular dynamics and are being used in two ways. One, to sample a diverse set of low-energy (local minima) conformations, where diversity, in this case, is usually measured by RMSD distances from all current selected conformations. Or two, in complement to another searching method, as a post-processing stage [163].

2.5.4 Scoring Function

Scoring functions are currently the bottle-neck of docking algorithms. A good scoring function should be able to differentiate active and non-active ligands, identify correct poses, and estimate binding energy. Free energy calculations using molecular dynamic simulations may provide better accuracy but they are extremely slow and thus inadequate for evaluating large molecular libraries. Scoring functions implemented in docking algorithm tend to simplify or ignore complex physical and chemical terms such as entropy or hydrophobicity [164].

Chapter

RESULTS AND DISCUSSION

Many researches have been carried out for the development of trypanocidal drugs; nevertheless, no drug could achieve a parasitological cure and some of them presented serious toxic side effects. While, **RESEARCHES NEVER END WITH NOTHING**, some compounds were found to be important candidates as inhibitors against the parasitic flagellate *T. Cruzain and T. Rhodsain*.

Therefore, following our interest in this field, we applied several computational methods on the most effective trypanocidal candidates in order to, study structures and properties of these ligands, evaluate their biological response and better understand Trypanosomes inhibition mode.

To simplify our ideas, we divide this chapter to three big titles as fellow:

- **Molecular Geometry, Electronic Properties, MPO Methods and Structure Activity/Property Relationship Studies of 1,3,4-Thiadiazole Derivatives.**

At first, we aim to study the molecular geometry, electronic properties and substitution effects of 1,3,4-thiadiazole nucleus using ab-initio and DFT methods, Afterward, we extend a dataset of 1,3,4-thiadiazole derivatives, in order to identify compounds with higher potency, we use rules of thumb, calculated metrics and structure activity/property relationships (SAR/SPR) respectively to their antitrypanosomal activity against Trypanosoma Cruzi.

- **Quantitative Structure-Activity Relationships of Antitrypanosomal Activities of Alkyldiamine Cryptolepine Derivatives.**

Next, in this part we develop QSAR models of Cryptolepine derivatives with respect to their antitrypanosomal activities against *Trypanosoma Cruzain* and *Trypanosoma Rhodesain*.

- **In Silico Investigation by Conceptual DFT and Molecular Docking of Antitrypanosomal Compounds for Understanding Trypanosomes Inhibition.**

Finally, we attempt to determine the active center sites of trypanocidal compounds using the Conceptual DFT, starting by studying the molecular behavior using Fukui function followed by a visualization of the reactivity at these sites using FMO theory and molecular electrostatic potential (MESP) surface. Afterward, we extend this study toward the Molecular Docking using an energy-based scoring function, to identify the most favorable ligand conformation among the trypanocidal compounds under study when bound to the target *T. Cruzi*, followed by a visualization of Ligand–receptor binding mode and a comparison with experimental results, to improve the affinity of this investigation in understanding the *T. Cruzi* active site.

3.1 Molecular Geometry, Electronic Properties, MPO Methods and Structure Activity/Property Relationship Studies of 1,3,4-thiadiazole Derivatives

As we discussed in the first chapter, the development of new drugs is an important priority to treat Trypanosoma disease. Therefore, 1,3,4-thiadiazole and its derivatives were selected as good candidates that have a considerable attention to enhance the Antiparasitic activity particularly against Trypanosomes [164, 165].

Drug discovery activities are producing ever-larger volumes of complex data that carry significant levels of uncertainty; multiparameter optimization (MPO) methods enable this data to be better utilized to quickly target compounds with a good balance of properties. Among these methods we carry out rules of thumb and calculated metrics.

Rules of thumb are the most common approach used to consider the quality of compounds relative to criteria beyond potency that provides guidelines regarding desirable compound characteristics. Several rules have been proposed; the most commonly used are Lipinski and Veber rules [122, 166]. On the other hand, calculated metrics aim to combine the potency with other parameters into a single metric which may be monitored during optimization. The earliest and most commonly applied metrics are the Ligand Efficiency (LE) and the Lipophilic Efficiency (LipE) [167].

The process of drug discovery balances a relentless search for molecules that have structural features that produce:

(1) Strong target binding using structure-activity relationship (SAR).

(2) High performance at *in-vivo* barriers, using structure-property relationship (SPR).

In the same way that design of structural features using SAR is known as structure-based design, the design of structural features using SPR has become known as property-based design. How a medicinal chemist goes about balancing these often disparate processes is a matter of experience and strategy [168].

Here, we carry out the Structure Activity/Property Relationship (SAR/SPR) studies which are attempting to enhance our understanding of fundamental processes and phenomena in medicinal chemistry and drug design [169-172] and to give us the correlation between molecular structures and properties [128, 173] such as lipophilicity, polarizability, electronic and steric parameters. The molecular properties, used in the correlations, relate as directly as possible to the key physicochemical processes taking place in the target activity.

3.1.1 Computational Methodology

The molecular modeling calculations for all the 1,3,4-thiadiazole derivatives are performed by HyperChem version 8.08 [89], Gaussian 09 [90] and MarvinSketch 6.2.1 software [88].

Firstly, the investigated molecules were pre-optimized by means of the Molecular Mechanics Force Field (MM+). After that, the resulted minimized structures were further refined using the semi-empirical PM3 method with a gradient norm limit of 0.01 kcal/Å for the geometry optimization.

In the next step, we realized the calculation of some geometric and electronic parameters, using various computational levels, HF/6-311 + + G (3d, 3p); HF/CC-PVTZ and B3LYP/6-311 + + G (3d, 3p); B3LYP/CC-PVTZ. This work also includes calculation of 3D MESP surface map and 2D MESP contour map to reveal the information regarding charge transfer within the molecule [174].

Finally, the PM3 optimized geometrical structures of nineteen derivatives of 1,3,4-thiadiazole were used to calculate the following physicochemical properties: Polar surface area (PSA), Surface Area Grid (SAG), molecular Volume (V), Hydration Energy (HE), Octanol-water partition coefficient (log P), Molecular Refractivity (MR), Polarizability (Pol) and Molecular Weight (MW).

3.1.2 Geometric and Electronic Structure of 1,3,4-thiadiazole

The optimized geometrical parameters of 1,3,4-thiadiazole (Fig. 24) are obtained using ab-initio/HF and DFT methods, listed in Table 3 with the experimental results [175]. Whereas, we notice that all the experimental results are approximately similar to the theoretical results, regarding bond length and valence angle values. Whereas, DFT/B3LYPE [6-311++G(3d,3p)] showed up to be the closest to the observed values.

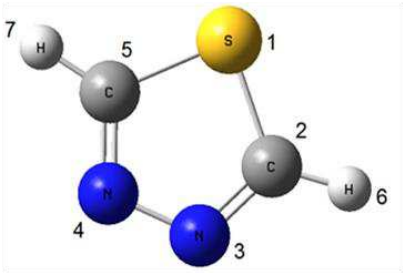


Fig. 24: 3D structure of 1,3,4-thiadiazole.

Table. 3: Bond lengths and valence angles of 1,3,4-thiadiazole.

Parameters		EXP [175]	Ab-initio/HF		DFT/B3LYP	
			6-311++G(3d,3p)	cc-PVTZ	6-311++G(3d,3p)	cc-PVTZ
Bond length (angstrom)	C5- S1	1.720	1.717	1.721	1.729	1.734
	C5-N4	1.302	1.269	1.268	1.297	1.294
	C5-H7	1.077	1.069	1.069	1.078	1.087
	N3-N4	1.371	1.361	1.361	1.368	1.367
Valence angle (degree)	S1-C2-N3	114.640	114.310	114.334	114.339	114.434
	C2-N3-N4	112.200	112.724	112.790	112.556	112.608
	C2-S1-C5	86.380	85.929	85.751	86.208	85.913
	S1-C2-H6	122.490	122.797	122.744	122.403	122.227

The theoretical dihedral angle values calculated by different methods and basis are practically equal to zero degree which explain that the geometry of 1,3,4-thiadiazole is planar, which make this conformation more stable. We can also note that charge densities calculated by ab-initio are similar to those calculated by the DFT method (Table 4).

Table. 4: Net charges distribution of 1,3,4-thiadiazole.

Atoms	HF / 6-311G++(3d,3p)	B3LYP / 6-311G++(3d,3p)
S1	- 0.069	- 0.040
C2	0.459	0.243
N3	- 0.486	- 0.342
N4	- 0.486	- 0.342
C5	0.459	0.243
H6	0.061	0.118
H7	0.061	0.118

The MESP surface map and contour map of 1,3,4-thiadiazole (Fig. 25) show the two regions characterized by red color (negative electrostatic potential) around the tow cyclic nitrogen atoms which explain the ability for an electrophilic attack on these positions, also by blue color (positive electrostatic potential) around the two hydrogen atoms which explain that these regions are susceptible for a nucleophilic attack. Finally for the green color located between the red and blue regions explain the neutral electrostatic potential surface.

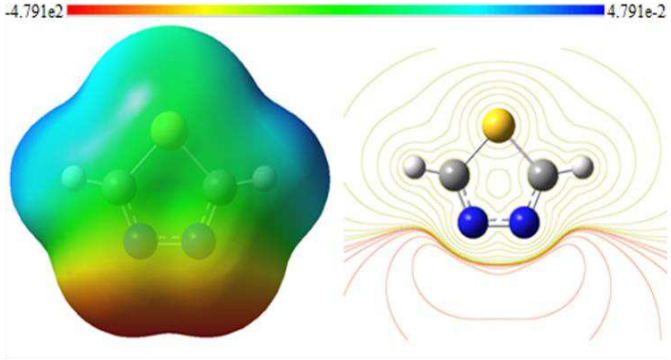


Fig. 25: 3D MESP surface map and 2D MESP contour map for 1,3,4-thiadiazole.

The variation in electrostatic potential produced by a molecule is largely responsible for binding of a drug to its active sites (receptor), as the binding site in general is expected to have opposite areas of electrostatic potential [176].

3.1.3 The Substitution Effect on 1,3,4-thiadiazole Structure

Calculated values such as heat of formation, dipole moment (μ), HOMO (Highest Occupied Molecular Orbital) and LUMO (Lowest Unoccupied Molecular Orbital) energies of 1,3,4-thiadiazole systems (Fig. 26) are given in Table 5, net atomic charges of 1,3,4-thiadiazole systems are also reported in Table 6 for the first series and in Table 7 for the second series.

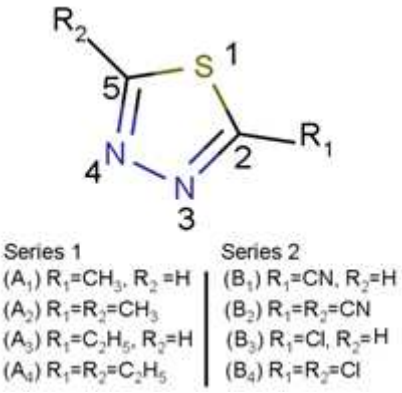


Fig. 26: 1,3,4-thiadiazole systems.

Table. 5: Electronic parameters of 1,3,4-thiadiazole systems.

Comp.	Systems	Heat of formation (Kcal/mol)	HOMO (a.u.)	LUMO (a.u.)	ΔE (a.u.)	μ (D)
X	1,3,4-thiadiazole	067.402	-0.285	-0.058	0.227	3.387
A_1	2-methyl-1,3,4-thiadiazole	059.564	-0.277	-0.050	0.227	3.430
A_2	2,5-dimethyl-1,3,4-thiadiazole	051.803	-0.269	-0.042	0.227	3.261
A_3	2-ethyl-1,3,4-thiadiazole	055.168	-0.276	-0.049	0.227	3.454
A_4	2,5-diethyl-1,3,4-thiadiazole	042.989	-0.268	-0.042	0.226	3.251
B_1	2-cyano-1,3,4-thiadiazole	108.478	-0.310	-0.107	0.203	4.777
B_2	2,5-dicyano-1,3,4-thiadiazole	150.517	-0.332	-0.143	0.189	2.151
B_3	2-chloro-1,3,4-thiadiazole	064.305	-0.293	-0.068	0.225	3.339
B_4	2,5-dichloro-1,3,4-thiadiazole	061.169	-0.290	-0.076	0.214	2.984

Table. 6: Mulliken charges of 1,3,4-thiadiazole systems series 1.

Comp.	T	A_1	A_2	A_3	A_4
S1	-0.040	-0.167	-0.301	-0.217	-0.431
N3	-0.342	-0.384	-0.403	-0.409	-0.475
N4	-0.342	-0.411	-0.513	-0.403	-0.475
C2	0.243	0.481	0.712	0.681	0.815
C5	0.243	0.564	0.686	0.554	0.817
C-methyl-2	-	-0.082	-0.218	-	-
C-methyl-5	-	-	0.039	-	-
C-ethyl-2	-	-	-	-0.205	-0.125
C-ethyl-5	-	-	-	-	-0.125

Table. 7: Mulliken charges of 1,3,4-thiadiazole systems series 2.

Comp.	T	B_1	B_2	B_3	B_4
S1	-0.040	-0.159	-0.133	-0.158	-0.238
N3	-0.342	-0.344	-0.374	-0.284	-0.310
N4	-0.342	-0.359	-0.374	-0.373	-0.310
C2	0.243	1.198	1.412	0.504	0.591
C5	0.243	0.593	1.412	0.416	0.591
Chloro-2	-	-	-	-0.103	-0.161
Chloro-5	-	-	-	-	-0.161
C-cyamide-2	-	-0.313	-0.367	-	-
N-cyamide-2	-	-0.616	-0.604	-	-
C-cyamide-5	-	-	-0.367	-	-
N-cyamide-5	-	-	-0.604	-	-

In this part, we have studied two series containing methyl and ethyl groups in the first one and chloride and cyanide groups in the second one for a comparative study between the effects of electron donors and acceptors.

For each addition of methyl, ethyl or chloride radical the heat of formation decreases of approximately 8, 12 or 3 (Kcal/mol) respectively but the addition of the cyanide group leads the increase of the heat of formation with 41 (Kcal/mol) approximately. Among the various substitutions that we have added each time to the 1,3,4-thiadiazole and by the calculations that we have performed, it was found that electron donors of compound A_4 (2,5-diethyl-1,3,4-thiadiazole) has the lowest energy gap HOMO-LUMO (0.226) for the first series and compound B_2 (2,5-dicyano- 1,3,4-thiadiazole) has the lowest energy gap (0.189) for the second series (Table 5).

From HSAB (Hard Soft Acid and Base) principle the lowest energetic gap allows an easy flow of electrons which makes the molecule soft and more reactive which means that A_4 and B_2 compounds are the most reactive in the two series of 1,3,4-thiadiazole derivatives [177]. The compound B_1 (2-cyano-1,3,4-thiadiazole) shows the maximum dipole moment value. It would originate from a resonance effect, involving a donor effect from the thiadiazole nucleus toward the electro-attractive group in position 2 [113].

The contour plots of the π-like frontier orbital for the ground state of the compound B_1 are shown in (Fig. 27). From the plots, we can observe that the HOMO mainly concentrates on S1 with some delocalization along N3-N4 and N8, whereas, the LUMO distributes over the whole molecule. These further demonstrate the existence of the delocalization of the conjugated π-electron system in the 2-cyano-1,3,4-thiadiazole molecule.

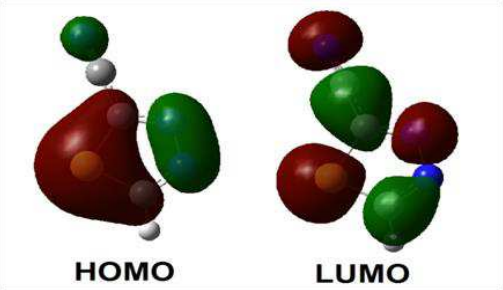


Fig. 27: π-like frontier orbitals of the compound B_1.

The negative atomic charge on S1, N3 and N4 is increased considerably for methyl and ethyl group substitutions (Table 6) for the first series and increased considerably for cyanide group substitution (Table 7) for the second series in except for chloride substitution in N3 and N4 atoms. The carbon C2 and C5 take the most important positive charges (0.815), (0.817) respectively in the compound A_4 (2,5-diethyl-1,3,4-thiadiazole) for the first series, also for compound B_4 (2,5-dichloro-1,3,4-thiadiazole) for the second series, the most important positive charges are in carbon C2 (0.591) and C5 (0.591) as shown in Table 7, these positions C2 and C5 with the important positive charges lead to preferential sites of nucleophilic attack.

3.1.4 Multi-Parameter Optimization of 1,3,4-thiadizole Derivatives

In this part, we have applied rules of thumb and calculated metrics on nineteen derivatives of 1,3,4- thiadiazole (Fig. 28) with respect to their antitrypanosomal activity (pIC_{50}) against *Cruzain* [165]. The properties involved are presented in Table 8.

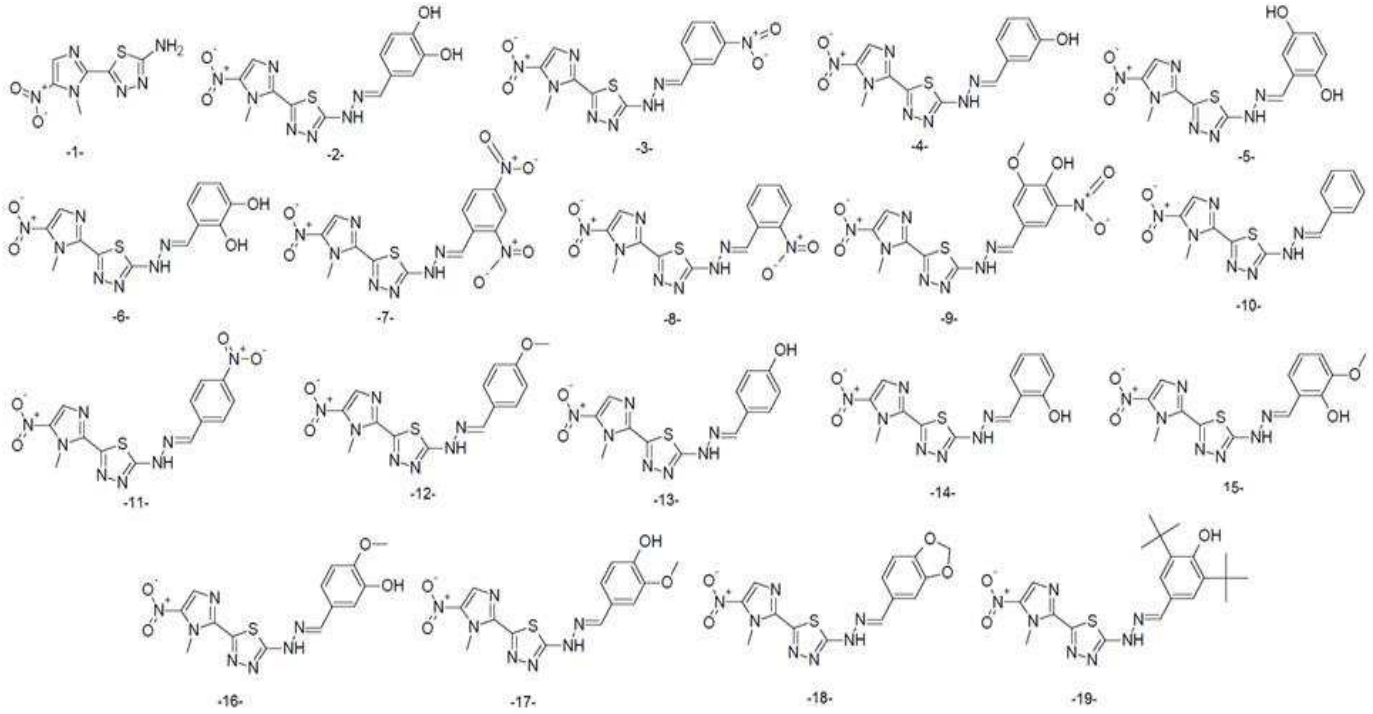


Fig. 28: 2D structures of 1,3,4-thiadiazole derivatives.

At first, we have studied Lipinski and Veber rules to identify "drug-like" compounds. Rich absorption or permeability is more likely when [122, 166]:

(1) There are less than 5 H-bond donors (expressed as the sum of OH_s and NH_s).

(2) The molecular weight is under 500 DA.

(3) The log P is under 5.

(4) There are less than 10 H-bond acceptors (expressed as the sum of N_s and O_s).

(5) Rotatable bonds are under 10.

(6) Polar surface area is under 140 $Å^2$.

We used the Lipinski rules to identify compounds posing problems of absorption and permeability if these compounds don't validate at least two of its rules [177, 178] in addition Veber rules suggest that molecular flexibility (Rotatable bonds) and polar surface area (PSA) are important determinants of oral bioavailability [130].

The rules are based on a strong physicochemical rationale. Hydrogen bonds increase solubility in water and must be broken in order for the compound to permeate into and through the lipid bilayer membrane. Thus, an increasing number of hydrogen bonds reduce partitioning from the aqueous phase into the lipid bilayer membrane for permeation by passive diffusion. 1,3,4-thiadiazole derivatives are in accordance with rule 1 but generally they don't validate rule 4 as shown in Table 8 so we can say that they are probably more polar and hardly absorbed.

Table. 8: Pharmacological Activities and properties involved in MPO methods for 1,3,4-thiadiazole derivatives.

No.	pIC_{50}	Log P	MW (amu)	HBD	HBA	Lipinski score of 4	NRB	PSA	LE	LipE
1	5.004	-0.627	226.213	2	08	4	2	112.760	0.467	5.631
2	5.275	-1.122	361.335	3	11	3	5	151.590	0.295	6.397
3	4.317	-1.819	374.334	1	12	3	6	154.270	0.232	6.136
4	4.131	-0.097	345.335	2	10	4	5	131.360	0.241	4.228
5	3.854	-1.122	361.335	3	11	3	5	151.590	0.216	4.976
6	3.509	-1.122	361.335	3	11	3	5	151.590	0.196	4.631
7	3.481	-4.566	419.331	1	15	3	7	197.410	0.168	8.047
8	3.139	-1.819	374.334	1	12	3	6	154.270	0.169	4.958
9	4.721	-3.837	420.359	2	14	3	7	183.730	0.228	8.558
10	3.699	0.928	329.336	1	09	4	5	111.130	0.225	2.771
11	4.657	-1.819	374.334	1	12	3	6	154.270	0.251	6.476
12	3.699	-0.065	359.362	1	10	4	6	120.360	0.207	3.764
13	4.935	-0.097	345.335	2	10	4	5	131.360	0.288	5.032
14	4.266	-0.097	345.335	2	10	4	5	131.360	0.249	4.363
15	3.699	-1.090	375.362	2	11	3	6	140.590	0.199	4.789
16	4.296	-1.090	375.362	2	11	3	6	140.590	0.231	5.386
17	3.699	-1.090	375.362	2	11	3	6	140.590	0.199	4.789
18	3.699	-0.869	373.346	1	11	3	5	129.590	0.199	4.568
19	4.478	2.530	457.550	2	10	4	7	131.360	0.196	1.948

Molecular weight (MW) is related to the size of the molecule. As molecular size increases, a larger cavity must be formed in water in order to solubilize the compound.

Increasing MW reduces the compound concentration at the surface of the intestinal epithelium, thus reducing absorption. Increasing size also impedes passive diffusion through the tightly packed aliphatic side chains of the bilayer membrane. We have all series compounds of 1,3,4- thiadiazole derivatives with molecular weights less than 500 Da (rule number 2), so they are likely soluble and easily pass through cell membranes.

Increasing Log P also decreases aqueous solubility, which reduces absorption. Thus, membrane transporters can either enhance or reduce compound absorption by either active uptake transport or efflux, respectively. 1,3,4- thiadiazole derivatives satisfy also the rule number 3 so it has a consequence of a better solubility in aqueous and lipidic solutions [168].

It is well known that high oral bioavailability is an important factor for the development of bioactive molecules as therapeutic agents. Reduced molecular flexibility (measured by the number of rotatable bonds) and low polar surface area are found to be important predictors of good oral bioavailability [179].

Whereas, rotatable bonds and polar surface area tend to increase with molecular weight may in part explain the success of these two parameters in predicting the oral bioavailability and the transport across membranes.

The low number of rotatable bonds (reduced flexibility) in the studied series indicates that these Ligands upon binding to a protein change their conformation only slightly. On the other hand concerning the polar surface area (PSA), in the studied series of 1,3,4-thiadiazole derivatives, the very high values of PSA result in worsening of the absorption of a drug. Indeed, compounds 1, 4, 10, 12, 13, 14, 15, 16, 17, 18 and 19 with PSA values between 111 and 140 belong to the compounds with reduced absorption (Table 8).

Secondly, we have studied the Ligand Efficiency (LE) to penalize large compounds over small compounds with similar potency because larger compounds tend to have poorer physicochemical and ADME properties [180, 181] In addition, we have studied Lipophilic Efficiency (LipE) to maximize potency while maintaining as low as possible the lipophilicity, due to the association between high lipophilicity and several issues including poor solubility, membrane permeation, metabolic stability, etc. [182, 125].

Respectively to the antitrypanosomal activity against *Cruzain* of the 1,3,4-thiadiazole derivatives series, high LE prefers compounds that gain to escape the affinity-biased selection and optimization towards larger ligands. The focus should be directed towards the generation

of compounds that use their atoms most efficiently. As regards LipE, it prefers compounds that gain a lot of their affinity through directed interactions, thus making the interaction with the receptor more specific. While one can say that LipE describes how efficient a Ligand exploits its lipophilicity, no explicit measure of molecular size is used.

From the results obtained in Table 8 we can say for ligands containing $pIC_{50} > 5$ we are able to penalize compound 2 with lowest LE value (0.295), for ligands containing $5 > pIC_{50} > 4$ we can penalize the compound 19 with lowest LE value (0.196) and for compounds containing $pIC_{50} < 4$ we can penalize compounds 7 and 8 with LE values 0.168 and 0.169 respectively.

As lipophilicity is the major factor for the promiscuity of compounds, LipE optimized compounds should be more selective. It is suggested to target a LipE in a range of 5–7 or even higher [183]. In the series studied LipE is changing during optimization (Table 8). The most hydrophilic compounds 7 and 9 reach a LipE of 8.047 and 8.558 respectively, which are even above the suggested range of 5–7. This explains that log P is the most important parameter for the inhibitory activity against *Cruzain*.

On the other side, Compounds 1, 2, 3, 11, 13 and 16 reach a LipE of 5.631, 6.397, 6.136, 6.476, 5.032 and 5.386 respectively, which are situated in the suggested range of 5–7, which indicates that these compounds were successfully optimized. Comparing the log P values of the compounds reveal that lipophilicity was kept fairly constant between (−0.097) and (−1.819) during the substitution in R_1 and R_2 radicals on 1,3,4-thiadiazole nucleus. In the other compounds none of them reach a LipE above 5. In this respect, optimization was not optimal; however log P values are various. This means that antitrypanosomal activity is mainly gained by the introduction of groups and other properties making specific directed interactions.

3.1.5 Structure Activity/Property Relationships for 1,3,4-thiadiazole Derivatives

For the series of 1,3,4-thiadiazole derivatives (Fig. 28) we have studied seven physicochemical properties with respect to their antitrypanosomal activity against *Cruzain* [165]. The properties involved are: Surface area grid (SAG), molar volume (V), hydration energy (HE), partition coefficient octanol/water (logP), molar refractivity (MR), polarizability (Pol) and molecular weight (MW). The results using HyperChem 8.0.8 software are shown in Tables 8 and 9.

Table. 9: Physicochemical properties of 1,3,4-thiadiazole derivatives.

No.	SAG (Å^2)	V (Å^3)	|HE| (kcal/mol)	MR (Å^3)	Pol (Å^3)
1	380.188	0585.185	20.491	053.126	20.389
2	570.438	0921.896	31.997	092.702	34.092
3	583.867	0944.192	24.134	095.207	34.659
4	555.226	0900.998	26.167	091.096	33.455
5	561.112	0916.949	28.235	092.702	34.092
6	560.454	0915.865	27.149	092.702	34.092
7	632.843	1023.244	28.856	100.920	36.500
8	591.812	0958.365	23.456	095.207	34.659
9	628.800	1036.228	29.557	103.190	37.768
10	538.730	0876.383	20.075	089.491	32.818
11	576.264	0937.894	25.011	095.207	34.659
12	588.042	0959.381	21.357	095.865	35.290
13	551.898	0896.813	26.895	091.096	33.455
14	545.528	0891.737	22.187	091.096	33.455
15	588.002	0975.319	23.205	097.471	35.927
16	607.021	0983.950	26.085	097.471	35.927
17	594.784	0980.812	27.033	097.471	35.927
18	582.191	0948.712	23.733	095.080	35.153
19	728.908	1258.720	17.069	126.910	48.135

From the results obtained in Tables 8 and 9 the molecular refractivity and polarizability increase relatively to the size and the weight of 1,3,4-thiadiazole derivatives. This explains the accordance of our results with Lorentz-Lorenz expression [184-187].

This relation show that the molecular refractivity and polarizability increase with the volume and molecular weight of molecules, for example, the compound 19 with bulky substituent which are OH and two ter-butyls groups has the important values of molecular refractivity and polarizability (126.910 Å^3), (48.135 Å^3) respectively.

However, the first compound is the smallest molecule in this 1,3,4-thiadiazole derivatives series, this compound has the smallest value of the molecular refractivity (53.126 Å^3) and polarizability (20.389 Å^3).

We have also noticed that the surface and the volume of distribution of these molecules are different of Benznidazole, Nifurtimox and Cryptolepine which acting on the Trypanosoma Cruzi [58, 188].

We found that their surfaces and volumes take: 424.58 Å^2 and 711.62 Å^3 values respectively for Benznidazole, 469.95 Å^2 and 771.18 Å^3 values respectively for Nifurtimox and 422.16 Å^2 and 701.56 Å^3 values respectively for Cryptolepine. By the contrary, the most active compound against Trypanosoma Cruzi is the second one in the 1,3,4- thiadiazole derivatives series, with a surface of 570.438 Å^2 and a volume of 921.896 Å^3.

So, the compound 2 in 1,3,4-thiadiazole derivatives series is bulkier than Benznidazole, Nifurtimox and Cryptolepine which explain that the molecular size is an important factor in the antitrypanosomal activity against Trypanosoma Cruzi.

Hydration energy in absolute value, the most important is that of the compound 2 (31.997 Kcal/mol) and the smallest value is that of the compound 19 (17.069 Kcal/mol). Indeed, in the biological environments the polar molecules are surrounded by water molecules. They are established hydrogen bonds between them.

Hydrophobic groups in 1,3,4-thiadiazole derivatives induce a decrease of hydration energy, however, the presence of hydrophilic groups as in compound 2 (E)-1-(benzylidene-3,4-diol)-2-(5-(1-methyl-5-nitro-1Himidazol- 2-yl)-1,3,4-thiadiazol-2-yl)hydrazine (Fig. 29) possessing three (HBD) hydrogen bond donors (NH, 2OH) and eleven (HBA) hydrogen bond acceptors (NO_2, four cyclic nitrogen, NH, N and 2OH), result the increase in the hydration energy.

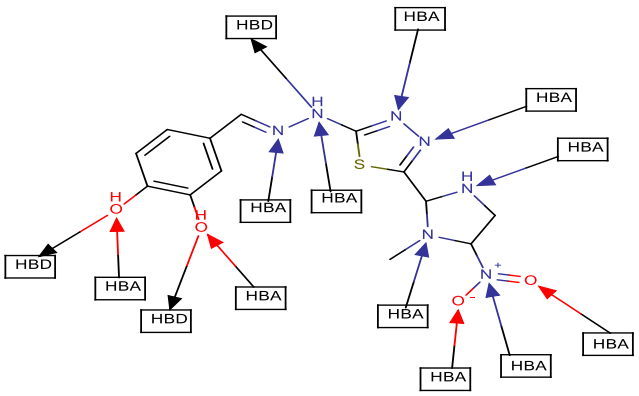


Fig. 29: Donor and acceptor sites of compound 2.

Otherwise, the lipophilie increase proportionally with the hydrophobic features of substituent. As seen in Tables 8 and 9 the compound 19 has the smallest value of hydration energy in absolute value (17.069) and the most important value of log P (2.530).

The results obtained by calculating Log P of 1,3,4-thiadiazole derivatives, show that the compounds 7 and 9 present small coefficients of log P (−4.566) and (−3.837) respectively. These compounds provide a good solubility and a bad absorption and penetration in cellular membranes, because of the weaker permeability of the passive diffusion but these molecules have a good permeability because of their small molecular weights [189].

3.2 Quantitative Structure-Activity Relationships of Antitrypanosomal Activities of Alkyldiamine Cryptolepine Derivatives

In this part, the inhibition of cysteine proteases, by a group of cryptolepine derivatives was investigated to predict a QSAR model using molecular descriptors and MLR analysis. A group of 22 cryptolepine derivatives inhibitors of cysteine proteases was selected for the study.

The structures and inhibitory activities expressed as values of IC_{50} obtained for the *in-vitro* [190] antitrypanosomal activities against *Cruzain* log ($1/IC_{Cruz}$) and *Rhodesain* log ($1/IC_{Rhod}$) were adopted as reported by João Lavrado et al [191].

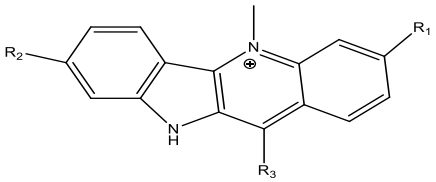
3.2.1 Computational Methodology

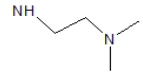
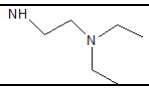
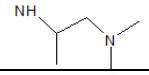
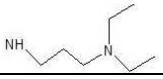
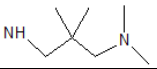
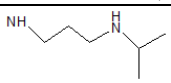
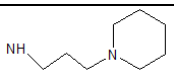
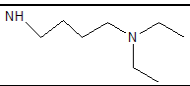
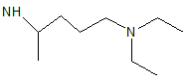
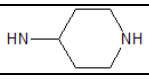
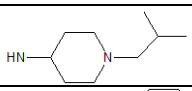
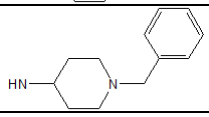
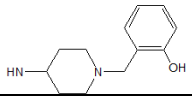
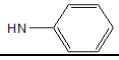
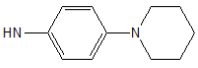
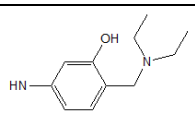
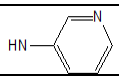
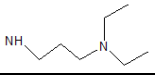
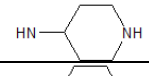
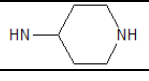
Firstly, twenty-two investigated molecules (table. 10) were preoptimized using the Molecular Mechanics Force Field (MM+) method included in HyperChem version 8.08 package [89]. The resultant minimized structures were further refined using the semi-empirical PM3 Hamiltonian as implemented also in HyperChem. The gradient norm limit of 0.01 kcal/Å was chosen for the geometry optimization.

Then, structures with a PM3 optimized geometry was used to calculate a number of physicochemical descriptors: surface area grid (SAG), molar volume (V), hydration energy (HE), partition coefficient octanol/water (log P), the molar refractivity (MR), molar polarizability (Pol) and molar weight (MW).

Finally, the multiple linear regression analysis method was carried out using the stepwise strategy in SPSS version 19 for Windows [91].

Table. 10: Chemical structures, experimental and predicted activities of the molecules under study.



No.	R_1	R_2	R_3	log (1/IC_{Cruz})			log (1/IC_{Rhod})		
				Obs.	Pred.	Res.	Obs.	Pred.	Res.
1	H	H	H	-1.447	-1.478	0.031	-1.362	-1.422	0.060
2	H	H		-1.361	-1.082	-0.278	-0.845	-0.781	-0.064
3	H	H		-1.361	-1.188	-0.172	-1.041	-0.807	-0.234
4	H	H		-1.000	-1.142	0.142	-0.699	-0.816	0.117
5	H	H		-1.146	-1.031	-0.115	-0.778	-0.689	-0.089
6	H	H		-1.531	-1.329	-0.201	-1.113	-0.942	-0.171
7	H	H		-1.414	-1.279	-0.135	-1.278	-0.929	-0.349
8	H	H		-1.623	-1.416	-0.207	-1.414	-1.052	-0.362
9	H	H		-1.041	-1.007	-0.033	-0.698	-0.585	-0.112
10	H	H		-1.342	-1.541	0.199	-0.778	-1.125	0.347
11	H	H		-1.431	-1.411	-0.021	-1.230	-1.036	-0.193
12	H	H		-0.477	-0.899	0.422	-0.301	-0.565	0.264
13	H	H		-0.903	-0.992	0.089	-0.301	-0.623	0.321
14	H	H		-0.903	-0.878	-0.025	-0.602	-0.449	-0.152
15	H	H		-0.778	-0.737	-0.040	0.045	-0.228	0.273
16	H	H		-1.176	-1.094	-0.081	-0.845	-0.838	-0.006
17	H	H		-0.477	-0.762	0.285	0.000	-0.316	0.316
18	H	H		-1.000	-0.884	-0.116	-0.698	-0.406	-0.292
19	H	H		-0.698	-0.798	0.100	-0.301	-0.473	0.172
20	Cl	H		-0.301	-0.581	0.280	0.301	-0.165	0.465
21	Cl	H		0.000	-0.266	0.265	0.301	0.109	0.191
22	Cl	Cl		0.000	0.388	-0.388	0.301	0.804	-0.502

3.2.2 Quantitative Structure Activity Relationships Studies

The several physical and chemical properties [192] known as physicochemical descriptors were used as independent variables and were correlated with biological activities of Cryptolepine derivatives for the generation of QSAR50 models by multiple linear regressions (MLR) analyzes.

Developing a QSAR model requires a diverse set of data, and thereby, a large number of descriptors have to be considered. Descriptors are numerical values that encode different structural features of the molecules. Selection of a set of appropriate descriptors from a large number of them requires a method, which is able to discriminate between the parameters. Pearson's correlation matrix has been performed on all descriptors by using SPSS statistics 19 Software. The analysis of the matrix revealed five descriptors for the development of MLR models. The values of descriptors used in MLR analysis are presented in table 11.

Table. 11: Values of physicochemical descriptors used in the regression analysis.

No.	SAG ($Å^2$)	V ($Å^3$)	HE (kcal/mol)	log P	MR ($Å^3$)	Pol ($Å^3$)	MW (amu)
1	422.167	0701.567	-4.936	0.442	083.764	28.426	233.293
2	569.146	0973.094	-4.250	-0.889	110.108	38.468	319.429
3	554.819	1014.380	-3.028	-0.204	119.604	42.138	347.483
4	582.190	1012.730	-3.223	-0.476	114.526	40.303	333.456
5	571.138	0996.869	-3.955	-0.837	114.973	40.303	333.456
6	619.425	1100.190	-3.203	-0.152	124.469	43.973	361.510
7	621.656	1095.420	-2.739	0.271	123.732	43.973	361.510
8	595.606	1051.120	-5.487	-0.443	118.844	42.138	347.483
9	577.396	1065.500	-3.013	-0.115	127.114	45.034	373.521
10	634.019	1145.680	-2.924	0.299	129.113	45.808	375.537
11	692.626	1209.140	-2.203	0.712	133.531	47.643	389.564
12	552.830	0959.055	-6.535	-1.254	112.244	39.529	331.440
13	654.401	1147.670	-2.892	0.322	131.283	46.869	387.548
14	683.514	1209.220	-5.614	0.121	146.290	51.024	421.565
15	687.937	1227.980	-8.551	-0.904	147.896	51.661	437.564
16	544.305	0937.700	-6.457	-0.237	116.971	39.437	324.405
17	661.175	1173.420	-4.671	-0.464	142.731	49.189	407.538
18	702.326	1228.090	-8.883	-1.004	145.856	50.600	425.553
19	535.456	0923.922	-7.845	-1.586	113.477	38.728	325.393
20	641.300	1135.030	-2.877	-0.375	129.185	45.901	395.955
21	575.195	1001.480	-6.256	-1.476	116.960	41.457	365.885
22	598.196	1043.570	-5.904	-1.699	121.677	43.385	400.330

The correlation between the biological activities and descriptors expressed by the following relations:

$$log(1/IC_{Cruz}) = -1.775 + 0.013\,SAG - 0.015\,V + 0.099\,HE - 0.230\,logP + 0.027\,MW$$

n = 22; R = 0.904; SE = 0.235; F = 14.300; Q = 3.85

$$log(1/IC_{Rhod}) = -2.030 + 0.010\,SAG - 0.013\,V + 0.100\,HE - 0.285\,logP + 0.027\,MW$$

n = 22; R = 0.869; SE = 0.309; F = 9.864; Q = 2.81

The values of fraction variance may lie between 0 and 1. QSAR model having $R^2 > 0.6$ will only be considered for validation. For example, the value R = 0.904 and $R^2 = 0.817$ (model 1) allowed us to indicate firmly the correlation between different parameters (independent variables) with antitrypanosomal activity against *Cruzain*.

The F value has found to be statistically significant at 95% level, since all the calculated F value is higher as compared to tabulated values. The positive value of quality factor (Q) for these two QSAR's models 1 and 2 suggests its high predictive power.

In models (1) and (2), the negative coefficients of log P explain that any increase in the lipophilie of the molecules causes a decrease in biological activity. Also the positive coefficient of HE indicates that any increase in hydration energy causes an increase in biological activity. From these two parameters it may be concluded that hydrophilic molecules are more important for antitrypanosomal activity against *Cruzain* and *Rhodesain*.

In order to test the validity of the predictive power of selected MLR models (model 1 and 2), the Leave-One-Out technique (LOO-technique) was used. The developed models were validated by calculation of the following statistical parameters: predicted residual sum of squares (PRESS), total sum of squares deviation (SSY), the predictive error of the coefficient of correlation (PE) and cross-validated correlation coefficients (R^2_{adj} and R^2_{cv}) (table 12).

Table. 12: Cross-validation parameters.

Model	PRESS	SSY	PRESS/SSY	S_{PRESS}	r^2_{cv}	r^2_{adj}	6PE
1	0.885	4.838	0.183	0.200	0.817	0.760	0.157
2	1.531	6.249	0.245	0.263	0.755	0.678	0.211

PRESS is an important cross-validation parameter as it is a good approximation of the real predictive error of the models. Its value being less than SSY points out that model predicts better than chance and can be considered statically significant. The smaller PRESS value means the better of the model predictability. From the results depicted in table 12, models 1 and 2 are statistically significant.

Further, for reasonable QSAR model, the PRESS/SSY ratio should be lower than 0.4 [193]. The data presented in Table 12 indicate that for the developed models this ratio is 0.183 for the first model and 0.245 for the second one. Our findings of R^2_{cv} for these QSAR models have been to be 0.817 for the first model and 0.755 for the second one. The high value of R^2_{cv} and R^2_{adj} are essential criteria for the best qualification of our QSAR models.

The predictive error of the coefficient of correlation (PE) is yet another parameter used to evaluate the predictive power of the proposed models [194]. We have calculated the PE value of the proposed models and they are reported in Table 12. For the two models developed the condition R > 6 PE is satisfied and hence they can be said to have a good predictive power.

However, the only way to estimate the true predictive power of developed model is to predict the by calculation of log ($1/IC_{Cruz}$) and log ($1/IC_{Rhod}$) values of the investigated cryptolepines using model 1 and 2 respectively (Table 12).

Figure 30 shows the plots of linear regression of predicted versus experimental values of the biological activity of cryptolepines outlined above. The plots for models 1 and 2 show a good deal of correspondence with experimentally reported data having $R^2 = 0.817$ and 0.755, respectively. The present QSAR study shows that models 1 and 2 can be successfully applied to predict antitrypanosomal activities against *Cruzain* and *Rhodesain* in these molecules generations.

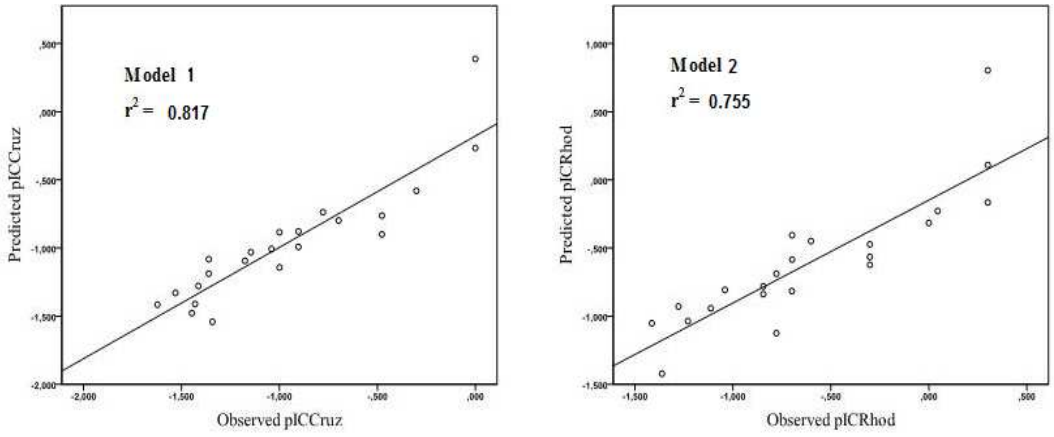


Fig. 30: Plots of predicted versus experimentally observed antitrypanosomal activities against *Cruzain* and *Rhodesain*.

To investigate the presence of a systematic error in developing the QSAR models, the residuals of predicted values of the biological activity log ($1/IC_{Cruz}$) and log ($1/IC_{Rhod}$) were plotted against the experimental values, as shown in Figure 31.

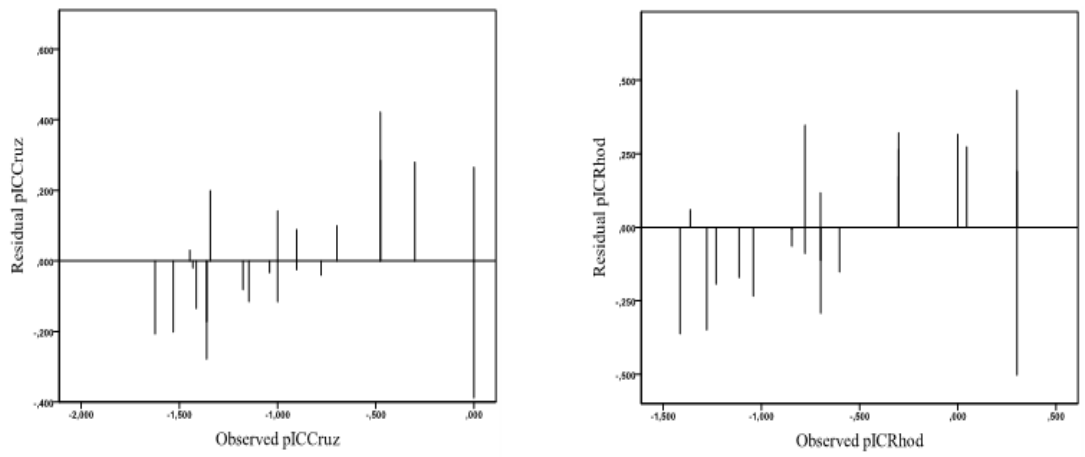


Fig. 31: Plots of residual against experimental values of log ($1/IC_{Cruz}$) and log ($1/IC_{Rhod}$).

The propagation of the residuals on both sides of zero indicates that no systemic error exists [195], as suggested by Jalali-Heravi and Kyani [196]. It indicates that these models can be successfully applied to predict the antitrypanosomal activity against *Cruzain* and *Rhodesain* of this class of molecules.

3.3 In Silico Investigation by Conceptual DFT and Molecular Docking of Antitrypanosomal Compounds for Understanding Trypanosomes Inhibition.

In this part, Three compounds were derived from literature to study the inhibition active site of the cysteine protease *T. Cruzi*, the molecular modeling and Molecular Docking calculations for Cryptolepine, 1,3,4-thiadiazol derivative and Nifurtimox are performed by Gaussian 09 [90], Molegro Virtual Docker [92] and LigPlot+ v.1.4.5 software [197].

3.3.1 Computational Methodology

Firstly, we start with a comparison between the different atomic charge models, Mulliken and CHELPG methods, carrying out the equilibrium geometries of the investigated molecules that have been determined and analyzed at the DFT level employing B3LYP/6-31 G and B3LYP/6-311 + + G basis sets.

In the next step, for understanding the molecular behavior of the neutral, mono-positive and mono-negative species ions under study we calculate Fukui functions as a site reactivity descriptor, followed by the visualization of LUMO + 1, LUMO, HOMO, HOMO - 1 and MESP using Gaussian 09 and GaussView 5.0 software.

Finally, the X-ray crystal structures of *Cruzain* bound with vinyl sulfone derived inhibitor (2OZ2), was downloaded from RCSB Database [87]. The structure of the protein was corrected for missing atoms or unknown units using MVD program. In addition, all the heteroatoms

including water molecules and any co-crystallized solvent were removed from the PDB file of *T. Cruzi* to simplify the simulation. For the visualization of the binding of trypanocidal compounds under study with *Cruzain* active site, we use LigPlot+ which is a graphical front-end to the LigPlot and DimPlot programs. Additionally, we superpose related LigPlots to highlight similarities and differences between related Ligands–Proteins binding.

3.3.2 Comparison of the Different Atomic Charge Methods

A main problem in comparing different point-charge models is the missing of clear criterion for the quality of the charges. This is probably the reason why so many charge models have been suggested. Furthermore, different applications put different demands on the charges. For example, in molecular dynamics, molecules move, so the charges must be able to describe the electrostatics properly in all accessible points in the phase space, and they should also be invariant to changes in the internal coordinates of the molecule [103-105].

The starting point of the present investigation was to obtain atomic charges of antitrypanosomal compounds to perform the DFT calculations for the prediction of the chemical behavior of Cryptolepine, 1,3,4-thiadiazol derivative and Nifurtimox as inhibitors of *T. Cruzi*. Therefore, we started to test the performance of two different partial charges available in the Gaussian software, Mullikan and CHELPG charges were calculated for all the studied trypanocidal compounds (Table 13) using different basis sets B3LYP/6-31 G and B3LYP/6-311 + + G. The charges quality was estimated from the residual sum of squares. From the results obtained in table 13, the residual sum of squares were equal to 25.60 and 0.61 for Mulliken and CHELPG methods, respectively. The CHELPG charges show the smallest variation whereas the variation is significantly larger for the Mulliken charges. Apparently, the electrostatic moments vary more than the electrostatic potential with the selected level of theory and the basis sets. In this case, we can say that CHELPG method gives results more accurate and it is the best approximation to perform the next part of studying the electronic properties of the antitrypanosomal compounds.

3.3.3 Fukui Functions as Site Reactivity Descriptor of Antitrypanosomal Compounds

The reactivity of the target molecules is an important aspect evaluated via molecular descriptors interpretation, whereas the most successful solution to better understand the active center sites of molecules is the frontier electron theory which has been developed by Fukui and coworkers [198].

The Fukui function is an index of considerable importance for understanding the molecular behavior of the species (neutral, mono-positive and mono-negative ions) under study. The evaluation of the Fukui function values is not straightforward due to the difficulties associated with solving the derivatives. Therefore, Yang and Mortier [199] have given a simple procedure to calculate the atomic condensed Fukui function indices, and can be written as:

$$f_k^-(r) = q_k(N) - q_k(N-1) \quad \text{..... (20)}$$

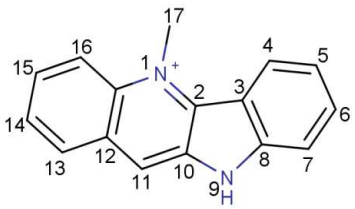
$$f_k^+(r) = q_k(N+1) - q_k(N) \quad \text{..... (21)}$$

$$f_k^0(r) = \frac{1}{2}(q_k(N+1) - q_k(N-1)) \quad \text{..... (22)}$$

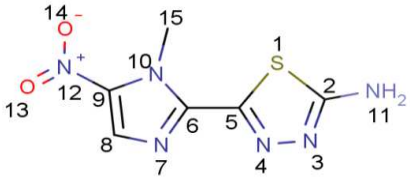
Where q_k (N) is the charge on atom k for N total electrons, f_k^- and f_k^+ describe the ability of an atom to accommodate an extra electron or to cope with the loss of an electron and f_k^0 is then considered as an indicator for radical reactivity [200].

The behavior of molecules as electrophiles/nucleophiles during reaction depends on the local behavior of molecules, i.e. on how the atomic sites of the molecule react toward the approaching reagent (electrophile, nucleophile and radical). Therefore, it is very important to infer the reactivity trends of local site during the electrophilic or nucleophilic attack.

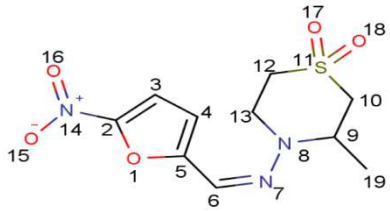
Table. 13: Structures and charges of studied antitrypanosomal compounds.



Cryptolepine



5-(1-methyl-5-nitro-1H-imidazol-2-yl) 1,3,4-thiadiazol-2-amine



Nifurtimox

	Atoms \ DFT/B3LYP	Mulliken			CHELPG		
		6-31G	6-311++G	Resd.	6-31G	6-311++G	Resd.
Cryptolepine	N1	-0.776	0.125	-0.901	-0.220	-0.272	0.052
	C2	0.213	1.150	-0.937	-0.057	-0.072	0.015
	C3	0.093	2.512	-2.419	0.014	0.038	-0.024
	C4	-0.133	-0.724	0.591	-0.135	-0.165	0.030
	C5	-0.141	-0.208	0.067	-0.155	-0.136	-0.019
	C6	-0.146	-0.657	0.511	-0.076	-0.087	0.011
	C7	-0.108	-1.683	1.575	-0.242	-0.258	0.016
	C8	0.257	-0.765	1.022	0.297	0.356	-0.059
	N9	-0.849	-0.571	-0.278	-0.688	-0.798	0.110
	C10	0.294	-0.488	0.782	0.373	0.455	-0.082
	C11	-0.109	-0.412	0.303	-0.432	-0.483	0.051
	C12	0.061	0.362	-0.301	0.162	0.179	-0.017
	C13	-0.195	-0.564	0.369	-0.190	-0.194	0.004
	C14	-0.119	-0.450	0.331	-0.102	-0.104	0.002
	C15	-0.139	-0.089	-0.050	-0.101	-0.098	-0.003
	C16	-0.152	-0.352	0.200	-0.210	-0.234	0.024
	C17	0.326	-0.348	0.674	0.156	0.196	-0.040
	C18	-0.252	-0.574	0.322	0.198	0.311	-0.113
5-(1-methyl-5-nitro-1H-imidazol-2-yl) 1,3,4-thiadiazol-2-amine	S1	0.393	0.210	0.183	-0.251	-0.197	-0.054
	C2	0.192	-0.648	0.840	0.796	0.943	-0.147
	N3	-0.261	0.422	-0.683	-0.406	-0.606	0.200
	N4	-0.175	-0.650	0.475	-0.315	-0.020	-0.295
	C5	-0.099	0.465	-0.564	0.370	0.063	0.307
	C6	0.398	-0.531	0.929	0.288	0.469	-0.181
	N7	-0.371	-0.275	-0.096	-0.588	-0.506	-0.082
	C8	0.040	0.915	-0.875	0.303	0.124	0.179
	C9	0.584	-0.031	0.615	-0.236	-0.006	-0.230
	N10	-0.625	-0.162	-0.463	0.078	-0.189	0.267
	N11	-0.745	-0.490	-0.255	-1.084	-1.176	0.092
	N12	0.020	-0.596	0.616	0.775	0.755	0.020
	O13	-0.289	0.038	-0.327	-0.452	-0.487	0.035
	O14	-0.310	0.062	-0.372	-0.454	-0.472	0.018
	C15	-0.233	-0.854	0.621	-0.147	0.109	-0.256

Table 13 (Continued)

		Mulliken			CHELPG		
	DFT/B3LYP / Atoms	6-31G	6-311++G	Resd.	6-31G	6-311++G	Resd.
Nifurtimox	O1	-0.494	-0.287	-0.207	-0.323	-0.345	0.022
	C2	0.556	-0.580	1.136	0.179	0.210	-0.031
	C3	-0.078	0.499	-0.577	-0.177	-0.172	-0.005
	C4	-0.158	0.255	-0.413	-0.153	-0.159	0.006
	C5	0.269	-0.098	0.367	0.300	0.315	-0.015
	C6	-0.046	-0.765	0.719	0.053	0.086	-0.033
	N7	-0.151	-0.492	0.341	-0.306	-0.325	0.019
	N8	-0.383	1.166	-1.549	-0.094	-0.142	0.048
	C9	0.021	-0.348	0.369	-0.093	-0.026	-0.067
	C10	-0.484	-0.980	0.496	0.124	0.036	0.088
	S11	1.142	0.553	0.589	0.724	0.831	-0.107
	C12	-0.515	-0.882	0.367	-0.115	-0.138	0.023
	C13	-0.109	-0.584	0.475	0.168	0.244	-0.076
	N14	0.029	-0.467	0.496	0.700	0.726	-0.026
	O15	-0.269	0.045	-0.314	-0.403	-0.424	0.021
	O16	-0.301	0.024	-0.325	-0.453	-0.482	0.029
	O17	-0.589	-0.276	-0.313	-0.524	-0.566	0.042
	O18	-0.585	-0.267	-0.318	-0.538	-0.587	0.049
	C19	-0.394	-1.031	0.637	-0.127	-0.068	-0.059

In this part, Fukui function indices for different antitrypanosomal compounds were calculated by CHELPG at B3LYPE/6-311 + + G level of theory, from the results obtained in Table 14, we can see that cryptolepine prefers nucleophilic, electrophilic and radical attack on C11 site, as shown from previous works of João et al [201]. where they have synthetized a series of 22 derivatives of cryptolepine by performing a substitution on this site C11 to enhance the biological activity against *T. Cruzi*, therefore, we can say that our results are in accordance to explain the possibility of this molecule to relate to its receptor via electrophilic, nucleophilic or radical susceptibility on C11 position to improve the affinity of the Ligand–receptor binding.

Table. 14: CHELPG charges and fukui parameters of Cryptolepine.

B3LYP/6311++G / Atoms	0	+1	-1	f°	f^{+}	f^{-}
N1	-0.272	-0.239	-0.296	0.028	0.033	0.024
C2	-0.072	0.121	-0.256	0.188	0.193	0.184
C3	0.038	-0.067	0.138	-0.102	-0.105	-0.100
C4	-0.165	-0.101	-0.259	0.079	0.064	0.094
C5	-0.136	-0.138	-0.124	-0.007	-0.002	- 0.012
C6	-0.087	0.018	-0.217	0.117	0.105	0.130
C7	-0.258	-0.268	-0.220	-0.024	-0.010	-0.038
C8	0.356	0.418	0.286	0.066	0.062	0.070
N9	-0.798	-0.749	-0.838	0.044	0.049	0.040
C10	0.455	0.366	0.562	-0.098	-0.089	-0.107
C11	-0.483	-0.267	-0.742	0.237	0.216	0.259
C12	0.179	0.102	0.343	-0.120	-0.077	-0.164
C13	-0.194	-0.131	-0.320	0.094	0.063	0.126
C14	-0.104	-0.071	-0.114	0.021	0.033	0.010
C15	-0.098	-0.010	-0.181	0.085	0.088	0.083
C16	-0.234	-0.242	-0.234	-0.004	-0.008	0.000
C17	0.196	0.242	0.139	0.051	0.046	0.057
C18	0.311	0.235	0.484	-0.124	-0.076	-0.173

From previous works of Noolvi et al [202]. Whereas, they synthetized a series of 29 compounds, their structure–activity relationship results showed that the Nitro group in all 1,3,4-thiadiazole-2-arylhydrazone derivatives is essential for the antitrypanosomal activity. Reduction or any chemical changes made to Nitro group will decrease the activity. In this investigation, we selected the most active compound from the studied 1,3,4-thiadiazole derivatives series against *T. Cruzi* which named 5-(1-methyl-5-nitro-1H-imidazol-2-yl)-1,3,4-thiadiazol-2-amine, from our results (Table 15) this molecule contains an amine group NH_2 characterized by an electrophilic site N_{11} which explains the nucleophilic susceptibility on this site where an incoming nucleophile is most likely to react. The favorable reactive site for nucleophilic attack explains that the hydrogen atoms of N11 are involved in intermolecular hydrogen bonding which can be directly related to the receptor.

Table. 15: CHELPG charges and fukui parameters of 5-(1-methyl-5-nitro-1H-imidazol-2-yl)-1,3,4-thiadiazol-2-amine.

B3LYP/6311++G / Atoms	0	+1	-1	f°	f^{+}	f^{-}
S1	-0.197	-0.184	-0.373	0.094	0.013	0.176
C2	0.943	0.915	0.891	0.012	-0.028	0.052
N3	-0.606	-0.362	-0.501	0.069	0.244	-0.105
N4	-0.020	-0.126	-0.369	0.121	-0.106	0.349
C5	0.063	0.251	0.401	-0.075	0.188	-0.338
C6	0.469	0.408	0.212	0.098	-0.061	0.257
N7	-0.506	-0.573	-0.665	0.046	-0.067	0.159
C8	0.124	0.342	0.241	0.050	0.218	-0.117
C9	-0.006	-0.024	-0.129	0.052	-0.018	0.123
N10	-0.189	0.027	0.107	-0.040	0.216	-0.296
N11	-1.176	-0.973	-1.176	0.101	0.203	0.000
N12	0.755	0.739	0.655	0.042	-0.016	0.100
O13	-0.487	-0.395	-0.659	0.132	0.092	0.172
O14	-0.472	-0.385	-0.672	0.143	0.087	0.200
C15	0.109	-0.424	-0.254	-0.085	-0.533	0.363

The nitro group NO_2 was found to be more nucleophile because of the higher electro-negativity in O14 and O13 sites which pulls the electron density to leaving the nitro group the favored site for electrophilic attack. Indeed, Fukui function indices values of these positions are relatively higher, according to the maximum values of the radical Fukui function indices, which explain that NO_2 atoms serve as potential sites of hydroxyl radicals HO° attack.

Besides, we found the same results for Nifurtimox as shown in table 16 where we observe the most nucleophilic sites are N_7 and N_8 which explain the electrophilic susceptibility in these positions. In addition, we note the nucleophilic susceptibility of the studied molecule identify C9, C4 and N8 sites where an incoming nucleophile is most likely to react.

Table. 16: CHELPG charges and fukui parameters of Nifurtimox.

Atoms \ B3LYP/6311++G	0	+1	-1	f°	f^+	f^-
O1	-0.345	-0.374	-0.444	0.035	-0.029	0.099
C2	0.210	0.388	0.408	-0.010	0.178	-0.198
C3	-0.172	-0.263	-0.370	0.053	-0.091	0.198
C4	-0.159	0.088	0.001	0.043	0.247	-0.160
C5	0.315	0.168	0.121	0.023	-0.147	0.194
C6	0.086	0.255	0.292	-0.018	0.169	-0.206
N7	-0.325	-0.385	-0.552	0.083	-0.060	0.227
N8	-0.142	0.060	-0.363	0.211	0.202	0.221
C9	-0.026	0.574	0.200	0.187	0.600	-0.226
C10	0.036	-0.365	-0.072	-0.146	-0.401	0.108
S11	0.831	0.891	0.872	0.009	0.060	-0.041
C12	-0.138	-0.119	-0.187	0.034	0.019	0.049
C13	0.244	0.110	0.381	-0.135	-0.134	-0.137
N14	0.726	0.676	0.552	0.062	-0.050	0.174
O15	-0.424	-0.356	-0.577	0.110	0.068	0.153
O16	-0.482	-0.397	-0.632	0.117	0.085	0.150
O17	-0.566	-0.615	-0.634	0.009	-0.049	0.068
O18	-0.587	-0.482	-0.613	0.065	0.105	0.026
C19	-0.068	-0.334	0.030	-0.182	-0.266	-0.098

Also, we observe the same preferable sites of radical attack in Nifurtimox at NO_2 which has been discussed in previews works of Juan et al. where they applied experimental research to study the mode of action of natural and synthetic drugs against *T. Cruzi* and their interaction with mammalian host, where they found that Nifurtimox acts through the formation of radical metabolites by Nitro group which is reduced to an Amino group by the action of Nitroreductases with the formation of various free radical intermediates and electrophilic metabolites [13].

Finally, our theoretical results are in good agreement with other experimental results [13, 201, 202], therefore, we can see the importance of Fukui function theory in understanding the reaction between antitrypanosomal compounds and its receptor *T. Cruzi*.

To visualize the reactivity at these sites, we provide the frontier molecular orbital theory and the molecular electrostatic potential surface in the next section.

3.3.4 FMO theory and MESP Surface Visualizations

The most important orbitals in a molecule are the frontier molecular orbitals, called highest occupied molecular orbital (HOMO) which explains the ionization energy and lowest unoccupied molecular orbital (LUMO) which explains the electron affinity in a molecule.

These orbitals determine the way the molecule interacts with other species [107] including the interactions between a ligand and its receptor. To enhance the information about the reactivity of a molecule, the molecular electrostatic potential is an important parameter. Whereas, the MESP lies in the fact that it simultaneously displays the molecular size and shape as well as positive, negative and neutral electrostatic potential regions in terms of the electrostatic surface, which explain the investigation of the molecular structure with its physiochemical property relationships [13, 203, 204].

In this part, we use at first, the HOMO and LUMO respectively to visualize the information about the reactivity of specific regions of the studied molecules. In the same way, we use the second Highest Occupied and Unoccupied (HOMO-1 and LUMO+1, respectively), whereas, these orbitals are lower in energy than HOMO and LUMO, then we use the 3D MESP surface map to visualize the variation in electrostatic potential of these trypanocidal compounds under study (Fig. 32).

Figure 32 displays a three-dimensional (3D) visualization of the frontier molecular orbitals and the MESP plots at B3LYP/6-311 + + G of the antitrypanosomal compounds. For Cryptolepine HOMO and LUMO localized in the same atoms except nitrogen atom in NH. Further, we demonstrate the existence of the delocalization of the conjugated-electron system. For 1,3,4-thiadiazol derivative and Nifurtimox, HOMO and LUMO are mainly localized in the same regions, and are in accordance with Fukui function analysis where we observe the shape density and localization of LUMO on the nitro group NO_2 and HOMO on nitrogen atoms for both molecules. However, the second HOMOs and LUMOs are slightly different, the HOMO-1 and LUMO+1 are located in proximity of the HOMO and LUMO.

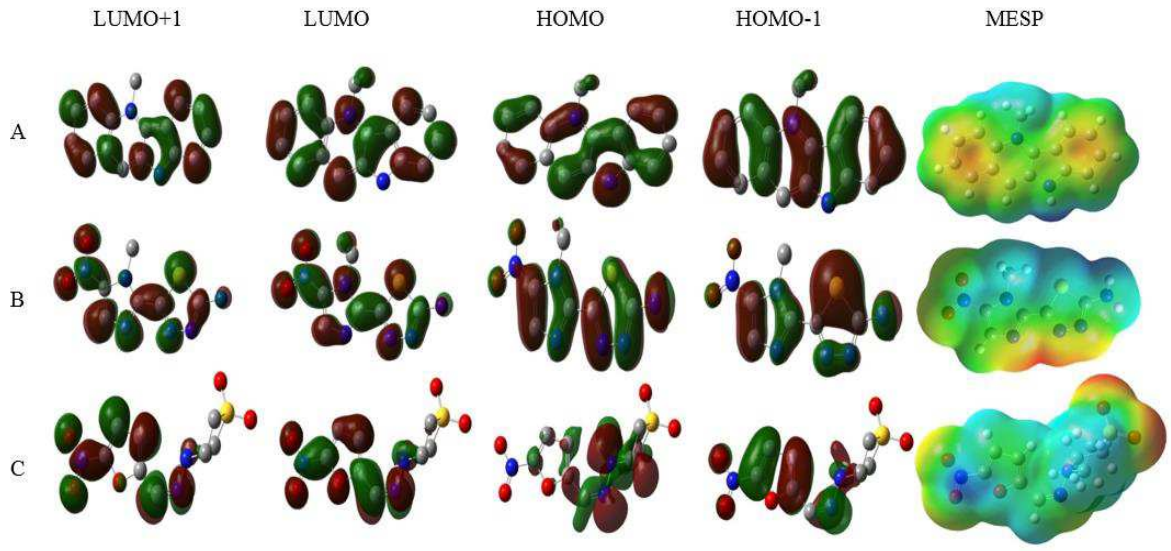


Fig. 32: Frontier molecular orbitals and MESP plots of antitrypanosomal compounds, A: Cryptolepine, B: 5-(1-methyl-5-nitro-1H-imidazol-2-yl)-1,3,4-thiadiazol-2-amine, C: Nifurtimox at B3LYP/6-311 + + G.

The MESP surface map of the studied molecules (Fig. 32) shows the two regions characterized by red color (negative electrostatic potential) on the two extremity rings of Cryptolepine, where the electrophilic attack in these sites is privileged, also by blue color (positive electrostatic potential) around the methyl and hydrogen atoms which explain that these regions are susceptible to a nucleophilic attack also the green color around the molecule which explains the neutral electrostatic potential surface for Cryptolepine molecule. This variation in the electrostatic potential, produced by a molecule, is largely responsible for binding of a drug to its active sites (receptor) [205], as the binding site of *T. Cruzi* in general is expected to have negative areas of electrostatic potential. However 1,3,4-thiadiazole derivative and Nifurtimox contain the negative electrostatic potential around the nitro group for both molecules and nitrogen atoms on 1,3,4-thiadiazole derivative molecule this explain the ability for an electrophilic attack on these positions as shown from Fukui function analysis. Finally, we can say that this visualization provides good information about the reactivity, which is in accordance with the experimental and theoretical results explained in Sec. 3.3.3.

Although, enhanced reactivity at these sites can be affected by steric hindrance effects, which is not encapsulated in the definition of the reactivity indicators, for this reason we perform the Molecular Docking study in the next section for the studied molecules with its receptor *T. Cruzi*.

3.3.5 Molecular Docking of T. Cruzi Inhibitors

To explore the protein–ligand interactions, we perform the Molecular Docking protocol to study the interactions at the active site of the crystal structure of *Cruzain* (PDB 2OZ2 with resolution: 1.95 Å and R-value: 0.159 (obs.)) [206] carrying out three different ligands that have been previously reported in the Conceptual DFT analysis. These compounds act as inhibitors of the cysteine proteases *T. Cruzi*, the responsible in pathogenesis and host cell invasion.

The Molecular Docking was performed by Molegro Virtual Docker software [113] using the crystal structure of *Cruzain*, as shown from figure 33, *Cruzain* contains two domain fold of papain superfamily cysteine proteases, characterized by a hydrophobic cleft and a very large hydrophobic pocket with a volume of 800.768 Å. Whereas, the reference ligand (vinyl sulfone derived inhibitor) is surrounded in the largest cavity (green color) in the active site of *T. Cruzi* (Fig. 33).

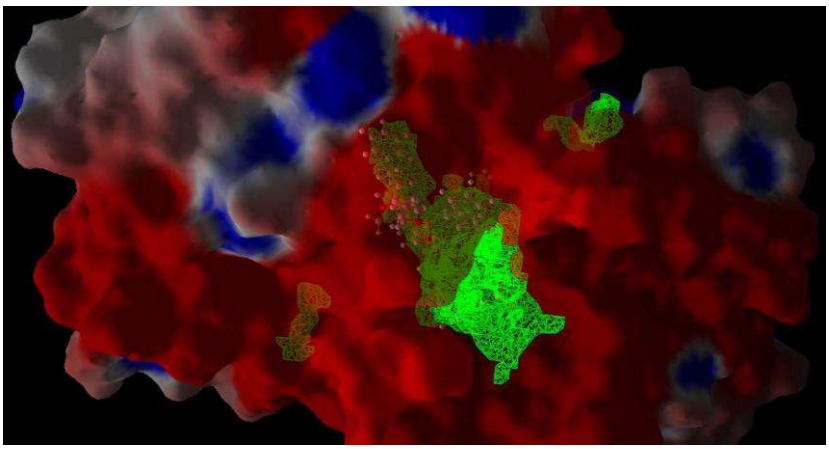


Fig. 33: The crystal structure of *Cruzain* bound to the reference ligand 'vinyl sulfone derived inhibitor'.

In order to perform optimally the Molecular Docking analysis of the studied compounds using MVD, the molecules in the workspace must be properly prepared before the docking begins; in this study, we use the built-in preparation method available in MVD, starting with the preparation of the protein, which is taken from the RCSB protein data bank [206, 207] where all structural information such as bond orders, flexible torsions in ligand and hybridization were assigned beforehand. In order to predict cavities that are responsible for the ligand–receptor binding, we create a protein surface based on the default settings that is an opaque solvent accessible surface colored according to the electrostatic potential.

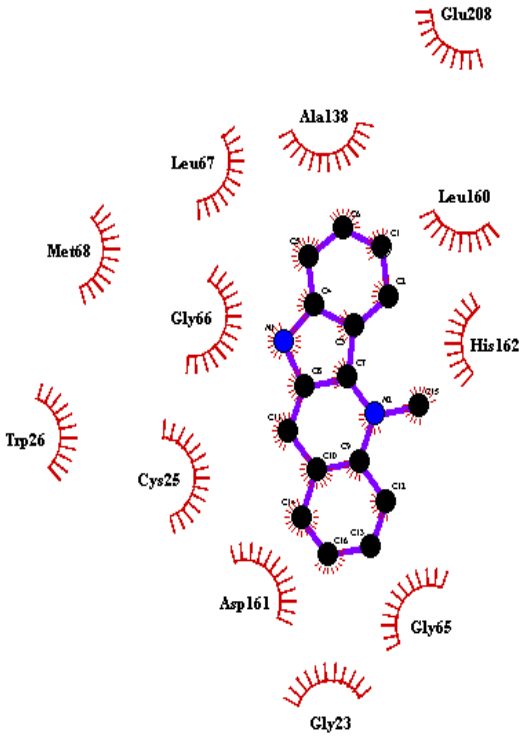
Finally, we use the Docking Wizard to determine the different interactions between the ligand and its receptor, after the preparation of the studied antitrypanosomal compounds, the

most important thing in this part is to set the binding site, which is the largest hydrophobic pocket (800.768 Å) where the best poses with the smallest values of MolDock score (GRID) were found to take -92.5603 for cryptolepine, -105.472 for 1,3,4-thiadiazol derivative and -99.309 for nifurtimox, results are shown in Fig. 34.

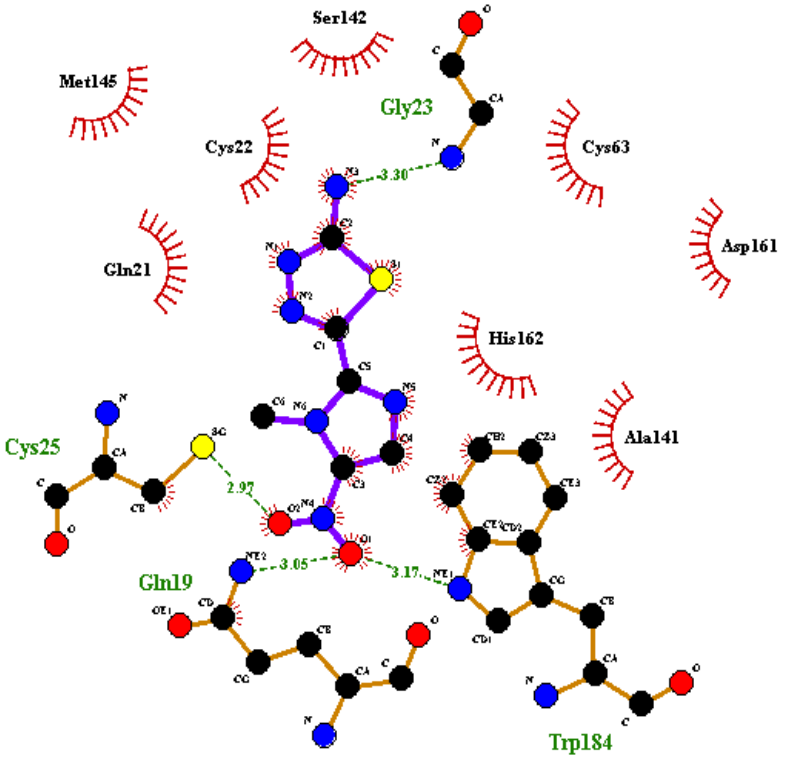
Analyses by Ligplot+ program for visualization of ligand–receptor interactions (Fig. 34) of the *Cruzain* structure with the studied antitrypanosomal compounds, which present the lowest docking energy, show that Cryptolepine molecule share generally the same amino acids residues as in Nifurtimox Molecule, where Cryptolepine molecule shows that Met68, Trp26, Gly66, Asp161, Gly65, Leu160 and Ala138 are involved in hydrophobic contacts. It is related to the active site only by hydrophobic interactions, this makes cryptolepine less selective and has a poor binding affinity with *Cruzain*, from Conceptual DFT results, we found that C11 prefers nucleophilic, electrophilic and radical attack which explain the ability of this Ligand to be enhanced in this concept against *T. Cruzi* by substitution on the C11 position. That was confirmed by João et al [201]. where the substitution on this site C11 was realized to enhance the biological activity against *T. Cruzi*.

Unlike, 1,3,4-thiadiazole derivative shows polar interactions involving strong hydrogen bonding [207] with different amino acids residues Gys25 (2.97 Å), Gln19 (3.05 Å) and Trp184 (3.17 Å) and medium hydrogen bonding with Gly23 (3.30 Å). From the DFT results, the favorable reactive site for nucleophilic attack on N11 explains that the hydrogen atoms of N11 are involved in intermolecular hydrogen bonding where we confirm this from Docking results by the hydrogen interactions with the amino acid Gly23. In the same way, from DFT results, the nitro group NO_2 has found to be more nucleophile because of the higher electro-negativity in O13 and O14 where, we notice from Molecular Docking results, this nitro group interacts with several amino acids by strong hydrogen interactions.

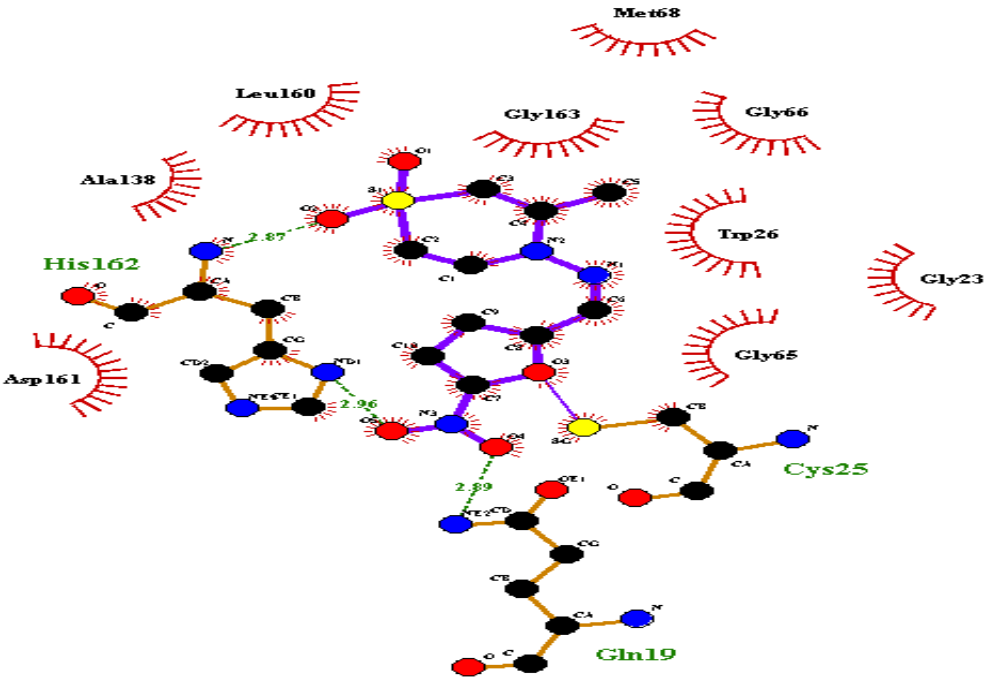
However, Nifurtimox is the special case in the studied compounds, this ligand shows strong hydrogen interactions with Gln19 and His162, and covalent bonding to Cys25. As known, several chemical forces may result in a temporary binding of the drug to its receptor. Covalent bonds would be very tight and practically irreversible. Since, by definition, the drug–receptor interaction is reversible, covalent bond formation is rather rare except in a rather toxic situation.



Cryptolepine



5-(1-methyl-5-nitro-1H-imidazol-2-yl) -1,3,4-thiadiazol-2-amine



Nifurtimox

Fig. 34: Visualization of *Trypanosoma Cruzain* interactions with best poses of studied trypanocidal compounds.

In recent literature, there has been a growing interest in the design of drugs forming a covalent bond with the target protein, with nearly 30% of the marketed drugs targeting enzymes known to act by covalent inhibition [208, 209]. These types of inhibitors derive their activity from both non-covalent interactions and the formation of the covalent bond between the inhibitor and the target protein [210-215].

Covalent interaction with the target protein has the beneffit of prolonged duration of the biological effect. However, these types of inhibitors tend to be associated with toxicity because of the difficulty of disassociation if off-target binding happens [216].

The sulfur atom of cysteine (Cys) provides a considerable range of chemical reactivity, a remarkable range of biologic functions is supported by Cys because sulfur is stable in multiple coordinate covalent bonds with the major atoms (C, H, O, N, P) [217].

The covalent adduct with the sulfur of the active site cysteine thiol in this enzyme is shown in Fig. 34. Whereas, Maya et al. proved experimentally that the covalent bonding to macromolecules is involved in trypanocidal effects [13].

GENERAL CONCLUSION

From the results discussed above, our work is subsequently focused on the study of molecular geometry and electronic properties, which are largely responsible for binding of a drug to its active sites; this allowed us to predict the influence of certain structural modifications on the biological activity against Trypanosomes.

The first part, offers the ability to guide design and selection to quickly identify compounds from the 1,3,4-thiadiazole derivatives series with higher potency and a good balance of properties. Also, it provides a discussion of several qualitative approximations of the structure activity/property relationship to search the preferred conformations and comparing the antitrypanosomal activities against *Cruzain* with 1,3,4- thiadiazole derivatives to establish correlations between structural parameters and the various properties of the investigated molecules. Whereas, the volume and lipophilicity are important parameters developing the response of Trypanocidal compounds. Therefore, we use the most effective parameters as descriptors in the second part to develop QSAR models.

In the second part, accurate mathematical models were developed for predicting the antitrypanosomal activity against *Cruzain* and *Rhodesain*. The validity of the models has been established by the determination of suitable statistical parameters. Henceforth the developed QSAR models were used to predict the antitrypanosomal activity of the investigated cryptolepines and close agreement between experimental and predicted values has been achieved. The low residual activity and high cross-validated values obtained suggests a good predictive ability of the developed QSAR models. It indicates that the activity studied on cryptolepine derivatives series can be successfully modeled using various molecular descriptors. The developed QSAR models show that hydrophilic derivatives of cryptolepines give a good antitrypanosomal activity against *Cruzain* and *Rhodesain*.

Afterward, our QSAR models were applied on different molecules. But, more informations about inhibition mechanism were needed. Therefore, we applied the conceptual DFT and Molecular Docking combination approach to extract more informations for understanding *Cruzain* inhibition.

In the third part of our investigation, we started with a comparison between Mulliken and CHELPG atomic charge methods, where we found that CHELPG method gives results more accurate and it is the best approximation for studying the electronic properties of the Trypanocidal compounds. The present part provides two different axes, Conceptual DFT and Molecular Docking, which were found to be in complementary for each other in understanding

the Ligand–receptor binding. Whereas, we applied the Fukui function theory as a site reactivity descriptor of Cryptolepine, where the C11 position is susceptible for nucleophilic, electrophilic or radical substitution to improve the affinity of the ligand–receptor binding. 1,3,4-thiadiazol derivative is characterized by an electrophilic site N11 which explains this amine group involved in intermolecular hydrogen bonding that can be directly related to its receptor. Also, this trypanocidal molecule is characterized by a nitro group that was found to be more nucleophilic, which explains that NO_2 serves as a potential site of hydroxyl radicals OH° attack. In the same way, we observe the same preferable sites of radical attack in Nifurtimox at NO_2. The FMO theory and MESP surface visualizations offer good information about the reactivity at these sites that are found to be in accordance with Fikui function theory results and previous experimental works derived from literature.

Although, enhanced reactivity at these sites can be affected by steric hindrance effects, therefore, the Molecular Docking analysis was performed in understanding the binding of these Trypanocidal compounds with the cysteine protease *T. Cruzi*. The best poses, with the smallest MolDock score, binding with *Cruzain* showed close agreement with Conceptual DFT and experimental results.

This explains the possibility of studying the receptor-binding mode using these ideas to guide design, to quickly identify more selective trypanocidal compounds that are likely to achieve outcomes in clinical trials, decrease the toxic side and achieve parasitological cure. In addition, this helps to select the most important properties that are needed to be enhanced and use them as descriptors in QSAR projects and virtual screening of large databases.

Finally, we can say that:

This atomic level information can provide insight into the design of corroborative laboratory research and thereby help investigators discern corresponding molecular sites and mechanisms of action.

References

[1]. S. C. Basak, Chemo-bio-informatics: the advancing frontier of computer-aided drug design in the post-genomic era, *Curr Comput Aided Drug Des* 8: 1, **2012**.

[2]. M. Malagoli, J. L. Bredas, Density functional theory study of the geometric structure and energetics of triphenylamine-based hole-transporting molecules, *Chemical Physics Letters* 327: 13, **2000**.

[3]. L. Eriksson, E. Johansson, Multivariate design and modeling in QSAR, *Chemometrics and Intelligent Laboratory Systems* 34: 1, **1996**.

[4]. S. Sakkiah, S. Thangapandian, S. John, Y. J. Kwon, K. W. Lee, 3D QSAR pharmacophore based virtual screening and molecular docking for identification of potential HSP90 inhibitors, *European Journal of Medicinal Chemistry* 45: 2132, **2010**.

[5]. D. Seeliger, B. L. de Groot, Ligand docking and binding site analysis with PyMOL and Autodock/Vina, *Journal of Computer-Aided Molecular Design* 24: 417, **2010**.

[6]. H. V. de Waterbeemd and E. Gifford, ADMET in silico modelling: towards prediction paradise?, *Nature Reviews Drug Discovery* 2: 192, **2003**.

[7]. J. Dargie, Bulletin d'information sur les glossines et les trypanosomoses, Organisation des nations unies pour l'alimentation et l'agriculture, Rome, **2009**.

[8]. F. Chappuis, S. Sundar, A. Hailu, H. Ghalib, S. Rijal, R. W. Peeling, J. Alvar and M. Boelaert, Visceral leishmaniasis: what are the needs for diagnosis, treatment and control?, *Nat Rev Microbiol* 5: 873, **2007**.

[9]. D. R. Moreira, A. C. Leite, R. R. dos Santos and M. B. Soares, Approaches for the development of new anti-Trypanosoma cruzi agents, *Curr Drug Targets* 10: 212, **2009**.

[10]. S. L. Croft, M. P. Barrett and J. A. Urbina, Chemotherapy of trypanosomiases and leishmaniasis, *Trends in Parasitol* 21: 508, **2005**.

[11]. T. Salah, S. Belaidi, N. Melkemi, I. Daoud and S. Boughdiri, In silico investigation by conceptual DFT and molecular docking of Antitrypanosomal compounds for understanding Cruzain inhibition, *Journal of Theoretical and Computational Chemistry* 15: 17, **2016**.

[12]. J. Lavrado, A. Paulo, J. Gut, P. J. Rosenthal and R. Moreira, Cryptolepine analogues containing basic aminoalkyl side-chains at C-11: Synthesis, antiplasmodial activity, and cytotoxicity, *Bioorganic & Medicinal Chemistry Letters* 18: 1378, **2008**.

[13]. D. M. Juan, K. C. Bruce, I. V. Patricio, F. Jorge, F. Mario, G. Norbel, F. Arturo, M. Antonio, Mode of action of natural and synthetic drugs against Trypanosoma Cruzi and their interaction with the mammalian host, *Comparative Biochem Physiol A* 146: 601, **2007**.

[14]. W. E. Sneader, Drug discovery (the history), John Wiley & Sons, New York Inc. **2007**.

[15]. J. Drews, In quest of tomorrow's medicines, Springer-verlag New York Inc. **1999**.

[16]. F. J. Moore, A history of chemistry, McGraw-Hill New York Inc. **1918**.

[17]. A. M. Roberts, Serendipity, Wiley New York Inc. **1989**.

[18]. P. Ehrlich, In gesammelte arbeiten, Springer-Verlag Inc. **1957**.

[19] F. W. Serturner, Ueber das morphium, und die mekonsäure, als hauptbestandteile des opiums, annalen der physik 55: 56, **1817**.

[20]. W. Sneader, Drug discovery, Wiley New York Inc. **1985**.

[21]. E. Farber, The evolution of chemistry; a History of its ideas, methods and Materials, Ronald Press New York Inc. **1952**.

[22]. J. Koch-Weser and P. J. Schechter, the making of modern pharmacology, *Life Sci* 22: 1361, **1978**.

[23]. H. C. Peyer, Roche-Geschichte einer unternehmung, Roche Basel Inc. **1996**.

[24]. J. drews, Drug Discovery: A historical perspective, *Drug discovery reviews* 287: 1960, **2000**.

[25]. European Federation of Pharmaceutical Industries and Associations, EFPIA Brussels, **2010**.

[26]. E. Czerepak and S. Ryser, Drug approvals and failures: implications for alliances. *Nature Rev Drug Discovery* 7: 197, **2008**.

[27]. A. V. Zakharov, A. A. Lagunin, D. A. Filimonov and V. V. Poroikov, Quantitative prediction of antitarget interaction profiles for chemical compounds, *Chem Res Toxicol* 25: 2378, **2012**.

[28]. W. L. Jorgensen, The many roles of computation in drug discovery, *Science* 303: 1813, **2004**.

[29]. S. Shailza, K. M. Balwant and K. S. Durlabh, Molecular drug targets and structure based drug design: A holistic approach, *Bioinformatics* 1: 2006, **2006**.

[30]. T. Bredow, Introduction to computational chemistry, Frank Neese Inc. **2007**.

[31]. I. M. Kapetanovic, Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach, *Chem Biol Interact* 30: 165, **2008**.

[32]. P.G. Jansens, M. Kivits and J. P. Vuylsteke, Medicine et hygiene en Afrique Centrale de 1885 à nos jours. *Revue française d'histoire d'outre-mer* 82 : 112, **1992**.

[33]. M. Wéry, Protozoologie médicale, Institut de médecine tropicale prince léopold, Antwerpen (Belgique), **1983**.

[34]. Official website of World Health Organization : http://www.who.int

[35]. C. Boda, Contribution des modèles expérimentaux dans l'étude des trypanosomoses africaines, Université de Limoges (France), **2005**.

[36]. A. Stich, P. M. Abel and S. Krishna, Human African trypanosomiasis, *BMJ* 325: 203, **2002**.

[37]. I. Balakrishnan and S. H. Gillespie, John Wiley & Sons, New York Inc. **2002**.

[38]. T. Al-Fozan, P. Manuel, I. Rajasingh, and R. S. Rajan, A New Technique to Compute Padmakar-Ivan Index and Szeged Index of Pericondensed Benzenoid Graphs, *J Comput Theor Nanosci* 11: 533, **2014**.

[39]. J. R. C. Carvalho, J. E. V. Ferreira, J. P. Barbosa, M. S. Lobato et al., Computational Modeling of Artemisinins with Antileishmanial Activity, *J Comput Theor Nanosci* 8: 2193, **2011**.

[40]. Pere P Simarro, Giuliano Cecchi, Massimo Paone, José R Franco1 et al., The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases, *International Journal of Health Geographics* 9: 57, **2010**.

[41]. G. Federica, Dissertation, New approaches to fluorescence-based diagnostics for Human African Trypanosomiasis. PhD thesis. University of Glasgow, **2010**.

[42]. P. N. Acha and B. Szyfres, Zoonoses and communicable diseases common to man and animals, Pan American Health Organization (PAHO), Washington, **2003**.

[43]. S. A. Kjos, K. F. Snowden, T. M. Craig, B. Lewis, N. Ronald and J. K. Olson, Distribution and characterization of canine Chagas disease in Texas, *Vet Parasitol* 152: 249, **2008**.

[44]. L. V. Kirchhoff, Chagas disease American trypanosomiasis, *Infect Dis Clin North Am* 7: 487, **1993**.

[45]. J. C. Dias, Elimination of Chagas disease transmission: perspectives. *Mem Inst Oswaldo Cruz* 104: 41, **2009**.

[46]. F. E. G. Cox, History of Human Parasitology, *Clin Microbiol Rev* 15: 595, **2002** .

[47]. C. Dennis, The human genome. *Nature* 409 :813, **2001**.

[48]. S. A. Tishkoff, A. J. Pakstis, M. Stoneking, J. R. Kidd, G. Destro-Bisol et al., Short tandem-repeat polymorphism/alu haplotype variation at the PLAT locus: implications for modern human origins. *Am J Hum Genet* 67: 901, **2001**.

[49]. A. R. **Templeton,** Out of Africa again and again. *Nature* 416**:** 45, **2002**.

[50]. M. M. Parida, S. D. Sannarangaiah, P. K. Dash, P. V. L. Rao and K. Morita, Loop mediated isothermal amplification (LAMP): a new generation of innovative gene amplification technique; perspectives in clinical diagnosis of infectious diseases, *Reviews in Medical Virology*, 18: 407, **2008**.

[51]. W. P. Price, General concepts on the evolutionary biology of parasites, *Society for the Study of evolution stable*, 31: 405, **1977**.

[52]. W. P. Price, Evolutionary biology of parasites, Princeton university press, the united kingdom, **1980**.

[53]. D. Assaf, E. Kibru, S. Nagesh, S. Gebreselassie and F. Deribe, Medical Parasitology, Lecture Notes, Jimma University, **2004**.

[54]. J. Moore, Parasites and the behavior of animals, Oxford Series in Ecology and Evolution, Oxford university press Inc. **2002**.

[55]. O. A. Bush, C. J. Fernandez, W. G. Esch and J. R. Seed, the diversity and ecology of animal parasites, Cambridge university Press, **2001**.

[56]. F. E. G. Cox, Taxonomy and classification of human parasitic protozoa and helminths. In: Manual of clinical microbiology. Section VIII. Parasitology. Washington, **2011**.

[57]. S. CAVALIER, Kingdom Protozoa and Its 18 Phyla, *microbiological reviews, American Society for Microbiology* 57: 953, **1993**.

[58]. T. Salah, Dissertation, Développement des modèles QSAR pour la prédiction des activités inhibitrices antitrypanosomiennes des dérivés de cryptolepine, MS thesis BISKRA University, **2013**.

[59]. W. C. Duffy, L. MacLean and L. Sweeney, A. Cooper, R. Turner, A. Tait and A. MacLeod, Population Genetics of *Trypanosoma Brucei Rhodesiense*: Clonality and Diversity within and between Foci, *PLoS Negl Trop Dis* 7: 2526, **2013**.

[60]. A. C. ROUZER and A. CERAMI, Hypertriglyceridemia associated with Trypanosoma Brucei infection in rabbits: Role of defective triglyceride removal, *Mol Biochem Parasitol* 2: 31, **1980**.

[61]. P. Truc, About Trypanosoma Brucei Gambiense, the causative agent of the chronic form of Human African Trypanosomiasis: some findings and proposals, *African Journal of Biotechnology* 2: 657, **2003**.

[62]. L. Alphey, M. Andreasen, Dominant lethality and insect population control, *Mol Biochem Parasitol* 121: 173, **2002**.

[63]. C. E. Reisenman, G. Lawrence, P. G. Guerenstein, T. Gregory, E. Dotson and J. G. Hildebrand, Infection of kissing bugs with Trypanosoma Cruzi, Tucson, Arizona, USA, *Emerging Infectious Diseases* 16: 400, **2010**.

[64]. Z. Brener, Biology of Trypanosoma Cruz! *Annu Rev Microbiol* 27: 347, **1973**.

[65]. A. D. McNaught and A. Wilkinson, IUPAC. Compendium of Chemical Terminology, Blackwell Scientific Publications, Oxford, **1997**.

[66]. G. L. Boye, O. Ampofo, Clinical uses of Cryptolepis sanguinolenta, University of Science and Technology, Ghana, **1983**.

[67]. A. Paulo and A. Duarte, In vitro antibacterial screening of Cryptolepis sanguinolenta alkaloids, *J Ethnopharm* 44: 127, **1994**.

[68]. A. Paulo, E. T. Gomes, J. Steele, D. C. Warhurst and P. Houghton, Antiplasmodial activity of Cryptolepis sanguinolenta alkaloids from leaves and roots, *J Planta Med* 66: 30, **2000**.

[69]. K. Bonjean, M. C De Pauw-Gillet, M. P. Defresne, P. Colson, C. Houssier et al., The DNA Intercalating Alkaloid Cryptolepine Interferes with Topoisomerase II and Inhibits Primarily DNA Synthesis in B16 Melanoma Cells, *Biochemistry* 37: 5136, **1998**.

[70]. L. Dassonneville, K. Bonjean, M. C. De Pauw-Gillet, P. Colson et al., Stimulation of Topoisomerase II-Mediated DNA Cleavage by Three DNA-Intercalating Plant Alkaloids: Cryptolepine, Matadine, and Serpentine, *Biochemistry* 38: 7719, **1999**.

[71]. J. N. Lisgarten, M. Coll, J. Portugal, C. W. Wright, The antimalarial and cytotoxic drug cryptolepine intercalates into DNA at cytosine-cytosine sites, *J Nat Struct Biol* 9: 57, **2002**.

[72]. A. K. Jain, S. Sharma, A. Vaidya, V. Ravichandran and R. K. Agrawal, 1,3,4-Thiadiazole and its Derivatives: A Review on Recent Progress in Biological Activities, *Chem Biol Drug Des* 81: 557, **2013**.

[73]. S. A. Carvalho, F. A. S. Lopes, K. Salomao, N. C. Romeiro, S. M. S. V. Wardell, S. L. De Castro, E. F. Da Silva and C. A. M. Fraga, Studies toward the structural optimization of new brazilizone-related trypanocidal 1,3,4-thiadiazole-2-arylhydrazone derivatives, *Bioorg Med Chem* 16: 413, **2008**.

[74]. M. N. Noolvi, H. M. Patel and N. S. Sethi, 2D-QSAR Study of 1,3,4-thiadiazole-2-arylhydrazone Derivatives: An Approach to Design Antitrypanosomal Agents. *International Journal of Drug Design and Discovery* 1: 195, **2010**.

[75]. J. Pepin, F. Milord, F. Meurice, L. Ethier, L. Loko and B. Mpia, High-dose Nifurtimox for arseno-resistant Trypanosoma Brucei gambiense sleeping sickness: An open trial in central Zaire, *Trans R Soc Trop Med Hyg* 86: 254, **1992**.

[76]. M. H. Tsuhako, M. J. M. Alves, W. Colli, L. S. Filardi, Z. Brener and O. Augusto, Comparative studies of Nifurtimox uptake and metabolism by drug-resistant and susceptible strains of Trypanosoma Cruzi. *Comp Biochem Physiol C*, 99: 317, **1991**.

[77]. R. Docampo, Sensitivity of parasites to free radical damage by antiparasitic drugs, *Chem Biol Interact* 73: 1, **1990**.

[78]. S. N. Moreno, R. Docampo, R. P. Mason, W. León and A. O. Stoppani, Different behaviors of benznidazole as free radical generator with mammalian and Trypanosoma Cruzi microsomal preparations. *Arch Biochem Biophys* 218: 585, **1982**.

[79]. N. J. Temperton, S. R. Wilkinson, D. J. Meyer and J. M. Kelly, Overexpression of superoxide dismutase in Trypanosoma Cruzi results in increased sensitivity to the trypanocidal agents gentian violet and benznidazole. *Mol Biochem Parasitol* 96: 167, **1998**.

[80]. E. G. Díaz de Toranzo, J. A. Castro, B. M. Franke de Cazzulo and J. J. Cazzulo, Interaction of Benznidazole reactive metabolites with nuclear and kinetoplastic DNA, proteins and lipids from Trypanosoma Cruzi, *Experientia* 44: 880, **1988**.

[81]. R. Docampo and A. O. Stoppani, Generation of superoxide anion and hydrogen peroxide induced by Nifurtimox in Trypanosoma Cruzi, *Arch Biochem Biophys* 197: 317, **1979**.

[82]. R. Docampo and S. N. Moreno, Free radical metabolites in the mode of action of chemotherapeutic agents and phagocytic cells on Trypanosoma Cruzi, *Rev Infect Dis* 6: 223, **1984**.

[83]. R. Docampo and A. O. Stoppani, Mechanism of the Trypanocidal action of Nifurtimox and other Nitro-derivatives on Trypanosoma Cruzi, *Medicina (B Aires)* 1: 10, **1980**.

[84]. T. J. Davin, Dissertation, Computational chemistry of organometallic and inorganic species. PhD thesis. University of Glasgow, **2009**.

[85]. A. Hinchliffe, Molecular Modelling for Beginners, Wiley Manchester Inc. **2003**.

[86]. Official website of Cambridge Crystallographic Data Centre (CCDC) database: http://www.ccdc.cam.ac.uk.

[87]. Official website of Protein Data Bank (PDB) : www.rcsb.org/pdb.

[88]. MarvinSketch 6.2.1 (2014), Chemaxon: http://www.chemaxon.com.

[89]. HyperChem 8.0 (Molecular Modeling System), Hypercube Inc. **2008**.

[90]. M. J. Frisch, G. W. Trucks, H. B. Schlegel et al. Gaussian 09, Gaussian Inc. Wallingford, CT, **2010**.

[91]. SPSS 19 For Windows, SPSS software packages, SPSS Inc., 444 North Michigan Avenue, Suite 3000, Chicago, Illinoi, 60611, USA.

[92]. T. Rene, M. H. Mikael, MolDock: A new technique for high-accuracy molecular docking, *J Med Chem* 49: 3315, **2006**.

[93]. W. J. Hehre, A Guide to Molecular Mechanics and Quantum Chemical Calculations, Wavefunction Inc. USA, **2003**.

[94]. P. Atkins, Molecular quantum mechanics, Oxford university inc. **2005**.

[95]. A. Kerassa, Dissertation, Etude par la modélisation moléculaire des relations structures-propriétés de quelques séries hétérocycliques bioactives, PhD thesis, BISKRA University, **2015**.

[96]. N. L. Allinger, Y. H. Yuh and J. H. Lii, Molecular mechanics, *J Am Chem Soc* 111: 8551, **1989**.

[97]. F. L. Hirshfeld, Electron density distributions in molecules, *Cryst Rev* 2: 169, **1991**.

[98]. H. Bernhard Schlegel, Geometry optimization, *Computational molecular science* 1: 790, **2011**.

[99]. D. C. Young, Computational chemistry, A practical guide for applying techniques to real-world problems, Cytoclonal pharmaceutics Inc. **2001**.

[100]. O. Prasad, L. Sinha and K. Naveen, Theoretical Raman and IR spectra of tegafur and comparison of molecular electrostatic potential surfaces, polarizability and hyerpolarizability of tegafur with 5-fluoro-uracil by density functional theory, *J At Mol Sci* 1: 201, 2010.

[101]. A. Chunzhi, C. Y. Li, W. Yonghua, C. Yadong and Y. Ling, Insight into the effects of chiral isomers quinidine and quinine on CYP2D6 inhibition, *J Bioorg Med Chem Lett* 19: 803, **2009**.

[102]. K. S. Rajesh and V. Narayan, Vibrational, structural and electronic properties of 6-methyl nicotinic acid by density functional theory, *J chem pharm res* 4, 3287 **2012**.

[103]. K. B. Wiberg and P. R. Rablen, Comparison of atomic charges derived via different procedures, *Comput Chem* 14: 1504, **1993**.

[104]. J. G. Angyan, C. Chipot, Choosing between alternative MP2 algorithms in the self- consistent reaction field (SCRF) theory of solvent effects, *Chem Phys Lett* 241: 51, **1995**.

[105]. E. Sigfridsson, U. Ryde, Comparison of methods for deriving atomic charges from the electrostatic potential and moments, *J Comput Chem* 19: 377, **1998**.

[106]. M. L. Lakhanpal, Fundamentals of chemical thermodynamics, Tata McGraw-Hill Inc. **1983**.

[107]. B. B. Howard and A. W. Ian, Selection of heterocycles for drug design, *J Mol Graphics Modelling* 23: 51, **2004**.

[108]. I. lessigiarska, Dissertation, development of structure-activity relationships for pharmaco toxicological endpoints relevant to european union legislation, Liverpool John Moores University, **2006**.

[109]. T. Ooi, M. Oobatake, G. Nemethy and H. A. Scheraga, Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides, *Proc Natl Acad Sci USA* 84: 3086, **1987**.

[110]. Y. Cohen, Pharmacologie moléculaire, Collection de biologie moléculaire, Masson, Paris, **1978**.

[111]. Official website: waterloo.ca/science.

[112]. S. Belaidi, A. Dibi, and M. Omari, A conformational exploration of dissymmetric macrolides antibiotics, *Turkish Journal of Chemistry*, 26: 491, **2002**.

[113]. S. Belaidi, N. Melkemi and D. Bouzidi, Molecular geometry and structure-property relationships for 1, 2-dithiole-3-thione derivatives, *Int J Chem Res* 4: 134, **2012**.

[114]. F. J. Luque, M. Orozco, P. K. Bhadane and S. R. Gadre, SCRF calculation of the effect of water on the topology of the molecular electrostatic potential, *J Phys Chem* 97: 9380, **1993**.

[115]. J. S. Murray, K. Sen, Molecular electrostatic potentials, concepts and applications, Elsevier, Amsterdam, **1996**.

[116]. R. K. Pathak, S. R. Gadre, Maximal and minimal characteristics of molecular electrostatic potentials, *J Chem Phys* 93: 1770, **1990**.

[117]. F. Bazooyar, M. Taherzadeh, C. Niklasson and K. Bolton, Molecular Modelling of Cellulose Dissolution, *J Comput Theor Nanosci* 10: 2639, **2013**.

[118]. Official website: formules-physique.com.

[119]. T. Salah, S. Belaidi, N. Melkemi and N. Tchouar, Molecular geometry electronic properties MPO methods and structure activity/property relationship studies of 1,3,4-thiadiazole derivatives by theoretical calculations, *Rev Theor Sci* 3: 355, **2015**.

[120]. K. Palm, P. Stenberg, K. Luthman and P. Artursson, Polar molecular surface properties predict the intestinal absorption of drugs in humans, *Pharm Res* 14: 568, **1997**.

[121]. P. Stenberg, K. Luthman, H. Ellens, Ch. Pin Lee, P. L. Smith, A. Lago, J. D. Elliot, P. Artursson, Prediction of the intestinal absorption of endothelin receptor antagonists using three theoretical methods of increasing complexity, *Pharm Res* 16: 1520, **1999**.

[122]. D. F. Veber, S. R. Johnson, H. Y. Cheng, B. R. Smith, K. W. Ward and K. D. Kopple, Molecular properties that influence the oral bioavailability of drug candidates, *J. Med. Chem.* 45: 2615, **2002**.

[123]. A. L. Hopkins, C. R. Groom and A. Alexander, Ligand Efficiency: a useful metric for lead selection, *Drug Discovery Today* 9: 430, **2004**.

[124]. J. D. Hughes, J. Blagg, D. A. Price, S. Bailey, G. A. Decrescenzo, R. V. Devraj, E. Ellsworth, Y. M. Fobian, M. E. Gibbs and R. W. Gilles, *J Bioorg Med Chem Lett* 18: 4872, **2008**.

[125]. M. P. Edwards and D. A. Price, Role of physicochemical properties and ligand lipophilicity efficiency in addressing drug safety risks, *Annu Rep Med Chem* 45: 380, **2010**.

[126]. P. D. Lesson and B. Springthrope, The influence of drug-like concepts on decision-making in medicinal chemistry, *Nat Rev Drug Discov* 6: 881, **2007**.

[127]. T. Görnitz, Simplest Quantum Structures and the Foundation of Interaction, *Rev Theor Sci* 2: 289, **2014**.

[128]. P. Rengin, O. Hüseyin and E. Sakir, Binary Alloy Clusters: Structures and Electronic Properties, *Rev Theor Sci* 2: 301, **2014**.

[129]. C. G. Wermuth, The Practice of Medicinal Chemistry, London, **2008**.

[130]. H. Mohsen and R. Abbas, Design of Novel Efficient XOR Gates for Quantum-Dot Cellular Automata, *J Comput Theor Nanosci* 10: 643, **2013**.

[131]. D. S. Paolo, Present and future of nanotechnologies: peculiarities, phenomenology, theoretical modelling, perspectives, *Rev Theor Sci* 2: 146, **2014**.

[132]. C. Chia and J. Abraham, Theoretical design of free radical "Sponges" in DNA, *J Comput Theor Nanosci* 10: 591, **2013**.

[133]. C. D. SELASSIE, History *of quantitative structure-activity relationships, Medicinal Chemistry and Drug Discovery*, John Wiley & Sons Inc. **2003**.

[134]. J. LESZCZYNSKI, Challenges and advances in computational chemistry and physic, Jackson State University, U.S.A. **2010**.

[135]. R. José, Dissertation, Développement de nouvelles méthodes chimiométriques d'analyse, application à la caractérisation spectroscopique de la qualité des aliments, PhD thesis, agro paris tech, **2009**.

[136]. F. J. Bryan, Multivariate statistitcal methods, Chap man & hall, **1994**.

[137]. D. Downing, J. Clark, Business statictics, Barron's Educational Series, **2010.**

[138]. O. Sanja, D. D. Cvetković and D. J. Barna, QSAR analysis of 2-amino or 2-methyl-1-substituted benzimidazoles against pseudomonas aeruginosa, *International Journal of Molecular Sciences* 14: 1670, **2009.**

[139]. Official website: transtutors.com/homework-help/**statistics**/correlation.aspx.

[140]. S. K. Purkayastha, T. Jha, D. K. Pal and A. U. De, Possible antineoplastic agents: Part XIII. Synthesis, biological evaluation and QSAR studies of some 1-(substituted benzenesulphonyl)-5-oxopyrrolidine-2-carboxylic acid derivatives, *Anticancer Drug Des* 8: 95, **1993**.

[141]. K. Srikanth, C. A. Kumar, D. Goswami, A. U. De and T. Jha, Quantitative structure activity relationship (QSAR) studies of some substituted benzenesulphonyl glutamines as tumour suppressors, *Indian J Biochem Biophys* 38: 120, **2001**.

[142]. K. Srikanth, B. Debnath and T. Jha, Syntheses, biological evaluation and QSAR study on antitumor activity of 1,5-N,N'-disubstituted-2-(substituted benzenesulphonyl) glutamamides, *Bioorg Med Chem* 10: 1841, **2002**.

[143]. S. Samanta, K. Srikanth, S. Banerjee, B. Debnath, S. Gayen and T. Jha, 5-N-Substituted-2-(substituted benzenesulphonyl) glutamines as antitumor agents. Part II: synthesis, biological activity and QSAR study. *Bioorg Med Chem* 12: 1413, **2004**.

[144]. I. V. Tetko, V. Y. Tanchuk and A. E. Villa, Prediction of n-octanol/water partition coefficients from PHYSPROP database using artificial neural networks and E-state indices, *J Chem Inf Comput Sci* 41: 1407, **2001**.

[145]. R. G. Parr, W. Yang, Density-Functional Theory of the Electronic Structure of Molecules, *Annu Rev Phys Chem* 46: 701, **1995**.

[146]. R. G. Parr, W. Yang, Density Functional Theory of Atoms and Molecules, Oxford University Press, New York, **1989**.

[147]. D. Legros, G. Ollivier, M. Gastellu-Etchegorry, C. Paquet, C. Barri, J. Jannin and P. Buscher, Treatment of human African trypanosomiasis present situation and needs for research and development, *Lancet* 2: 437, **2002**.

[148]. S. Belaidi and N. Melkemi, Conformational analysis and physical-chemistry property relationship for 22-membered macrolides, *Asian J Chem* 25: 4527, **2013**.

[149]. S. Belaidi, R. Mazri, M. Mellaoui, A. Kerassa and H. Belaidi, Electronic structure and effect of methyl substitution in oxazole and thiazole by quantum chemical calculations, *J Pharm Biol Chem Sci* 5: 811, **2014**.

[150]. S. Belaidi, R. Mazri, H. Belaidi, T. Lanez and D. Bouzidi, Electronic structure and physical chemistry property relationship for thiazole derivatives, *Asian J Chem* 25: 9241, **2013**.

[151]. P. Geerlings, F. De Proft and W. Langenaeker, Conceptual density functional theory, *Chem Rev* 103: 1793, **2003**.

[152]. K. H. Manoj and B. Arup, Many-electron problem in terms of the density: from thomas–fermi to modern density-functional theory, *J Theor Comput Chem* 2: 301, **2003**.

[153]. O. Tamer, D. Avci and Y. Atalay, Spectroscopic study, NLO properties and homo–lumo analysis on different donor and acceptor substituents of thiazolylazopyrimidine chromophores, *J Theor Comput Chem* 12: 1, **2013**.

[154]. F. De Proft and W. Langenaeker, Conceptual Density Functional Theory, *Chem Rev*, 103: 1793, **2003**.

[155]. H. M. Berman, The protein data bank, *Nucleic Acids Res 28*: 235, **2000**.

[156]. J. Weigelt, Structural genomics-Impact on biomedicine and drug discovery, *Exp Cell Res 316*: 1332, **2010**.

[157]. S. F. Sousa, P. A. Fernandes, M. J. Ramos, Protein-ligand docking: Current status and future challenges, *Proteins Struct Funct Bioinform 65*: 15, **2006**.

[158]. N. Brooijmans and I. D. Kuntz, Molecular recognition and docking algorithms, *Annu Rev Biophys Biomol Struct* 32: 335, **2003**.

[159]. M. Rarey, B. Kramer, T. Lengauer and G. Klebe, A fast exible docking method using an incremental construction algorithm, J Mol Biol 261: 470, **1996**.

[160]. Z. Zsoldos, D. Reid, A. Simon, S. B. Sadjad and A. P. Johnson, eHiTS: A new fast, exhaustive exible ligand docking system, *J Mol Graph Model* 26: 198, **2007**.

[161]. G. Jones, P. Willett and R. C. Glen, Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation, *Journal of Molecular Biology* 245: 43, **1995**.

[162]. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, and A. J. Olson, Automated docking using a lamarckian genetic algorithm and an empirical binding free energy function, *J Comput Chem* 19: 1639, **1998**.

[163]. S. Makino and I. D. Kuntz, Automated flexible ligand docking method and its application for database search, *J Comput Chem* 18: 1812, **1997**.

[164]. A. Jamloki, C. Karthikeyan, H. N. Moorthy and P. Trivedi, QSAR analysis of some 5-amino-2-mercapto-1,3,4-thiadiazole based inhibitors of matrix metalloproteinases and bacterial collagenase, *J Bioorg Med Chem Lett* 16, 3847, **2006**.

[165]. N. M. Noolvi, M. H. Patel and S. N. Sethi, 2D-QSAR study of 1,3,4-thiadiazole-2-arylhydrazone derivatives: an approach to design antitrypanosomal agents, *International Journal of Drug Design and Discovery* 1: 195, **2010**.

[166]. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *J Adv Drug Delivery Rev* 46: 4, **2012**.

[167]. D. M. Segall, Multi-parameter optimization: identifying high quality compounds with a balance of properties, *J Curr Pharm Des* 18: 1292, **2012**.

[168]. E. Kerns, Drug-like properties: concepts, structure design and methods: From ADME to toxicity optimization, Elsevier Science, **2008**.

[169]. Z. Wang, F. Wang, C. Su and Y. Zhang, Computer Simulation of Polymer Delivery System by Dissipative Particle Dynamics, *J. Comput. Theor. Nanosci.* 10: 2323, **2013**.

[170]. A. Eghdami and M. Monajjemi, quantum modeling of alpha interferon subunits in point of nano anticancer drug, *Quantum Matter* 2: 324, **2013**.

[171]. Y. Chen, D. Xu and M. Yang, Quantitative study on longitudinal strain of left ventricle in patients with myocardial ischemia by two-dimensional speckle tracking imaging, *J Comput Theor Nanosci* 10 : 2916, **2013**.

[172]. B. Wu, X. Kong, Z. Cao, Y. Pan, Y. Ren, Y. Li, Q. Yang and F. Lv, Structural Characterization and Statistical Modeling of Nanopeptide Collision Cross-Sections in Ion Mobility Spectrometry, *J Comput Theor Nanosci* 10: 2403, **2013**.

[173]. D. S. C. B. Rodrigues, V. J. Braga, D. S. F. Adriane et al., Validation of Computational Methods Applied in Molecular Modeling of Artemisinin with Antimalarial Activity, J Comput Theor Nanosci 11: 553, **2014**.

[174]. M. Karabacak, L. Sinha, O. Prasad, A. M. Asiri and M. Cinar, An experimental and theoretical investigation of Acenaphthene-5-boronic acid: Conformational study, NBO and NLO analysis, molecular structure and FT-IR, FT-Raman, NMR and UV spectra, *J Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 115: 753, **2013**.

[175]. J. S. Kwiatkowski, J. Leszczynski and I. Teca, molecular structure and infrared spectra of furan, thiophene selenophene and their 2,5-N and 3,4-N derivatives: density functional theory and conventional post-Hartree-Fock MP2 studies, *J Molecular Structure* 45: 436, **1997**.

[176]. J. Pathak, V. Narayan, L. Sinha, O. Prasad, Theoretical Raman and FTIR vibrational analysis of 2-phenyl-1H-indene-1,3(2H)-dione by ab initio method, *J At Mol Sci* 3: 95, **2012**.

[177]. G. L. Miessler and D. A. Tarr, Inorganic Chemistry, Prentice-Hall Upper Saddle River, USA, **1999**.

[178]. H. Koster, T. Craan, S. Brass, C. herhaus, M. Zentgraf, L. Neumann et al., a small nonrule of 3 compatible fragment library provides high hit rate of endothiapepsin crystal structures with various fragment chemotypes, *J Med Chem* 25: 7784, **2011**.

[179]. M. Remko, M. Swart and F. M. Bickelhaupt, Theoretical study of structure, pKa, lipophilicity, solubility, absorption, and polar surface area of some centrally acting antihypertensives, *J Bioorg Med Chem* 14: 1715, **2006**.

[180]. A. L. Hopkins, C. R. Groom, and A. Alexander, Ligand efficiency: a useful metric for lead selection, *J Drug Discov Today* 9: 430 **2004**.

[181]. P. D. Leeson and B. Springthorpe, The influence of drug-like concepts on decision-making in medicinal chemistry. *J Nat Rev Drug Discov* 6: 881, **2007**.

[182]. J. D. Hughes, J. Blagg, D. A. Price, S. Bailey, G. A. Decrescenzo, R. V. Devraj, E. Ellsworth, Y. M. Fobian, M. E. Gibbs and R. W. Gilles, Physiochemical drug properties associated with in vivo toxicological outcomes, *J Bioorg Med Chem Lett* 18: 4872, **2008**.

[183]. S. Schultes, C. Graaf, E. Haaksma, J. P. Iwan and O. Kramer, Ligand efficiency as a guide in fragment hit selection and optimization, *J Drug Discovery Today: Technologies* 7: 157, **2010**.

[184]. V. Dimitrov and T. Komatsu, An interpretation of optical properties of oxides and oxide glasses in terms of the electronic ion polarizability and average single bond strength, *Journal of the University of Chemical Technology and Metallurgy* 45: 219, **2010**.

[185]. N. Melkemi and S. Belaidi, Structure-property relationships and quantitative structure-activity relationship modeling of detoxication properties of some 1,2-dithiole-3-thione derivatives, *J. Comput. Theor. Nanosci.* 11: 801, **2014**.

[186]. M. Roberto, P. Andrea and C. Fabio, Geometrical Lorentz Violation and Quantum Mechanical Physics, *Quantum Matter* 3: 219, **2014**.

[187]. B. Ersfeld, B. U. Felderhof, Retardation correction to the Lorentz-Lorentz formula for the refractive index of a disordered system of polarizable point dipoles, *Pysical Review E* 57: 1118, **1998**.

[188]. D. M. Juan et al., Mode of action of natural and synthetic drugs against Trypanosoma Cruzi and their interaction with the mammalian host, *J Comparative Biochemistry and Physiology* 146: 601, **2007**.

[189]. D. R. Moreira, A. C. Leite, R. R. dos Santos, and M. B. Soares, Approaches for the development of new anti-Trypanosoma cruzi agents, *J Curr Drug Targets* 10: 212, **2009**.

[190]. T. Ke, B. Xiao-Zhi, L. Xiao-Qiang, Z. Yue, S. Ji-Hong, T. Chao-Wu and H. Da-Hai, In vitro induction of adipose-derived mesenchymal stem cells differentiated into sweat gland epithelial cells, *J. Comput. Theor. Nanosci.* 11: 1785, **2014**.

[191]. J. Lavrado, Z. Mackey, E. Hansell, J. H. McKerrow, A. Paulo, and R. Moreira, Antitrypanosomal and cysteine protease inhibitory activities of alkyldiamine Cryptolepine derivatives, *Bioorg Med Chem Lett* 22: 6256, **2012**.

[192]. S. Nithiyanantham and L. Palaniappan, physico-chemical and structural studies on maltose with amylase in aqueous media at 298.15 K, *J Comput Theor Nanosci* 11: 1721, **2014**.

[193]. S. O. Podunavac-Kuzmanovic, D. D. Cvetkovic and D. J. Barna, QSAR analysis of 2-amino or 2-methyl-1-substituted Benzimidazoles against pseudomonas aeruginosa, *Int. J. Mol. Sci.* 10: 1670, **2009**.

[194]. D. Baumann and K. Baumann, Reliable estimation of prediction errors for QSAR models under model uncertainty using double cross-validation, *J cheminform* 6: 47, **2014**.

[195]. M. Mellaoui, S. Belaidi, D. Bouzidi and N. Gherraf, Electronic structure and physical-chemistry property relationship for cephalosporin derivatives, *Quantum Matter* 3: 435, **2014**.

[196]. M. Jalali-Heravi and A. Kyani, Use of computer-assisted methods for the modeling of the retention time of a variety of volatile organic compounds: a PCA-MLR-ANN approach, *J Chem Inf Comput Sci* 44: 1328, **2004**.

[197]. R. A. Laskowski, M. B. Swindells, LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery, *J Chem Inform Modelling* 51: 2778, **2011**.

[198]. R. G. Parr and Y. Weitao, Density functional approach to the frontier-electron theory of chemical reactivity, *J Am Chem Soc* 106: 4049, **1984**.

[199]. W. Yang and W. J. Mortier, The use of global and local molecular parameters for the analysis of the gas-phase basicity of amines, *J Am Chem Soc* 108: 5708, **1986**.

[200]. P. Kolandaivel, G. Praveena and P. Selvarengan, Study of atomic and condensed atomic indices for reactive sites of molecules, *J Chem Sci* 117: 591, **2005**.

[201]. L. João, M. Zachary, H. Elizabeth, H. M. James, P. Alexandra and R. Moreira, Antitrypanosomal and cysteine protease inhibitory activities of alkyldiamine cryptolepine derivatives, *Bioorganic Med Chem Lett* 22: 6256, **2012**.

[202]. N. M. Noolvi, M. H. Patel and S. N. Sethi, 2D-QSAR study of 1,3,4-thiadiazole-2-arylhydrazone derivatives: an approach to design antitrypanosomal agents, *Int J Drug Des Discovery* 1: 195, **2010**.

[203]. A. Chunzhi, C. Li, W. Yonghua, C. Yadong and Y. Ling, Insight into the effects of chiral isomers quinidine and quinine on CYP2D6 inhibition, *Bioorganic Med Chem Lett* 19: 803, **2009**.

[204]. R. K. Srivastava, V. Narayan, O. Prasad, L. Sinha, Vibrational, Structural and electronic properties of 6-methyl nicotinic acid by density functional theory, *J Chem Pharm Res* 4: 3287, **2012**.

[205]. J. Pathak, V. Narayan, L. Sinha, O. Prasad, Theoretical Raman and FTIR vibrational analysis of 2-phenyl-1H-indene-1,3(2H)-dione by ab initio method, *J Atomic Mol Sci* 3: 95, **2012**.

[206]. I. D. Kerr, J. H. Lee, C. J. Farady, R. Marion, M. Rickert et al., Vinyl sulfones as antiparasitic agents and a structural basis for drug design, *J Biol Chem* 284: 25697, **2009**.

[207]. A. Imberty, K. D. Hardman, J. P. Carver, S. Perez, Molecular modelling of protein-carbohydrate interactions. Docking of monosaccharides in the binding site of concanavalin A, *Glycobiology* 1: 631, **1991**.

[208]. J. G. Robertson, Enzymes as a special class of therapeutic target: Clinical drugs and modes of action, *Curr Opin Struct Biol* 17: 674, **2007**.

[209]. J. G. Robertson, Mechanistic basis of enzyme-targeted drugs, *Biochemistry* 44: 8918, **2005**.

[210]. T. Doane and C. Burda, Nanoparticle mediated non-covalent drug delivery, *Adv Drug Deliv Rev* 65: 607, **2013**.

[211]. N. K. Jain and U. Gupta, Application of dendrimer-drug complexation in the enhancement of drug solubility and bioavailability, *Exp Opin Drug Metab Toxicol* 4: 1035, **2008**.

[212]. A. S. Kalgutkar and D. K. Dalvie, Drug discovery for a new generation of covalent drugs, *Expert Opin Drug Discov* 7: 561, **2012**.

[213]. A. E. F. Nassar and A. Lopez-Anaya, Strategies for dealing with reactive intermediates in drug discovery and development, *Curr Opin Drug Discov Dev* 7: 126, **2004**.

[214]. Y. Pommier, Drugging topoisomerases: Lessons and challenges, *ACS Chem Biol* 8: 82, **2013**.

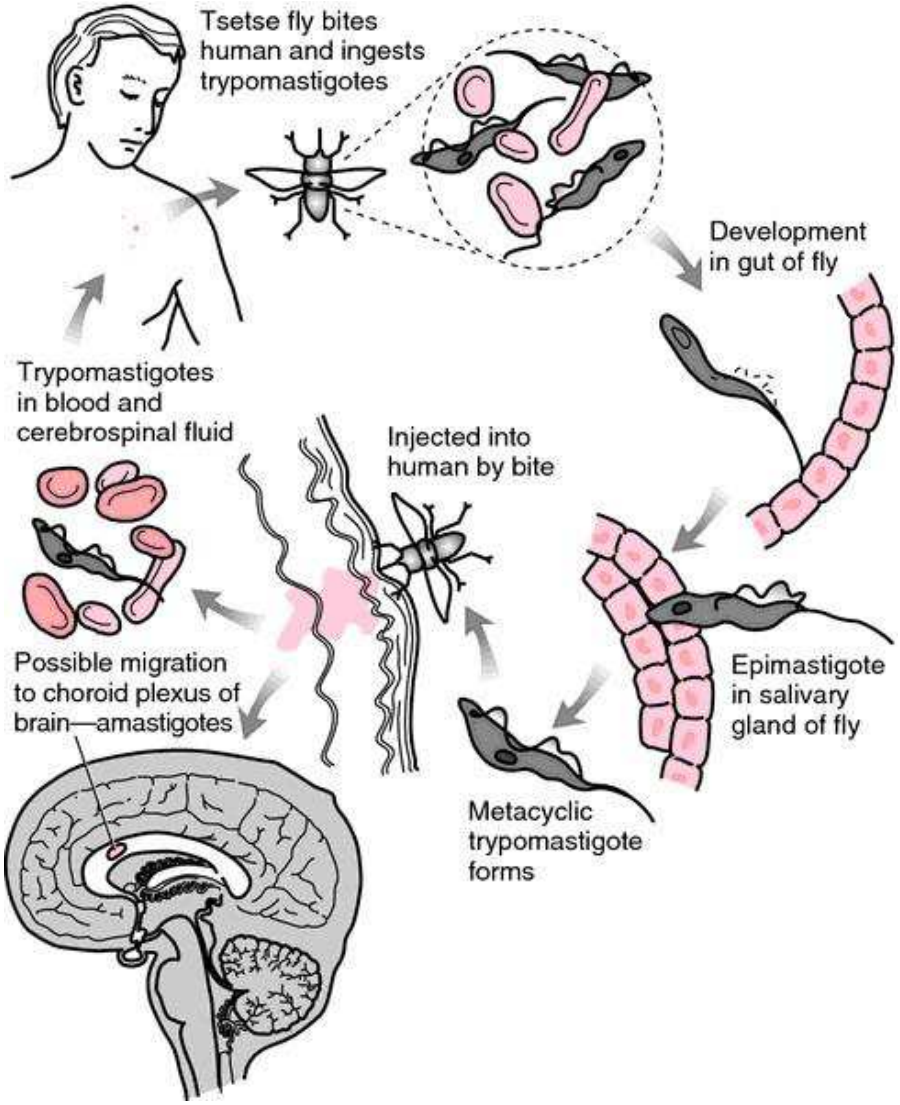
[215]. S. F. Zhou, E. Chan, W. Duan, M. Huang and Y. Z. Chen, Drug bioactivation, covalent binding to target proteins and toxicity relevance, *Drug Metab Rev* 37: 41, **2005**.

[216]. M. K. Hezekiel, B. Soumendranath and E. S. S. Mahmoud, Theory and applications of covalent docking in drug discovery: Merits and pitfalls, *Molecules* 20: 1984, **2015**.

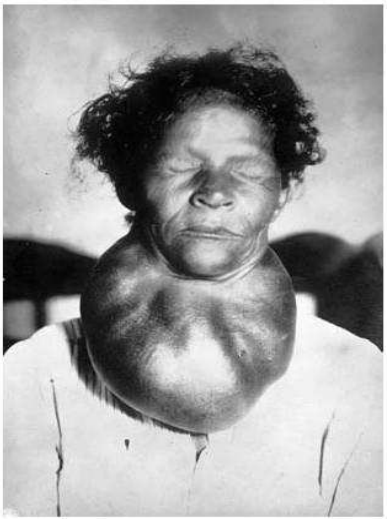
[217]. G. Young-Mi, D. C. Joshua and O. J. Dean, The cysteine proteome, *Free Radical Biol Med* 84: 227, **2015**.

Appendices

Appendix A: Life cycle of the etiologic agents of sleeping sickness (*T. Gambiense and T. Rhodesiense*).

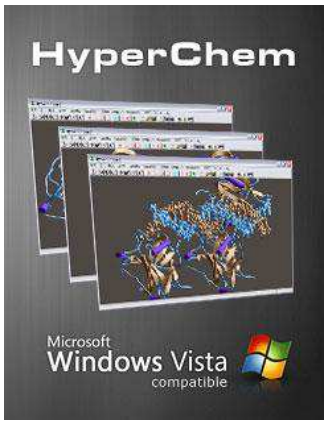


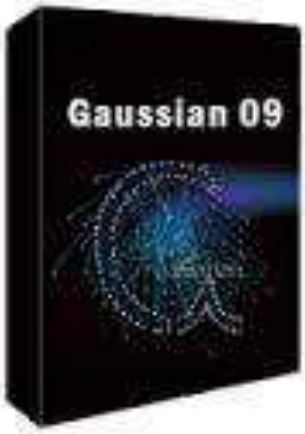
Appendix B: Nifurtimox drug (Lampit®).

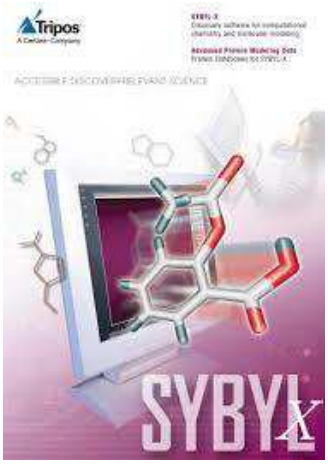


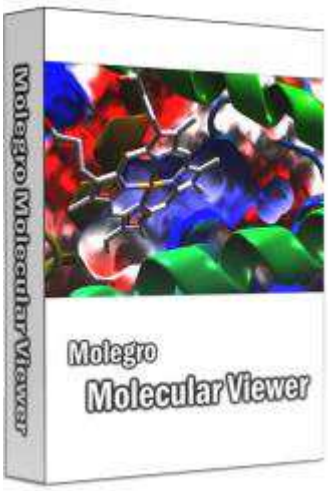
Appendix C: Infected people by Trypanosoma.

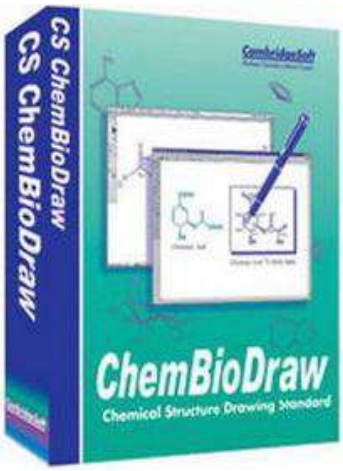
Appendix D: Some Computer-Aided-Drug-Design software.











Appendix E: Table of F variable (Fisher-Snedecor).

V_1 = degrees of freedom for numerator $F_{95}(V_1, V_2)$ $\alpha = 0.05$

V_2 = degrees of freedom for denominator

v_2/v_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.44	199.50	215.70	224.58	230.16	233.98	236.76	238.88	240.54	241.88	243.90	245.94	248.01	249.05	250.09	251.14	252.19	253.25	254.31
2	18.512	19.000	19.164	19.246	19.296	19.329	19.353	19.371	19.384	19.395	19.412	19.429	19.445	19.454	19.462	19.470	19.479	19.487	19.495
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5494	8.5264
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6877	5.6581	5.6281
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.3650
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705	2.5309	2.4901	2.4480	2.4045
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410	2.2962
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6022	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9604
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.8780
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660	2.3210	2.2504	2.1757	2.0960	2.0540	2.0102	1.9645	1.9165	1.8657	1.8117
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.9380	1.8894	1.8380	1.7831
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201	2.2747	2.2036	2.1282	2.0476	2.0050	1.9605	1.9139	1.8648	1.8128	1.7570
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002	2.2547	2.1834	2.1077	2.0267	1.9838	1.9390	1.8920	1.8424	1.7896	1.7330
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.7110
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655	2.2197	2.1479	2.0716	1.9898	1.9464	1.9010	1.8533	1.8027	1.7488	1.6906
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7306	1.6717
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360	2.1900	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6376
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401	1.9926	1.9174	1.8364	1.7480	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.4290	1.3519	1.2539
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.3940	1.3180	1.2214	1.0000

Abstract

Current knowledge about Trypanosomiasis, the potentially life-threatening disease caused by protozoan parasites (Trypanosomes), has led to the development of new drugs more effective (enhanced activity) with reduced side effects (good balance of properties) and to the understanding of their mode of action.

Many researches have been carried out for the development of Trypanocidal drugs; but no drug could achieve a parasitological cure and some of them presented serious toxic after effects.

Therefore, the present research aim to discuss the different computational methods and approaches; used in Computer-Aided-Drug-Design. Whereas, three antitrypanosomal compounds have been used to study the inhibition activity against trypanosomes. Results such as CHELPG charges, bond length, dipole moment, Fukui functions, MESP, QSAR models, Lipinski's parameters, Lipophilic Efficiency (LipE), Molecular Docking analysis, etc., are reported and discussed in the present investigation. A close agreement with experimental results was found, which, improves the affinity of the present work.

Keywords: Conceptual DFT; Fukui function; Molecular Docking; MPO; QSAR; SAR/SPR; Trypanocidal compounds (Cryptolepine, 1,3,4-thiadiazole, Nifurtimox).

Résumé

Les connaissances actuelles sur la trypanosomiase, la maladie potentiellement mortelle causée par des parasites protozoaires (les trypanosomes) a conduit à la mise au point de nouveaux médicaments plus efficace (activité amélioré) avec des effets indésirables réduits (un bon équilibre des propriétés) et à la compréhension de leur mode d'action.

Plusieurs recherches ont été effectuées pour le développement des médicaments trypanocides; mais aucun médicament ne pourrait atteindre une guérison parasitologique et certains d'entre eux ont présenté des effets secondaires toxiques graves.

Par conséquent, la présente recherche, a pour objectif de discuter les différentes méthodes et approches computationnelles; utilisées dans la conception des médicaments assistée par ordinateur. Alors que, trois composés antitrypanosomiens ont été utilisés pour étudier l'activité d'inhibition contre les trypanosomes. Les résultats tels que les charges de CHELPG, longueur de liaison, moment dipolaire, fonctions de Fukui, MESP, les modèles de QSAR, les paramètres de Lipinski, l'efficacité lipophile (LipE), analyses de Docking moléculaire, etc., sont présentés et discutés dans la présente enquête. Tandis qu'un bon accord avec les résultats expérimentaux a été trouvé, ce qui améliore l'affinité de ce travail.

Mots clés: DFT conceptuelle; Fonction de Fukui; Docking Moléculaire; MPO; QSAR; SAR/SPR; composés Trypanocides (Cryptolepine, 1,3,4-thiadiazole, Nifurtimox).

ملخص

المعرفة الحالية حول داء المثقبيات، المرض الاكثر تهديدا للحياة الذي تسببه الطفيليات (المثقبيات)، أدت إلى تطوير أدوية جديدة أكثر فعالية (تعزيز النشاط) مع انخفاض الآثار الجانبية (توازن جيد للخصائص) وفهم طريقة عملها.

العديد من التجارب أجريت لتطوير مضادات المثقبيات؛ ومع ذلك، لا يوجد أي دواء استطاع تحقيق علاج الطفيليات و البعض منهم يعرض لتأثيرات جانبية سامة جدا.

وبالتالي، فالبحث الحالي يهدف إلى مناقشة الطرق والأساليب الحسابية المختلفة؛ المستخدمة في تصميم الأدوية بمساعدة الحاسوب. في حين، ثلاثة مركبات مضادة للمثقبيات تم استعمالها لدراسة نشاط التثبيط ضد المثقبيات. النتائج ك: شحنات CHELPG، طول الرابطة، عزم ثنائي الاقطاب، وظائف فيكوي، MESP، نماذج QSAR، فعالية المواد المحبة للدهون، تحليل الالتحام الجزيئي، إلخ.، ذكرت ونوقشت في البحث الحالي. في حين، تم العثور على اتفاق قريب مع النتائج التجريبية، مما يزيد من ألفة هذا العمل.

كلمات البحث: DFT المفاهيمية ؛ وظيفة FUKUI ؛ الالتحام الجزيئي ؛ MPO ؛ QSAR؛ SAR/SPR؛ مركبات مضادة للمثقبيات (Cryptolepine, 1,3,4-thiadiazole, Nifurtimox).