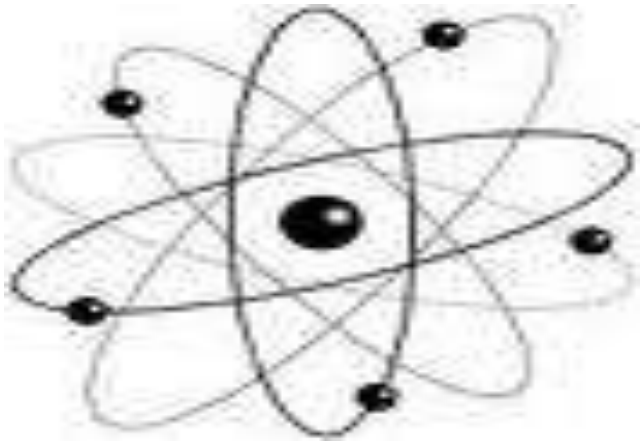
République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Med Khider Biskra
Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie

Département des Sciences de la Matière
Domaine : Sciences de la Matière
Filière : Chimie
Spécialité : Chimie Pharmaceutique



Mémoire de fin d'étude en Master
Intitulé :

Application du criblage virtuel, du « drug-likness » au modèle QSAR dans une série hétérocyclique bioactive.

Présenté par : BENAMOR Ahlam

Devant le jury :

MELLAOUI Malika	*Président*
BELAIDI Salah	*Rapporteur*
YOUCEF Oum Hani	*Examinateur*

2017-2018

Summary

Tables list

Chapter III

Chapter IV

Chapter V

Figures list

Chapter I

Chapter II

Figures list

Chapter III

Chapter IV

Chapter V

Abreviation list

ADMET	Absorption, Distribution, Metabolisme, and Excretion- Toxicity
AM1	Auserin Model 1
AMBER	Assisted Model Building with Energy Refinement
B3LYP	Becke 3-parameter lee-yang-parr
CHARMM	Chemistry Harvard Macromolécular Mechanic
CNDO	Complete Neglect of Differntial Overlap
DFT	Density- functional theory
ENR	Enoyl acyl carrier protein Reductase
GROMOS	Groningen Molecular Simulation Program Package
HBA	Hydrogen Bond Acceptor
HBD	Hydrogen Bond Donnor
HE	Hydration Energy
HF	Hartree-Fock
HIV	Human immunodeficiency virus
HOMO	High Occupied Molecular Orbital
INDO	Intermediate Neglect of Differential Overlap
INH	Isoniazid or isonicotinic acid Hydrazide
InhA	Enoyl-acyl carrier protein reductase HPLC
LLE	Lipophilic Ligand Efficiency
LOG D	Distribution coefficient
LOG P	Partition coefficient
LOO	Leave-one-out
LUMO	Lower Unoccupied Molecular Orbital
MPO	Multi-Parameter Optimization
MM	Molecular Mechanic
MLR	Multople Linear Regression
MR	Molar Refractivity
MW	Molecular Weigt
NADH	Nicotinamide Adenine Dinucleotide Hydride
NDDO	Neglect of Differential Diqto ;ic Overlap
NRB	Nomber of Rotatable Bonds
OPLS	Optimized Potentials for Liquid Simulation
PM3	Parametric Method 3
POL	Polarizability
PRESS	Predicted Residual Sum of Squares
PSA	Polar Surface Area
QSAR	Quantitative structure-activity relationships
QSPR	Quantitative structure-property relationships
SAG	Surface Area Grid
SPASIBA	Spectroscopic Potential Algorithm for Simulating Biomolecular conformational Adaptability
TB	Tuberculosis
VS	Virtuel Screening
2D, 3D	Two dimension, Three dimension

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. The cost benefit of using computational tools in the lead optimization phase of drug development is substantial **[1].**

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 **[2]** .

Molecular modeling has become a valuable and essential tool to medicinal chemists in the drug design process.

Molecular modeling describes the generation, manipulation or representation of three dimensional structures of molecules and associated physico-chemical properties. It involves a range of computerized techniques based on theoretical chemistry methods and experimental data to predict molecular and biological properties **[3]**.

Theoretical studies are currently moving towards rational design "Rational design "which means knowledge of the relationships between physicochemical properties and the molecular structure of known molecules allows scientists to develop new molecules, with a good anticipation **[4].**

Among the chemo informatics techniques we can mention the QSAR techniques which consist in finding a correlation between a biological activity measured for a panel of compounds and certain molecular descriptors.

QSAR, a quantum chemical technique is known to relate the biological activity of compounds with their molecular structure and has been extensively used as predicting tool in rational drug design **[5].** Quantitative structure – activity relationships (QSARs), as one of the most important areas in chemometrics, QSAR models are mathematical equations relating chemical structure to their biological activity. QSAR are attempts to correlate

molecular structure, or properties derived from molecular structure with a particular kind of chemical or biochemical activity **[6].**

Our work is part of a fundamental and original research on pyrazol and their derivatives, the main objective of this work is the application of different methods of molecular modeling to predict the activities expected biologics in relation to their anti mycobacterium tuberculosis activities.

In this work, the molecular modeling approach was used to study the electronic structure of the nucleus of pyrazole. On the other hand a qualitative study on the structure-activity / property relationship, in new bioactive molecules for a series of pyrazole derivatives.

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

- The first part concerns a bibliographical synthesis, contains two chapters containing respectively:
  - **Chapter 1:** pyrazole and its derivatives in the treatment of the disease and biological activities of pyrazole derivatives.

General information on the chemical aspect, some synthetic methods and some biological activities of the pyrazole nucleus, and generalities on mycobacterium tuberculosis were presented.

  - **Chapter 2:** Methods used in molecular modeling and methods of selection of drug candidates.

This chapter contains a theoretical study of molecular modeling that presents the different methods of computation.

- For the second part, in which we analyze the results of our calculations, it is composed of three chapters containing respectively:
  - **Chapter 3:** Structural and electronics study of pyrazole using several quantum calculation methods.

This chapter compares the results obtained by the three calculation methods (PM3, HF, DFT) on the molecule of pyrazole and experimental data.

- **Chapter 4:** Qualitative study of the QSAR properties of a series of pyrazol derivatives as new inhibitors of Enoyl Acyl Carrier Protein Reductase from Mycobacterium tuberculosis and application of selection methods (MPO, drug-likeness, Golden triangle,).

This chapter is devoted to the study of the structure-activity / property relationship and the drug likeness properties of a bioactive series of pyrazole derivatives.

- **Chapter 5:** Quantitative study of the QSAR properties of this series of pyrazole derivatives and application of chemometric methods

This chapter includes a quantitative study that aimed to describe the relationships between physicochemical properties and the biological activity of a series of bioactive derivatives of pyrazole by the use of a QSAR Model.

In the end of this manuscript, we finished with a general conclusion that summarizes our work.

References:

1. Basak SC. Chemo bioinformatics: the advancing frontier of computer-aided drug design in the post-genomic era. Curr. Comput. Aided. Drug. Des. **2012**; 8:1-2.
2. Giulia Palermo and Marco De Vivo, Computational Chemistry for Drug Discovery, Springer Science+Business Media Dordrecht **2015**.
3. Rama Rao Nadendla, Molecular Modeling: A Powerful Tool for Drug Design and Molecular Docking,Resonance, May **2004.**
4. J. Acar, La recherche, 314,50 , **1998** .
5. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 **(1985**).
6. N. Melkemi, S. Belaidi, J. Comput. Theor. Nanosci. 801-806m11 **(2014)**.

I.1 Introduction:

Pyrazole is five membered heterocyclic rings which is versatile lead compound for designing potent bioactive agent. The interesting groups of this compound has diverse biological activities such as antimicrobial, anti -inflammatory, anticancer, analgesic, anticonvulsant, anthelmintic, antioxidant and herbicidal. Given data represents that pyrazole being heterocyclic planar five membered rings have various pharmacological actions **[1].**

Pyrazoles have illustrious history; in 1883, a German chemist Ludwig Knorr was the first to discover antipyreti action of pyrazole derivative in man, he named the compound antipyrine. When he attempted to synthesize quinoline derivatives with antipyretic activity, accidentally obtained antipyrine (2,3- dimethyl-1-phenyl-3-pyrazolin-5-one) which has analgesic, antipyretic and antirheumatic activity; whichstimulated interest in pyrazole Chemistry[**2].**

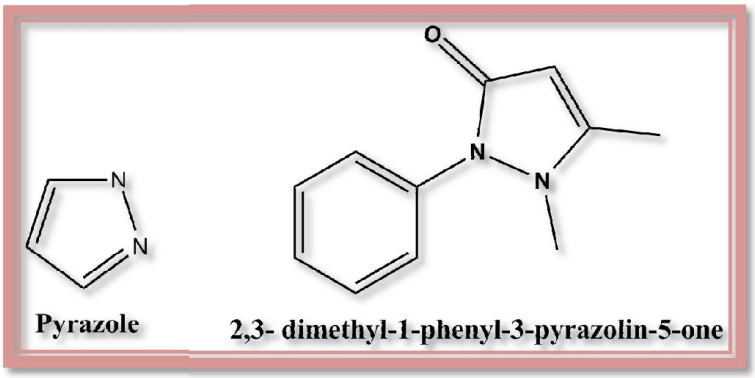


Figure I.1: Structure of pyrazole and 2, 3- dimethyl-1-phenyl-3-pyrazolin-5-one

The wide range of biological activities associated with pyrazoles has made them popular synthetic targets. Numerous methods have been developed for preparation of substituted pyrazoles. In general, pyrazoles are synthesized by the reaction of 1,3-diketones with hydrazines, 1,3-dipolar cycloaddition of diazo compounds with alkynes and the reaction of α,β-unsaturated aldehydes and ketones with hydrazines **[3].**

Pyrazole derivatives have a long history of application in agrochemicals as herbicides and insecticides and in pharmaceutical industry as antipyretic and anti-inflammatory. Antipyrine is the one of the earliest synthetic drugs and is named after its antipyreticproperties.

Butazolidine, another pyrazolone is a powerful anti- inflammatory drug used in rheumatic conditions. Many pyrazole derivatives are associated with anti-fungal, antidiabetic and anti-inflammatory properties **[4].**

Pyrazole derivatives have been reported to show a broad spectrum of biological activity including antimicrobial **[5, 6]**, anti-inflammatory **[7, 8]**, antituberculosis **[9, 10],** antiviral **[11, 12],** hypoglycemic **[13, 14],** anti-tumor **[15,16],** antihypertensive **[17-18].** Due to its wide range of biological activity, pyrazoles have received a considerable interest in the field of drug discovery and therefore pyrazole ring constitutes a relevant synthetic target in pharmaceutical industry. In fact, such a heterocyclic moiety represents the core structure of a number of drugs.

I.2 chemical appearance of pyrazole nucleus:

I.2.1 Generality:

It is a structural isomer of imidazole, the name pyrazole comes from the nucleus pyrrole to which a nitrogen atom has been added: "azole". The two atoms of nitrogen have different properties: one behaving like that pyridine can undergo protonation in acidic medium; the other has the property of the nitrogen of pyrrole, the doublet participating in the aromaticity of the cycle.

In official nomenclature, the pyrazole motif is called 1,2-diazole. Pyrazole is an aromatic π-surplus aromatic heterocycle. The reactions of Electrophilic substitutions are preferentially in position 4 and the nucleophilic attacks in position 3 and 5 **[19].**

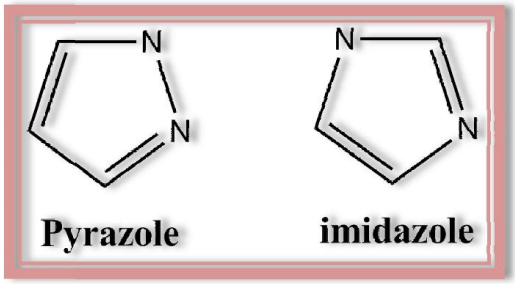


Figure I.2: Structure of pyrazole and imidazole

I.2.2 Natural appearance of pyrazole:

The first natural pyrazole derivative was isolated by Japanese workers Kosuge and Okeda in the year 1954, till their discovery it was thought that pyrazoles could not be obtained naturally. They isolated 3-nonylpyrazole from Houttuynia Cordata, a plant of the "piperaceae" family from tropical Asia; which showed antimicrobial activity. They also isolated levo-β-(1-pyrazolyl) alanine an amino acid from watermelon seeds (Citrullus Vulgaris)**[2].**Withasomnine has also been isolated from a plant (withania somnifera Dum) used in medicine Indian traditional medicine for the treatment of mild analgesics and antidepressants).

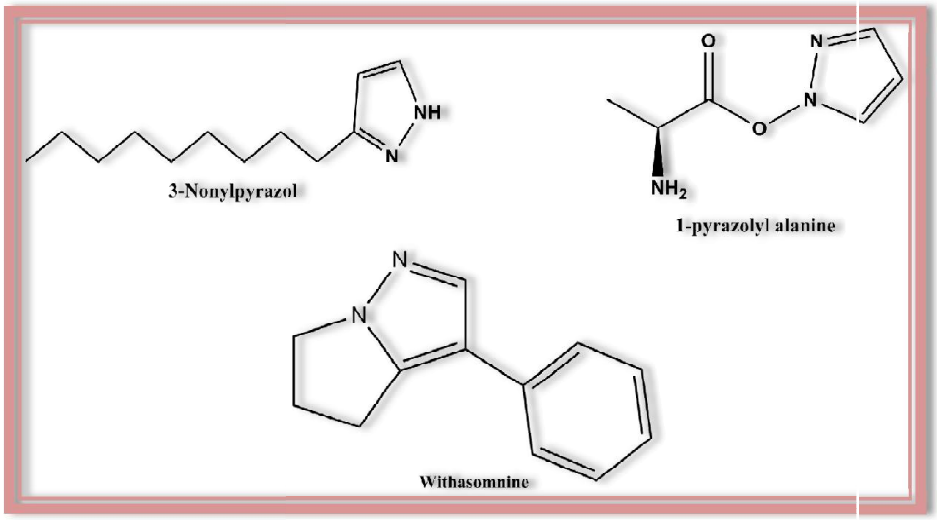


Figure I.3: Natural products contains pyrazole

I.2.4 Synthesis methods:

The various types of access to the pyrasole nucleus have undergone numerous modifications since the first ones described by Knorr, Pechmann , or Huisgen**[19].**

I.2.4.a From 1, 3-diketones:

The cyclocondensation of 1, 3-dicarbonyl compounds with hydrazine derivatives is a simple and rapid approach to obtain polysubstituted pyrazole. The method developed by Knorr [**20]** at the end of the 19th century generally results, in the presence of 1,3-dicarbonylated substrates no symmetrical, with the formation of two regioisomers which can be difficult to separate.

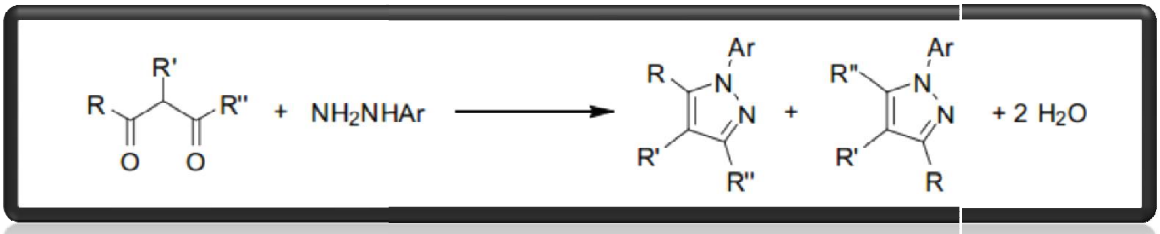


Figure I.4: Synthesis of pyrazole from 1, 3-diketones

I.2.4.b Diazomethane:

The 1, 3-dipolar cycloaddition reactions between an alkyne and a diazo compound was first used by Pechmann **[21]** (inventor of diazomethane) in 1889This reaction initially led to the formation of the intermediate.3,H-pyrazole, which then undergoes a sigmatropic rearrangement reaction to lead to the corresponding 1H-pyrazol.

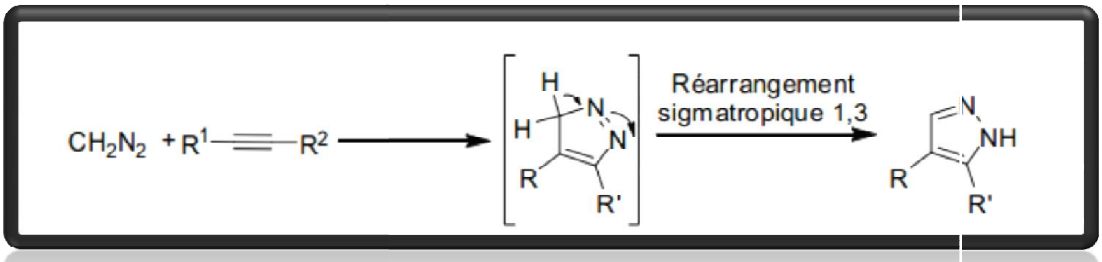


***Figure* I.5:** Synthesis of pyrazole by 1, 3-dipolar cycloaddition reactions between an alkyne and diazomethane

I.2.4.C InSitu Formation of Hydrazines:

Since substituted hydrazines are sometimes difficult to obtain, it is also possible to generate them in situ by cupro-catalyzed coupling between di-tert-butyl diazodicarboxylate and various aryl boronic acids. Hydrazine derivatives thus formed can react with 1,3-diketones to form the pyrazole ring . The high commercial availability of boronic acids makes this reaction very interesting for the synthesis of pyrazole libraries**[22].**

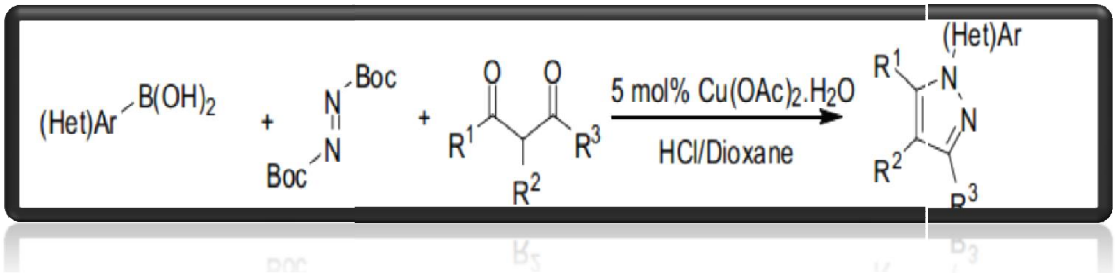


Figure I.6: Synthesis of pyrazole by InSitu formation of Hydrazines

I.3 Biological activity of pyrazole derivatives:

These Pyrazole skeletons comprise various ranges of pharmacological activities such as analgesic, antipyretic, anticancer, antiviral, anti-inflammatory, antioxidants, antimicrobial, anti-diabetic **[23] …**

Figure I.7 shows some of these activities:

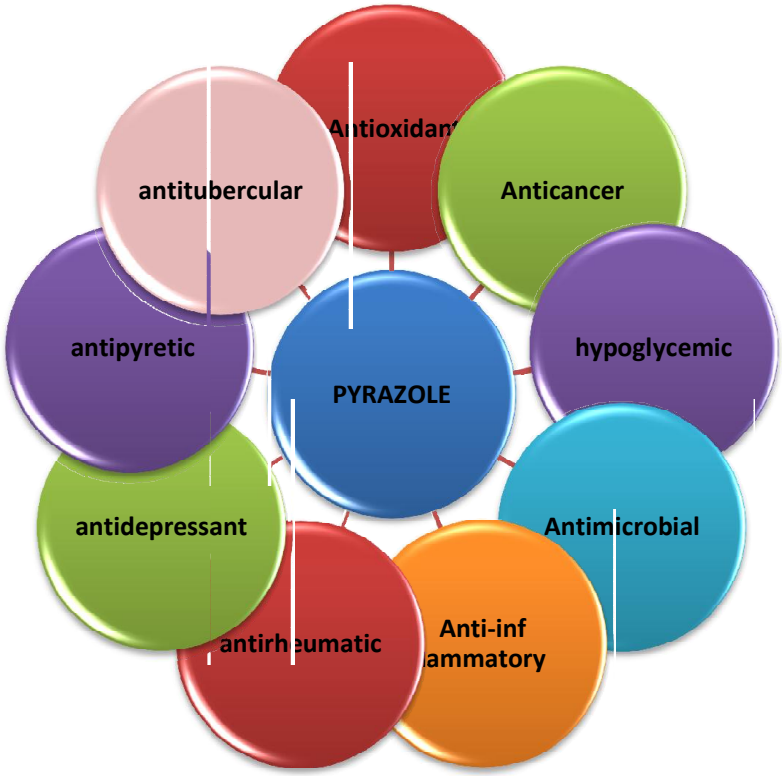


Figure I.7: Biological activities of pyrazole

I.4 Uses of pyrazole:

I.4. 1 Pyrazole as drugs:

The recent success of pyrazole COX-2 inhibitor has further highlighted the importance of these heterocyclic rings in medicinal chemistry. A systematic investigation of this class of heterocyclic lead revealed that pyrazole containing pharmacophore active agents play important role in medicinal chemistry. The prevalence of pyrazole cores in biologically active molecules has stimulated the need for elegant and efficient ways to make these heterocyclic lead **[23].**

Pyrazoles are found in the structure of several molecules biologically active agents such as Allopurinol which is used in the treatment of joint-related diseases such as gout, Allopurinol inhibits the enzyme Xanthine oxidase (XO) that converts hypo xanthine and xanthine in uric acid, it also has anticancer activity and antibacterial **[24].**

Figure I.8: Structure of Allopurinol

Pyrazolynones and pyrazilidine-3,5-diones are widely derivedused in the pharmaceutical field. Phenazone, for example, is a antipyretic used in the treatment of rheumatism and against fever, then that Phenylbutazone is an antiphlogistic agent used against inflammatory **[25].**

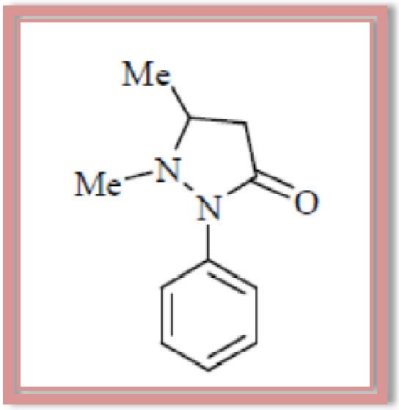


Figure I.9: Structure of Phénazone

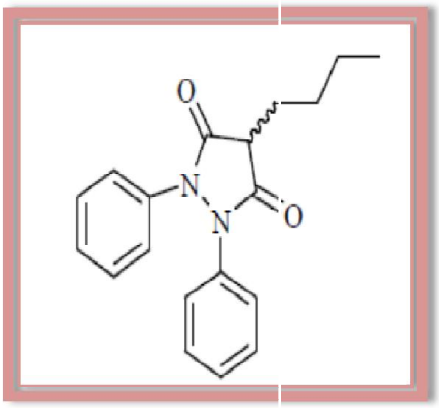


Figure I.10: Structure of phénylbutazone

Among the COX-2 inhibitors already launched on the market Celecoxib occupies a unique position as an effective anti-inflammatory agent **[26].**

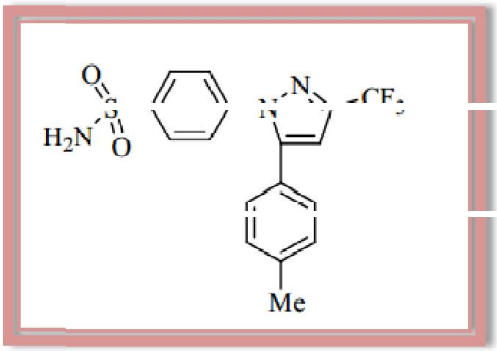


Figure I.11: Structure of Celecoxib

The pyrazole ring is present as the core in a variety of leading drugs such as Celebrex, Sildenafil (Viagra), Ionazlac, Rimonabant and Difenamizole etc **[2].**

Figure I.12 shows some of these drugs:

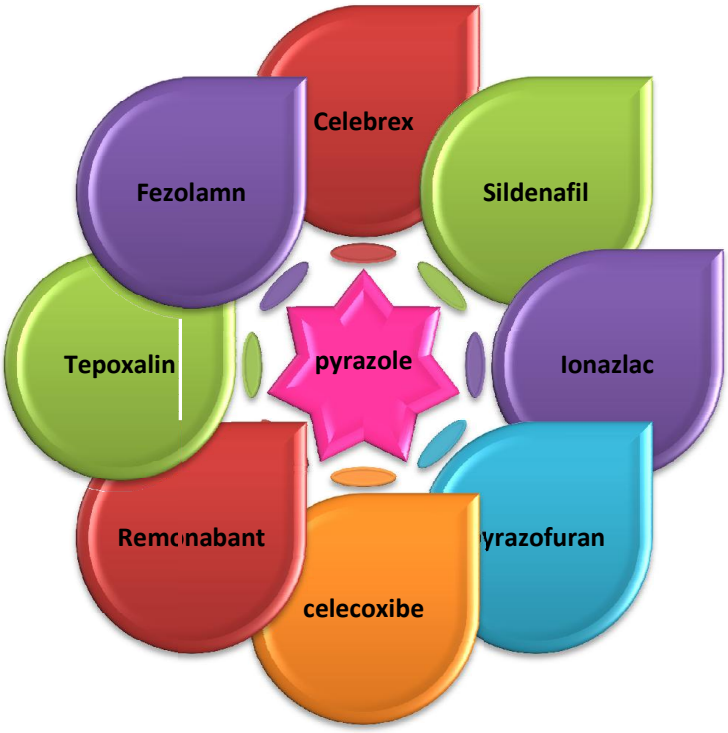


Figure I.12: Example of some drugs

I.5 Inhibitors of Enoyl Acyl Carrier Protein Reductase from mycobacterium tuberculosis:

I.5.1 Mycobacterium tuberculosis:

Mycobacterium tuberculosis, which is the main agent for tuberculosis (TB), is a persistent pathogen that has infected many people. With the growing appearance of TB resistant and HIV infection, many individuals are vulnerable to TB **[27,28].** Mycobacteria are ubiquitous organisms that are becoming increasingly important intracellular pathogens that establish an infection in oxygen-rich macrophage of the lung **[29].**

TB is one of the major causes of morbidity and mortality throughout the world. Approximately 32% of the world's population is currently living with this infectious disease **[30].** Tuberculosis is still a major cause of death from an infectious agent in developing countries, and constitutes a serious threat in large cities of industrialized countries. The resurgence of tuberculosis during the last decade has been partly caused by the emergence of multi-drug resistant Koch bacilli **[31].**

Most drug-resistant Mycobacterium tuberculosis clinical isolates are resistant to isoniazid (INH), which is a first-line drug used to treat tuberculosis.

I.5.2 Enoyl Acyl Carrier Protein Reductase and its inhibitors:

The NADH-dependent enoyl-ACP reductase of the Mycobacterium tuberculosis is one of the important molecular targets. The target for the first line anti tubercular drug isoniazid (INH) is also the same **[32].** Enoyl-ACP reductase (ENR) is a key regulatory enzyme in fatty acid elongation, and it catalyses the NADH-dependent stereo specific reduction of α, β - unsaturated fatty acids bound to the acyl carrier protein **[33].** Inhibition of ENR disrupts the biosynthesis of the mycolic acids that are central constituents of the mycobacterial cell wall.

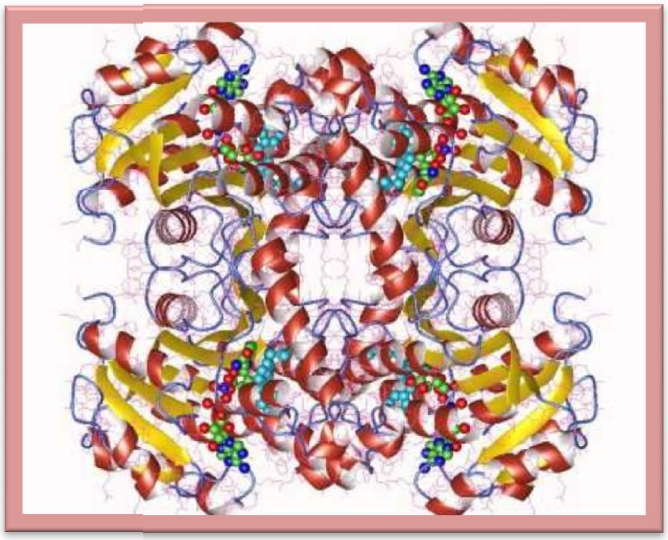


Figure I.13: Structure of Enoyl Acyl Carrier Protein Reductase

The isoniazid (INH)-NADH adduct functions as a potent inhibitor of InhA as well as InhA inhibitors like diazaborines, triclosan, pyrazole derivatives and indole-5-amides have been reported earlier.

Pyrazolines are important heterocycles containing N-N bond linkage that exhibit biological activities **[34, 35].** Various other nitrogen-containing heterocycles such as pyrazolines, pyridazinones and pyrrolones have also been developed as active antimicrobial drugs that are pharmacologically active. Earlier reports on pyrazoline derivatives exhibited

antimicrobial, antibacterial, antifungal **[36-39],** antiamoebic **[40],** antitubercular **[41],** antiHIV **[42],** anticancer **[43],** antidepressant and anticonvulsant **[44]** activities.

Thus, we are claiming for a new class of anti tubercular agent with a different mode of action and it has led medicinal chemists to explore a wide variety of chemical structures**[45].** Where as many reports **[46-50**] were proved the emergence of pyrazoles as potent antitubercular agents. Imidazole derivatives are known to possess antitubercular, antifungal, anti-neoplastic activities **[51-54].**

I.6 Conclusion:

We have been able, through this chapter on the studied molecule of1,2-diazol, on the one hand, to demonstrate the importance of pyrazole-containing compounds that have been developed to treat a wide assortment of medical conditions including mycobacterium, s, cancer, inflammatory response, hypoglycemic . As well as the techniques and processes that make it possible to transform and functionalize this type of molecule in order to enhance its biological and pharmacological advantage. Therefore drugs that inhibit enoyl acyl Carrier protein reductase activity have the potential for the treatment of Mycobacterium tuberculosis.

References:

1. A. Jamwal, A. Javed, V. Bhardwaj, A review on Pyrazole derivatives of pharmacological potential, J. Pharm. BioSci. 3**(2013)** 114-123
2. K. Ajay Kumar, P. Jayaroopa , Pyrazoles: Synthetic Strategies andTheir Pharmaceutical Applications-An Overview, International Journal of PharmTech Research,5**(Oct-Dec 2013)** 1473-1486.
3. Ph.D. Thesis: Synthesis, Characterization andBiological Studies of Five Membered Nitrogen Heterocycles .
4. S.R. Pattan , P.V. Patel, G.S. Athare , A.B. Jagnar , S.A. Nirmal, J.S.Pattan, Synthesis and evaluation of some substituted pyrazole derivatives of biological interest, Bulgarian Chemical Communications, 46(pp. 125 –134) **2014.**
5. V.K. Ahluwalia, B. Mittal, Indian J. Chemistry,8,17. **(1989).**
6. G. Menozzi, L. Merello, P. Fossa, S.Schenone, A. Ranise, L. Mosti et al. Bioorganic & Medicinal Chemistry, 12, 5465 **(2004).**
7. B. Coli, Silverstrini et al; Ep Med. Pathol, 6, 68.**(1967).**
8. L. Gomez, M. D. Hack, K. McClure, C. Sehon, L. Huang, M. Morton et al. Bio-org & Med Chem Letters, 17, 6493 **(2007).**
9. K.A. Jensen, L. Henrikesen, Acta. Chem. Scand.,:22,1170 **(1968).**
10. Singaravelmohan, SarkkariAnanthan , IJSPR, 1(9), 391 **(2010).**
11. Sh. R. Pattan, P. A. Rabara, J. S. Pattan, A. A.Bukitagar, V. S. Wakale, D. S. Musmade, Indian Journal of Chemistry, 48B, 1453 **(2009).**
12. G. A. Idrees, O. M. Aly, A. A. G. El-Din, AbuoRahma, M.F. Radwan, Eur. J. Med. Chem., 44, 3973 **(2009).**
13. P. Kamaria, N. Kawathekar, P. Chaturvedi; EJournal of Chemistry, 8(1), 305.
14. Smith, D L., forist, A A., et al; J. Med. Chem,8: 350 **(1965).**
15. P.-Ch. Lv., H.-L. Zhu, H.-Q. Li, J. Sun, Y. Zhou,. Bioorg. Med. Chem., 18: 4606 **(2010).**
16. H. R. Brunner, K. Eichenbergen, et.al; Expermentia, 22, 208 **(1966).**
17. L. Kuczynski, A. E. Morzik, W. Banasszkicz, K. Poreba, Pol J Pharmacol Pharm, 31(3), 217**(1979)**

18. M. Bonesi, M. R. Loizzo, G. A. Statti, S. Michel, F. Tillequin, Menichini, Bioorg. Med. Chem. Lett, 20, 1990 **(2010).**

19. M.DELAUNAY. Thierry, Synthèses concises de pyrazoles et pyridones diversement fonctionnalisées dans le but d'effectuer des réactions de couplages croisés sélectifs,17 décembre 2010.

20. L.Knorr,Ber. Dtsch. Chem. Ges.,1883, 16, 2597.

21. H.v. Pechmann, Chem. Ber.,1898, 31, 2950.

22. R.E.Beveridge, D.Fernando, B. S. Gerstenberger, Tetrahedron Lett. 2010,51,5005.

23. Md. Jahangir Alam, Ozair Alam, Perwaiz Alam, Mohd Javed Naim, A Review on Pyrazole chemical entity and Biological Activity, International Journal of Pharma Sciences and Research,6(12 Dec 2015).

24. L. Navidpour; H. Shafaroodi; K. Abdi; M. Amin; M. H. Ghahreman; A. R. Dehpour; A. Shafiee, Bio. Med. Chem. Lett., 2006, 14, 2507.

25. H. J. Roth; A. Liemann,<<Pharmaceutical Chemestry>> , Ed 2, Vol. 1, 1988, p.213.

26. A. Palomer; F. Cabre; J. Pascual; J.Campos; M. A. Trujillo; A. Entrena; M. A. Callo; L. Garcia; D. Mauleon; A. Espinosa, J. Med. Chem., 2002, 45, 1402.

27. Corbett, E.L.; Watt, C.J.; Walker, N.; Maher, D.; Williams, B.G.; Raviglione, M.C. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. Arch. Intern. Med.,2003,163, 1009-1021

28. Nachega, J.B.; Chaisson, R.E. Tuberculosis drug resistance: a global threat. Clin. Infect. Dis., 2003, 36, 24-30.

29. R.J. O'Brien, P.P. Nunn, Am. J. Respir. Crit. Care Med. 163 (2001)1055-1058.

30. Aftab Ahmad, Asif Husain, Shah Alam Khan, Mohd. Mujeeb,Anil Bhandari, Synthesis, antimicrobial and antitubercular activities of some novel pyrazoline derivatives, Journal of Saudi Chemical Society (2014).

31. Hedia Marrakchi, Gilbert Lanéelle and AnnaÏkQuémard, InhA, a target of the antituberculous drug isoniazid, is involved in a mycobacterial fatty acid elongation system, FAS-II, Microbiology 146(2000),1, 289–296.

32. A. Banerjee, E. Dubnau, A. Quemard, V. Balasubramanian, Science,1994, 263, 227.

33. Suvarna. G. Kini, Anilchandra.R.Bhat, Byron Bryant,John.S.Williamson, Franck.E. Dayan, European Journal of Medicinal Chemistry, 2009,44, 492.

34. [9] Abunada, N.M.; Hassaneen, H.M.; Kandile, N.G.; Miqdad, O.A. Synthesis and biological activity of some new pyrazoline and pyrrolo [3,4-c]pyrazole-4,6-dione derivatives: reaction of nitrilimines with some dipolarophiles. Molecules, 2008, 13,1011-1024.

35. Shaaban, M.R.; Mayhoub, A.S.; Farag, A.S. Recent advances in the therapeutic applications of pyrazolines. Expert. Opin. Ther. Pat.,2012, 22, 253-291.

36. Kini, S.; Gandhi, A.M. Novel 2-pyrazoline derivatives as potential antibacterial and antifungal agents. Ind. J.Pharm. Sci., 2008, 70, 105-108.

37. Sharshira, E.M.; Hamada, N.M. Synthesis and antimicrobial evaluation of some pyrazole derivatives. Molecules, 2012, 17, 4962-4971.

38. Baseer, M.; Ansari, F.L.; Ashraf, Z. Synthesis, in vitroantibacterial and antifungal activity of some N-acetylated and non-acetylated pyrazolines. Pak. J. Pharm. Sci., 2013, 26, 67-73.

39. Kumar, S.; Meenakshi, K.S.; Kumar, P. Synthesis and anti- microbial activity of some (3-phenyl-5-(1-phenyl-3-aryl-1Hpyrazol-4-yl)-4,5-dihydro-1H-pyrazol-1-yl)(pyridin-4-yl)methanones: new derivatives of 1,3,5-trisubstituted pyrazolines. Med. Chem. Res.2013, 22, 433-439.

40. Abid, M.; Bhat, A.R.; Athar, F.; Azam, A. Synthesis, spectral studies and antiamoebic activity of new 1-N-substitutedthiocarbamoyl-3-phenyl-2-pyrazolines. Eur. J. Med. Chem., 2009, 44,417-425.

41. Sharma, S.; Sharma, P.K.; Kumar, N.; Dudhe, R. A review on various heterocyclic moieties and their antitubercular activity. Biomed. Pharmacother., 2011, 65, 244-251.

42. Rizvi, S.U.F.; Siddiqui, H.L.; Johns, M.; Detorio, M.; Schinazi, R.F. Anti-HIV-1 cytotoxicity studies of piperidyl-thienyl chalcones and their 2-pyrazoline derivatives. Med. Chem. Res., 2012, 21, 3741-3749.

43. Bashir, R.; Ovais, S.; Yaseen, S.; Hamid, H.; Alam, M.S.; Samim, M.; Singh, S.; Javed, K. Synthesis of some new 1,3,5-trisubstituted pyrazolines bearing

benzene sulfonamide as anticancer and antiinflammatory agents. Bioorg. Med. Chem. Lett., 2011, 21, 4301-4305.

44. Ozdemir, Z.; Kandilci, H.B.; Gumusel, B.; Calis, U.; Bilgin, A.A. Synthesis and studies on antidepressant and anticonvulsant activities of some 3-(2-furyl)-pyrazoline derivatives. Eur. J. Med. Chem.2007, 42, 373-379.

45. Xin He, Akram Alian, Paul R Ortiz de Montellano, Bioorganic & Medicinal Chemistry, 2007, 15, 6649.

46. Mack.R.Kuo, Hector.R.Morbidoni, David Alland, Scott.F.Sneddon, Brian.B.Gourlie, Mark.M. Stavesk, The Journal Of Biological Chemistry,2003, 278, 23, 20851.

47. Joseph.A.Maddry, Subramanian Ananthan, Robert.C.Goldman, Tuberculosis, 2009, 1.

48. D.H.Vyas, M.F. Dhakuk, S.Tala, J.D. Akbari, H.S. Joshi, Indian Journal of Heterocyclic Chemistry,2007, 16, 169.

49. [9] H.M.Kanjariya, T.V. Radhakrishnan, Ramachandran, P.Hansa, Indian Journal of Chemistry, 2004, 43B, 1569.

50. Marco radi, Vincenz O Brenardo, Beat Rice, Bechi Daniele, Tetrahydron letters, 2009, 50, 6572.

51. Preeti Gupta, Shahul Hameed, Rahul Jain, European Journal of Medicinal chemistry, 2004, 39, 805.

52. Jyoti Pandey, Vinod. K. Tiwari, Shyam.S.Verma, S.Vinita Chaturvedi, European Journal of Medicinal Chemistry, 2009, 44, 3350.

53. Afshin Fassihi, Zahra Azadpour, Neda Delbari , Lotfollah Saghaie et al, European Journal of Medicinal Chemistry, 2009, 44, 3253.

54. Sang-Ho Lee; Suhyun Kim; Min-Han Yun; Yong Sup Leeet al, Bioorganic & Medicinal Chemistry Letters, 2011, 21, 1515.

II.1 Introduction:

Computational chemistry is comprised of a theoretical (or structural) modeling part, known as molecular modeling, and a modeling of processes (or experimentations) known as molecular simulation. The term theoretical chemistry may be defined as the mathematical description of chemistry. The term computational chemistry is generally used when a mathematical method is sufficiently well developed that it can be automated for implementation on a computer **[1, 2].**

Molecular modeling involves the development of mathematical models of molecules that can be used to predict and interpret their properties. Computational Chemistry is the modeling of chemical phenomenon using computers rather than chemicals. The models used vary in their sophistication, chemoinformatics, molecular mechanics, Semi-empirical methods and ab initio quantum chemistry. All these methods, except the last, rely on empirical information (parameters, energy levels etc.), ab initio means “from the beginning” or “from first principles”, (quantum mechanics). Over the last four decades powerful molecular modeling tools have been developed which are capable of accurately predicting properties of molecules. These developments have come about largely due to the dramatic increase in computer speed and the design of efficient quantum chemical algorithms **[3, 4].**

Molecular modeling is concerned with ways to describe the behavior of molecules and molecular systems. Computational techniques have revolutionized molecular modeling to the extent that calculations could not be performed without the use of a computer molecular modeling is invariably associated with computer modeling or computational chemistry. This discipline encompasses not only quantum mechanics (QM) but also molecular mechanics (MM), conformational analyses and other computer-based methods for understanding and predicting the behavior of molecular systems **[5].**

Molecular modeling is a tool for researchers concerned about structure and reactivity of molecules. Knowledge of the structure of buildings molecules makes it possible to understand what is achieved in a physical transformation, chemical or biological. It can also make it possible to envisage such transformations. The understanding like forecasting is greatly facilitated when one can visualize the structures. The essential question is to represent a molecule on the screen as close as possible to the "reality"**[6].**

II.2 Representation of calculation methods:

The development of molecular modeling techniques has opened up a new highway to a more detailed picture of molecular level information; molecular modeling is a rapidly evolving discipline that has undoubtedly benefited a lot from advances in computer science **[7].**

All molecular computing techniques can be classified into three general categories:

- ab initio and calculations of electronic structures operating in density
- semi-empirical methods
- Empirical Methods and Molecular Mechanics

II.2.1 Molecular mechanic:

Molecular mechanics appeared in 1930, but developed from 1960s, with advances in computer accessibility and performance **[8].** The "mechanical" molecular model was developed out of a need to describe molecular structures and properties in as practical a manner as possible, molecular mechanics methods are based on the laws of classical physics. They do not explicitly treat the electrons: the electronic effects, such as chemical bonds, are included implicitly in the energy function, known as the force field, which describes the interactions between the nuclei only.

If a molecule is too big to effectively use a semi empirical treatment, it is still possible to model its behavior by totally avoiding quantum mechanics. The methods, referred to as molecular mechanics, set up a simple algebraic expression for the total energy of a compound, with no necessity to compute a wave function or total electron density. Molecular mechanics allows the modeling of very large molecules, such as proteins and segments of DNA, making it the primary tool of computational biochemists **[9].**

The molecular mechanics “energy” of a molecule is described in terms of a sum of contributions arising from distortions from “ideal” bond distances (stretch contribution), bond angles (“bend contributions”) and torsion angles (torsion contribution), together with contributions due to “non-bonded” (vander 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.

$$E_{strain} = \sum_A^{bonds} E_A^{stret} + \sum_A^{bond\ angie} E_A^{bend} + \sum_A^{tortion\ angle} E_A^{tor} + \sum_A^{non\text{-}bonded} \sum_B E_{AB}^{non\text{-}bonded} \quad \textbf{(II.1)}$$

The first three summations in equation 1 are over all “bonds”, all “bond angles” and all “torsion angles”, respectively. Thus, information about bonding is “part of the input” to a molecular mechanics calculation, in contrast to a quantum chemical calculation where it is “part of the output”. The last summation in equation 1 is over all pairs of atoms which are not bonded **[10].**

II.2.1.1 Force field:

Very briefly, a force field is a mathematical function which returns the energy of a system as a function of the conformation of the system. But a better idea may be obtained by considering the situation physically. Consider a molecule as a collection of atoms held together by elastic forces. (If you want to get even simpler than one could consider a molecule to be a collection of point masses connected by elastic springs). Now the forces can be written in terms of potential energy functions of various structural features such as bond lengths, bond angle, non bonded interactions etc. The force field is the combination of these potential energy terms. Hence force fields are also sometimes referred to as potentials **[11].**

The writing and parameterization of these force fields differ, but they all have valence terms and non-bonded atom interaction terms **[12].**

The force field of MM+ is an extension of MM2 that has been developed by Allinger his collaborators. **[13]** And is designed primarily for small molecules organic although it is extended to peptides **[14]** and other systems **[15].** The efforts of Allinger's group have been focused more on very precise results for certain classes of molecules than on developing a generic (but less precise) method that can be applied to almost all situations in organic chemistry.

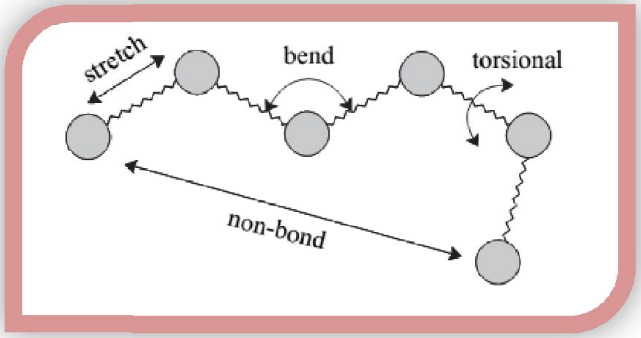
In this work, we used the MM + force field which we are now going to present individual contributions.

II.2.1.2 Field of force in molecular mechanics:

The force field energy is written as a sum of terms, each describing the energy required for distorting a molecule in a specific fashion.

$$E_{strain} = E_{Str} + E_{bend} + E_{tors} + E_{vdw} + E_{el} \quad \textbf{(II.2)}$$

E_{str} is the energy function for stretching a bond between two atoms ,E_{bend} represents the energy required for bending an angle ,E_{tors} is the torsional energy for rotation around a bond E_{vdw} and E_{el} describe the non-bonded atom–atom interactions**[16].**



FigureII.1: illustration of the fundamental force field energy terms

II.2.1.2.a Energy of interaction between the bound atoms:

- ✓ **The stretch energy :**

The potential energy expression for this change is given as:

$$\mathbf{E(r) = \frac{1}{2[\,k_1(r-r_0)^2]}} \quad \textbf{(II.3)}$$

Where **K1:** is the constant of elongation or constant of Hooke

L_o: the length of the reference link.

L: the length of the link in the model **[17].**

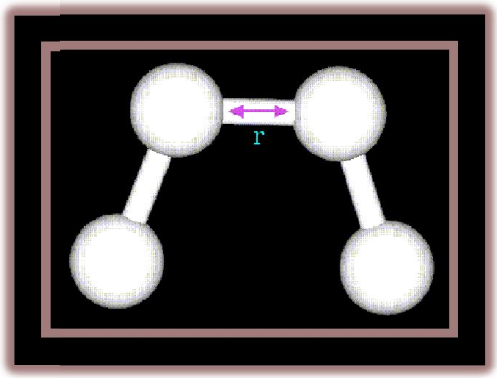


Figure II.2: elongation between two atoms

✓ **Angular bending energy (inflection):**

The potential energy expression associated with bending is given by:

$$\mathbf{E(\theta) = 1/2[K_f\,(\theta - \theta_0)^2}\ \textbf{(II.4)}.$$

Kf: bending constant.

θ_o: binding reference angle.

θ: connection angle **[17].**

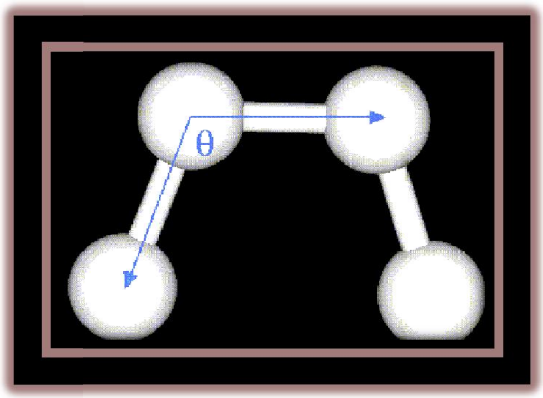


Figure II-3: Deformation of the valence angles

✓ **The torsional energy:**

The torsion potential is a Fourier series that accounts for all 1–4 through-bond relationships**:**

$$\boldsymbol{E(\tau) = 1/2[V_1(1 + \cos\tau) + V_2(1 - \cos 2\tau) + V_3(1 + \cos 3\tau)]}\ \textbf{(II.5)}$$

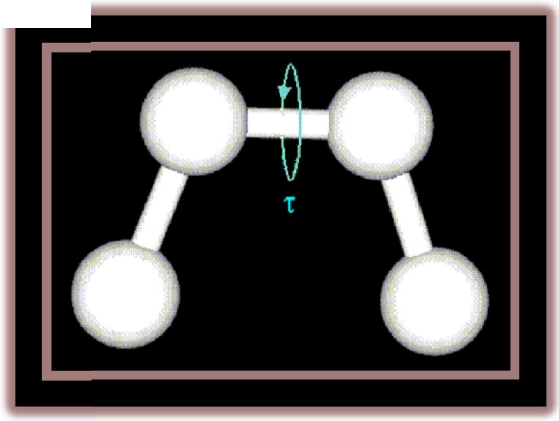


Figure II-4: Dihedral angle formed by the 1-2-3-4 atoms.

II.2.1.2.b Energy of interaction between the none bound atoms:

- ✓ **The van der Waals energy:**

The energy expression takes the form of:

$$E_{ij} = \sum_i \sum_j -\frac{A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} \quad \textbf{(II.6)}$$

rij: distance between the two unbound atoms.

Aij and Bij are constants of Van Der Waals **[17].**

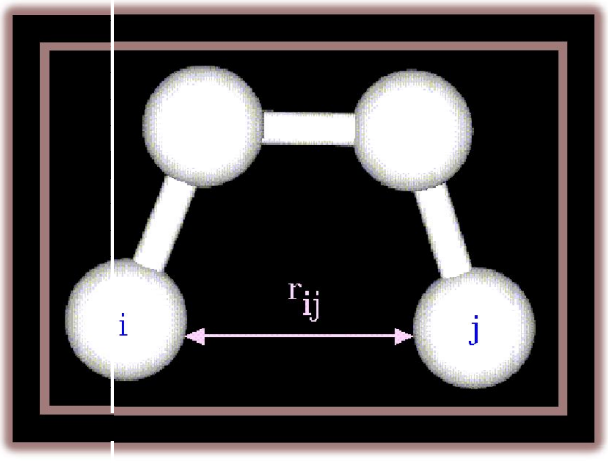


Figure II.5: VDW interaction energies

✓ **The electrostatic energy: charges and dipoles**

$$\cup \mathbf{coulomb}\,(\mathbf{r}) = \frac{\mathbf{q_i q_j}}{\mathbf{4\pi\varepsilon_0 r}} \textbf{ (II.7)}$$

qi and qj are the partial charges of two atoms and ε0 is the permittivity of the free space.

✓ **The hydrogen bond energy:**

The phenomena of repulsion and electronic delocalisation intervene. Several types of potential energy functions have been developed to account for the directivity of the hydrogen bond. Currently, the most used functions for expressing these interactions in important molecular systems are often simplified:

$$\mathbf{E_H = A/r_{ij}^{12} - B/r_{ij}^{10}} \textbf{ (II.8)}$$

II.2.1.3 Existing force field in MM:

➢ **AMBER:**

Assisted model building with energy refinement (AMBER) is the name of both a force field and a molecular mechanics program. It was parameterized specifically for proteins and nucleic acids. AMBER uses only five bonding and nonbonding terms along with a sophisticated electrostatic treatment. No cross terms are included. Results are very good for proteins and nucleic acids, but can be somewhat erratic for other systems **[2]**

➢ **CHARMM:**

CHARMM (Chemistry at Harvard Macromolecular Mechanics, developed by Mackerell and Karplus, et al., 1995) was parameterized by experimental data. It has been used widely for simulations ranging from small molecules to solvated complexes of large biological macromolecules **[1].**

➢ **Gromos:**

(Groningen Molecular Simulation) developed at the University of Groningen and the ETH (Eidgenössische Technische Hochschule) of Zurich**[17].** A force field that comes as part of the GROMOS software, a general-purpose molecular dynamics computer simulation package for the study of biomolecular systems.

- **MM2, MM3 and MM4 :**

It was developed by Allinger in 1976 and it's the field of force most used by the community of organic chemists**[18].**

MM2 is the first force field developed by Allinger and colaborator**[19-21].** He wasdesigned for simple molecules (alkanes, alkenes, non-conjugated alkynes, amines ...), but its improved versions MM3 (1989) **[22]** and MM4 (1996) **[23]** allows to treat organic molecules more and more complex.

- **MM+:**

Is an extension of the force field MM2, with the addition of some parameters additional **[24]**. MM+ is a robust force field, it has the ability to take in considering the parameters neglected in other force fields and so can apply for more complex molecules such as inorganic compounds **[25]**.

- **OPLS:**

(Optimised Potentials for Liquid Simulation), as his name indicates, is designed to optimize the potential that allows the description of solvation properties. He is written by W. L. Jorgensen and J. Tirado Rives **[26].** It was developed by Karplus and coliaborator **[27].**

- **SPASIBA :**

(Spectroscopic Potential Algorithm for Simulating Biomolecular conformational Adaptability). Elaborated by Gerard Vergoten and al (1995). It combines the benefits of Urey-Bradley-Shimanouchi's modified spectroscopic field and field AMBER Molecular Mechanics **[28].** It allows to find at the same time the structures, conformational energies and vibrational frequencies to a minimum energy of a molecule **[29].**

II.2. 2 Quantum mechanic:

II. 2. 2. 1 Basics of quantum mechanics:

The work carried out at the beginning of the twentieth century by Planck, Einstein, Bohr, De Broglie, Schrödinger and Heisenberg led to the development of the mechanics of microsystems. In 1925, thanks to the efforts of W. Heisenberg and E. Schrödinger and P.

Dirac, J. von Neumann, N. Bohr, M. Born and others, a new Mechanics was created: Quantum Mechanics **[30-32]**, which has explained many properties physical properties, such as the chemical properties of the elements and the formation of chemical **[33].**Modeling methods based on quantum mechanics **[34]** are intended to describe the system is studying by a wave function that can theoretically be determined by Solving the Schrödinger equation **[35]**. This equation connects stationary states of a molecular system and the energies associated with it to a Hamiltonian operator and their wave function.

$$\hat{H}\Psi=E\Psi \quad \textbf{(II.9)}$$

where $\hat{H}$ is the Hamiltonian operator and E is the energy of the system.

The total Hamiltonian of a molecule with N nuclei and n electrons is defined by the sum of five terms (kinetic term of the electrons, kinetic term of the nuclei, term of electron - electron repulsions, term of nuclei - nuclei repulsions and term of attractions electrons - nuclei).

$$\boldsymbol{H = T_e + V_{ee} + V_{eN} + V_{NN} + T_N} \quad \textbf{(II.10)}$$

Where $T_e$and T_N is the kinetic energy operator for electrons and nuclei. V_{ee}, V_{eN} and is columbic energy operator for electrons only, between electrons and nuclei, and nuclei only, respectively**[36]**.

- **Ab initio methods:**

Ab initio methods are non-empirical methods, all integrals are rigorously calculated and there is no approximation to make except that of Born Oppenheimer and the OM-CLOA approximation. In ab-initio methods, all particles (nucleus and electrons) are treated explicitly. No parameters are used empirical in the calculation of energy.

The ab initio methods are divided into two sub-families: the Hartree-Fock methods (HF, RHF, UHF, ROHF) (Hartree, 1928, Fock, 1930), and post Hartree-Fock methods, (MPn, CAS,) (Moller, 1934). The main difference between these two methods is that the Electronic interactions are neglected in HF methods and reintroduced into post HF methods. These methods can only be applied to systems of some dozens of atoms for HF methods and a dozen atoms only for post HF methods **[37]**.

- **DFT methods:**

The density functional theory is based on the postulate proposed by Thomas and Fermi who says that properties can be described in terms of functionalities of electronic density, by applying locally appropriate relationships to a homogeneous electronic system **[38].** Hohenberg and Kohn, in 1964 **[39,40],** took up Thomas-Fermi's theory and showed that there is a functional energy E [ρ (r)] associated with a variational principle, which has allowed to lay the groundwork for the theory of the functional density The density functional theory is based on the Hohenberg-Kohn theorem **[41],** which establishes that the energy of a system in its ground state is a functional the electronic density of this system, ρ (r), and that any density, ρ '(r), other than density real energy necessarily leads to a higher energy. So unlike the methods previous, the theory of the density functional is not to look for a complex wave function, ψ, with 3N-dimensions describing the system to be studied, but rather a simple function with three dimensions: the total electronic density ρ **[42].** There are three types functional exchange-correlation energies: the local functional, the gradient-correcting functional and hybrid functional.

For the past 30 years density functional theory has been the dominant method for the quantum mechanical simulation of periodic systems. In recent years it has also been adopted by quantum chemists and is now very widely used for the simulation of energy surfaces in molecules. In this lecture we introduce the basic concepts underlying density functional theory and outline the features that have lead to its wide spread adoption. Recent developments in exchange correlation functionals are introduced and the performance of families of functional reviewed **[43].**

II. 2. 3 Semi-empirical methods:

Most molecular computations done by organic chemists, especially those examining minimum energy geometries, are done using this method because it provides the best compromise between speed and accuracy. This method can be thought of as a hybrid of molecular mechanics-type models based on experimentally measured empirical data and pure theory quantum chemical, thus the name semi-empirical. It uses the Schrödinger equation approximations, but in order to make the calculations less time-consuming, it only

calculates the locations of valence electrons, not all electrons. For the inner shell electrons, empirical data from typical organic molecules is used to estimate their locations **[44].**

Semi empirical methods modify Hartree-Fock (HF) calculations by introducing functions with empirical parameters. These parameters are adjusted with experimental conclusions to improve the quality of computation. The real cost of computation is due to the two-electron integrals in the Hamiltonian that has been simplified in this method **[4].**

The most difficult energy terms to calculate are estimated from the data The calculation times are considerably shortened, but the method depends on the compounds used to calibrate it. According to the nature of the approximations used **[45],** there are several variants:

- **Complete Neglect of Differential Overlep: (CNDO)**

First semi empirical method, it was proposed by Pople, Segal and Santry in 1965. Method with some flaws, among others: it does not take into account Hund's rule.

- **Intermediate Neglect of Differential Overlap: (INDO)**

Proposed by Pople, Beveridge and Dobosh in 1967. It distinguishes between singlet and the triplet states of a system while keeping the exchange integrals.

- **Neglect of Diatomic Differential Overlap: (NDDO)**

Proposed by Poplein 1965. All bicentrées bielectronic integrals are retained

- **Austrin Model 1: (AM1)**

Proposed by Dewar in 1985. He attempted to correct the MNDO defects.

- **Parametric Method 3: (PM3)**

Proposed by Stewart in 1989. Presents a lotin common with AM1, there is still a debate about the relative merits of parametrisation of each of them.

II. 2. 4 Metric methods of drug selection " virtual screening " and drug-likeness:

Discovery of new bioactive leads for subsequent optimization into drugs is both time consuming and expensive process. Two main approaches are currently available for lead discovery, namely, high throughput (in vitro) screening and computer-aided virtual (in silico) screening. Normally, in silico techniques are implemented aspre-filters to enrich the success rates of high throughput screening campaigns **[46]**.

Computational screening of databases has become increasingly popular in the pharmaceutical research. Virtual screening uses computer based methods to discover new ligands on the basis of biological structures. Virtual screening is divided into structural based screening (docking) and screening using active compounds as templates (ligand based virtual screening) **[50]**.

II. 2. 4. 1 Virtual screening:

Virtual screening(VS) is a computational technique used in drug discovery to search libraries of small molecules in order to identify those structures which are most likely to bind to a drug target, typically a protein receptor or enzyme**[47][48].**Virtual screening has been defined as the "automatically evaluating very large libraries of compounds" using computer programs**[49].**As this definition suggests, VS has largely been a numbers game focusing on how the enormous chemical space of over 1060conceivable compounds**[50]**can be filtered to a manageable number that can be synthesized, purchased, and tested. Although searching the entire chemical universe may be a theoretically interesting problem, more practical VS scenarios focus on designing and optimizing targeted combinatorial libraries and enriching libraries of available compounds from in-house compound repositories or vendor offerings. As the accuracy of the method has increased, virtual screening has become an integral part of the drug discovery process **[51]**.

II.2. 4.2 MPO: Multi-Parameter Optimization:

A high quality drug must strike a balance between many often conflicting properties, including power propertie and ADME. MPO methods integrate multi-property data to effectively identify the chemicals that are most likely to achieve an appropriate balance for a therapeutic project, therapeutic discovery project **[52]**.

In researching the needs of an ideal MPO method for drug discovery, the following factors should be considered:

- **Interpretation:** Ownership criteria and their impact on composite priority should be easy to understand.
- **Flexibility:** Each project will have a different set of ownership criteria depending on the therapeutic goals of the project, the intended route of administration and the competitive conditions in the market. The project team should be able to define appropriate criteria based on their experience or historical evidence.
- **Weighting:** the project team should be able to assign different weights to each property criterion, as different criteria will have different degrees of importance for the project outcome.
- **Uncertainty:** It is important to avoid rejecting potentially valuable compounds based on a property value that does not meet a criterion if this value presents a high degree of uncertainty.

II.2. 4.3 Qualitative method: SAR/SPR

Generality:

The process of drug development is time-consuming and cost-intensive. Several years are required for lead identification, optimization, in vitro and in vivo testing before starting the first clinical trials [**53, 54].** A new strategy introduced into drug discovery is structure–property relationships (SPRs). This is complementary to SAR. The structures of compounds are correlated to their property performance. SPR allows medicinal chemists to understand how structural modifications improve properties for their scaffold. Thus, the established strategy of structure-based design is supplemented with the new strategy of "property-based design" by van de Waterbeemd and col **[55]** the study and modification of structure to achieve property improvement.

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

- Strong target binding using structure-based design and the structure–activity Relationship (SAR).

❖ High performance at in vivo barriers, using property-based design and the Structure–property relationship (SPR).

b. Structure property properties:

- **Molecular Volume And Surface Area :**

Molecular volumes are often calculated by a numerical integration grid technique **[56]** that we can illustrates by considering the trivial problem of finding the volume of an atom whose van der Waals radius is R (the volume is of course $\frac{4}{3}\pi R^3$)Figure II.6 shows a two-dimensional representation of the atom whose van der Waals radius is R, surrounded by a three-dimensional grid of equally spaced points.

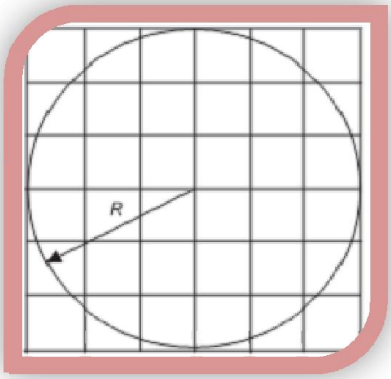


Figure II.6: Grid around atom

$$V / 8R^3 = n_a / n \quad \text{(II.11)}$$

For a polyatomic, we have to give special consideration to grid points that lie in the overlap region. Figure II.7 shows two atoms, A and B, with radii R_A and R_B. The overlap region is labeled X.

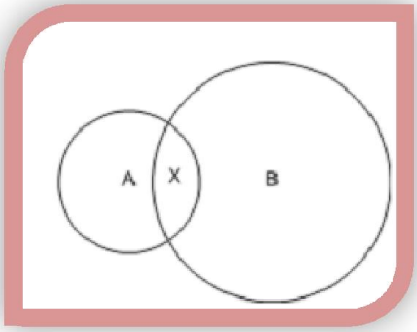


Figure II.7: Atoms A, B and overlap region X.

- **Molecular Refractivity (MR) :**

The molar refractivity is a steric parameter that is dependent on the spatial array of the aromatic ring in the synthesized compounds. The spatial arrangement also is necessary to study the interaction of the ligand with the receptor **[57]**.

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

$$\mathbf{MR} = \frac{\mathbf{n^2-1}}{\mathbf{n^2+2}} \times \frac{\mathbf{Mw}}{\boldsymbol{\rho}} \quad \textbf{(II.12)}$$

where n is the refraction index, Mw is the molecular weight, and ρ is the density.

The (n^2-1/n^2 +2)term provides a correction factor by defining how easily the substituent can be poralized, whereas the Mw/ρ term defines a volume.

- **Molecular Polarizability (Pol):**

Molecular Polarizability of a molecule characterizes the capability of its electronic system to be distorted by the external field, and it plays an important role in modeling many molecular properties and biological activities **[58]** .

Highly polarizable molecules can be expected to have strong attractions with other molecules.

$$\mathbf{P(e) = \varepsilon_o\, \alpha\, E} \quad \textbf{(II.13)}$$

- **Molecular Weight (MW):**

Molecular weight descriptor has been used as a descriptor in systems such as transport studies where diffusion is the mode of operation. It is an important variable in QSAR studies pertaining to cross resistance of various drugs in multi-drug resistant cell lines **[64].**Molecular weight is correlated with the size of the molecule **[59]**. Additionally, the systemic clearance of a compound is inversely proportional to the molecular weight **[60]**.

- **Hydration Energy (HE):**

Hydration energy is very important in the selectivity filter of ion channels where the drug is almost entirely strip from the hydration water**[61]** .

Indeed, in the biological environments the polar molecules are surrounded by water molecules.

- **Partition Coefficient (Log P):**

One of the most important physicochemical properties much interest in QSAR studies is lipophilicity (or hydrophobicity). Because it directly relates to solubility in aqueous phase, to membrane permeation (an important factor contributing to the toxicity of Chemicals) , and to its contribution to ligand binding at the receptor site.

The ability of a molecule to cross the biological membranes (permeability) is a very important bio-pharmaceutic parameter that governs the absorption, distribution, metabolism and excretion (pharmacokinetics) of a drug.

The partition coefficient P, defined as the ratio of molar concentration of a chemical dissolved at equilibrium in octanol phase C_{oct} to its molar concentration in aqueous phase C_{aq} **[62-64],** and is given by the equation:

$$P = \left(\frac{Coct}{Caq}\right)_{equilibrium} \quad \textbf{(II.14)}$$

II.2. 4.4 Quantitative Methods: QSAR / QSPR:

a. Generality :

Early work using the QSPR / QSAR methodology as employed currently are due to Hansch **[65]** and Free and Wilson **[66].** On the one hand, Hansch proposed models directly linking the biological activity of compounds with hydrophobic, electronic and steric at the molecular level. On the other hand, Free and Wilson have developed empirical models, called group contributions, for the study of biological activity.

b. QSAR/ QSPR models:

In the past decades, the Quantitative Relations Structure-Activity / Property (QSAR / QSPR) have become a powerful theoretical tool, alternative to mechanics quantum, for the description and prediction of the properties of molecular systems complex in different environments. The QSAR approach proceeds from the hypothesis of a unambiguous correspondence between any physical property, chemical affinity, or biological activity of a chemical compound and its molecular structure **[67].**

QSAR is an attempt to eliminate the chance factor from the drug's design by establishing a relationship in the form of a mathematical equation between the biological activity and measurable physico-chemical parameters of a drug that represents its properties such as lipophilicity , shape and electronic distribution, which have major effects on activity **[68,69].**

The relationship between these numerical properties and activity is described by a general equation

Propriété/ Activité = f (D1, D2,… Dn…) (II.15)

In the equation, the biological activity is normally expressed as log (1 / C), where C is the concentration of the compound needed to produce a standard response in a given time), D1, D2, Dn are descriptors of structures Molecular (properties derived from the structure of the molecule). These properties that influence the activity of a drug are very diverse, the most important being lipophilicity, steric effects and electronic effects **[70]**.

The mathematical relationship between a parameter or several physicochemical parameters and the biological activity of a compound can be expressed through the Hansch equation:

$$\textbf{Log } 1/C = -K_1(\log P)2 + K_2 \log P + K_3 \sigma + K_4 Es + K5 \quad \textbf{(II.16)}$$

This equation is obtained by regression analysis. The value of the coefficient r obtained after this analysis gives information on the concordance of the parameters that are used to form the Hansch equation. If the result obtained for r is> 0.9 it means that the parameters used are valid and that the equation can be exploited for the rest of the compounds.

The Hansch equation is used to predict the biological activity of compounds whose structure is similar to those used to form this equation **[71]**.

c. QSAR tools and techniques :

1. Molecular Descriptors :

The crucial point in the QSPR approach is an appropriate description of the molecular structures. The chemical descriptors take account of the different aspects of the chemical information. The molecular descriptor expresses chemical information transformed and encoded from a molecule and effectively solves chemical, pharmaceutical, and toxicological problems.

The advantage of theoretical molecular descriptors is in developing compounds that have never been synthesized or explored experimentally. Molecular descriptors are well known for their ability to establish linear regression relationships with physicochemical and biological properties **[72]**

2. Statistical method:

QSAR/QSPR is basically a statistical approach correlating the response property or activity data with descriptors encoding chemical information. Such correlation may be derived either in a regression-based approach (in cases where the response property is quantitative and available in a continuous scale) or a classification-based approach (in cases where the response property is graded or semi-quantitative) **[73].**

The most commonly used regression-based approaches are as follows:

• Multiple linear regressions (MLR)

• Partial least squares (PLS)

Some of the common classification-based approaches are as follows:

• Linear discriminant analysis (LDA)

• Cluster analysis

3. Multiple Linear Regressions:

The Multiple Linear Regression (MLR) **[74]** is an extension of the classical regression method to more than one dimension. MLR calculates QSAR equations by performing standard Multivariable regression calculations using multiple variables in a single equation. MLR expresses a single dependent variable (y) as a linear combination of multiple independent variables (x):

$$Y = Y_0 + a_1 X_1 + a_2 X_2 + \cdots + a_n Xn \quad \text{(II.17)}$$

- **Description of the method:**

Multiple linear regression analysis (MLR) is a statistical method that examines the cause-and-effect relationships between dependent and independent variables, in MLR the relationship between the input variable more than one (x1, x2, ...xn) and a dependent variable) is examined **[75].**

If we assume that the relation is well represented by a linear model in the regressed variables, an appropriate model can be :

$$y = b_0 + b_1 x_1 + b_2 x_2 \ldots.+e \quad \text{(II.18)}$$

In equation II.18, the b's are unknown constants called regression coefficients and the goal of the regression analysis is to estimate these constants, the MLR algebraic model is defined in equation II.18 and in matrix notation **[76]:**

$$y = X\beta + \varepsilon \quad \text{(II.19)}$$

X matrix is called the design matrix:

$$Y = \begin{bmatrix} Y1 \\ Y2 \\ . \\ . \\ Yn \end{bmatrix} \quad X = \begin{bmatrix} 1 & X1 \\ 1 & X2 \\ . & . \\ . & . \\ 1 & Xn \end{bmatrix} \quad \beta = \begin{bmatrix} \beta 0 \\ \beta 1 \\ . \\ . \\ \beta n \end{bmatrix} \quad \text{et} \quad \varepsilon = \begin{bmatrix} \varepsilon 0 \\ \varepsilon 1 \\ . \\ . \\ \varepsilon n \end{bmatrix}$$

Multiple Linear Regression (MLR) techniques based on least squares procedures are widely used to estimate the coefficients involved in the model equation.

4. Model validations:

Whatever the ultimate goal of the QSAR model, it must be validated before being interpreted or used for predictive purposes. There are different solutions for ensure the validity of a model. LOO (leave-one-out) cross-validation is a process that tests the predictive accuracy of a model, this method based on the calculation of some statistical parameters such as: the sum of the residual squares RSS (PRESS), Sum of total squares TSS, adjustment quality R^2_{adj}, coefficient of cross-validation correlation R^2_{CV}, standard validation of prediction errors (SPRESS) and the prediction error (PE). These statistical parameters are calculated from following relationships **[77]**:

PRESS (residual sum predicted from squares) **PRESS** = $\sum(\mathbf{Yobs} - \mathbf{Ycalc})^2$

TSS (total sum of squares) **SSY** = $\sum(\mathbf{Yobs} - \mathbf{Ymean})^2$

R2adj (adjusted R-squared) $\mathbf{r^2_{adj}} = \left([1 - (\mathbf{r^2})]\left(\frac{n-1}{n-p-1}\right)\right)$

R2CV (cross-validated correlation coefficient) $\mathbf{r^2_{CV}} = 1 - \frac{PRESS}{SSY}$

SPRESS (standard validation of prediction errors) $\mathbf{S_{PRESS}} = \sqrt{\frac{PRESS}{n}}$

PE (prediction error) **PE =** $\mathbf{0.67\ 45(1-r^2)/\sqrt{n}}$

II.2.5. Several Purposes and Applications of QSAR Models:

- ✓ QSAR models are used to predict the activity of new (hypothetical) chemical compounds, even before their synthesis. Thus, QSARs can save time and experimental resources for synthesizing and biological testing of large numbers of compounds. QSARs offer possibilities for reduction or replacement of animal use in research and toxicity testing.
- ✓ (Q)SARs can lead to better understanding of the mechanisms of interaction between compounds and biological systems. They may reveal important structural features for the biological effect.
- ✓ QSAR models provide useful information about a dose range for a biological effect of a compound, thus helping the experimental design (selection of doses and tests) in drug research and toxicity testing **[78]**.

II.3. Used programs and materials:

This work from this thesis was done within the computer chemistry team and pharmaceutical laboratory LMCE (Laboratory of Molecular Chemistry and Environment) at the University of Biskra.

The study of the electronic and structural properties 1, 2-diazole and its derivatives, was realized by the molecular modeling (molecular mechanics, Ab-initio, PM3 and QSAR), using the software HyperChem (8.0.7) and the software Gaussian (09) in a station (HP Intel® Xeon® CPU X3430 microprocessor, 4GB RAM) and in a PC (ACER Intel® Core ™ i3 processor 3217U CPU @ 1.80 GHz, 4GB ofRAM).

The qualitative study was done using HyperChem Software and MarvinSketch in ChemAxon products, and the online Molinspiration website.

The study of multilinear regression (RML), was established using SPSS 20 software.

In addition, we have used other software such as: GaussView (5.0), ChemDraw Ultra (8.0).

References:

1. K. I. Ramachandran ,G. Deepa ,K. Namboori, Computational Chemistry and Molecular Modeling: Principles and Applications 4, Springer.
2. David C. Young, Computational Chemistry: A Practical Guide for Applying Techniques to Real-World Problems, John Wiley & Sons 2001.
3. Modeling Molecular Structures with HyperChem, Computational Chemistry Laboratory.
4. Dr Andrew Gilbert, Introduction to Computational Chemistry: Theory, the austrilian national university,2011.
5. Barmak Mostofian, Filipp Frank, Basic Principles in Molecular Modeling, University of Heidelberg, Molecular Biotechnology (Summer 2004).
6. F.Lerbet-Sereni,Expériences de la modélisation ,modélisation de l expérience, Edition
   L'Harmattan,2004.
7. M.Shahamat, A.D.Rey. Chemical Engineering Science,121, 100,2015.
8. D.H.Andrew, Physical Review .36(1930)544.
9. E. Stuper, W. Brugger, and P. Jurs, Computer-Aided Analysis of the Relation Between Chemical Structure and Biological Activity, Mir, Moscow, 1982.
10. Warren J. Hehre, A Guide to Molecular Mechanics and Quantum Chemical Calculations, Wavefunction, Inc 2003.
11. Rajarshi Guha (9915607) and Rajesh Sardar (9915610), Force Fields in Molecular Mechanics, March 21, 2001.
12. Fabien Archambault ; 2011 ; « Vers une modélisationplus réaliste des systemes biologiques » thèse de doctorat, Université de Nancy ; PP 4.
13. N.L .Allinger, Journal of the Américan Chemical Society. 99(1977) 8127.
14. J. Lii et. Al. ; 1989; J. Comp. Chem. 10, 503.
15. K. B. Lipkowitz; 1992; QCPE Bulletin; India University; 12, 1.
16. Frank Jensen, Introduction to Computational Chemistry, John Wiley & Sons 2007.
17. N.L.Allinger;K. Chen. J.H . Lii, Journal of Computational Chemistry. 17 **(1996)** 642**.**
18. Frankie Tristram, Viviane Durier, Gérard Vergoten, Journal of Molecular Structure., 378, 249-256, **1996.**

19. B.J. Alder, T. E. Wainwright, J. Chem. Phys, 27, 1208., **1957.**

20. A. Rahman, Phys. Rev, 136, A 405. **1964.**

21. A. Rahman, F. H. Stillinger, J. Chem. Phys., , 5, 3336, **1971**.

22. A. Hoquet .M. Langgard, Molecular modeling annual. 4**(1988)**94.

23. G.B. Jones.B.J. Chapman, Synthesis. 1995 **(1995)** 475.

24. D. Jackson, Annu. Rep. Prog. Chem., Sect. B: Org. Chem. 85 **(1988)** 17.

25. J. Smith.M. Karplus, Journal of the American Chemical Society. 114 **(1992)** 801.

26. S. Weiner; P. Kollman; D. Nguyen.D. Case, J. Comput. Chem. 7 **(1986)** 200.

27. F. Tristram; V. Durier.G. Vergoten, Journal of Molecular Structure: THEOCHEM. 378 **(1996)** 249

28. P.A.M. Dirac, Lectures on Quantum Mechanics, Belfer graduate School of Science, Yeshiva University, New York, **1964.**

29. W. Heisenberg, The Physical Principles of the Quantum Theory, Dover, New York, **1949.**

30. L. Landau et E. Lifchitz, Mecanique quantique, Editions Mir, Moscou, **1967**.

31. C.M. Marle, De la mecanique classique à la mecanique quantique : pourquoi et comment quantifier?, Universite Pierre et Marie Curie Paris, France, **2003.**

32. A.R. Leach, Quantum Mechanical Models, in Molecular modelling: Principles and applications Addison Wesley Longman Ltd., Harlow**, 1996.**

33. E. Schrodinger, Ann. phys. Leipzig, 76, 361 **(1926).**

34. W. C. Ripika & J. M. Blaney , " Topics In Sptereochemistry ", Eds .Eliel & Wilen, **1991.**

35. C. Piron, Mécanique quantique: Bases et applications, PPUR presses polytechniques, **1998.**

36. L.R. Domingo; E.Chamorro.P.Perez, J.Org. Chem. 73 4615.

37. A.D. Becke, The Journal of Chemical Physics. 98 **(1993)** 5648.

38. C. Lee; W. Yang.R. Parr, Phys. Rev. A. 38 **(1988)** 3098.

39. P. Hohenberg.W. Kohn, B864. (1964).

40. W. Kohn, phys. Chem. 100 **(1996)** 12974.

41. N. M. Harrison, Department of Chemistry, Imperial College of Science Technology and Medicine, SW7 2AY, London and CLRC, Daresbury Laboratory, Daresbury, Warrington, WA4 4AD.

42. . Houmam BELAIDI, These de doctorat, UNIVERSITE MOHAMED KHIDER BISKRA , **2016.**

43. C.S. Tsai, An Introduction to Computational Biochemistry, Wiley-Liss, New York,**2002.**

44. Mutasem O. Taha Mixing Pharmacophore Modeling and Classical QSAR Analysis as Powerful Tool for Lead Discovery Dept. of Pharmaceutical Sciences, Faculty of Pharmacy, University of Jordan Jordan

45. Vivek VYAS *, Anurekha JAIN, Avijeet JAIN, Arun GUPTA, Virtual Screening: A Fast Tool for Drug Design. Sci Pharm. **2008**; 76: 333–360.

46. Rester U **(July 2008)**. "From virtuality to reality - Virtual screening in lead discovery and lead optimization: a medicinal chemistry perspective".Current Opinion in Drug Discovery & Development.11(4): 559–68.PMID 18600572.

47. Rollinger JM, Stuppner H, Langer T **(2008)**. "Virtual screening for the discovery of bioactive natural products".Progress in Drug Research. Progress in Drug Research.65(211): 211, 213 49.doi:10.1007/978-3-7643-8117-2_6.ISBN 978-3-7643-8098-4.PMID 18084917.

48. Walters WP, Stahl MT, Murcko MA(1998). "Virtual screening – an overview".Drug Discov. Today.3(4): 160–178.doi:10.1016/S1359-6446(97)01163-X.

49. Bohacek RS, McMartin C, Guida WC **(1996).** "The art and practice ofstructure-based drug design: a molecular modeling perspective".Med. Res. Rev.16: 3–50.doi:10.1002/(SICI)1098- 1128(199601)16:1.

50. McGregor, Malcolm J; Luo, Zhaowen; Jiang, Xuliang **(June 11, 2007).** "Chapter 3: Virtual screening . in drug discovery". In Huang, Ziwei.Drug Discovery Research. New Frontiers in the Post-Genomic Era. Wiley-VCH: Weinheim, Germany. pp. 63–88.ISBN 978-0-471-67200-5.

51. S. Durdagi, C. T. Supuran,; A. T. Srom ; N. Doostdar; M.K. Kumar ; A. R. Barron,; T. Mavromoustakos; M. G. Papadopoulos, J. Chem. Inf. Model. 49, 1139-1143 (2009).

52. A. Zakharov, A. Lagunin, D. Filimonov, and V. Poroikov, J. Chem.Res. Toxicol.25, 2378 (2012).

53. T. Ke, B. Xiao-Zhi, L. Xiao-Qiang, Z. Yue, S. Ji-Hong, T. Chao-Wu, and H. DaHai, J. Comput. Theor. Nanosci.11, 1785 (2014).

54. H. van de Waterbeemd, D. A. Smith, K. Beaumont, & D. K. Walker. J. Med. Chem.44, 1313–1333 (2001).

55. A. Hinchliffe, Molecular Modeling for Beginners, John Wiley & Sons Ltd, UK, (2008).

56. P. Pankaj, Ch. R. Kapupara, A. S. Matholiya, Dedakiya, R. D. Tusharbindu . Inter. Int. bull. Drug. res. 1(1), 1-10 (2011).

57. J. Wang, X. Q. Xie, T. Hou, X. Xu, Fast J. Phys. Chem. A. 24, 111, 4443-4448(2007).

58. Y. Madhu , Bioinformation . 7, 388-392 (2011).

59. P. D. Leeson, B. Springthorpe, Nat. Rev. Drug. Discov. 6(11), 881-90 (2007).

60. Edward H. Kerns and Li Di. Drug-like properties: Concepts, Structure Design and Methods: From ADME to Toxicity Optimization, Elsevier Science. 1–552, (2008).

61. [18] P. Lalitha and S. Sivakamasundari, Calculation of molecular lipophilicity and drug likeness for few heterocycles, Orient. J. Chem. 26(1), 135-141 (2010)

62. A. Leo, C. Hanch, D. Elkins, Chem. Rev. 71, 525- 526 (1971).

63. R. Done, G. Mîndril I. St nculescu, Anal. Univ. Buc. Chimie, XVI, 1, 6-83 (2007).

64. J. sangster, j. phys. Chem. Ref. Data,18, 1113-1116 (1989).

65. G.L. Patrick,AnIntroduction to Medicinal Chemistry, 2nd edition, Oxford University Press Inc., New York, **(2001)** .

66. G.Thomas, "Fundamentales of Medicinal Chemistry". "The SAR and QSAR approaches to drug deseign".**2003** : John Wiley & Sons, Ltd.

67. S. Durdagi; C. T. Supuran ; A. T. Strom, ; N. Doostdar; M.K. Kumar ; A. R. Barron,; T. Mavromoustakos; M. G. Papadopoulos, J. Chem. Inf. Model. 49, 1139-1143 **(2009).**

68. H.J. Bohm, G. Schneider, " Virtuel Screening for bioactive Molecules." Germany: Wiley-VCH 2000.

69. K. Naveen, Mahobia et al, Der. Pharma. Chemica, 2(5), 260-271 (**2010).**

70. A. T. C. Goh,W.G.Zhang . Engineering Geologie, 170, 1, **2014**.

71. H.Golchoubian, G. Moayyedi, H. Fazilati. Spectrochimica Acta PartnA , 85, 25**, 2012.**

72. M. Karelson, Molecular Descriptors in QSAR/QSPR, John Wiley & Sons, New Yor**(2000).**

73. K. Roy et al.,A Primer on QSAR/QSPR Modeling,SpringerBriefs in Molecular Science, DOI 10.1007/978-3-319-17281-1_2 **(2015).**

74. R. H. Myers, Classical and modern regression with applications. PWS-KENT Publishing Company: Boston, **(1990**).

75. I. Charif, Elaboration des corrélations quantitatives structure-activité des acidescarbonés. Etude théorique des effets de solvants sur les équilibres libres cétoénoliques des composés b-dicarbonyles cycliques, Thèse de doctorat, Université de Tlemcen, **2012.**

76. J.N. Miller, J.C. Miller, Statistics and Chemometrics for Analytical Chemistry, Prentice Hall, London**, 2000.**

77. MELKEMI Nadjib, Etude des corrélations entre les paramètres structurauet les propriétés bioactives de quelques composés organiques, THESE de doctorat, UNIVERSITE MOHAMEDKHIDER BISKRA, **2015.**

78. I. lessigiarska, PhD thesis, development of structure-activity relationships for pharmacotoxicological endpoints relevant to european union legislation, Liverpool John Moores University 4, **(2006)**

III.1 Introduction:

Molecular modeling is concerned with ways to describe the behavior of molecules and molecular systems. Computational techniques have revolutionized molecular modeling to the extent that calculations could not be performed without the use of a computer molecular modeling is invariably associated with computer modeling or computational chemistry. This discipline encompasses not only quantum mechanics (QM) but also molecular mechanics (MM), conformational analyses and other computer-based methods for understanding and predicting the behavior of molecular systems **[1]**.

The precise theoretical determination of the geometrical parameters of molecules at the minima of their potential energy surface and of the corresponding vibrational properties is of fundamental importance for the interpretation of vibrational spectroscopy experiments **[2].**

Therefore, these calculations can be performed at different levels of accuracy depending on the purpose of the theoretical study. Substituents attached to the molecular structure can increase or decrease the reactivity. The substituents were variable donors and with a design to study the effect of this change on the geometric, electronic and vibratory properties of the studied molecules. As a result, changes in reactivity in a series of reactions caused by substitution changes are related to changes in equilibrium or reactivity in another series caused by the same substitution changes **[3, 4].**

Molecular geometric determination can be a difficult experimental task. Historically, and even today, a great deal of our empirical understanding of electronic structure and bonding is based on the knowledge of molecular geometry. The recent development of analytical energy gradient methods **[5]** and efficient solution of the perturbed HF equations **[6]** have opened the way for computational quantum chemistry to provide calculated equilibrium structures on an almost routine basis, within computer resource limitations. The functional density theory methods offer another use of inexpensive computational methods that could manipulate relatively large molecules **[7-15].**

The energies of the highest occupied and lowest unoccupied molecular orbitals (HOMO and LUMO) have long been used as descriptors in QSAR **[16].**

From early in the development of quantitative structure-activity relationships (QSAR) descriptors derived from quantum chemistry have been correlated with biological activity. The descriptors used have included orbital energies, orbital coefficients, Mulliken charges on atoms and super delocalizabilities **[17].** The negative of the highest occupied molecular orbital (HOMO) energy is used as an estimate of the ionization potential (IP) and that of the lowest unoccupied molecular orbital (LUMO) energy is used as a measure of the electron affinity (EA). More recently, the identification **[18, 19]** that half the sum of the IP and EA is the electronegativity of the molecule, and half their difference is its hardness has made these two quantities attractive QSAR descriptors.

III.2 Study of structural and electronic properties of the basic nucleus of 1,2-diazol :

In this part we have studied in detail the structural and electronic parameters (distance, valence angle, charge) of the preferred conformation of 1,2-diazol nucleus.

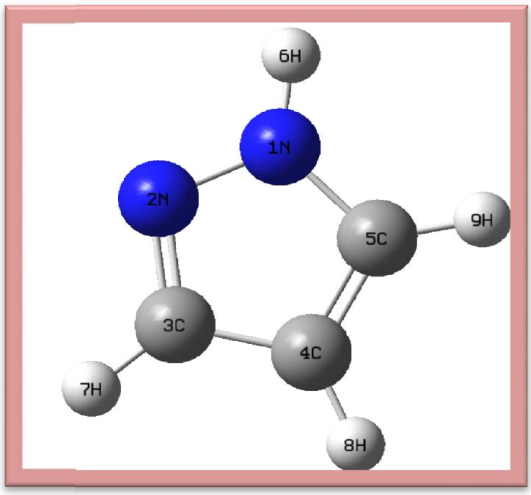


Figure III.1: 3D structure of 1, 2-diazol (GaussView 5.0)

The purpose of our study is to calculate the structural and electronic characteristics with different theoretical calculation methods and to obtain a possible similarity between the results of calculations obtained and the experimental results.

The optimized geometric parameters of 1,2-diazole by PM3 and ab initio / HF, and DFT listed in Table III.1, Table III.2, Table III.3 are consistent with the numbering (Figure III.1).

The calculations are for the following characteristics:

- ✓ The distances between the bound atoms
- ✓ The valence angles formed by three bound atoms
- ✓ The charges of each atom by the PM3 method and the DFT method and the Ab initio method.

The following theoretical calculation methods have been used:

- **Molecular mechanics:**
  - **(MM +):** (HyperChem8.0.7 software).
- **Semi-empirical:** PM3: (HyperChem 8.0.7 software)
- **Quantum mechanics:**
  - **DFT / B3LYP:** (6-31G), (6-31G ++ (d, p)) (Gaussian09 software).
  - **Ab initio / HF :** (6-31G), (6-31G ++ (d, p)) (Gaussian09 software).

Table III.1: Calculated values of length of bond

Bond length (Angstrom)		Ab initio/HF		DFT/B3LYP		
	PM3	6-31G	6-31G++(d, p)	6-31G	6-31G++(d, p)	EXP
N1-N2	1.35	1.35	1.32	1.37	1.35	1.35
N2-C3	1.35	1.31	1.30	1.34	1.33	1.33
C3-C4	1.41	1.41	1.41	1.41	1.41	1.41
C4-C5	1.39	1.36	1.36	1.38	1.38	1.37
C5—N1	1.39	1.35	1.34	1.36	1.35	1.36

PM3 (Hyperchem8.0.6), **ab initio/HF** and **DFT** (Gaussian09)

Table III.2: Calculated values of valence angle

Valence angle (°)		Ab initio/HF		DFT/B3LYP		
	PM3	6-31G	6-31G++(d,p)	6-31G	6-31G++ (d,p)	EXP
N1-N2-C3	107.18	105.25	105.14	103.83	104.18	104.1
N2-C3-C4	109	111.09	111.54	111.81	111.9	111.9
C3-C4-C5	106.74	104.88	103.91	105.3	104.53	104.5
C4-C5-C1	105.88	106.59	106.61	106.27	106.15	106.4
C5-N1-N2	111.18	112.15	112.77	112.76	113.21	113

PM3 (Hyperchem8.0.6), **ab initio/HF** and **DFT** (Gaussian09)

Table III.3: Net atomic charges of pyrazole

Charge	Ab initio/HF		DFT/B3LYP	
	6-31G	6-31G++ (d, p)	6-31G	6-31G++ (d, p)
N1	-0.6542	-0.2629	-0.5087	-0.1580
N2	-0.2627	-0.2222	-0.2187	-0.1580
C3	0.0262	-0.0784	0.0183	-0.2123
C4	-0.3540	-0.1608	-0.2232	0.0874
C5	0.1465	-0.0741	0.1293	-0.2567

Results interpretation:

From these results, a good correlation can be observed between the ab initio, PM3 and DFT for the bond lengths, and the angles calculated by these methods are also approximately similar.

The theoretical results obtained by the various calculation methods PM3, HF and DFT and with the different bases are illustrated as follows:

For the bond lengths, the difference varies from 0.02 Å to 0.07 Å between the theoretical results obtained by the PM3, HF and DFT method and the experimental values.

For the valence angles, the difference varies from 0.57 ° and 10.56 ° and between the results obtained by the methods PM3, DFT and HF and the experimental values. From the results obtained, it can be said that there is a similarity of the results obtained by different methods of calculation. The effectiveness of the DFT / B3LYP method can be controlled by comparison with the results obtained by more elaborate calculations such as Ab initio and PM3, it can be deduced that the best method to deepen our study is the DFT method. If we compare the experimental results with our calculation results, we find that all results obtained by calculation are close to the experimental results. The results obtained by the DFT method with the base 6-31G $^{++}$ (d, p) are the closest ones so it can be deduced that the DFT method is the best method to deepen our study on the structural and electronic properties of the 1, 2-diazol and its derivatives.

The important aspect of frontier electron theory is the emphasis on the highest occupied and lowest vacant molecular orbitals (HOMO and LUMO), instead of thinking of the total electron density in a nucleophile, we should think to the location of the HOMO orbital because the electrons of this orbital are more free to participate in the reaction. Similarly, the theory of frontier orbitals predicts that a site where the most unoccupied low orbital is localized is a good electrophilic site **[20].**

According to the Mulliken population in Table III.3, we see that the N1, N2 and C3and C5 atoms have negative charges that lead to electrophilic substitution; while the C4atom has a positive charge that leads to a preferential nucleophilic site attack. This is clearly seen in (Figure III.2) which represents the location of the HOMO and LUMO orbitals.

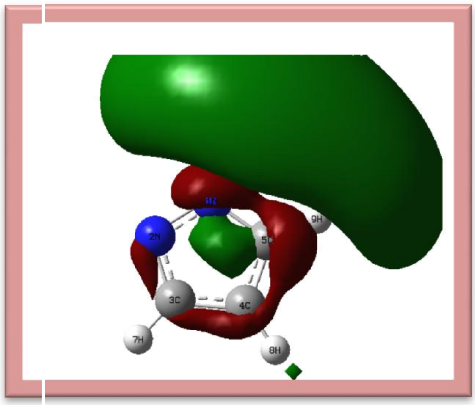


Figure III.2: The HOMO and LUMO frontier orbitals of the 1,2-diazol nucleus.

III.3 Conclusion:

A structural and electronic comparison was made for typical examples using different theoretical methods (PM3, Ab initio, DFT) and a similarity was found between their results.

By total comparison of the theoretical results of all these methods, it has been found that the DFT method is the most appropriate method for making future calculations on the 1,2-diazol nucleus.

References:

1. Barmak Mostofian, Filipp Frank, Bachelor Thesis: Basic Principles in Molecular Modeling , University of Heidelberg, Molecular Biotechnology **(Summer 2004).**
2. Andrea Zen, Delyan Zhelyazov, and Leonardo Guidoni, Optimized Structure and Vibrational Properties by Error Affected Potential Energy Surfaces, Chem. Theory Comput**.2012.**
3. E. Tarcan, A. Pekparlak, D. Avci, Y. Atalay The Arabian Journal for Science and Engineering; 34:55-62**(2009).**
4. R S. Borges, O AS. Romero J. Comput. Theor. Nanoci;8: 1129-1131**(2011).**
5. P. Pulay, Mol. Phys., 17 **(1969)** 197.
6. J. Gerratt and I.M. Mills, J. Chem. Phys., 49 **(1968)** 1719.
7. Ch. M. Chang, H.L. Tseng, A. F. Jalbout,and Aned de Leon, J. Comput. Theor. Nanosci. 10, 527 **(2013).**
8. T. L. Jensen, J. Moxnes, and E. Unneberg, J. Comput. Theor. Nanosci, 10, 464**(2013).**
9. M. Narayanan and A. John Peter,Quatum Matter, 1, 53**(2012).**
10. G.H. Cocoletzi and N. Takeuchi, Quantum Matter.2,382-387 **(2013).**
11. M. Ibrahim and H. Elhaes, Rev. Theor. Sci. 1, 368-376 **(2013).**
12. E. Chigo Anota, H. Hernandez Cocoletzi, and M. Castro, J. Comput. Theor. Nanosci.10, 2639-2646 **(2013).**
13. F. Bazooyar, M. Taherzadeh, C. Niklasson, and K. Bolton, J. Comput. Theor. Nanosci. 10, 2639-2646**(2013).**
14. H. Langueur, K. Kassali, and N. Lebgaa, J. Comput. Theor. Nanosci.10, 86,**(2013).**
15. N. Melkmi and S. Belaidi, J. Comput. Theor. Nanosci. 11, 801-806 **(2014).**
16. Brian W. Clare, Frontier orbital energies in quantitative structure-activity relationships: A comparison of quantum chemical methods, Theor Chim Acta 87:415-430**(1994).**
17. Franke R, Theoretical drug design methods. Elsevier, NY, pp 115-123**(1984).**
18. Parr RG, Pearson RG, J Am, Chem Soc 105:7512**(1983).**
19. Pearson RG, Palke WE J, Phys Chem 96:3283**(1992).**

20. Dr. Kalju Kahn, Departement of chemistry and biochemistry, UC. Santa Barbara, **2007-2010.**

IV.1 Introduction:

Qualitative and Quantitative Structure-Activity Relationships (SAR and QSAR) are attempts to correlate molecular structure or properties derived from molecular structure with a particular kind of chemical or biochemical activity **[1-3].** The kind of activity is a function of the user's **[4]** interest. QSAR is a predictive tool for a preliminary evaluation of the activity of chemical compounds by using computer aided models **[5-8].**

The molecular properties used in the correlations relate as directly as possible to key physical or chemical processes taking place in the target activity **[4]** .Quantitative structure activity relationship techniques increase the probability of success and reduce time and cost in drug discovery process.

Drug-likeness is a qualitative concept used in drug design, which is estimated from the molecular structure before the substance is even synthesized and tested. The calculation of drug-like property can give us better assumption of biological activity of certain molecule. The theoretical calculation of certain properties of a molecule can fill the parameters, which are essential to show certain biological activity **[9].**

The term drug-like captures the concept that certain properties of compounds are most advantageous in their becoming successful drug products. The term became commonly used following the pivotal work of Lipinski and his colleagues at Pfizer **[10].**Their work examined the structural properties that affect the physicochemical properties of solubility and permeability and their effect on drug absorption. The term drug-like property has expanded and has been linked to all properties that affect ADME/Tox.

IV.2 Study of the QSAR properties of the 1,2-diazole derivatives series:

IV.2.1 Chemical structures and nomenclature of 1, 2-diazole derivatives:

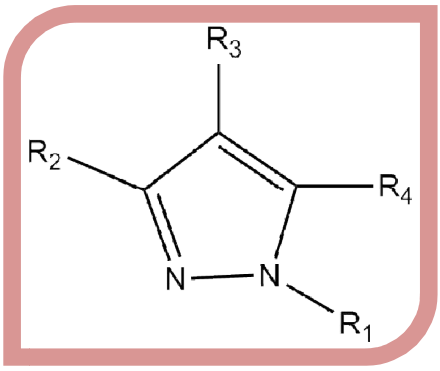
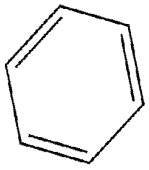
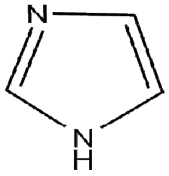
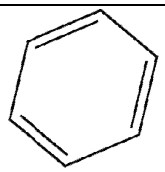
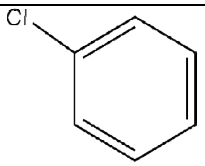
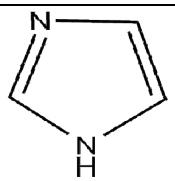
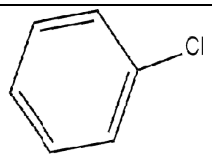
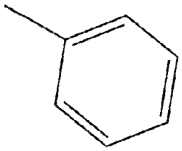
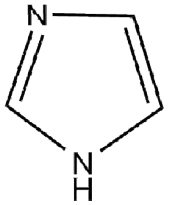
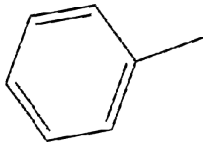
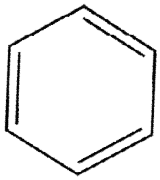
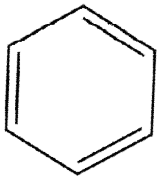
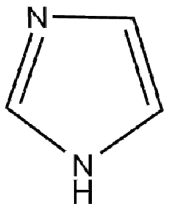
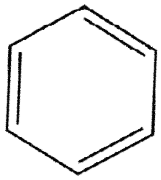
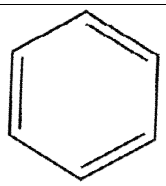
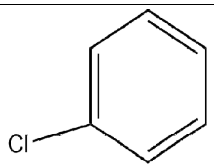
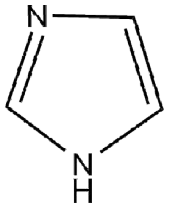
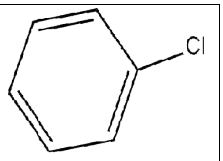
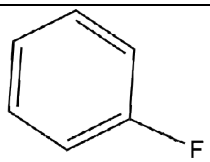
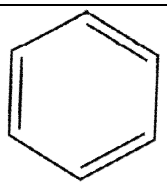
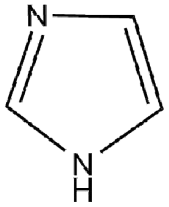
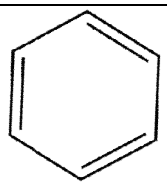
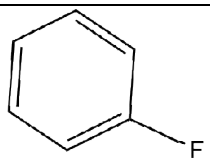
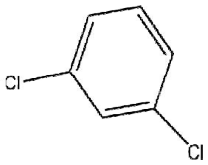
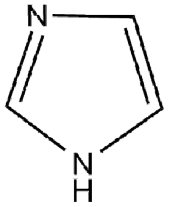
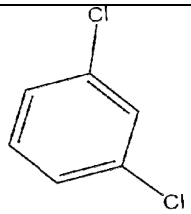
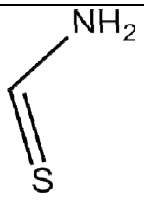
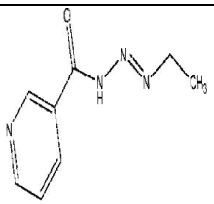
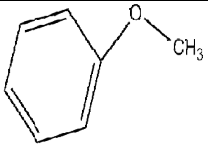
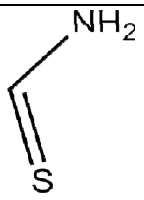
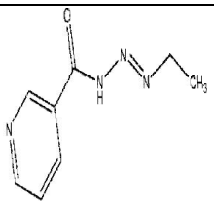
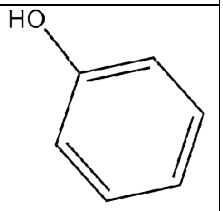
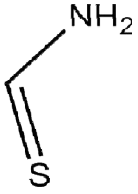
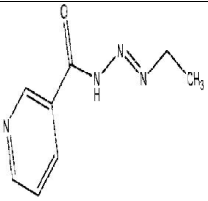
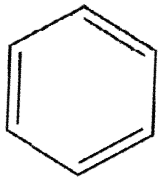
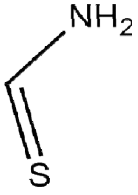
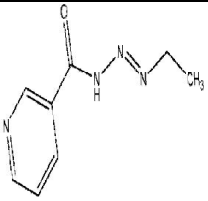
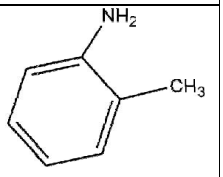
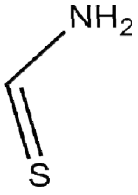
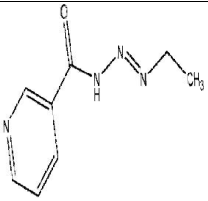
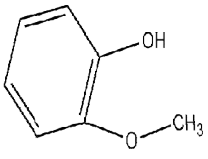
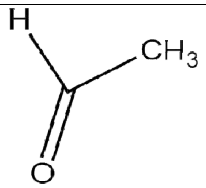
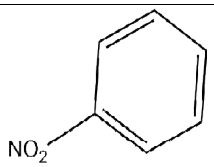
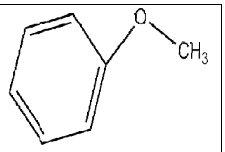
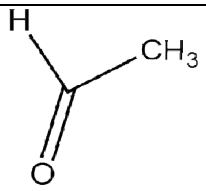
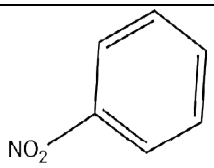
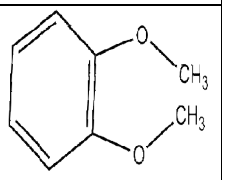
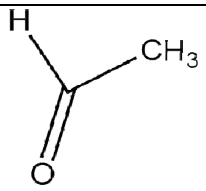
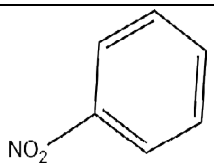
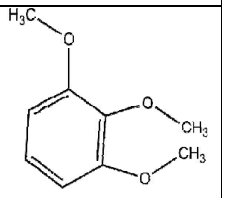
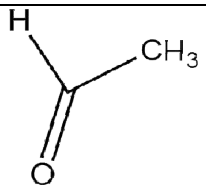
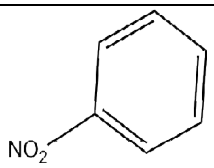
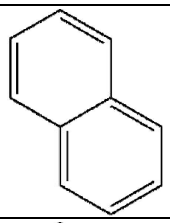
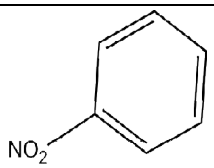
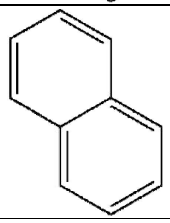
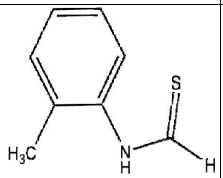
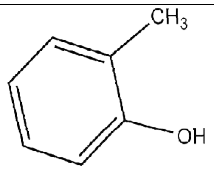
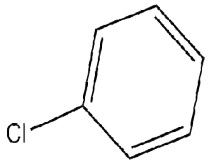
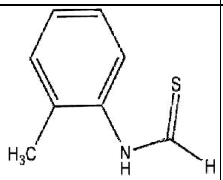
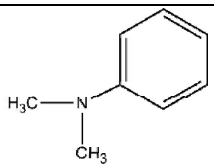
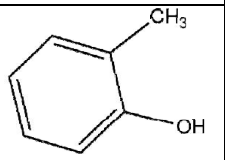
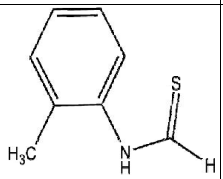
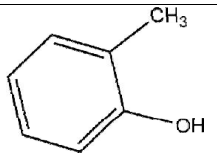
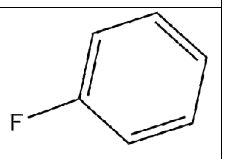


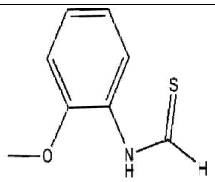
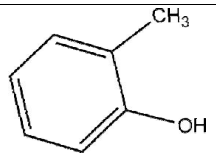
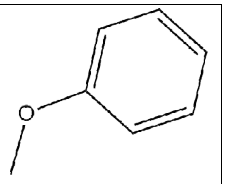
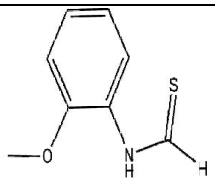
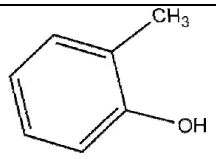
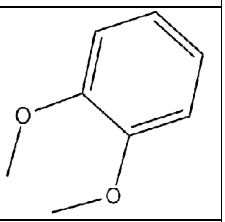
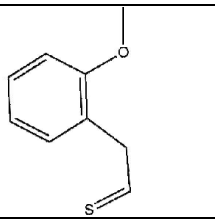
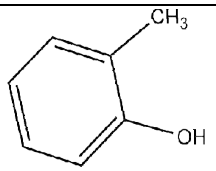
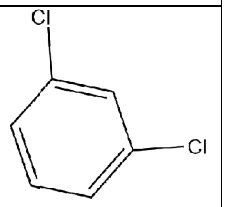
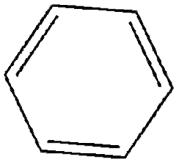
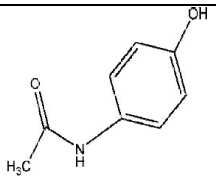
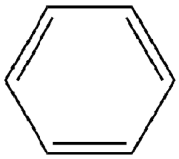
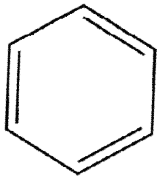
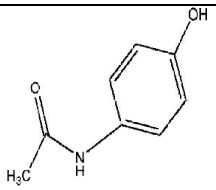
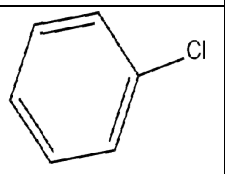
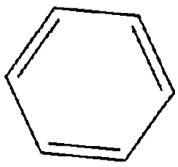
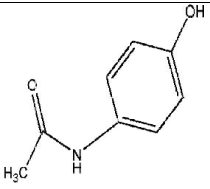
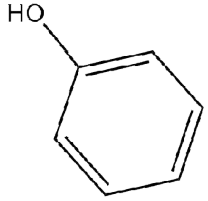
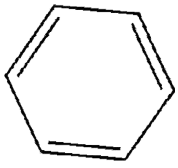
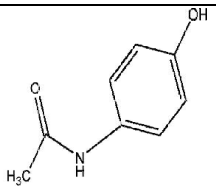
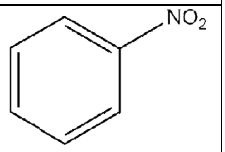
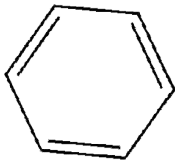
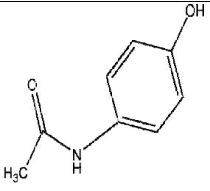
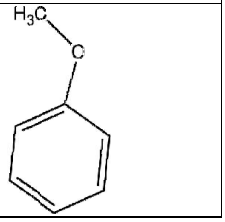
Figure IV.1: General structure and derivatives of 1, 2-diazole

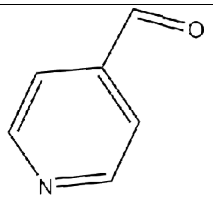
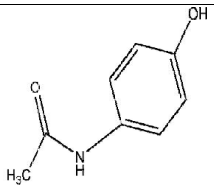
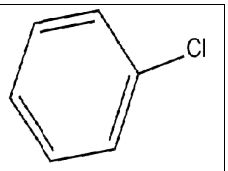
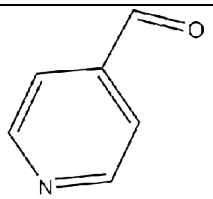
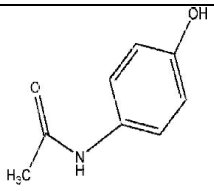
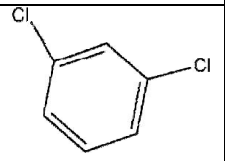
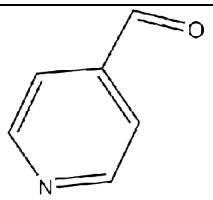
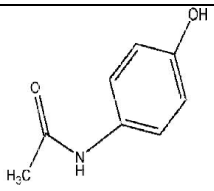
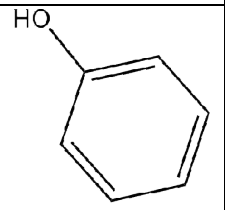
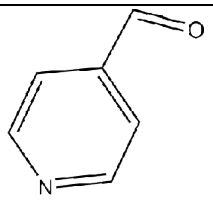
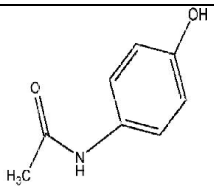
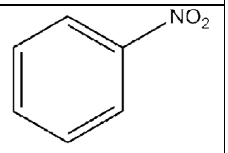
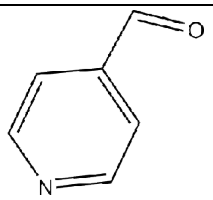
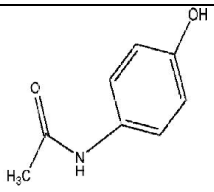
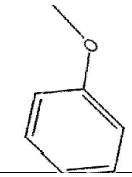
Table IV.1: structural comparison of 1, 2-diazole derivatives

Compound	*R1*	*R2*	*R3*	*R4*
1	H			
2	CH_3			
3	CH_3			
4				

5				
6				
7				
8			H	
9			H	
10			H	
11			H	
12			H	

13			H	
14			H	
15			H	
16			H	
17	H		H	
18			H	
19			H	
20			H	

21			H	
22			H	
23			H	
24			H	
25			H	
26			H	
27			H	
28			H	

29			H	
30			H	
31			H	
32			H	
33			H	

IV.2.2 Study of the amphiphilic parameters of the series of1, 2-diazole:

We studied seven physicochemical properties of a series of thirty-three of 1, 2-diazole derivatives using HyperChem 8.03 software **[11].**

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

Table IV.2: QSAR parameter of 1, 2-diazole derivatives

Compound	Molecular surface area(Å^2)	Molar volume (Å^3)	Molecular weight (uma)	Polarisability (Å^3)	Hydration energy (Kcal/mol)	Refractivity (Å^3)	LogP
1	**502.30**	**840.79**	**288.35**	**33.93**	-7.63	**95.31**	1.99
2	588.54	1003.20	330.43	39.44	**-2.82**	108.77	2.21
3	631.42	1106.90	433.34	47.45	-4.81	134.22	1.95
4	599.44	1074.39	417.89	45.82	-5.00	131.05	2.43
5	652.84	1171.25	520.22	51.22	-3.93	143.78	0.90
6	585.49	1001.2	362.43	43.40	-6.12	127.15	1.03
7	561.27	963.07	371.27	39.63	-3.56	109.64	1.46
8	595.68	1038.93	397.45	42.47	-15.31	118.71	0.69
9	558.46	976.94	383.43	40.64	-18.92	113.95	0.66
10	546.75	960.19	367.43	40.00	-13.76	112.34	1.68
11	594.05	1037.61	396.47	43.19	-16.80	120.17	0.12
12	602.13	1054.53	413.45	43.11	-21.24	120.32	-0.33
13	568.02	940.99	339.35	35.18	-8.89	98.48	-3.13
14	607.41	1015.79	369.38	37.65	-9.25	104.85	-4.13
15	645.81	1091.11	399.40	40.12	-9.84	111.23	-5.12
16	574.40	970.95	359.38	38.89	-7.65	110.30	-2.06
17	527.96	869.54	317.35	35.13	-10.38	100.73	-1.73
18	620.28	1127.96	435.97	49.44	-10.53	138.22	3.19
19	602.591	1094.48	419.52	47.42	-10.59	133.63	**2.81**
20	671.11	1192.79	447.55	50.62	-12.73	141.97	1.27
21	688.33	**1252.18**	447.58	**53.09**	-13.35	**148.35**	0.28
22	642.49	1188.04	**485.43**	52.49	-9.03	145.43	1.96
23	615.61	1050.98	371.44	42.51	-10.18	121.40	0.36
24	638.70	1094.50	405.88	44.44	-9.90	126.12	0.14
25	618.28	1068.27	387.44	43.15	-11.84	123.01	-0.66
26	648.45	1107.98	416.48	44.36	-15.02	126.62	-4.32
27	660.25	1129.85	401.46	44.99	-11.33	127.78	-0.63
28	644.81	1103.65	401.44	44.17	-13.52	123.12	-0.75
29	669.76	1146.93	435.89	46.10	-13.21	127.84	-0.97
30	692.54	1187.84	470.33	48.03	-12.95	132.56	-1.20
31	653.57	1120.47	417.44	44.81	-17.44	124.73	-1.78
32	679.45	1160.63	446.44	46.01	**-18.40**	128.34	**-5.43**
33	**691.95**	1181.43	431.47	46.65	-14.96	129.50	-1.74

Blue: great values **Green:** small values

Results interpretation:

In light of these results, we note that the values of polarizability are generally proportional to the values of surfaces and volumes.

Polarizability and molar refractivity increase relatively with the size and molecular weight of the pyrazole studied. This result is consistent with the Lorentz-Lorenz formula which gives a relation between polarizability, molar refractivity and molecular size **[12, 13].**

As can be seen, compound 21 substituted with a bulky radical has high polarization (53.09Å^3) and molar refractivity (148.35Å^3) values. In contrast, compound 1 is a small molecule in the series studied above; It has small values of polarizability (33.93Å^3) and molar refractivity (95.31Å^3).

In addition, the important value of the refractivity and the polarizability corresponding to the compound 21, it also has the important values of volume and area (1252.18Å^3, 688.33 Å^2) respectively.

Lipophilicity is a property that has a major effect on solubility, absorption, distribution, metabolism and excretion properties, as well as pharmacological activity. Hansch and Leo have explained that highly lipophilic molecules will be distributed in the lipid interior of membranes and will be retained **[14].** For good oral bioavailability, the log P must be greater than zero and less than 3 (0 <log P <3). For log P too high, the drug has low solubility and a low log P; The drug has difficulty in penetrating the lipid membranes **[15].**

For the Log P values of the compounds of the series, they belong to the interval -5.43 <Log P <3.19. Compound 32 has the low coefficient of division (-5.43), comes after compound 15 (-5.12). When the division coefficient is rather low, it has consequently better gastric tolerance. Compounds 4 and 19 which have higher values, respectively, 2.43 and 2.81, have abilities to depend on plasma proteins.

IV.3 QSAR Theoretical and Multi-Parameter Optimization (MPO):

Nowadays, various approaches to simultaneously optimize many factors in drug design are broadly described under the term ‘multi-parameter optimization’ (MPO) .

Starting with rules of thumb, we use Lipinski , Veber and as well as the rule of Ghoset and col rules to study the high oral bioavailability at the target site. On the other hand the oldest and most commonly used metrics are the efficiency of the ligand (LE) and the lipophilic efficiency of the ligand (LLE) **[16].**

IV.3. 1 Representation of "drug-like" calculations based on Lipinski:

Although medicinal chemists and pharmaceutical scientists had used structural properties in various ways for many years, rules became more prominent and defined in the field with the report by Lipinski and col **[17]** of the “rule of 5,” or what has become known as the “Lipinski rules.” These rules are a set of property values that were derived from classifying the key physicochemical properties of drug-like compounds. The rules were used at Pfizer for a few years prior to their publication and since then have become widely used.

It is important to keep in mind the intended purpose of the rule of 5. The article states: poor absorption or permeation is more likely when:

- H-bond donors >5 (expressed as the sum of all OHs and NHs)
- MW > 500
- logP > 5 (or MlogP > 4.15)
- H-bond acceptors >10 (expressed as the sum of all Ns and Os)
- Substrates for biological transporters are exceptions to the rule

The toxicity of various parameters of isolated compounds n (1-33) can be found in Table IV.3

Table IV.3: Lipinski rules of pyrazole derivatives

Compound	MW (u.m.a)	Log p	HBA	HBD	Violation N°
1	288.35	1.99	4	1	0
2	330.43	2.21	4	0	0
3	433.34	1.95	4	0	0
4	417.89	2.43	4	0	0
5	520.22	0.90	4	0	1
6	362.43	1.03	4	0	0
7	371.27	1.46	4	0	0

8	397.45	0.69	9	3	0
9	383.43	0.66	9	4	0
10	367.43	1.68	8	3	0
11	396.47	0.12	9	5	0
12	413.45	-0.33	10	4	0
13	339.35	-3.13	7	0	0
14	369.38	-4.13	8	0	0
15	399.40	-5.12	9	0	0
16	359.38	-2.06	6	0	0
17	317.35	-1.73	5	1	0
18	435.97	3.19	4	2	0
19	419.52	2.81	4	2	0
20	447.55	1.27	6	2	0
21	447.58	0.28	7	2	0
22	485.43	1.96	4	1	0
23	371.44	0.36	5	2	0
24	405.88	0.14	5	2	0
25	387.44	-0.66	6	3	0
26	416.48	-4.32	8	2	0
27	401.46	-0.63	6	2	0
28	401.44	-0.75	7	2	0
29	435.89	-0.97	7	2	0
30	470.33	-1.20	7	2	0
31	417.44	-1.78	8	3	0
32	446.44	-5.43	10	2	0
33	431.47	-1.74	8	2	0

HBA: hydrogen bond acceptors **HBD:** hydrogen bond donors

Results interpretation:

We used Lipinski's rules to identify compounds that pose absorption and permeability problems if these compounds do not validate at least two of its rules.

The Lipinski rule is the most used to characterize drug-like compounds. We recall that this rule is intended to identify compounds presenting absorption and permeability problems. Compounds that do not validate at least two of their criteria are very likely to have problems with absorption or permeability **[18].**

It can be seen from Table IV.3 that all compounds have values less than 5 for lipophilicity, ranging from -5.43 to 3.19; therefore these compounds are better solubilized in aqueous and lipid solutions. On the one hand, a negative value for log P indicates that the compound is too hydrophilic. It has good solubility in water, better gastric tolerance and effective elimination by the kidneys. On the other hand, a positive value for logP indicates that the compound is too lipophilic. Thus, it has good permeability across the biological membrane, better binding to plasma proteins, elimination by metabolism, but low solubility and gastric tolerance **[19],** for our series studied we have for example the compounds (15, 26, 32) take the values of log P respectively (-5.12, -4.32, -5.43) indicates that these compounds have good solubility and gastric tolerance. Compound 18, which has a maximum log P (3.19), is also taken so that it has a better permeability across the biological membrane.

It can also be seen in Table **IV.3** that all these compounds have hydrogen acceptor numbers of less than 10 (O) and number of hydrogen donors less than 5 (OH, NH). And for molecular weights there are compounds that have values below 500 Da, so they are easily cross cell membranes.

For the number of violations one notices that some compounds have a null violation for the Lipinski rules and others have a single violation.

Thus, all compounds meet the rules of Lipinski (rules of five), suggesting that these compounds theoretically would not have problems with oral bioavailability.

IV. 3. 2 Veber rules:

Additional rules were proposed by Veber and col**[20]** .They studied structural properties that increase oral bioavailability in rats. They concluded that molecular flexibility, polar surface area (PSA), and hydrogen bond count are important determinants of oral bioavailability. Rotatable bonds can be counted manually or using software. PSA is calculated using software and is closely related to hydrogen bonding. Veber rules for good oral bioavailability in rats are as follows:

- rotatable bonds ≤10
- PSA ≤140 $Å^2$

The number of rotatable bonds (NRB) has been defined as a single bond, not in a ring bound to a non-terminal atom (i.e, non-hydrogen). Amide C-N bonds are excluded from the account because of their high rotational energy barrier **[21].**

On the other hand, the polar surface (PSA) which is formed by polar atoms of a molecule. It is a descriptor that shows a good correlation with passive molecular transport across membranes, and thus allows estimating the transport properties of drugs.

Note: The calculation of these parameters is calculated using Molinspiration.Also, the PSA was used to calculate the percent absorption (% ABS) according to the equation:

% ABS = 109 ± 0.345 x PSA [22].

Table IV.4: Veber rules of pyrazole derivatives

Compound	NRB	PSA(Å²)	Veber score	%ABS
1	3	42.22	2	94.43
2	3	33.43	2	97.46
3	4	33.43	2	97.46
4	4	33.43	2	97.46
5	4	33.43	2	97.46
6	4	33.43	2	97.46
7	3	33.43	2	97.46
8	7	117.58	2	68.43
9	6	128.57	2	64.64
10	6	108.34	2	71.62
11	6	134.37	2	62.64
12	7	137.81	2	61.45
13	4	87.73	2	78.73
14	5	96.97	2	75.54
15	6	106.20	2	72.36
16	3	78.20	2	81.31
17	3	70.22	2	84.77
18	5	47.86	2	92.48
19	5	47.86	2	92.48
20	7	66.33	2	86.11
21	8	75.56	2	82.93
22	6	45.06	2	93.45

23	4	64.93	2	86.59
24	4	64.93	2	86.59
25	4	85.16	2	79.61
26	5	110.75	2	70.79
27	5	74.16	2	83.41
28	4	94.89	2	76.26
29	4	94.89	2	76.26
30	4	94.89	2	76.26
31	4	115.12	2	69.28
32	5	140.12	1	60.65
33	5	104.12	2	73.07

NRB: The number of rotatable bound **PSA:** the polar surface area

Results interpretation

From the results of the table it will be noted that the values of the number of rotational bonds are all less than10. The low number of rotational bonds (reduced flexibility) in the test compounds indicates that these ligands during binding to a protein only slightly change their conformation.

For the results of PSA, it is noted that all the compounds of the series studied have values lower than 140 Å² (only the compound 32), which shows the good prediction of the oral bioavailability and the transport through the membranes.

So we notice that the majority of the studied compounds are in agreement with the rule of Veber.

For percent absorption values (% ABS) It can obviously be observed that all compounds had a high% ABS ranging from 64.581 to 91.802%, indicating that these compounds should have good cell membrane permeability (Table **VI**).).

IV. 3. 3 The rules of Ghose-Viswanadhan-Wendoloski:

An earlier analysis **[23]** of known drugs concerned molecular frameworks and used form description methods to prepare a list of common drug forms. Another analysis **[24]** of known medicines (the Comprehensive Medical Chemistry (CMC) database) and other databases such as the Available Chemicals Directory (ACD) was devoted to identifying criteria to be used in the selection of compounds for screening.

A group of eight drug-like indices has been proposed by Ghose-Viswanadhan-Wendoloski, Analysis of the Distribution of Certain Physicochemical Properties (logP, AMR, MW, nAT) and Chemical Constitutions of Drug Molecules Available in the Setting The Medicinal Comprehensive program was proposed by Ghose-Viswanadhan-Wendoloski Chimie (CMC) **[25].**

This filter defines the medical likeness constraints as follows:

- $0.4 \leq \log P \leq 5.6$
- $160 \leq$ mass ≤ 480
- $20 \leq$ number of atoms ≤ 70
- $40 \leq$ refractivity ≤ 130

Despite having only four properties in the filter, each property must be included twice as two conditions are applied for each property**.**

Furthermore, the compound should be a combination of some of the following functional groups: a benzene ring. A heterocyclic ring (both aliphatic and aromatic), an aliphatic amine, a carboxamide group, an alcoholic hydroxyl group, a carboxyl ester, and a keto group. For example, according to the CMC-80 index. An organic compound is a drug-like molecule if: the calculated ALOGP is between -0.4 and 5.6, the AM R molar refractivity between 40 and 130, the molecular weight MW between 160 and 480. The total number of atoms between 20 and 70, and it comprises at least one of the functional groups mentioned above **[26].**

Table IV.5: The rules of Ghose-Viswanadhan-Wendoloski of pyrazole derivatives.

Compound	Log P	MW	Molar Refractivity	Atoms nomber	Ghose score
1	1.99	288.35	95.31	38	**4**
2	2.21	330.43	108.77	47	**4**
3	1.95	433.34	134.22	48	3
4	2.43	417.89	131.05	48	3
5	0.90	520.22	143.78	48	2
6	1.03	362.43	127.15	46	**4**
7	1.46	371.27	109.64	41	**4**
8	0.69	397.45	118.71	47	**4**

9	0.66	383.43	113.95	44	**4**
10	1.68	367.43	112.34	43	**4**
11	0.12	396.47	120.17	48	**4**
12	-0.33	413.45	120.32	48	**4**
13	-3.13	339.35	98.48	42	3
14	-4.13	369.38	104.85	46	3
15	-5.12	399.40	111.23	50	3
16	-2.06	359.38	110.30	44	3
17	-1.73	317.35	100.73	39	3
18	3.19	435.97	138.22	52	3
19	2.81	419.52	133.63	52	3
20	1.27	447.55	141.97	57	3
21	0.28	447.58	148.35	61	3
22	1.96	485.43	145.43	54	2
23	0.36	371.44	121.40	49	**4**
24	0.14	405.88	126.12	49	**4**
25	-0.66	387.44	123.01	50	3
26	-4.32	416.48	126.62	53	3
27	-0.63	401.46	127.78	50	3
28	-0.75	401.44	123.12	50	3
29	-0.97	435.89	127.84	51	3
30	-1.20	470.33	132.56	50	2
31	-1.78	417.44	124.73	51	3
32	-5.43	446.44	128.34	53	3
33	-1.74	431.47	129.50	55	3

Results interpretation:

From the results of Table IV.5 we obtain log P values between -5.43and 3.19 and with an average value of -0.27, for the molecular weights the values are between 288.35and 485.43with an average value of 402.68, also for molar refractivity values that are between 95.31 and 148.35 and average value 123.01 and finally for atom numbers belongs to a range from 22 to 33 and takes a mean value 28.96.

So we can notice that one third of compounds validate all the rules of GVW (score 4) so the values of Log P, the molar mass, the molar refractivity as well as the number of atoms belong to the GVW range and compound 1, which carries the values of log P, MW, molar refractivity and the number of atoms, is taken as an example respectively (1.99, 288.35 uma, 95.31 $Å^3$, 22), so these compounds can be classified as drugs (a drug-like molecule),

also these compounds have a combination of different functional groups such as: the benzene ring, a nitrogenous heterocycle as well as two carbonyl groups.

IV. 3. 4 Ligand efficiency:

Lipophilia is a physico-chemical property that plays a fundamental role in determining the properties of ADME (absorption, distribution, metabolism and excretion). Lipophilicity correlates with too many other properties, such as tissue storage, bioavailability, permeability, toxicity, volume of distribution, plasma protein binding and enzyme receptor binding **[27, 28].** The smallest compound tends to have the best physicochemical properties and a good ADME in terms of ligand efficiency **[29, 30].**

In this section we have studied the effectiveness of ligand to penalize large compounds on small compounds with a similar power, because the larger compounds have lower physicochemical properties and ADME **[26, 30].**

The effectiveness of the ligand (LE) is most often defined as the ratio of the affinity of a ligand divided by the number of heavy atoms (non-hydrogen) in the molecule.

$$LE = \frac{\Delta G}{HAC} = \frac{\mathbf{RT \ln K}}{\mathbf{NH}}$$

This report was first described in 1999 [38] and gained widespread popularity in drug discovery circles shortly after **[30].**

In more common units, this can be expressed as follows:

LE =1.4pIC50/NH

Where: NH is the number of heavy atoms.

pIC50 = -log (IC50)

Table IV.6: The ligand efficiency of 1,2-diazole derivatives -

Compound	pIC50	Refrence(PIC50)	NH	LE
1	4.79	[31]	22	0.309
2	4.79	[31]	25	0.268
3	5.39	[31]	30	0.251

4	5.09	[31]	29	0.245
5	4.79	[31]	33	0.203
6	5.09	[31]	28	0.254
7	5.09	[31]	25	0.285
8	5.60	[31]	28	0.280
9	5.42	[31]	27	0.281
10	5.90	[31]	26	0.317
11	5.20	[31]	28	0.260
12	5.60	[31]	29	0.270
13	3.60	[31]	25	0.201
14	3.30	[31]	27	0.171
15	4.90	[31]	29	0.236
16	3.00	[31]	25	0.168
17	4.20	[31]	22	0.267
18	5.20	[31]	30	0.242
19	5.20	[31]	30	0.242
20	5.38	[31]	32	0.235
21	5.24	[31]	34	0.215
22	5.78	[31]	32	0.252
23	4.60	[31]	28	0.230
24	4.90	[31]	29	0.202
25	4.60	[31]	29	0.222
26	5.20	[31]	31	0.234
27	4.90	[31]	30	0.228
28	4.30	[31]	30	0.200
29	5.50	[31]	31	0.248
30	5.50	[31]	32	0.240
31	4.90	[31]	31	0.221
32	5.50	[31]	33	0.233
33	5.20	[31]	32	0.227

Results interpretation:

From the results of the table it is noted that the values of the ligand (LE) efficiency decreases with the increasing number of heavy atoms, obtaining high ligand efficiency requires compounds with weak heavy atoms and that can be explained by the correlation that exists between the size of the compound and their physicochemical properties.

It has been suggested that a typical fall-contributing factor for LE in larger ligands may be due to less favorable binding entropies for larger and more flexible ligands **[32].** And that

clearly see in Table IV.6 as for example for the compounds (1,10, 17) having the small values of heavy atoms respectively take the efficiency (0.309,0.317,0.267), while the compound (21) which has the largest heavy atom value 34 corresponding to the small 0.215 ligand efficiency values, may have poor physicochemical and ADME properties. On the other hand, the efficiency varied proportionally with the values of pIC50.

IV. 3. 5 Lipophilic efficiency of ligand:

On the other hand we have studied lipophilic efficiency (LipE) to maximize potency while keeping lipophilicity as low as possible, due to the association between high lipophilicity and several problems, including low solubility, membrane permeability metabolic stability, etc **[33,34].**

However, to obtain optimal ADMET properties, molecular size and lipophilicity are important factors to consider. If lipophilicity is too high, the likelihood of a compound binding to multiple targets increases. To facilitate the optimization of lipophilicity affinity, Leeson and Springthorpe **[29]** have defined the efficacy of lipophilic ligand-lipophilicity (LLE), also known as lipophilic efficiency (Lipophilic Efficiency) (LipE)**:**

$$\mathbf{LLE = pIC50 \_ logP}$$

High LLE promotes compounds that gain a lot of their affinity through directed interactions thus making the interaction with the receptor more specific.

Table IV.7: The lipophilic ligand efficiency of 1,2-diazole derivatives -

Compound	pIC50	Log P	LLE
1	4.79	1.99	2.8
2	4.79	2.21	2.58
3	5.39	1.95	3.44
4	5.09	2.43	2.66
5	4.79	0.90	3.89
6	5.09	1.03	4.06
7	5.09	1.46	3.63
8	5.60	0.69	4.91
9	5.42	0.66	4.76
10	5.90	1.68	4.22

11	5.20	0.12	5.08
12	5.60	-0.33	5.93
13	3.60	-3.13	6.73
14	3.30	-4.13	7.43
15	4.90	-5.12	10.02
16	3.00	-2.06	5.06
17	4.20	-1.73	5.93
18	5.20	3.19	2.01
19	5.20	2.81	2.93
20	5.38	1.27	4.11
21	5.24	0.28	4.96
22	5.78	1.96	3.82
23	4.60	0.36	4.24
24	4.90	0.14	4.76
25	4.60	-0.66	5.26
26	5.20	-4.32	9.52
27	4.90	-0.63	5.53
28	4.30	-0.75	5.05
29	5.50	-0.97	6.47
30	5.50	-1.20	6.7
31	4.90	-1.78	6.68
32	5.50	-5.43	10.93
33	5.20	-1.74	6.94

Results interpretation:

LLE provides a means of assessing the affinity of a compound for its lipophilicity. The challenge is to increase power without increasing lipophilicity at the same time. Since lipophilicity is the main promissory factor for compounds, optimized LLE compounds should be more selective. It is suggested to target an LLE in a range of 5-7 or even higher **[30].** If LipE is between 5 and 7 or more than 7, the optimized compounds are more selective **[28].**

From the lipophilic ligand efficiency values obtained in Table IV.7, the LLE is found to change values of 2.01 and 10.93. For example, the compounds (11,13, 33) having LLE values (5.08, 6.73, 6.94) which belong to the range 5-7, indicate that these compounds have been successfully optimized.

On the other hand, the values of the lipophilic ligand efficiency for the compounds (2, 7, 18) are respectively (2.58, 3.63, 2.01). It is noted that none of the compounds reaches an

LLE greater than 5. In these cases, the affinity gain is accompanied by an increase in lipophilicity. In this respect, the optimization was not as optimal as in the first example.

IV.3. 6 Golden triangle:

The Golden Triangle is a visualization tool developed from in vitro permeability, in vitro clearance and computational data designed to aid medicinal chemists in achieving metabolically stable, permeable and potent drug candidates. Classifying compounds as permeable and stable and plotting molecular weight (MW) versus octanol:buffer (pH 7.4) distribution coefficients (logD) or estimated octanol:buffer (pH 7.4) distribution coefficients (elogD) reveals useful trends. Analysis of at least two orthogonal trends, such as permeability and clearance, can be extremely effective in balancing and optimizing multiple properties. In addition, molecular weight and logDimpact potency-efficiency calculations, allowing potency, clearance and permeability to be optimized simultaneously **[35].**

LogD and molecular weight (MW) have been identified as properties showing the correlation with available permeability and stability data for large data sets all combined trends lead to the observation that the polarity and molecular weight permeabilities in vitro and low clearance compounds are concentrated in an area with a baseline from log D = -2.0 to log D = 5.0 to MW = 200 and a log vertex D = 1.0- 2.0 and MW = 450. These trends lead to a shaped area known as the Golden Triangle and molecules within this area that are low clearance and permeable are expected to obey a golden triangle rule. .

Note: Log D calculated with MarvinSketch

Table IV.8: Distribution coefficients of 1, 2-diazole

Compound	Molecular weight (uma)	Log D (PH = 7.4)
1	288.35	3.02
2	330.43	3.99
3	433.34	6.45
4	417.89	5.38
5	520.22	7.80
6	362.43	5.41
7	371.27	4.17
8	397.45	2.02

9	383.43	1.87
10	367.43	2.18
11	396.47	1.86
12	413.45	1.72
13	339.35	2.57
14	369.38	2.41
15	399.40	2.25
16	359.38	3.72
17	317.35	4.01
18	435.97	7.02
19	419.52	6.56
20	447.55	5.59
21	447.58	5.43
22	485.43	6.88
23	371.44	4.32
24	405.88	4.93
25	387.44	4.02
26	416.48	4.17
27	401.46	2.45
28	401.44	3.05
29	435.89	3.05
30	470.33	3.56
31	417.44	2.14
32	445.43	3.22
33	431.47	2.29

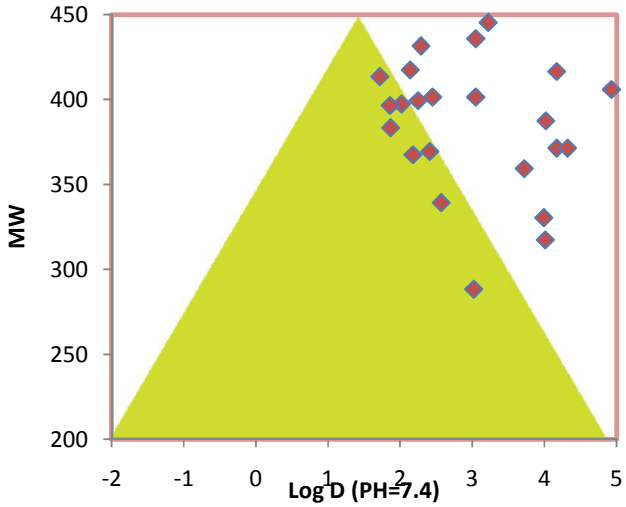


Figure IV.2: Golden triangle

Results interpretation:

Johnson and co-workers reported that molecular weight (MW) and lipophilicity (LogD at pH 7.4) act as surrogates of many different molecular descriptors and have been used to develop a useful visualization tool of the Golden Triangle. Compounds that reside inside the Golden Triangle are more likely to be both metabolically stable and to possess good membrane permeability than those outside. The golden Triangle (Fig. 2) shows that the compounds 1,4, 8,10, 11, 12, 13 and 14 sit inside of the triangle therefore these derivatives have good permeability and clearance. Also, the other compounds are the inverse **[36].** In general, lower log D and higher molecular weight molecules fail due to low permeability while higher log D and higher MW compounds fail due to elevated in vitro clearance **[37].**

IV.4 Conclusion:

The present study offers a structural comparison between the thirty-three 1,2-diazole. It also provides a discussion of several qualitative approximations of the relationship structure activity / property.

The values of polarizability are generally proportional to the values volumes and surfaces and refractivity. Compound 21 takes the values important for the polarizability 53.09 $Å^3$ and the refractivity 148.35$Å^3$ as well as the volume and surface values which are respectively 1252.8 $Å^3$, 688.33$Å^2$

The hydration energy in absolute value, the most important is that of compound 32(18.40 kcal / mol). They are therefore the best distribution in the tissues.

In this study all compounds with log P values less than 5 are soluble in aqueous solution and therefore able to reach the surface of membranes and have gastric tolerance.

Compound 32 has the lowest partition coefficient (Log P) (-5.43), this molecule is the most hydrophilic product. It results in better gastric tolerance. While compound 19 has the highest value (2.81), so it has abilities to depend on plasma proteins.

The application of the Lipinski rules on the 1, 2-diazole derivatives studied shows that the compounds studied, theoretically, have no problem of oral bioavailability.

Another study based on the rules of Veber shows that the compounds of the series studied are all in correlation with the first rule of Veber, that is to say that the number of rotational bond less than 10 is applied only compound 32. All the compounds, on the other hand the PSA values are all lower than 140 Å². Also the compound 2, 3,4,5,6 and 7 having the highest values of% ABS (97.46%) therefore they have the best membrane permeability.

The application of the GVW rules shows that two-third of the compounds studied in this series do not correspond to the associated criteria.

For the study of ligand efficiency it is found that compound 10 has the highest LE value which is 0.317 which allows it a good physico-chemical and ADME properties.

Compound 15 had the highest LipE value in the data set of 10.02 and was considered the most optimal compound.

The compounds 1, 4, 8, 10, 11, 12, 13 and 14 sit inside of the golden triangle , so these derivatives have good permeability and clearance.

References:

1. H. Langueur, K. Kassali, and N. Lebgaa, J. Comput. Theor. Nanosci, 10, 86 **(2013).**
2. N. Melkemi and S. Belaidi, J. Comput. Theor. Nanosci, 11, 801 **(2014).**
3. M. Narayanan and A. John Peter, Quantum Matter, 1, 53 **(2012).**
4. Y. C. Martin, Quantitative Drug Design, Marcel Dekker, New York, NY, USA, **(1978).**
5. A.I. Ageenko and Y.E. Vitorgan, Vopr. Virusol, 2, 1599 **(1975).**
6. V.R. Talageri, S.N. Revankar, B.N. Mashelkar and K.I. Ranadive, Biochem. Biophys, 8, 79 **(1971).**
7. F.B. Hershey, G. Johnson, S.M. Murphy and M. Schmidt, Cancer Res, 26, 265 **(1966).**
8. T.T. Otani and H.P. Morris, J. Natt. Cancer Inst. (US), 47, 1247 **(1971).**
9. M.Mellaoui, S. Belaidi, D. Bouzidi, and N.Gherraf,Quantum Matter, 3, 435-441 **(2014)**
10. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney,Adv. Drug Deliv. Rev. 4, 17
11. HyperChem (Molecular Modeling System) Hypercube, mc, 1115 MV 4th Street, Gainesville, EL 32601; USA, (2007), http://www.hyperchem.com/
12. N. I. Zhokhova, I. I. Baskin, V. A. Palyulin, A. N. Zefirov, N. S. Zefirov, Russ. Chem. 52: 1061 (2003).
13. R. Mignani, A. Petrucci, F. Cardone, Quantum Matter; 3: 219 (2014).
14. A. K. Ghose, G. M. Crippen, J. Chem. Inf. Comput. Sci.27:21 (1987).
15. G. Burgot, N. Bona, M. V. Christen, J. L. Burgot, Int. J. Pharm.129:295 (1996).
16. D. M. Segall, *J. Curr. Pharm. Des.* 18, 1292 **(2012).**
17. Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings.Advanced Drug Delivery Reviews, 23, 3–25.
18. MONGE Aurélien; 2006; “Création et utilisation de chimiothèques optimisées pour la recherche « in silico » de nouveaux composés bioactifs. » ; Thèse de Doctorat deL’UNIVERSITE D’ORLEANS; pp. 29, 30
19. Schultes S, Graaf C, Haaksma E, Iwan J P and Kramer O, J. Drug Discovery Today: Technologies, **2010**;7:157.

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

21. D. F. Veber, S. R. Johnson, H. Y. Cheng, B. R. Smith, K. W. Ward, and K. D. Kopple, *J. Med. Chem.* 45, 2615 **(2002).**

22. E. P., R. B., and S. P., *J. MedChem.* 43, 3714 **(2000).**

23. Bemis, G. W.; Murcko, M. A. The Properties of Known Drugs. I. Molecular Frameworks. *J. Med. Chem.* **1996**, *37*, 2887-2893.

24. McGregor, M. J.; Pallai, P. V. Clustering Large Databases of Compounds: Using MDL "Keys" as Structural Descriptors. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 443-448.

25. Ghose, A.K.; Viswanadhan, V.N.; Wendoloski, J.J. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J. Comb. Chem.* **1999**, *1*, 55–68.

26. TODESCHINI, Roberto et CONSONNI, Viviana. *Molecular descriptors for chemoinformatics, volume 41 (2 volume set)*. John Wiley & Sons, **2009**.

27. Kerns E, Elsevier Science, **2008**;1–552.

28. Pliska V, Testa B, van de Waterbeemd H, Mannhold R, Kubinyi H, and Timmerman H, Lipophilicity in Drug Action and Toxicology, Wiley-VCH, Federal Republic of Germany, **1996.**

29. Leeson, P.D. and Springthorpe, B. **(2007)** The influence of drug-like concepts on decision-making in medicinal chemistry. Nat. Rev. Drug Discov. 6, 881–890

30. Hopkins A L, Groom C R, and Alexander A, J. Drug. Discov. T, **2004**;9:430.

31. Sheshagiri R. Dixit, Shrinivas D. Joshi, V.H. Kulkarni and Tejraj M. Aminabhavi , 3D-QSAR and Molecular Docking Studies of Pyrazole Derivatives as Inhibitors of Enoyl Acyl Carrier Protein Reductase from Mycobacterium tuberculosis ,Letters in Drug Design & Discovery, **2017**, 14, 414-433.

32. Loving, K.; Alberts, I.; Sherman, W. Computational approaches for fragment-based and de novo design. *Curr. Top. Med. Chem.* **10**, 14-32 **(2012).**

33. 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).**

34. M. P. Edwards and D. Price, *Annu. Rep. Med. Chem.* 45, 381 **(2010).**

35. Ted W. Johnson, Klaus R. Dress, Martin Edwards, Using the Golden Triangle to optimize clearance and oral absorption, Bioorg. Med. Chem. Lett. 19 **(2009)** 5560–5564.

36. J. M. Beale, J. Block, and R. Hill, Organic Medicinal and Pharmaceutical Chemistry, Philadelphia: Lippincott Williams and Wilkins,New York**(2010)**

37. T. W. Johnson, K. R. Dress, and M. Edwards,Bioorg. Med. Chem.Lett.19, 5560**(2000).**

V.1 Introduction:

QSAR is a statistical model that relates a set of structural descriptors of a chemical compound to its biological activity **[1].** Such molecular structural information is encoded in molecular descriptors and a QSAR model defines mathematical relationships between descriptors and biological activities of known ligands to predict unknown ligands activities **[2].**

Many statistical methods have been employed to generate QSAR models from ,descriptive variables. Simple and multiple linear regressions is one of the more successful techniques use by many researcher in construct of QSAR models **[3-5].** The multiple linear regression model is used to study the relationship between a dependent variable and one or more independent variables. Multiple linear regression or MLR **[6]** is a commonly used method in QSAR due to its simplicity, transparency, reproducibility, and easy interpretability.

In this study, we have tried to describe the structure-property / activity relationship studies on pyrazoles and to develop the best QSAR model on these compounds with respect to their enoyl acyl carrier protein reductase inhibitory activities. Suite to our interest in this area, a representative set of thirty three 1, 2-diazole derivatives was studied by the QSAR method with different physicochemical descriptors.

V.2. Quantitative studies on structure-activity relationships:

In this section, the inhibition of enoyl acyl carrier protein reductase by a group of33 of 1, 2-diazole derivatives was studied to predict a QSAR model using molecular descriptors. The different physical and chemical properties **[7],** known as physicochemical descriptors, as well as Lipinski, Veber parameters were used as independent variables and were correlated with biological activities of pyrazole derivatives for generation of QSAR model by multiple linear regression analyzes (MLR).

Table V.1: Values of molecular descriptors used in regression analyze.

Compound	pIC50	Molecular surface area ($Å^2$)	Molecular volume ($Å^3$)	Molecular weight (uma)	Polarisability ($Å^3$)	Hydration energy (Kcal/mol)	Refractivity ($Å^3$)	Log P	HBA	HBD	NRB	LogD
1	4.79	502.30	840.79	288.35	33.93	-7.63	95.31	1.99	4	1	3	3.02
2	4.79	588.54	1003.20	330.43	39.44	-2.82	108.77	2.21	4	0	3	3.99
3	5.39	631.42	1106.90	433.34	47.45	-4.81	134.22	1.95	4	0	4	6.45
4	5.09	599.44	1074.39	417.89	45.82	-5.00	131.05	2.43	4	0	4	5.38
5	4.79	652.84	1171.25	520.22	51.22	-3.93	143.78	0.90	4	0	4	7.80
6	5.09	585.49	1001.2	362.43	43.40	-6.12	127.15	1.03	4	0	4	5.41
7	5.09	561.27	963.07	371.27	39.63	-3.56	109.64	1.46	4	0	3	4.17
8	5.60	595.68	1038.93	397.45	42.47	-15.31	118.71	0.69	9	3	7	2.02
9	5.42	558.46	976.94	383.43	40.64	-18.92	113.95	0.66	9	4	6	1.87
10	5.90	546.75	960.19	367.43	40.00	-13.76	112.34	1.68	8	3	6	2.18
11	5.20	594.05	1037.61	396.47	43.19	-16.80	120.17	0.12	9	5	6	1.86
12	5.60	602.13	1054.53	413.45	43.11	-21.24	120.32	-0.33	10	4	7	1.72
13	3.60	568.02	940.99	339.35	35.18	-8.89	98.48	-3.13	7	0	4	2.57

14	3.30	607.41	1015.79	369.38	37.65	-9.25	104.85	-4.13	8	0	5	2.41
15	4.90	645.81	1091.11	399.40	40.12	-9.84	111.23	-5.12	9	0	6	2.25
16	3.00	574.40	970.95	359.38	38.89	-7.65	110.30	-2.06	6	0	3	3.72
17	4.20	527.96	869.54	317.35	35.13	-10.38	100.73	-1.73	5	1	3	4.01
18	5.20	620.28	1127.96	435.97	49.44	-10.53	138.22	3.19	4	2	5	7.02
19	5.20	602.591	1094.48	419.52	47.42	-10.59	133.63	2.81	4	2	5	6.56
20	5.38	671.11	1192.79	447.55	50.62	-12.73	141.97	1.27	6	2	7	5.59
21	5.24	688.33	1252.18	447.58	53.09	-13.35	148.35	0.28	7	2	8	5.43
22	5.78	642.49	1188.04	485.43	52.49	-9.03	145.43	1.96	4	1	6	6.88
23	4.60	615.61	1050.98	371.44	42.51	-10.18	121.40	0.36	5	2	4	4.32
24	4.90	638.70	1094.50	405.88	44.44	-9.90	126.12	0.14	5	2	4	4.93
25	4.60	618.28	1068.27	387.44	43.15	-11.84	123.01	-0.66	6	3	4	4.02
26	5.20	648.45	1107.98	416.48	44.36	-15.02	126.62	-4.32	8	2	5	4.17
27	4.90	660.25	1129.85	401.46	44.99	-11.33	127.78	-0.63	6	2	5	2.45
28	4.30	644.81	1103.65	401.44	44.17	-13.52	123.12	-0.75	7	2	4	3.05
29	5.50	669.76	1146.93	435.89	46.10	-13.21	127.84	-0.97	7	2	4	3.05
30	5.50	692.54	1187.84	470.33	48.03	-12.95	132.56	-1.20	7	2	4	3.56
31	4.90	653.57	1120.47	417.44	44.81	-17.44	124.73	-1.78	8	3	4	2.14
32	5.50	679.45	1160.63	446.44	46.01	-18.40	128.34	-5.43	10	2	5	3.22
33	5.20	691.95	1181.43	431.47	46.65	-14.96	129.50	-1.74	8	2	5	2.29

➢ **Regression analyzes:**

A relationship between independent and dependent variables (physicochemical descriptors and biological activities, respectively) were determined statistically using regression analysis. In the present work, Multiple Linear Regression MLR analysis of molecular descriptors was carried out using the stepwise strategy in SPSS version 20 for Windows **[8].**

V.3 QSAR model:

In the present study we tried to develop best QSAR model to explain the correlations between the physicochemical parameters and the biological activities IC50 values of 1, 2-diazole derivatives. Among several QSAR equations the best QSAR models were selected on the basis of various statistical parameters such as:

correlation coefficient R which measures the degree of line association between two variables.

Squared correlation coefficient (R^2> 0.6) which is relative measure of quality of fit. Standard error of estimate representing absolute measure of quality of fit, Fischer's value (F), F is the Fisher ratio, reflects the ratio of the variance explained by the model and the variance due to the error in the regression. High values of the F-test indicate that the model is statistically significant **[9].**

After multiple regression analysis on the software SPSS the best statistically significant correlations generated along with pertinent statistical parameters are given below:

Log (I/IC50) = 2,057+0,031MR+0,305LogP-0,259LogD

n=26 **R=0.887** **R^2=0.786** **SE=0.339** **F=26.993** **Q=2.61**

Or: n is the number of compounds. R is the coefficient of correlation, F is the Fischer statistic and SE is the standard error of estimation and Q is the quality of the adjustment or the adaptation.

Interpretation of the model:

The values of the variance of the fractions can be between 0 and 1 The OSAR model having R^2> 0.6 will only be considered for the validation. For example, like our case, the R= 0,887 and R^2= 0.786 values allowed us to strongly indicate the correlation between different parameters (independent variables) with the inhibitory activity of enoyl acyl carrier protein reductase.

The calculated F value for the QSAR model generated exceeds the F value tabulated by large margin as desired for a significant regression. In addition, the F-value was found to be statistically significant at the 95% level for this model.

The positive value of quality factor (Q) for this OSAR model suggests high predictive power and its lack of adaptation as well as the low standard deviation demonstrates the accuracy of this model.

In the model obtained in equation **(v.1)** the negative coefficient of log P explains that any increase of the lipophilicity of the molecules causes an increase in biological activity, which corresponds to the hydrophobicity of the molecule.

Also it can be observed that the increase of positive coefficient of the molar refractivity causes an increase of the biological activity, ie that our nucleus must be substitute with one or the other electron or free electronic pairs.

Thus log D with the negative coefficient causes a decrease in the biological activity for the big compounds.

The correlation matrix for the biological activity plC50 and the descriptors selected to build the model QSAR-2D is presented in the table **(V.2)** .The parameters used in this model are almost independent, which can be seen from the correlation matrix.

The table **V.2** indicate the importance of physicochemical parameters which the lipophilicity log P is the most important (68.7%) in the description of the specific activity of 1, 2-diazole.

Table V.2: Correlation matrix of model

	pIC 50	**MR**	**LogP**	**LogD**
pIC 50	**1**			
MR	**0,573**	**1**		
LogP	**0,687**	**0,441**	**1**	
LogD	**0,223**	**0,641**	**0,609**	**1**

V.4. Model validation:

The predictive powers of the equations were validated by the "leave-one-out" (LOO) **[10, 11]** cross validation method. Cross-validation is a convenient and reliable way to test the meaning of a model. Therefore, to validate the final models generated individually for different activities properties. A one-out method is used to perform cross-validation In order to test the validity of the predictive power of the selected MLR model (eq log (l /IC50), the technique of "Leave-One-Out" (LOO-technique) was used.

The developed model has been validated by the calculation of the following statistical parameters: PRESS (sum of predicted residual squares), SSY (sum of squares of response value), R^2_{CV} (global predictive ability), R^2 adjusted, PE predictive error correlation coefficient) and S_{PRESS} (prediction uncertainty) **(Table V.3).**

Table V.3: Cross-validation parameters

Model	P_{RESS}	SSY	P_{RESS}/ SSY	S_{PRESS}	r^2_{cv}	r^2_{adj}	6PE
1	2,540	11,888	0,213	0,311	0,773	0,746	0,169

P_{RESS} is an important cross validation parameter because it is a good approximation of the actual predictive error of the model. Its value being lower than SSY indicates that this model predicts better than chance and can be considered statistically significant. According to the results presented in **(Table V.3)** this value is equal to **2.540** the model is statistically significant.

In addition, for a reliable QSAR model, the P_{RESS}/ SSY ratio should be less than 0.4 **[12].** From the data presented in **(Table V.3)** indicate that for the model developed, this ratio is **0.213**.

The indication of the performance of the model is obtained from $\mathbf{R^2_{cv}}$ (the overall prediction capability). The high value of $\mathbf{r^2_{cv}}$ and $\mathbf{r^2_{adj}}$ are essential criteria for the best qualification of the QSAR model. Our result of these two values for this QSAR model was **0.773** and **0.746** respectively.

S_{PRESS} (Predictive Uncertainty) is a good parameter used to decide the uncertainty in the prediction. The lower the value of this parameter, the predictive power of the model will be better in our case this parameter carries a small value **0.311** which explains that the ability of prediction is the best for this model.

The predictive error of the correlation coefficient **(PE)** is another parameter used to evaluate the predictive power of the proposed models. The calculation of the proposed model **PE** value is present in **(Table V.3)** .For this model, the condition **R> 6PE** is satisfied and can therefore be said to have good predictive power. However, the only way to estimate the real predictive power of the developed model is to predict the calculation of the log (I/ IC50) values of the 1, 2-diazole studied using this model. The experimental, predicted and residual ENR inhibitory activity of 1, 2-diazole and its derivatives shown in **(Table V.4)** were deduced by SPSS software **[13].**

Table V.4: Experimental, predicted and residual values of (log (I/IC_{50})) of 1, 2-diazol derivates.

Number	pIC50	Predicted value	Residual
1	4,79	4,8208	-0,30780
2	4,79	5,0518	-0,28178
3	5,39	5,1205	0,26955
4	5,09	5,4460	-0,35599
5	4,79	4,7457	0,44340
6	5,09	4,8911	0,19886
7	5,09	4,8033	0,28666
8	5,60	5,4046	0,19540

9	5,42	5,2875	0,13247
10	5,20	5,3172	-0,11723
11	5,60	5,2209	0,37911
12	3,60	3,4739	0,12609
13	3,30	3,4068	-0,10680
14	3,00	3,8669	-0,86692
15	4,20	3,5975	0,60254
16	5,20	5,4743	-0,27432
17	5,20	5,3360	-0,13600
18	5,38	5,3746	0,00539
19	5,24	5,3109	-0,07086
20	5,78	5,3578	0,44219
21	4,60	4,7917	-0,19169
22	4,90	4,7123	0,18774
23	4,90	5,1705	-0,27048
24	4,30	4,8350	-0,53495
25	4,90	4,8061	0,09395
26	5,20	4,9265	0,38431

Figure **V.1** shows the plot of linear regression predicted versus experimental values of biological activity of 1, 2-diazole derivatives outlined above. The plot for model shows to be more convenient with R^2= 0.786. The evaluation set has a good distribution along the range of values of the training set. The present QSAR study shows that this model can be applied successfully to predict inhibitory activities against mycobacterium tuberculosis disease in these generations of molecules.

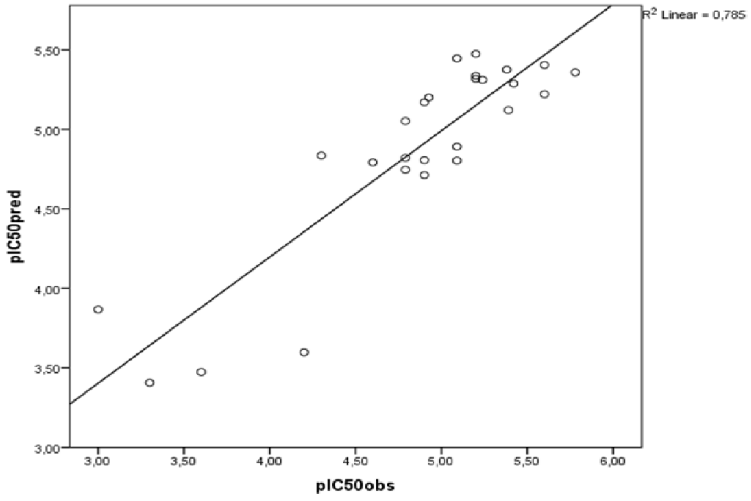


Figure V.1: The curve of the predicted values according to the experiment values of log (I/IC_{50})

To investigate the presence of a systematic error in developing the QSAR models, the residuals of predicted values of the biological activity (log (I/IC50) was plotted against the experimental values Table **V.4**, as shown in Figure **V.2.**

The propagation of the residuals on both sides of zero indicates that no systemic error exists, as suggested by Jalali-Heravi and Kyani **[14].** It indicates that this model can be successfully applied to predict the anti-tubercular activity of this class of molecules.

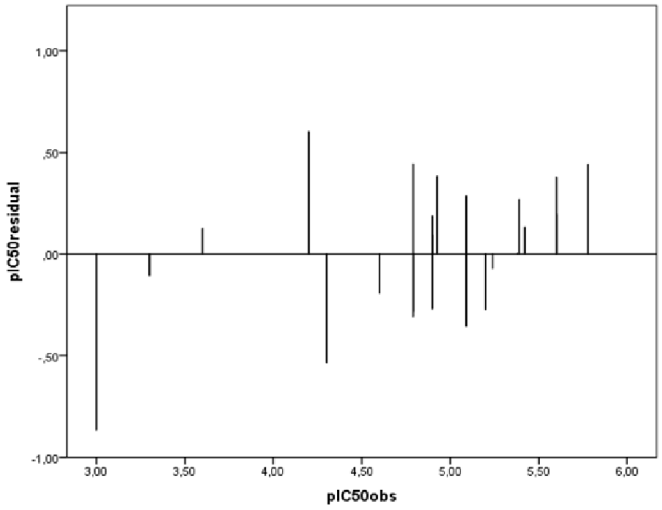


Figure V.2: Curve of residual values in relation to the experimentally observed

V.5 Leads identification:

Computer-aided drug design uses computational chemistry for the discovery, enhancement and study of biologically active drugs and molecules **[15].** All the stages of intervention of the computer tool are presented in the following summary diagram:

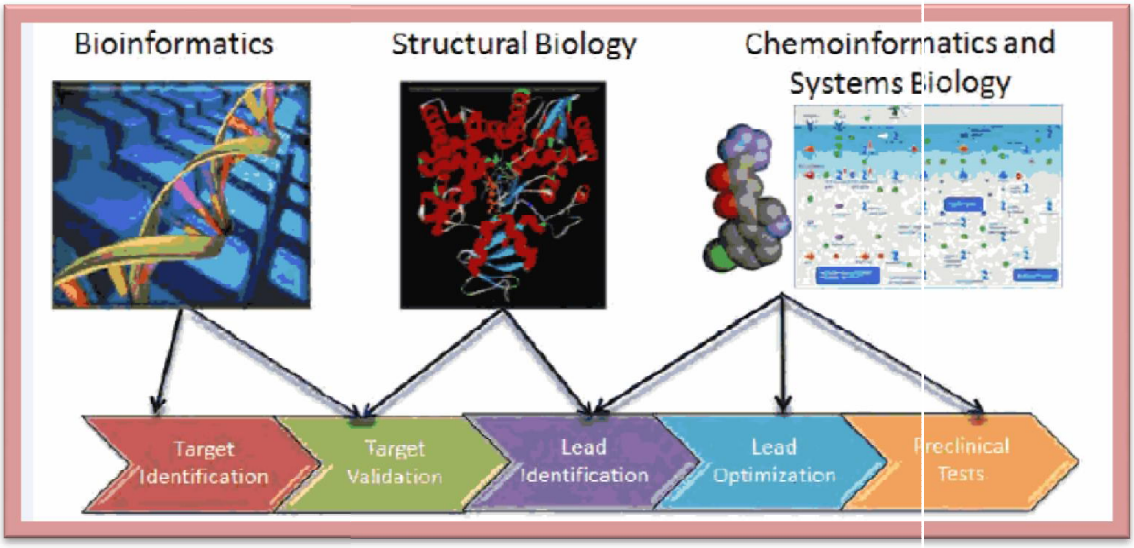


Figure V.3: Computer-aided drug design

Among the keys obtained at the end of a screening campaign, it is necessary to identify those that represent leads relevant for pathology to be treated **[16].**

A "lead" molecule is a molecule that will interact with the target, for to inhibit, activate or modify its activity in any way. So these are general ligands of the target. So that these molecules have a chance to become a drug candidate, they must additionally possess ADME-T properties meeting certain criteria specific to the life of a drug in the human body or animal.Extensive studies of structure-activity relationships (RSA) are conducted by varying the chemical structures of the hits by modifying the functional groupings while keeping their basic skeleton.

Functional groups while keeping their basic skeletons **[17].** The best leads will then be optimized taking care to preserve the favorable properties of activity and drug-like while trying to improve the affinity, the selectivity as well as the permeability. Moreover, during the lead phase, it is important to look for evidence that the observed biological effect is indeed induced by the lead's interaction with the target **[18].** In our case, the lead molecules

have been identified. Among all the compounds studied which corresponds to the compounds 11 and 12 which present on the one hand, an activity important, and on the other hand, optimal physico-chemical and biological properties that will make the molecule a drug that is both effective and little (or not) toxic.

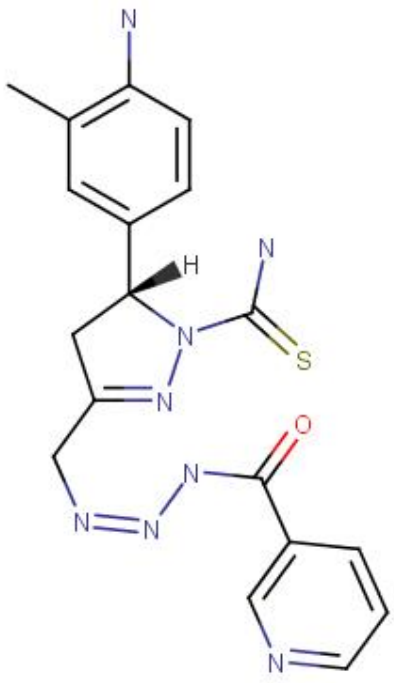


Figure V.4: Compound 11 (5R)-5-(4-amino-3-methylphenyl)-3-{[(Z)-{[(pyridin-3-yl)formamido]imino}amino]methyl}-4,5-dihydro-1H-pyrazole-1-carbothioamide

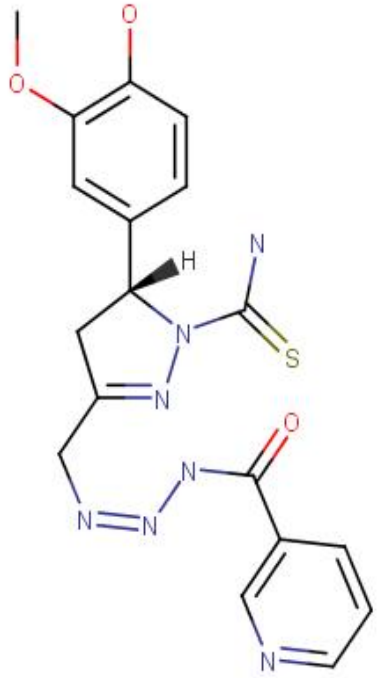


Figure V.5: Compound 12 (5R)-5-(4-hydroxy-3-methoxyphenyl)-3-{[(Z)-{[(pyridin-3-yl)formamido]imino}amino]methyl}-4,5-dihydro-1H-pyrazole-1-Carbothioamide

By comparison, of the two structures of the 1,2-diazole derivatives the compound 11and 12 after evaluation of the results obtained in the SAR and QSAR we can conclude that these two compounds presents the highest possibility for becoming molecules.

V.6 Conclusion:

The QSAR analysis was carried out to determine the quantitative relationship between the molecular structure of the compounds with their activities was carried out. Different physicochemical parameters were used; in particular the partition coefficient (Log P), RM, (Log D), can be used successfully to model anti tubercular activities derivatives.

The QSAR model indicates that these descriptors have significant relationships with the observed bioactivity. We observed a high similarity between the experimental values and the predicted values of the activity, which indicates the excellent quality of the QSAR model. By comparison of the different selection methods applied to the compounds studied, two molecules were found in the series with high ADME properties.

References:

1. Ceyda Oksel, Xue Z. Wang, Quantitative Structure-Activity Relationship (QSAR) models, University of Leeds 27.03.2013.
2. S. P. Bradbury. Toxicol. Lett.79(1-3), 229 (1995)
3. C. Hansch, A. Leo, R. Stephen, Eds. Heller, Exploring QSAR, Fundamentals and Applications in Chemistry and Biology, ACS professional Reference Book, American Chemical Society, Washington, D.C., (1995).
4. M. Zahouily, M. Lazar, A. Elmakssoudi, J. Rakik, S. Elaychi, A. Rayadh, J. Mol. Model. 12 (4), 398-405 (2006).
5. G.-Z. Liang, H. Mei, P. Zhou, Y. Zhou, Z.-L. Li. Acta. Phys-Chim. Sin 22(3), 388-390 (2006).
6. Christophe Hurlin, The Multiple Linear Regression Model, University of OrlÈans, November 23, 2013.
7. 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.**
8. SPSS software packages , SPSS Inc., 444 North Michigan Avenue, Suite 3000, Chicago, Illinoi, 60611, USA.
9. K . J. Sanmati, J. Rahul, S. Lokesh, K. Y. Arvind, J. Chem. Pharm. Res. 4, 3215-3223 (2012).
10. S. K. Purkayastha, T. Jha, D. K. Pal and A. U. De, Anticancer. Drug. Des. 8, 95-100(1993).
11. I. V. Tetko, V. Y. Tanchuk and A. E. Villa, J. Chem. Inf. Comput. Sci. 41, 1407-1421(2001).
12. S O. Podunavac-Kuzmanovi c, D.D. Cvetkovi c, and D.J.Barna. Int. J. Mol. Sci. 10, 1670, 2009.
13. SPSS 20 pour Windows.
14. M. Jalali-Heravi, A. Kyani, J. Chem. Inf. Comput. Sci. 44, 1328 (2004).
15. FARIDA, Louacheni. Developpement d'un portail web pour ie criblage virtuel sur la grille de calcul, 2014. Thèse de doctorat.

16. HAZAN, Corinne. Recherche d'inhibiteurs de haute affinité de l'ADN polymérase bêta par criblage virtuel et RMN. 2008. Thèse de doctorat. Université de Toulouse, Université Toulouse III-Paul Sabatier.

17. B.C. Duffy; L.Zhu; H. Decornez ; et al. Early phase drug discovery cheminformatics and computational techniques in identifying lead series. Bioorg Med Chem, 2012.**20**(18): p.5324-42.

18. Hubbard, R.E. 3D structure and the drug-discovery process. Mol. BioSyst, 2005. **1**: p.391-406.

GENERAL CONCLUSION

Our work concerns the use of computational chemistry for the evaluation of biological activity and the physicochemical characterization of pyrazole and its derivatives against mycobacterium tuberculosis disease.

Various molecular modeling methods have been used in this work on 1, 2-diazole molecules; so our work contains:

- a conformational analysis on the pyrazole nucleus.
- a qualitative study on the structure-properties / activity relationships of a bioactive series of 1,2-diazole derivatives
- a quantitative study of structure-activity relationships on a series of thirty-three 1, 2-diazole derivatives, of which a QSAR model was successfully developed.

First, the results demonstrate that the structural and electronic comparison of the 1, 2-diazole nucleus shows similar results using different calculation methods: semi-empirical method (PM3) and quantum methods Ab initio and DFT. The effectiveness of the method (DFT) has been demonstrated to make calculations on the nucleus of pyrazole.

Pyrazoles are an important class of heterocyclics in medicinal chemistry because many derivatives can identify activities of interest in a wide range of biological targets.

The qualitative study of the structure-properties relationship was carried out on a series of 33 of 1, 2-diazole derivatives. The molecules used in this study have pharmacological activities. The diversity of the groups which bind to the base nucleus ie with studied structural diversities which influences the physico-chemical properties of the pyrazole derivatives and consequently on their pharmacological properties.

One of the objectives of this work in this study was to develop and evaluate quantitative structure- activity (QSAR) relationships for the prediction of ENR inhibition by pyrazole derivatives, as anti-tubercular agents.

A multiple linear regression analysis was performed to derive activity and structure quantitative relationship models that were internally evaluated for prediction of activity from derived molecular descriptors belonging to the series pyrazole.

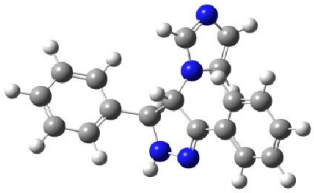
GENERAL CONCLUSION

Our results suggest the best QSAR model with the following descriptors: MR, log P and log D for each IC50 biological activity with (R^2 = 0.786, Fischer test value F =26.993).

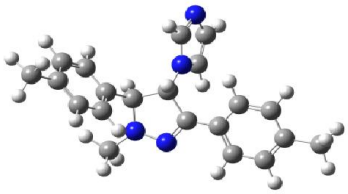
The developed models were validated by the "leave one out" technique as well as by the calculation of the statistical parameters.

The developed QSAR model can be useful for predicting the inhibitory activity of enoyl acyl carrier protein reductase againstmycobacterium tuberculosis . The QSAR studies were performed on 33 pyrazole derivatives.

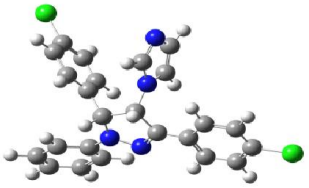
3D presentation of the structures of derivatives of 1, 2-diazole



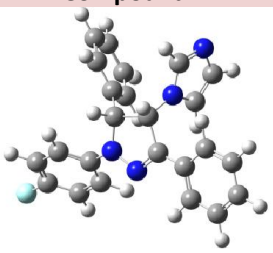
Compound 1



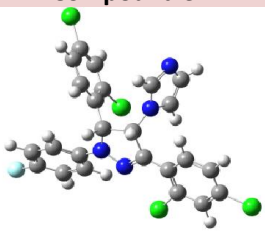
Compound 2



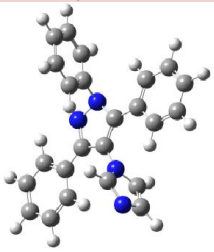
Compound 3



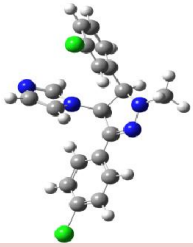
Compound 4



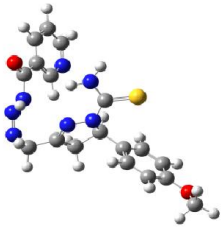
Compound 5



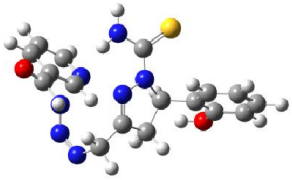
Compound 6



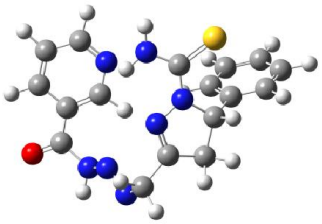
Compound 7



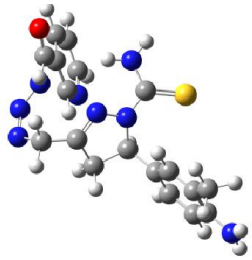
Compound 8



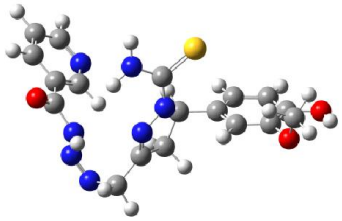
Compound 9



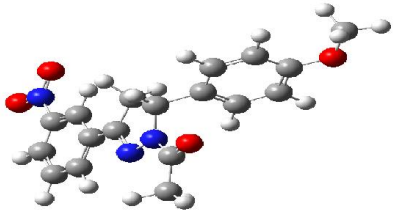
Compound 10



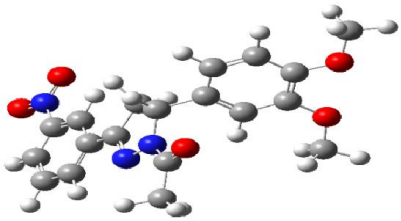
Compound 11



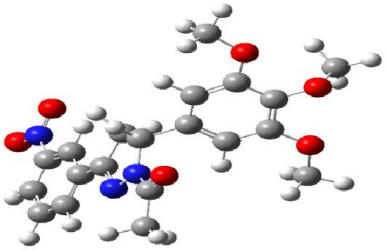
Compound 12



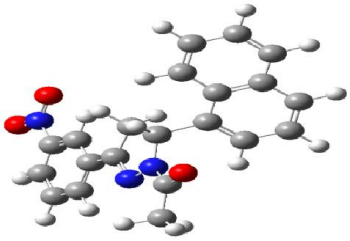
Compound 13



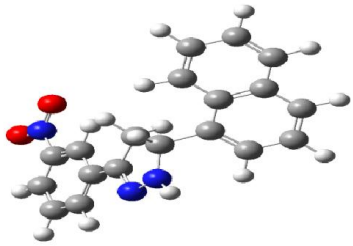
Compound 14



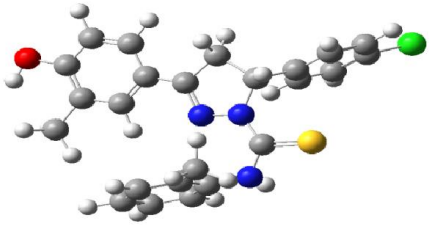
Compound 15



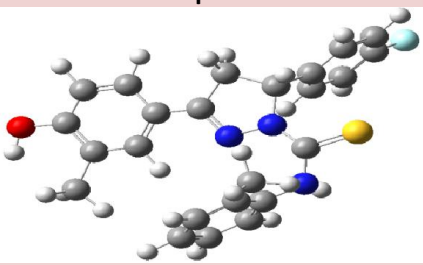
Compound 16



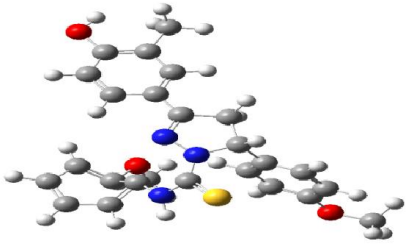
Compound 17



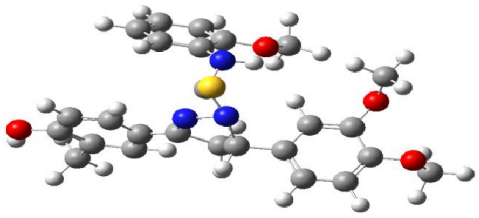
Compound 18



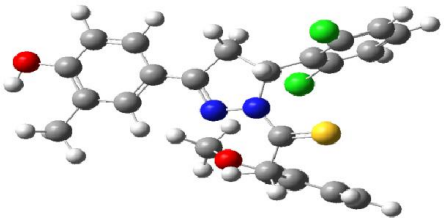
Compound 19



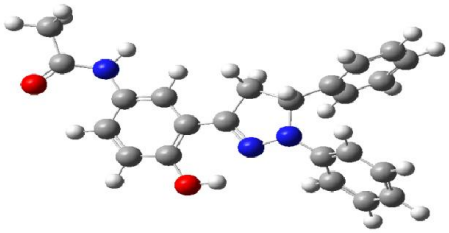
Compound 20



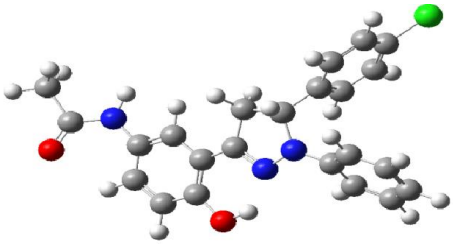
Compound 21



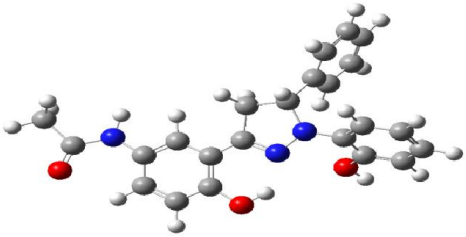
Compound 22



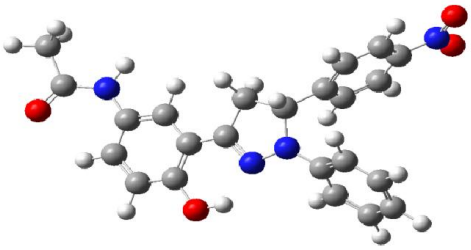
Compound 23



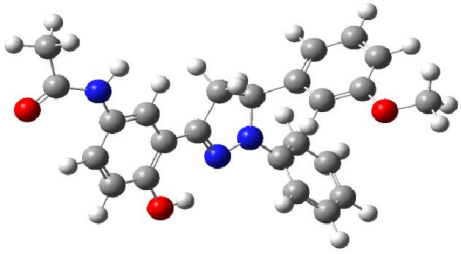
Compound 24



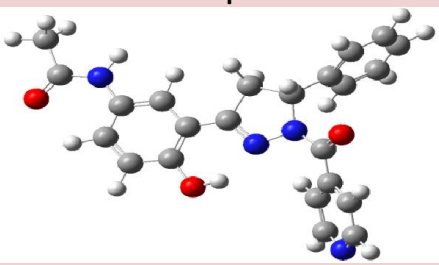
Compound 25



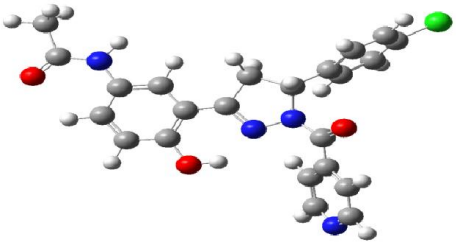
Compound 26



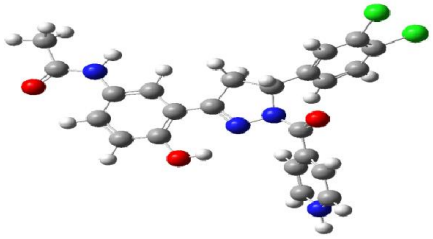
Compound 27



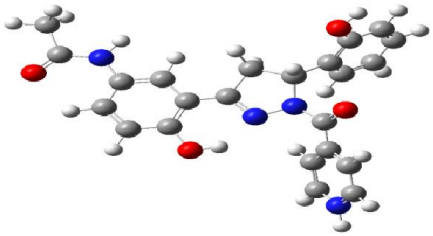
Compound 28



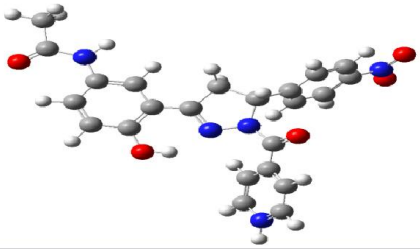
Compound 29



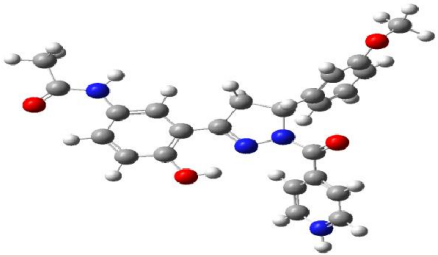
Compound 30



Compound 31



Compound 32



Compound 33

Abstract

The work done in this thesis deals with a general research that aims to discuss the activity of a series of thirty-three of 1,2- diazole derivatives as inhibitors of ENR that may be potential agents for treating mycobacterium tuberculosis.

To begin with, various theoretical calculation methods were used to carry out our work: PM3, ab initio HF and DFT / B3LYP. These methods were used to determine the structural and electronic parameters associated with the studied molecules. Then, a qualitative study of the structure-activity relationship (SAR) was also performed for a bioactive series of pyrazole derivatives using different MPO methods. Finally, a quantitative study was also performed to predict the biological activity of the compounds studied and its derivatives by suggesting the best QSAR model. The model prediction obtained was confirmed by the LOO cross validation method.

Keyword: pyrazole, DFT, ENR, SAR, MPO, QSAR

Résumé

Le travail effectué au cours de ce mémoire concernent une recherche générale qui vise à discuter de l'activité d'une série de trente trois dérivés de 1,2- diazole comme des inhibiteurs de ENR qui peuvent être des agents potentiels pour traiter le mycobactéries TB.

Pour commencer, des différentes méthodes de calcul théorique ont été utilisés pour réaliser notre travail: PM3, ab initio/ HF et DFT/B3LYP. Ces méthodes ont été utilisées pour déterminer les paramètres structuraux, électroniques associés aux molécules étudiées. Ensuite, une étude qualitative de la relation structure-activité (SAR) a été effectuée également pour une série bioactive de dérivés de pyrazole en utilisant des différentes méthodes MPO.

Enfin, une étude quantitative a été effectuée également pour prédire l'activité biologique des composés étudiés et ses dérivés en suggérant le meilleur modèle QSAR. La prédiction du modèle obtenue a été confirmé par la méthode de validation croisée LOO.

Mot-clé: pyrazole, DFT, ENR, SAR, MPO, QSAR

الملخص

العمل المنجز في هذه المذكرة يتعلق ببحث عام يهدف الى مناقشة نشاط سلسلة من ثلاثة و ثلاثين مشتقات 1,2-diazol كمثبطات لENR و التي قد تكون محتملة لعلاج داء السل الميكوبكتيري .

بداية تم استخدام طرق مختلفة للحساب النظري لأداء عملنا : B3LYP/ DFT/, ab initio/HF,PM3

و قد استخدمت هذه الطرق لتحديد العوامل الهيكلية و الالكترونية المرتبطة بالجزيئات المدروسة. ثم تم اجراء دراسة نوعية في العلاقة بين الهيكل و النشاط (SAR)لسلسلة نشطة بيولوجيا من مشتقاتpyrazole باستخدام اساليب MPOمختلفة .

و اخيرا اجريت دراسة كمية للتنبؤ بالنشاط البيولوجي للمركبات المدروسة و مشتقاتها من خلال اقتراح افضل نموذج QSAR. ثم تاكيد نموذج التنبؤ الذي تم الحصول عليه من خلال طريقة التحققLOO .

الكلمات المفتاحية: بيرازول، DFT ،ENR، SAR، MPO ،QSAR