

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE

UNIVERSITE MOHAMED KHIDER BISKRA

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

Département des sciences de la matière

THESE

Présentée par

HAZHAZI Halima

En vue de l'obtention du diplôme de :

Doctorat en chimie

Option :

Chimie Moléculaire

Intitulée:

**Etude par la modélisation moléculaire de la réactivité
chimique et l'activité biologique de quelques composés
organiques**

Soutenue le : 2017/2018

Devant la commission d'examen :

M. OMARI Mahmoud	Prof.	Université de Biskra	Président
M. BOUMEDJANE Youcef	MC/A	Université de Biskra	Directeur de thèse
M. BELAIDI Salah	Prof.	Université de Biskra	Examineur
M. MESSAOUDI Abdelatif	MC/A	Université de Batna	Examineur
M. MELKEMI Nadjib	MC/A	Université de Biskra	Examineur

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA
RECHERCHE SCIENTIFIQUE

UNIVERSITE MOHAMED KHIDER BISKRA

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

Département des sciences de la matière

THESE

Présentée par

HAZHAZI Halima

En vue de l'obtention du diplôme de :

Doctorat en chimie

Option :

Chimie Moléculaire

Intitulée:

**Etude par la modélisation moléculaire de la réactivité
chimique et l'activité biologique de quelques composés
organiques**

Soutenue le : 2017/2018

Devant la commission d'examen :

M. OMARI Mahmoud	Prof.	Université de Biskra	Président
M. BOUMEDJANE Youcef	MC/A	Université de Biskra	Directeur de thèse
M. BELAIDI Salah	Prof.	Université de Biskra	Examineur
M. MESSAOUDI Abdelatif	MC/A	Université de Batna	Examineur
M. MELKEMI Nadjib	MC/A	Université de Biskra	Examineur

To my beloved parents

To my Brothers and sisters

To my husband

To all my friends anywhere

Acknowledgements

First of all I should be so grateful and thankful for Allah, the Most Gracious and the Most Merciful for giving me the patience, power and will to complete this work.

Appreciation to Mr. Youcef Boumedjane, MC/A at Med Khider University of Biskra for his supervises, his precious advices, encouragement and his motivation over the course of the research to make my thesis complete.

I am deeply grateful to Prof. Nadjib Melkemi for his support and wise advice and make everything easy. I am also thanks to Prof. Boulanouar Messaoudi at University of Tlemcen, for all their encouragement, help and support.

I hereby address my sincere thanks to Mr. Mahmoud OMARI president of the Jury. Professor at University of Biskra, and to Mrs: Salah BELAIDI, Professor at the University of Biskra, Nadjib MELKEMI, MC/A at the University of Biskra, Abdelatif MESSAOUDI, MC/A at the University of Batna, or accepting to examine and evaluate this work.

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

And finally, thanks to my dear parents and my brothers, sisters and my best friends for their unconditional support and advice.

Contents

LIST OF ABBREVIATIONS	I
LIST OF TABLES	III
LIST OF FIGURES & SCHEMAS.....	IV

General Introduction

General Introduction.....	02
References.....	06

Chapter I

Quantum methods in chemistry

I.1. Introduction.....	09
I.2. Theoretical basis.....	10
I.2.1. Schrödinger equation.....	10
I.2.2. Born-Oppenheimer Approximations.....	10
I.2.3. Hartree-Fock approximation.....	11
I.3. Density-functional theory (DFT).....	12
I.3.1. Hohenberg and Kohn theorems.....	12
I.3.2. Kohn-Sham theorems.....	14
I.3.3. Exchange-correlation (XC) functional.....	15
I.3.3.1. Local density approximation (LDA).....	16
I.3.3.2. Generalised Gradient Approximation (GGA).....	17
I.4. Hybrid functional.....	18
I.5. Basis sets.....	19
I.6. solvation methods.....	20
I.7. Comparison between HF and DFT.....	22
I.8. References.....	24

Chapter II

Theoretical study of the regio- and stereoselectivity of the 1,3-DC reaction of 2,3,4,5-tetrahydropyridine-1-oxide with methyl crotonate

II.1. Introduction.....	28
II.2. 1,3-Dipolar cycloaddition reactions.....	30
II.2.1. The 1,3-dipole.....	30
II.2.2. The dipolarophile.....	32
II.3. Mechanistic aspects.....	33
II.4. Regioselectivity.....	34
II.5. Stereoselectivity.....	35
II.6. Theoretical analysis of 1,3-dipolar cycloaddition reactions.....	36
II.6.1. Frontier molecular orbital theory and Sustmann classification.....	36
II.6.2. Transition state theory (TST).....	37
II.6.3. Chemical Reactivity Indexes.....	38
II.6.3.1. Global properties.....	39
II.6.3.1.1. Electronic chemical potential.....	39
II.6.3.1.2. Chemical hardness.....	40
II.6.3.1.3. Electrophilicity.....	41
II.6.3.1.4. Nucleophilicity.....	41
II.6.3.2. Local properties.....	42
II.7. Materials and methods.....	43
II.8. Results and Discussion.....	44
II.8.1. Regiochemistry study based on FMO and reactivity indices.....	44
II.8.2. Mechanistic study of the cycloaddition reaction based on activation energy..	49
II.8.2.1. Energies of the Transition Structures.....	49
II.8.2.2. Geometry analysis.....	53
II.9. Conclusion.....	54
II.10. References.....	55

Chapter III

DFT-based reactivity and QSAR modeling of 1,2,4,5-tetrazine inhibitors

III.1. Introduction.....	60
III.2. Objectives of QSAR.....	61
III.3. Molecular Descriptors.....	62
III.3. 1. Constitutional descriptors.....	62
III.3. 2. Topological descriptors.....	62
III.3. 3. Electrostatic descriptors.....	62
III.3. 4. Geometrical descriptors.....	63
III.3. 5. Quantum chemical descriptors.....	63
III.4. Biological Parameters.....	63
III.5. Statistical methods.....	64
III.5.1. Multiple linear regression.....	65
III.6. Chemometric Tools.....	65
III.6. 1. Determination coefficient (R^2).....	65
III.6. 2. Adjusted R^2 (R_a^2).....	66
III.6. 3. Variance ratio (F).....	66
III.6. 4. Standard error of estimate (s).....	67
III.7. Validation of QSAR Models.....	67
III.8. Material and methods.....	69
III.8.1. Computational methods.....	69
III.8.2. QSAR modeling.....	70
III.9. Results and discussion.....	70
III.9.1. Analysis of the DFT reactivity indices of 1,2,4,5-tetrazine.....	70
III.9.2. Study of Quantitative structure-activity relationship (QSAR) for 1,2,4,5 - tetrazine derivatives.....	73
III.10. Conclusion.....	79
III.11. References.....	80
General conclusion.....	85
Appendix.....	88
Abstract	

List of Abbreviations

A	Electron Affinity
B3LYP	Becke 3-Parameter Lee-Yang-Parr
CDFT	Conceptual DFT
CV	Cross-Validation
DFT	Density Functional Theory
DC	Dipolar Cycloaddition
DCM	Dichloromethane
E	Electronic Energy
FMO	Frontier Molecular Orbital
GGA	Generalized Gradient Approximation
GTO	Gaussian Type Orbital
HF	Hartree-Fock
HOMO	Higher Occupied Molecular Orbital
HSAB	Hard and Soft Acids and Bases
I	Ionization Potential
IRC	Intrinsic Reaction Coordinate
IED	Inverse Electron Demand
IC50	Half maximal Inhibitory Concentration
KS	Kohn-Sham
LDA	Local Density Approximation
LSD	Local Spin-Density Approximation
LCAO	Combination of Atomic Orbital
LUMO	Lower Unoccupied Molecular Orbital
LOO	Leave One Out
MM	Molecular Mechanics
MO	Molecular Orbital
MLR	Multiple Linear Regression
NPA	Natural Population Analysis
NBO	Natural Bond Orbital
N	Number of Electrons
NED	Normal Electron Demand

PCM	Polarized Continuum Model
PES	Potential Energy Surface
QM	Quantum Mechanics
QSAR	Quantitative Structure Activity Relationship
SCF	Self Consistent Field
SCRf	Self-Consistent Reaction Field
STO	Slater-Type Orbital
TCE	Tetracyanoethylene
TST	Transition State Theory
XC	Exchange-Correlation
2D, 3D	Two-Dimensional, Three-Dimensional

List of Tables

Table I.1: Different kinds of salvation models

Table II.1: Classification of the parent 1,3-dipoles

Table II.2: FMOs Molecular Coefficients of the dipole and dipolarophile

Table II.3: FMO energies (a.u), electronic chemical potential (a.u), chemical hardness (a.u), electrophilicity index (eV) and nucleophilicity index (eV)

Table II.4: Local properties of dipole and dipolarophiles calculated at B3LYP/6-31G (d) level of theory

Table II.5: Total energies (a.u), Relative energies ΔE (in kcal/mol), relative free energies ΔG (in kcal/mol) and enthalpies ΔH (in kcal/mol) in gas phase and in DCM, of the stationary points involved in the 1, 3-DC reaction between 2,3,4,5-tetrahydropyridine-1-oxide and methyl crotonate

Table II.6: Values of $|\Delta d|$ in TS-men, TS-mex, TS-oen and TS-oex of the 1, 3-DC reaction of dipolarophile with dipole in the gas phase and solvent DCM

Table III.1: Reactivity descriptors for 1,2,4,5-tetrazine at the B3LYP/6-311++G(d,p) level

Table III.2: Fukui function values of 1,2,4,5-tetrazine in gas and aqueous phases

Table III.3: Chemical structures and experimental activity of the 1,2,4,5-tetrazine derivatives under study

Table III.4: Quantum chemical descriptors of 1,2,4,5-tetrazine derivatives in both gas and aqueous phases

Table III.5: Cross-validation parameters in both gas and aqueous phases

Table III.6: Experimental, predicted and residual activity of 1,2,4,5-tetrazine derivatives in gas and aqueous phases

List of Figures & Schemes

Figure 1: Timeline from quantum mechanics to computational chemistry

Figure II.1: Classification of the parent 1,3-dipoles

Figure II.2: Examples of dipolarophiles in 1,3-dipolar cycloaddition reactions

Figure II.3: Sustmann classification of 1,3-dipolar cycloadditions

Figure II.4: Reaction co-ordinate diagram

Figure II.5: Optimized geometries and visualized FMOs for the reactants

Figure II.6: The interactions between HOMO and LUMO orbitals of a 1,3 dipole/dipolarophile

Figure II.7: Prediction of the favoured interactions between dipole and dipolarophile using DFT based indices

Figure II.8: Optimized transition structures of the 1-3 DC reaction between tetrahydropyridine-1-oxide and methyl crotonate

Figure II.9: Energy profiles, in kcal/mol, for the 1,3-DC reactions of stationary point in gas phase

Figure II.10: Energy profiles, in kcal/mol, for the 1,3-DC reactions of stationary point in solvent phase

Scheme II.1: 1,3-Dipolar cycloaddition of 2,3,4,5 tetrahydropyridine 1-oxide with methylcrotonate

Scheme II.2: Concerted mechanism for DCR

Scheme II.3: Stepwise diradical mechanism for DCR

Scheme II.4: The exo and endo approaches of tetrahydropyridine-1-oxide to methyl crotonate

Figure III.1: Schematic overview of the QSAR process

Figure III.2: Electron-density mapped $f(+)$ and $f(-)$ Fukui function for 1,2,4,5-tetrazine in both gas and aqueous phases (the blue regions show the areas of the molecules most susceptible to nucleophilic attacks and the red regions show the areas of the molecules most susceptible to electrophilic attacks)

Figure III.3: Predicted plots versus experimental observed antitumor activity for models in both gas and aqueous phase

Figure III.4: Plots of residual against experimental observed in gas and aqueous phase

GENERAL INTRODUCTION

General introduction

Until the beginning of the 1960's, the fundamentals of Classical Physics and Quantum Chemistry were slowly described and implemented in computer models. In the last two decades of the 20th century this field developed rapidly. The made progress led to the creation of a new field of study for the scientists that investigate matter and its properties and interactions: the computational chemistry [1].

The computational chemistry field has become an important research tool in chemistry, physics, and biology; the last began with quantum theory, which is the study of the interaction of atoms with each other and with energy at the subatomic and atomic level. Quantum theory aims to predict the behavior of atoms based on the physics principles that apply at such a small scale. Calculations that using quantum mechanics-based equations such as the wave function and Schrödinger equation permit scientists to make prediction and analysis various characteristics and behaviors of both atoms and molecules [2].

In the modern times, The main relevance of theoretical and computational chemistry has been recognized by the Nobel Prizes Foundation, which awarded to Walter Kohn and John A. Pople in 1998 and, more recently, to Karplus, Levitt and Warshel in 2013 (Figure 1) [3].

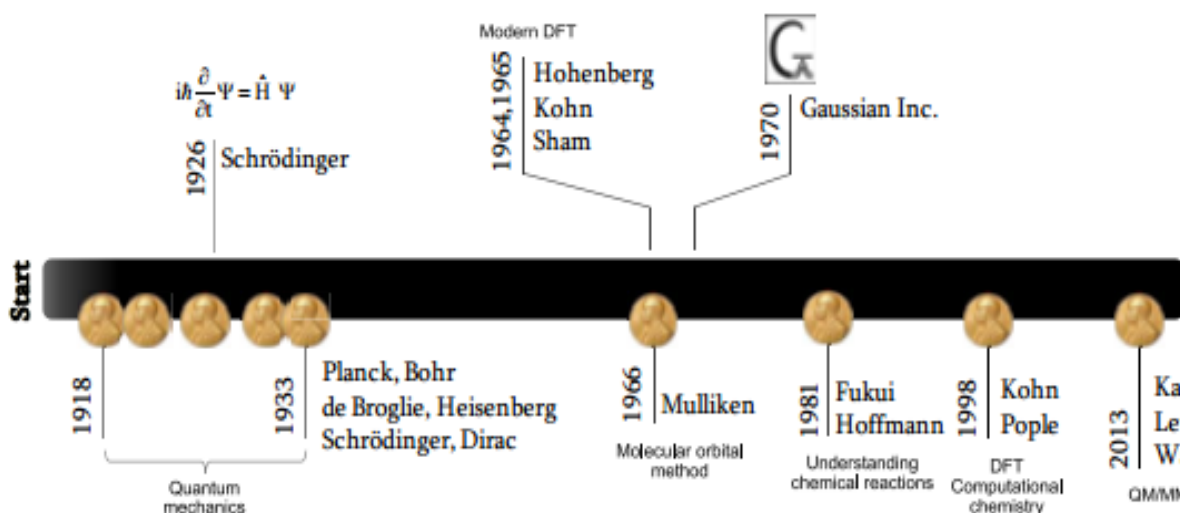


Figure 1: Timeline from quantum mechanics to computational chemistry

General introduction

Computational chemistry (also called molecular modelling; both terms mean about the same thing) is a techniques set for resolving chemical problems using a computer. Among the common questions investigated in this field are:

- Molecular geometry: the shapes of molecules – bond lengths, angles and dihedrals. The energies of molecules and transition states: this tells us which isomer is favored at equilibrium, and (from transition state and reactant energies) how fast a reaction should go.
- Chemical reactivity: for example, knowing where the electrons are concentrated (nucleophilic sites) and where they want to go (electrophilic sites) enables us to predict where various kinds of reagents will attack a molecule [4].

During the past two decades, density functional theory (DFT) method has revolutionized the theoretical studies of chemical reactivity in numerous domain of chemistry, such as, from inorganic to organic chemistry and from material science to biochemistry [5]. This important branch of the DFT, (called conceptual DFT, also known as chemical reactivity theory) is a powerful tool for the prediction, analysis, and interpretation of the outcome of chemical reactions [6].

Recently, Density functional theory (DFT) based descriptors calculated such as global quantities: (The electronic chemical potential (μ), the electrophilicity (ω) and the nucleophilicity indices (N)) and the local condensed indices: like the electrophilic (P_k^+) and nucleophilic (P_k^-) Parr functions, ect. have been widely utilized for the prediction of the regioselectivity and the reactivity of atoms in molecules [7-9]). In fact, many studies relating to the theoretical prediction of regioselectivity and stereoselectivity in 1,3-dipolar cycloaddition reactions have been the subject of theoretical studies [10].

1,3-dipolar cycloadditions (13DC) are one of the most important classes of organic reactions and are among the most versatile and powerful preparative methods for the synthesis of cyclic compounds. These cycloadditions have been used for the preparation of compounds that are of fundamental importance in various areas of chemistry [11, 12].

General introduction

In addition The chemistry of 1,3-dipoles has created great interest and application over more than a century [13]. A historical study of their cycloaddition reactions has been taking a major importance in both academia and industry. Such as, the nitrones reactions leading to 5-membered isoxazolidine rings are of particular interest in bio-organic chemistry. Considering the need for stereo-specific synthesis of the cycloadducts, theoretical prediction of the probable adducts along with their preferred reaction path would be of much importance [14, 15].

In the other hand, Quantum chemical descriptors like orbital energies, frontier orbital densities, chemical potential (μ), hardness (η) etc. have been extensively used in developing different quantitative structure activity relationships (QSARs) for predicting reactivity in terms of the structure and physicochemical properties of molecules [16]. Numerous reviews have been published on the applications of quantum chemical descriptors in QSAR. Recently the uses of quantum chemical descriptors in the development of QSAR model have received attention due to reliability and versatility of prediction by these descriptors [17].

The main objective of QSAR is to look for new molecules with required properties using chemical intuition and experience transformed into a mathematically quantified and computerized form [17].

In this case, several statistic techniques can be utilized for developing a QSAR model. In our study, the QSAR model is developed by using a multiple linear regression (MLR) technique. The advantage of MLR is that it is simple to use and the derived models are easy to interpret [18].

The objective of this work is the study of some organic compounds using conceptual DFT methods. First of all, we focus on our investigation and interpretation of the regio- and stereoselectivity of 1,3-DC reactions (The work is published in: Moroccan Journal of Chemistry , 2016). Also, we study a quantitative structure activity relationship (QSAR) of a series of 1,2,4,5-tetrazine as antitumor activity against lung cancer cell lines (A-549) (The work will be published sooner).

General introduction

The manuscript of this thesis is divided into three parts:

- *Chapter I:* Quantum methods in chemistry

In this chapter, we present theoretical background of the methods of quantum chemistry (HF, and DFT methods, ect), as well as, the bases description of atomic orbitals.

- *Chapter II:* Theoretical study of the regio- and stereoselectivity of the 1,3-DC reaction of 2,3,4,5-tetrahydropyridine-1-oxide with methyl crotonate

For this part, we present general informations about 1,3 dipolar cycloaddition reaction, and the different theories which used to study the reactivity and the selectivity: Frontier molecular orbital theory (FMO), Transition state theory (TST), conceptual DFT. Then we discuss the important results of 1,3- DC reaction between nitron and methyl crotonate.

- *Chapter III:* DFT-based reactivity and QSAR modeling of 1,2,4,5-tetrazine inhibitors

This chapter contains information about QSAR (objectives of QSAR, molecular descriptors, statistical parameters), and we present the results into two points: firstly, we analyze the molecular reactivity of 1,2,4,5-tetrazine, While in the second point, we establish a quantitative relationship between physiochemical properties and biological activity of a series of 1,2,4,5 –tetrazine using QSAR modeling.

References

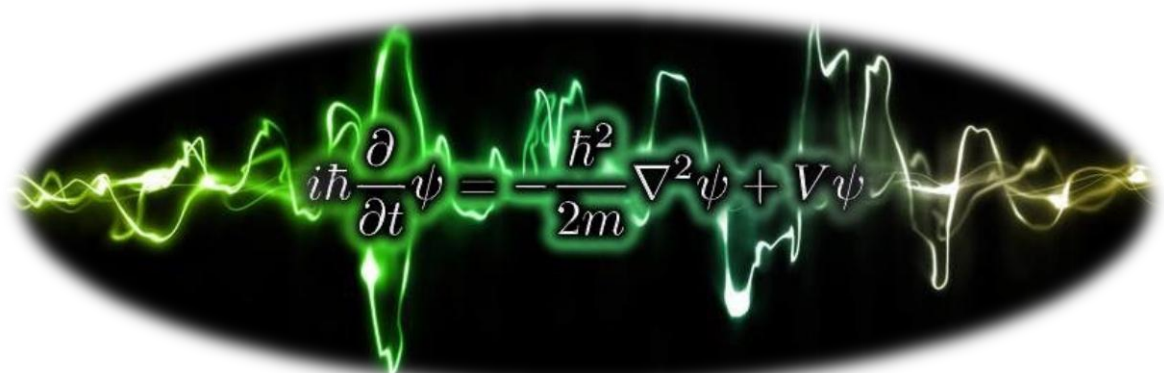
- [1] J. A. Sousa, P. P. Silva, A. E. H. Machado, M. H. M. Reis, L. L. Romanielo, C. E. Hori, *Braz. J. Chem. Eng.*, **30**, 84 (2013).
- [2] k. j. devaney, Ch. R. Hango, J. lu, D. sigalovsky, computational chemistry in the high school classroom, report for the bachelor of science degree, worcester polytechnic institute (2004).
- [3] M.A.O. Maqueda. Scope of computational organometallic chemistry; Doctoral thesis: Univ. of Barcelona; (2014).
- [4] E. G. lewars, computational chemistry, introduction to the theory and applications of molecular and quantum mechanics, 2nd edition, springer (2011).
- [5] T.Labbé, A. *Theoretical Aspects of Chemical Reactivity*; Elsevier Science: Amsterdam, the Netherlands, **19**, 19 (2007).
- [6] P. Chattaraj, *Chemical Reactivity Theory—A Density Functional View*; CRC Press, Taylor & Francis Group: Boca Raton, FL, USA, (2009).
- [7] S. Rajkhowa, R. C. Deka, *Curr Pharm Des.*, **20**, 4325 (2014).
- [8] R. G. Parr, R.G. Pearson, *J. Am. Chem. Soc.*, **105**, 7512 (1983).
- [9] P. Geerlings, F. De Proft, W. Langenaeker, *Chem Rev.*, **103**, 1793 (2003).
- [10] W. Lwowski, A. Padwa, Ed, « In 1,3- Dipolar Cycloaddition Chemistry», vol. **1**, Chapter 5, Wiley-Interscience, New York (1984).
- [11] K. Asit, Chandra, T. Uchimar, M. T. Nguyen, *J. Chem. Soc., Perkin Trans.*, **2**, 2117 (1999).
- [12] W.J. Choi, Zh. D. Shi, K.M. Worthy, L. Bindu, R.G. Karki, M.C. Nicklaus, R.J. Fisher Jr, T.R. Burke, *Bioorg. Med. Chem. Lett.*, **16**, 5262 (2006).
- [13] T. Curtius, *Ber. Dtsch. Chem. Ges.*, **16**, 2230 (1883).

General introduction

- [14] K.V. Gothlf, K.A. Jorgensen, *Chem. Rev.*, **98**, 863 (1998).
- [15] T. K. Das, S. Salampuria, M. Banerjee, *J. Mol. Struct. THEOCHEM.*, **959**, 22 (2010).
- [16] D. R. Roy, U. Sarkar, P. K. Chattaraj, A. Mitra, J. Padmanabhan, R. Parthasarathi, V. Subramanian, S. Van Damme. P. Bultinck. *Mol. Divers.*, **10**, 119 (2006).
- [17] R. Parthasarathi, V. Subramanian, D. R. Roy, P. K. Chattaraj, *Bioorg. Med. Chem.*, **12**, 5533 (2004).
- [18] N. Frimayanti, M. L. Yam, H. B. Lee, O. Rozana, Sh. M. Zain, A. R. Noorsaadah. *Int. J. Mol. Sci.*, **12**, 8639 (2011).

CHAPTER I

Quantum Methods in Chemistry


$$i\hbar \frac{\partial}{\partial t} \psi = -\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi$$

I.1. Introduction

Nowadays computers have become an integral part in the world of science such as physics and chemistry, especially when it comes to calculatory problems. For problems which cannot be solved analytically, computers and numerical methods are of crucial importance [1].

Quantum chemistry is the study of chemical processes, spectroscopic constants, interactions, reactions and all the other things we wish to know about molecules studied by methods of quantum mechanics. Since they were introduced in the first half of this century, the quantum chemical models have had a large conceptual impact on chemistry. Researchers in many areas of chemistry use terms adopted from quantum chemistry such as molecular orbitals, energy levels, hybridization and resonance theory, and use quantum chemical models to explain their observations. At the same time the development of efficient algorithms and the enormous increase in computational resources due to the development of modern computers have actually made it possible to perform ab initio calculations, that is, calculations from first principles, of chemical properties [2].

Density functional theory (DFT) [3-5] is one of the most successful quantum chemistry tools. Compared to other quantum chemistry methods, because it has a high accuracy per unit computer processor (CPU) time ratio in comparison to other methods [6]. The premise behind DFT is that the energy of a molecule can be determined from the electron density instead of a wave function. This theory originated with a theorem by Hohenberg and Kohn that stated this was possible (Hohenberg and Kohn 1964). Kohn and Sham who formulated a method similar in structure to the Hartree-Fock method (Kohn and Sham 1965) developed a practical application of this theory [7].

I.2. Theoretical basis

I.2.1. Schrödinger equation

The ultimate goal of most approaches in solid state Physics and quantum Chemistry is the solution of the time-independent, non-relativistic Schrödinger equation:

$$\hat{H}\Psi = E\Psi \quad (1)$$

E: is the energy of the molecule

Ψ : is the wave function (eigenfunction for a given Hamiltonian)

The Hamiltonian (\hat{H}) is a sum of all possible interactions between electrons and nuclei, \hat{H} can be expanded as:

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \frac{1}{2} \sum_{A=1}^M \frac{1}{M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} \quad (2)$$

Here, A and B run over the M nuclei while i and j denote the N electrons in the system.

The first two terms describe the kinetic energy of the electrons and nuclei. The other three terms represent the attractive electrostatic interaction between the nuclei and the electrons and repulsive potential due to the electron-electron and nucleus-nucleus interactions.

I.2.2. Born-Oppenheimer Approximations

Due to their masses the nuclei move much slower than the electrons so we can consider the electrons as moving in the field of fixed nuclei therefore the nuclear kinetic energy is zero and their potential energy is merely a constant. Thus, the electronic Hamiltonian reduces to

$$\hat{H}_{elec} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} = \hat{T} + \hat{V}_{Ne} + \hat{V}_{ee} \quad (3)$$

the solution of the Schrödinger equation with \hat{H}_{elec} is the electronic wave function Ψ_{elec} and the electronic energy E_{elec} . The total energy E_{tot} is then the sum of E_{elec} and the constant nuclear repulsion term E_{nuc} [8].

$$\hat{H}_{elec}\psi_{elec} = E_{elec}\psi_{elec} \quad (4)$$

$$E_{tot} = E_{elec} + E_{nuc} \quad (5)$$

Where

$$E_{nuc} = \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}$$

I.2.3. Hartree-Fock approximation

The electronic Schrödinger equation is still intractable and further approximations are required. The most obvious is to insist that electrons move independently of each other. In practice, individual electrons are confined to functions termed molecular orbitals, each of which is determined by assuming that the electron is moving within an average field of all the other electrons. The total wavefunction is written in the form of a single determinant (a so-called Slater determinant). This means that it is antisymmetric upon interchange of electron coordinates.

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(1) & \chi_2(1) & \dots & \chi_n(1) \\ \chi_1(2) & \chi_2(2) & \dots & \chi_n(2) \\ \chi_1(N) & \chi_2(N) & \dots & \chi_n(N) \end{vmatrix} \quad (6)$$

Here, χ_i is termed a spin orbital and is the product of a spatial function or molecular orbital, ψ_i , and a spin function, α or β .

The set of molecular orbitals leading to the lowest energy are obtained by a process referred to as a “self-consistent-field” or SCF procedure. The archetypal SCF procedure is the Hartree-Fock procedure, but SCF methods also include density functional procedures. All SCF procedures lead to equations of the form.

$$f(i) \chi(x_i) = \epsilon \chi(x_i) \quad (7)$$

Here, the Fock operator $f(i)$ can be written.

$$f(i) = -\frac{1}{2} \nabla_i^2 + v^{\text{eff}}(i) \quad (8)$$

x_i are spin and spatial coordinates of the electron i , χ are the spin orbitals and v^{eff} is the effective potential “seen” by the electron i , which depends on the spin orbitals of the other electrons. The nature of the effective potential v^{eff} depends on the SCF methodology [9].

I.3. Density-functional theory (DFT)

The conceptual root of density functional theory (DFT) is in the Thomas-Fermi model, developed by Thomas and Fermi in 1927, who used a statistical model to approximate the distribution of electrons in an atom. This Thomas-Fermi model is inaccurate for most applications. With the Hohenberg-Kohn theorems, DFT was put on a firm theoretical footing. With The Kohn-Sham computational scheme, using approximate functionals, it was possible to study molecular systems of chemistry successfully. Now DFT has established itself as one of the most popular quantum chemistry tools due to a rapid development after 1990 [10].

I.3.1. Hohenberg and Kohn theorems [11]

The Hohenberg-Kohn theorems relate to any system consisting of electrons moving under the influence of an external potential $V_{\text{ext}}(r)$. Stated simply they are as follows:

Theorem 1

Their first theorem states that the external potential $V_{\text{ext}}(\vec{r})$ applied on the system (V_{ext} is an external potential to the system which is due to the presence of the nuclei) is defined as a unique functional of the electronic density, $\rho(r)$. One particular external potential can be defined by one and only one particular electron density and vice versa. In turn V_{ext} fixes \hat{H} which is therefore a unique functional of $\rho(\vec{r})$

$$\rho_0 \Rightarrow \{N, Z, R\} \Rightarrow V_{\text{ext}} \Rightarrow \hat{H} \Rightarrow \Psi_0 \Rightarrow E_0 \quad (9)$$

where the “₀” index represents the system in its ground state. So, for the ground state, the energy of the system is written:

$$E_0 = T[\rho_0] + E_{\text{ee}}[\rho_0] + E_{\text{Ne}}[\rho_0] \quad (10)$$

The $T[\rho]$ and $E_{\text{ee}}[\rho]$ part of the equation are independent of the variables N , R and Z (respectively: number of electrons, electron-nucleus distance and the nuclear charge) whereas E_{Ne} is dependent upon those variables. So the previous equation can be re-written as following:

$$E_0[\rho_0] = \underbrace{T[\rho_0]}_{\text{universally valid}} + \underbrace{E_{\text{ee}}[\rho_0]}_{\text{system dependent}} + \underbrace{\int \rho_0(\vec{r}) V_{\text{Ne}} d(\vec{r})}_{\text{system dependent}} \quad (11)$$

The independent parts are gathered into a new quantity: the Hohenberg-Kohn functional:

$$F_{\text{HK}}[\rho] = T[\rho] + E_{\text{ee}}[\rho] \quad (12)$$

If the functional F_{HK} was known exactly, it would allow the calculation of E_0 . However the explicit forms of the two terms which compose this functional are unknown. The E_{ee} term can be separated in two terms: a Coulomb part and a term containing all the non-classical contributions to the electron-electron interaction.

$$E_{\text{ee}}[\rho] = J[\rho] + E_{\text{ncl}}[\rho] \quad (13)$$

Theorem 2

The second HK is simply the use of the variational theory applied to the electronic density. When an approximate electronic density $\tilde{\rho}(\vec{r})$, associated with an external potential V_{ext} , is used, the resulting energy, as in HF, will always be greater than or equal to the exact ground state energy:

$$E[\tilde{\rho}] = T[\tilde{\rho}] + E_{\text{Ne}}[\tilde{\rho}] + E_{\text{ee}}[\tilde{\rho}] \geq E_{\text{exact}} \quad (14)$$

I.3.2. Kohn-Sham theorems

The Hohenberg–Kohn theorems state that the electron density can be used to determine the energy and properties of a molecular system, but does not give a form for the energy functional $E[\rho]$. An approach to calculate the electron density was formulated in 1965 by Kohn and Sham [12]. They proposed to substitute the exact but unknown kinetic energy functional $T[\rho]$ with the known kinetic energy functional for a system of non-interacting particles, $T_s[\rho]$. The relatively small correction to the kinetic energy ($T[\rho] - T_s[\rho]$) is then taken together with the exchange and the correlation parts in the exchange–correlation functional $E_{\text{xc}}[\rho]$. The energy functional $E[\rho]$ may then be written as

$$E[\rho] = T_s[\rho] + J[\rho] + E_{\text{ext}}[\rho] + E_{\text{xc}}[\rho] \quad (15)$$

Kohn and Sham introduced one-electron functions ψ_i called KS orbitals from which the electron density can be calculated as

$$\rho = \sum_{i=1}^n |\psi_i|^2 \quad (16)$$

Where the sum is over all occupied KS orbitals i . The kinetic energy of the non-interacting system can be calculated exactly using KS orbitals as

$$T_s[\rho] = \sum_{i=1}^n \langle \psi_i | -\frac{1}{2} \nabla_i^2 | \psi_i \rangle \quad (17)$$

The KS orbitals can be obtained by solving the KS equations [12],

$$\left[-\frac{\nabla_i^2}{2} + \int \frac{p(r_j)}{|r_i - r_j|} dr_j + v_{\text{ext}}(r_i) + v_{\text{xc}}(r_i) \right] \psi_i(r_i) = \varepsilon_i \psi_i(r_i) \quad (18)$$

with the exchange–correlation potential $v_{\text{xc}}[\rho]$ defined as a functional derivative

$$v_{\text{xc}}[\rho] = \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho} \quad (19)$$

Note the similarity between the HF equations in Eq.7 and the KS equations in Eq.18. The three last terms on the left-hand side of Eq.18 are collectively referred to as the effective potential or the KS potential,

$$v_{\text{eff}}(\mathbf{r}_i) = \int \frac{\rho(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_j + v_{\text{ext}}(\mathbf{r}_i) + v_{\text{xc}}(\mathbf{r}_i) \quad (20)$$

The KS approach thus gives a practical way of calculating the electron density from the KS orbitals similar to the calculation of HF orbitals. First, a set of KS orbitals is chosen as an initial guess. Next, a new set of KS orbitals is obtained by solving Eq.18 using an appropriate exchange–correlation functional $E_{\text{xc}}[\rho]$. This SCF procedure is repeated until convergence. The optimized electron density can then be calculated from the occupied KS orbitals using Eq.16. If the exact form of the exchange–correlation functional $E_{\text{xc}}[\rho]$ were known, the KS approach would give the exact electronic energy including electron correlation. Various approximations to $E_{\text{xc}}[\rho]$ have been proposed and the quest towards an exact density functional is an important area of research in theoretical chemistry.

I.3.3. Exchange-correlation (XC) functional

Having emphatically derived such a formally exact theory, one might be tempted to assume the problem is solved—we have a working method, which includes correlation and is no more computationally difficult than Hartree–Fock. There is, however, a remaining problem, in that the exact form of the exchange–correlation energy functional $E_{\text{xc}}[\rho]$ (hereafter often simply referred to as “functional” or “ E_{xc} ” for brevity) is still unknown, and an approximation must still be made. Although the contribution of E_{xc} is relatively small, the quality of the approximation directly affects the accuracy of any calculations. Understandably, finding a universal functional form appropriate for all systems ranging from a single proton to the largest of proteins, in terms of the simple, three-dimensional density, is not a trivial task, and a large part of the last fifty years of research in DFT has been dedicated to creating and improving approximations to the elusive E_{xc} functional. Before discussing in detail the common schemes of functional development, we will briefly comment on the extension of DFT to a spin-dependent form, necessary for modelling the effect of a magnetic field on the spins of electrons, and in general improving the approximations in the absence of a magnetic field. We consider separately the α -electron and β -electron densities

$$\rho_{\alpha}(r) = \sum_i^{n_{\alpha}} |\Psi_i(r, \alpha)|^2 \quad (21)$$

And

$$\rho_{\beta}(r) = \sum_i^{n_{\beta}} |\Psi_i(r, \beta)|^2 \quad (22)$$

with the total density given by

$$\rho(r) = \rho_{\alpha}(r) + \rho_{\beta}(r) \quad (23)$$

n_{α} and n_{β} are the number of α - and β -spin electrons, respectively. In the case of closed-shell systems, $\rho_{\alpha} = \rho_{\beta} = \rho/2$ [13].

I.3.3.1. local density approximation (LDA)

The local density approximation (LDA) is the basis of all approximate exchange-correlation functionals. At the center of this model is the idea of a uniform electron gas. This is a system in which electrons move on a positive background charge distribution such that the total ensemble is neutral.

The central idea of LDA is the assumption that we can write E_{XC} in the following form

$$E_{XC}^{LDA}[\rho] = \int \rho(r) \epsilon_{XC}(\rho(r)) dr \quad (24)$$

Here, $\epsilon_{XC}(\rho(\mathbf{r}))$ is the exchange-correlation energy per particle of a uniform electron gas of density $\rho(\mathbf{r})$. This energy per particle is weighted with the probability $\rho(\mathbf{r})$ that there is an electron at this position. The quantity $\epsilon_{XC}(\rho(\mathbf{r}))$ can be further split into exchange and correlation contributions,

$$\epsilon_{XC}(\rho(r)) = \epsilon_X(\rho(r)) + \epsilon_C(\rho(r)) \quad (25)$$

The exchange part, ϵ_X , which represents the exchange energy of an electron in a uniform electron gas of a particular density, was originally derived by Bloch and Dirac in the late 1920's [8].

$$\epsilon_x = -\frac{3}{4} \left(\frac{3\rho(r)}{\pi} \right)^{1/3} \quad (26)$$

I.3.3.2. Generalised Gradient Approximation (GGA)

The generalized gradient approximation adds information about the density gradient at a particular point to the functional by introducing the dimensionless reduced density gradient,

$$x = \frac{|\nabla\rho|}{\rho^{4/3}} \quad (27)$$

Perdew and Wang (PW86) [14] modified the LDA exchange functional to include x with different exponents and three constants a , b and c [15],

$$E_x^{PW86}[\rho] = E_x^{LDA}[\rho] (1 + ax^2 + bx^4 + cx^6)^{1/15} \quad (28)$$

with x defined in Eq. 27. A popular GGA exchange functional is Becke's one-parameter functional known as B88,

$$E_x^{B88}[\rho] = \int \rho^{4/3} \left(C_x + \frac{\beta x^2}{1 + 6\beta x \cdot \text{arcsinh}(x)} \right) dr \quad (29)$$

This yields an electron density with the proper asymptotic limit. The single parameter β was obtained by fitting to exchange energies of noble gas atoms [16]. When comparing the B88 exchange functional with Eq.24, one can see that it is made up of $E_x^{LDA}[\rho]$ plus a correction that can be called ΔE_x^{B88}

Lee, Yang and Parr (LYP) proposed a GGA correlation functional with four parameters that were determined by fitting to the helium atom [17]. The B88 exchange and LYP correlation functionals are semi-empirical in the sense that they have parameters that are fit to experimental data. Alternatively, a functional can be *ab initio* by satisfying theoretically exact conditions [15].

Also higher-order derivatives of the density can be introduced in the density functional, giving rise to what is known as a *meta*-GGA functional. This is usually accompanied by a large increase in the number of semiempirical parameters [15].

I.4. Hybrid functionals

The previous functional types all present a problem because the exchange part is very poorly described due to a problem of electronic self-interaction. On the other hand, the exchange part in HF is defined exactly. So an alternative approach would be to use a mix of DFT and HF to describe the exchange energy. However, taking the correlation part from DFT and the exchange part from HF gives poor results (worse than GGA). A first approach to this problem would be to regroup the exchange and correlation parts, so a functional that describes the system better than the GGA functionals can be obtained [18]. The final solution to this problem is the use of a combination of HF, GGA and LSDA functionals to describe the exact exchange and correlation part of the hybrid functional. Usually hybrid functionals are composed of a mixture of exact and DFT exchange. The main element of these functional come from GGA functionals, so they are often called GGA hybrid functionals.

- **B3** contains exact exchange, and is an exchange functional developed by Becke [19]. It is a combination of LSDA and GGA functionals.
- **PBE0** also called PBE1PBE, has been developed by Adamo and Barone [20]. It is a combination of 75% PBE GGA exchange functional and 25% of HF exchange.
- **B97** and **B98** were developed first by Becke (B97) [21], then modified by Becke and Schmider (B98) [22]. Unlike PBE0 and B3 functionals, B98 and B97 are meta-GGA hybrid functionals instead of GGA hybrids. They contain an exchange part taken from HF method.

To describe correctly the exchange-correlation term, it is necessary to combine exchange and correlation functionals to obtain an hybrid functional such as: B1B95 [23], B1LYP [24] or B3P86 [25].

A famous example of exchange and correlation combination is the most often used hybrid functional: B3LYP [26]. The B3LYP functional is a mix between LDA and GGA functionals taken from the DFT and HF methods, to a certain extent, as shown below:

$$E_{XC}^{B3LYP} = E_{XC}^{LDA} + a_0(E_X^{HF} - E_X^{LDA}) + a_X \left(\underbrace{E_X^{GGA}}_{B88(B)} - E_X^{LDA} \right) + a_C \left(\underbrace{E_C^{GGA}}_{LYP} - E_C^{LDA} \right) \quad (30)$$

where $a_0 = 0.20$, $a_X = 0.72$ and $a_C = 0.81$ are three empirical parameters determined by fitting the predicted values to a set of atomisation energies, ionization potentials, proton affinities, and total atomic energies.

1.5. Basis sets

The spatial part of spin-orbital $\Psi(r,w) = X(r) \alpha(w)$ can be expanded as a Linear Combination of Atomic Orbitals (LCAO),

$$X_1(r) = \sum_m C_{lm} \Phi_m \quad (31)$$

where c_{lm} are the coefficients and Φ_m are the basis functions [27,28]. Φ_m are not necessarily orthogonal, but may be normalized. During SCF iteration, energy of the system is minimized by varying the coefficients until thresholds for change in energy and coefficients are reached. Earlier basis functions were analogous in formulation to the hydrogenic atomic orbitals, and were called as Slater-type orbitals (STOs) [29]. However, Gaussian-type of orbitals (GTOs) were later preferred [30], due to the complications in solving two-electron integrals. The GTOs can be expressed as,

$$\Phi(r_1 - r_c) = N(X_1 - X_c)^i (y_1 - y_c)^j (Z_1 - Z_c)^k e^{-\alpha(r_1 - r_c)^2} \quad (32)$$

Here N is the normalization constant, X_1 , y_1 and Z_1 are the co-ordinates of electron, X_c , y_c and Z_c are, co-ordinates of the center, i , j and k are the non-negative integers and α is a positive exponent. Multiple GTOs are often added together to approximate a single STO, a process known as 'contraction' [31,27]. For example, a STO approximated by n GTOs can be written as,

$$\Phi(r)_i^{STO-nG} = \sum_n d_{in} \psi_n(r) \quad (33)$$

The contraction coefficients d_{in} and the exponential constant α are usually obtained after iterative refinement procedure [27]. When compared to the simpler and less accurate STO-nG type of basis sets, split-valence (SV) type of basis sets are currently used more often. In SV basis sets, core and valence electrons are treated separately. One such commonly used basis set is 6-31G* (Pople notation [32-34]). In this notation the first number corresponds to the number of basis functions (or gaussian primitives) used to define the core electrons *e.g.* the $1s$ orbital. The last two digits denote the number of two separate contraction schemes that are used to treat the valence electrons. For example, 3 Gaussian primitives, and separately 1 Gaussian primitive for the valence electrons ($2s$ orbital onwards). These basis functions are often added with polarization (*) and diffuse functions (+) in order to provide a more realistic description of chemical bonding [34].

I.6. solvation methods




Understanding the effect of solvation on chemical reactions was one of the main objectives of research in the last century [35– 44] as evident from the wealth of literature available on this subject; numerous experimental and theoretical studies have been made in the past to assess the importance of solvent environment on the structure, stability, spectra, and reactivity. It is a well-known fact that accurate treatments of solvent effect on various chemical systems are particularly challenging for theoretical chemistry [36– 38].

In modeling the effect of solvation, different methods based on classical and quantum mechanics (QM) are used. Methods of treating solvation range from a detailed description at the atomic level to reaction field-based continuum methods. Some of the salient features of the various models are described in Table I.1.

It can be seen from Table I.1 that various methods used to model the effect of a solvent can be broadly classified into three types:

- 1- Those which treat the solvent as continuous medium,
- 2- Those which describe the individual solvent molecules (discrete/explicit solvation),
- 3- Combinations of (1 and 2) treatments. The following section provides a brief introduction to continuum models.

Table I.1: Different kinds of solvation models

Features	Models		
	Explicit Model	Continuum Model	Combined Model
Representation	Solvent molecules are explicitly treated	Represented as a solvent continuous medium	Combined representation of solute-solvent
			
Merits	It is generally more accurate and detailed Atomic interaction of solute-solvent	Simple, inexpensive to calculate	Atomic interaction of solute-solvent and bulk effects are included. Generally it gives better results than pure continuum models
Disadvantages	Computationally expensive	Ignore specific hydrogen bonding interactions	Computationally expensive when compared to continuum models and less demanding when compared to full explicit treatment

Continuum models are included in molecular mechanics (MM) and classical molecular dynamics approaches (force field - based simulations) to understand the solvent effect on large systems, specifically biomolecules [38]. These methods are successful to understand the effect of solvation on structure, stability, spectra, and reactivity [41–44].

In continuum model, the solvent is described as a uniform polarizable dielectric medium (ϵ) and the solute of suitable shaped cavity is placed in the dielectric medium. By definition, the continuum can be considered as a configuration-averaged or time-averaged solvent environment, where the averaging is Boltzmann-weighted at the temperature of interest. The creation of the solute cavity in the solvent continuum is a destabilization process.

Self-consistent reaction field (SCRF) models are the most efficient way to include condensed-phase effects into quantum mechanical calculations. This is accomplished by using SCRF approach for the electrostatic component. By design, it considers only one physical effect accompanying the insertion of a solute in a solvent, namely, the bulk polarization of the solvent by the mean field of the solute. This approach efficiently takes into account the long range solute–solvent electrostatic interaction and effect of solvent polarization. However, by design, this model cannot describe local solute–solvent interactions.

Pisa group of Tomasi and coworkers have made significant contributions to the development and implementation of these solvation models [41– 44]. They have developed a variety of quantum formulations of continuum models. In polarizable continuum method (PCM), more realistic cavity shape based on the van der Waals radii of the atoms in solute is defined [45]. In this method, cavity surface is divided into a large number of small surface elements with point charges. This system of point charges represents the polarization of solvent, and the magnitude of each surface charge is proportional to the electric gradient at that point. The total electrostatic potential at each surface element equals to the sum of potential due to solute and that of other surface charges. Various versions of PCM approach have been developed [41– 44]. Dielectric PCM (DPCM) [46], isodensity PCM (IPCM) [47], and self-consistent isodensity PCM (SCIPCM) [48] this some of the versions of PCM method.

1.7. Comparison between HF and DFT

The calculation schemes of these two methods are almost the same, which share common elements—the variational principle and the concept of single Slater determinant. However, there are some important differences between these two methods.

Dynamic Correlation

The dynamic correlation, caused by the electron-electron Coulomb interaction, is of great importance to the quality of a computational scheme. A very important difference recognized is that DFT handles this term properly while HF doesn't. Despite the fact that the exact exchange-correlation functional is not available, DFT generally performs better than HF when applied to the study of the electronic structure of a ground state.

Molecular Orbitals vs Total Density

The molecular orbitals from the HF calculation are related to the ground state of a real system. All the molecular properties can be calculated from these molecular orbitals. The molecular orbitals of Kohn-Sham DFT are those of the reference non-interacting system. The function of these orbitals is to provide a correct total density. In DFT, all the molecular properties are closely related to the total density.

This difference leads to the fact that HF performs better than DFT when applied to the study of some physical process that are closely related to the motion of an unpaired electron. For example, orbital energies vs experimental ionization potential, dissociation process associated with a singly occupied orbital, chemical reactions involving the charge transfer of an unpaired electron. These cases are generally recognized as the problems of SIE in DFT [10].

I.8. References

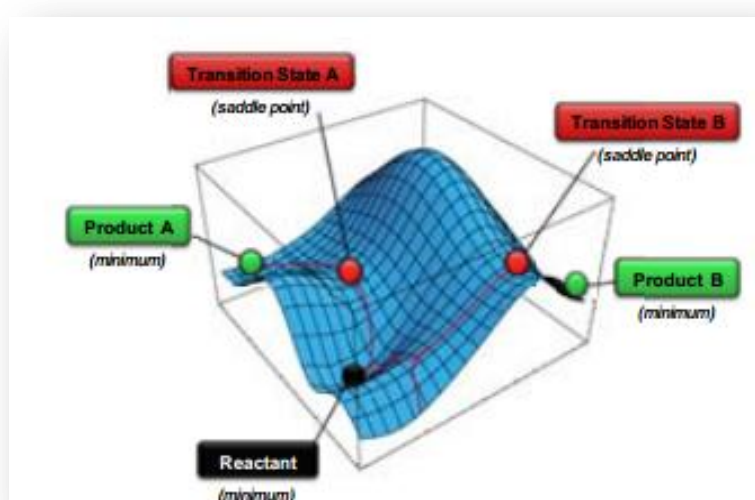
- [1] A. Szabo, N.S. Ostlund, *Modern Quantum Chemistry*. McGraw-Hill, (1989).
- [2] J. K. Laerdahl, *Development and Application of Relativistic Quantum Chemical Methods*; Doctoral thesis: Univ. of Oslo; (1997).
- [3] P. Hohenberg, W. Kohn, *Phys. Rev. B.*, **136**, 864 (1964).
- [4] W. Kohn, L. Sham, *Phys. Rev. A.*, **140**, 1133 (1965).
- [5] R. G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules*. Oxford University, Oxford, (1989).
- [6] k. j. devaney, Ch. R. Hango, J. lu, D. sigalovsky, *computational chemistry in the high school classroom*, report for the bachelor of science degree, worcester polytechnic institute (2004).
- [7] W. SNOR, *Molecular Modelling on Cyclodextrin Inclusion Complexes*; Doctoral thesis: Institute for Theoretical Chemistry, Univ. of Vienna; (2009).
- [8] A. Kh. Navid, *Linear response functions of solids within time-dependent density functional theory (TDDFT)*; Master Thesis: Univ. of Basque; (2011).
- [9] W. J. Hehre, *A Guide to Molecular Mechanics and Quantum Chemical Calculations*. Wavefunction, Inc (2003).
- [10] G. Tu, *Studies of Self-interaction Corrections in Density Functional Theory*; Royal Institute of Technology. Stockholm (2008).
- [11] T.J. Davin, *Computational chemistry of organometallic and inorganic species*, Doctoral thesis: Univ. of Glasgow; (2009).
- [12] W. Kohn, L. J. Sham. *Phys. Rev.*, **140**, A1133 (1965).
- [13] G. J. DAVID, *Correcting deficiencies in approximate density functionals*, Doctoral thesis: Univ. of Durham; (2015).
- [14] J. P. Perdew and Y. Wang. *Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation*. *Phys. Rev. B*, **33**, 8800–8802 (1986).

- [15] D. J. Tozer, Density functional theory. In R. Bast and P.-O. Widmark, editors, *European Summerschool in Quantum Chemistry. Book II*, pages 525–568. ESQC committee, eighth edition, (2013).
- [16] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, **38**, 3098 (1988).
- [17] C. Lee, W. Yang, R. G. Parr. Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, **37**, 785 (1988).
- [18] J. P. Perdew, M. Ernzerhof, K. Burke, *J. Chem. Phys.*, **105**, 9982 (1996).
- [19] A. D. Becke, *J. Chem. Phys.*, **98**, 5648 1993.
- [20] C. Adamo, V. Barone, *Chem. Phys. Lett.*, **298**, 113 (1998).
- [21] A. D. Becke, *J. Chem. Phys.*, **107**, 8554 (1997).
- [22] H. L. Schmider, A. D. Becke, *J. Chem. Phys.*, **108**, 9624 (1998).
- [23] A. D. Becke, *J. Chem. Phys.*, **104**, 1040 (1996).
- [24] C. Adamo, V. Barone, *Chem. Phys. Lett.*, **274**, 242 (1997).
- [25] J. P. Perdew, *Phys. Rev.*, B **33**, 8822 (1986).
- [26] C. Lee, W. Yang, R. G. Parr, *Phys. Rev.*, B **37**, 785 (1988).
- [27] P. Atkins, R. Friedman, *Molecular Quantum Mechanics*, Oxford Press (2005).
- [28] L. Jones, *J. E. Trans. Faraday Soc.*, **25**, 668 (1929).
- [29] J. C. Slater, *Phys. Rev.*, **36**, 57 (1930).
- [30] S. F. Boys, *Proc. R. Soc. Lond.*, A, **200**, 542 (1950).
- [31] W. J. Hehre, R. F. Stewart, J. A. Pople, *J. Chem. Phys.*, **51**, 2657 (1969).
- [32] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.*, **56**, 2257 (1972).
- [33] P. C. Hariharan, J. A. Pople, *Theoret. Chim. Acta.*, **28**, 213 (1973).

- [34] E. R. Davidson, D. Feller, *Chem. Rev.*, **86**, 681 (1986).
- [35] O. Tapia, J. Bertran, (eds.), *Solvent Effects and Chemical Reactivity*, Kluwer Academic, Dordrecht (1996).
- [36] F. Jensen, *Introduction to Computational Chemistry*, Wiley, Chichester (1999).
- [37] I. Levine, *Quantum Chemistry*, 5th ed., Prentice Hall, Upper Saddle River, NJ (1999).
- [38] A. R. Leach, *Molecular Modelling: Principles and Applications*, 2nd ed., Prentice Hall, Upper Saddle River, NJ (2001).
- [39] F. Hirata, *Molecular Theory of Solvation*, Kluwer Academic Press, New York (2003).
- [40] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, 2nd ed., Wiley, Chichester (2004).
- [41] J. Tomasi, M. Persico, *Chem. Rev.*, **94**, 2027 (1994).
- [42] C. J. Cramer, D. G. Truhlar, *Chem. Rev.*, **99**, 2161 (1999).
- [43] M. Orozco, F. J. Luque, *Chem. Rev.*, **100**, 4187 (2000).
- [44] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.*, **105**, 2999 (2005).
- [45] S. Miertus, E. Scrocco, J. Tomasi, *Chem. Phys.*, **55**, 117 (1981).
- [46] R. Cammi, J. Tomasi, *J. Comput. Chem.*, **16**, 1449 (1995).
- [47] J. B. Foresman, T. A. Keith, K. B. Wiberg, J. Snoonian, M. J. Frisch, *J. Phys. Chem.*, **100**, 16098 (1996).
- [48] K. B. Wiberg, T. A. Keith, M. J. Frisch, M. Murcko, *J. Phys. Chem.*, **99**, 9072 (1995).

CHAPTER II

Theoretical Study of the Regio- and Stereoselectivity of the 1,3-DC reaction of 2,3,4,5 Tetrahydropyridine -l-oxide with Methyl Crotonate

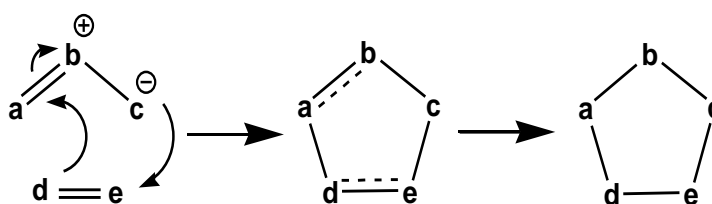


II.1. Introduction

Cycloaddition reactions are among the most powerful tools for synthesis and mechanistic interest in organic chemistry area because of their capacity to construct in region- and /or stereoselectively method [1].

There are different types of cycloaddition reaction. Concerted pericyclic cycloadditions involve reorganization of the π -electron systems of the reactants to form two new σ -bonds. Examples might include cyclodimerization of alkenes, cycloaddition of allyl cation to an alkene (1, 3-dipolar reaction), and the addition reaction between alkenes and dienes (Diels-Alder reaction). The cycloadditions can be characterized by specifying the number of π -electrons involved for each species, and for the above three cases, this would be [2+2], [3+2] and [4+2] respectively [2].

In recent years, 1, 3-dipolar reaction as a versatile method for preparing five-membered heterocyclic compounds is a classical reaction in organic chemistry and has been studied extensively. These cycloadditions have been utilized for the preparation of compounds that are of fundamental importance in diverse fields of chemistry [3]. The general concept of 1,3-dipolar cycloadditions was introduced by Huisgen and coworkers in the early 1960s [4]. Here, a five membered ring is formed by the cycloaddition of 1,3-dipole molecule (a three atom entity; $a-b-c$) and dipolarophile (a two atom entity; $d-e$) [5].



Several experimental and theoretical studies continued to populate the literature over different manners of the 1,3 dipolar cycloaddition[6]. Reactions between nitrones and alkenes to obtaining isoxazolidines are well-known due to their great importance in construction processes [7]. Substituted isoxazolidines are interesting biological active compounds [8] that could be used as enzyme inhibitors, antibiotics [9], and applied as synthetic intermediates of a variety of compounds [10]. Many theoretical investigations

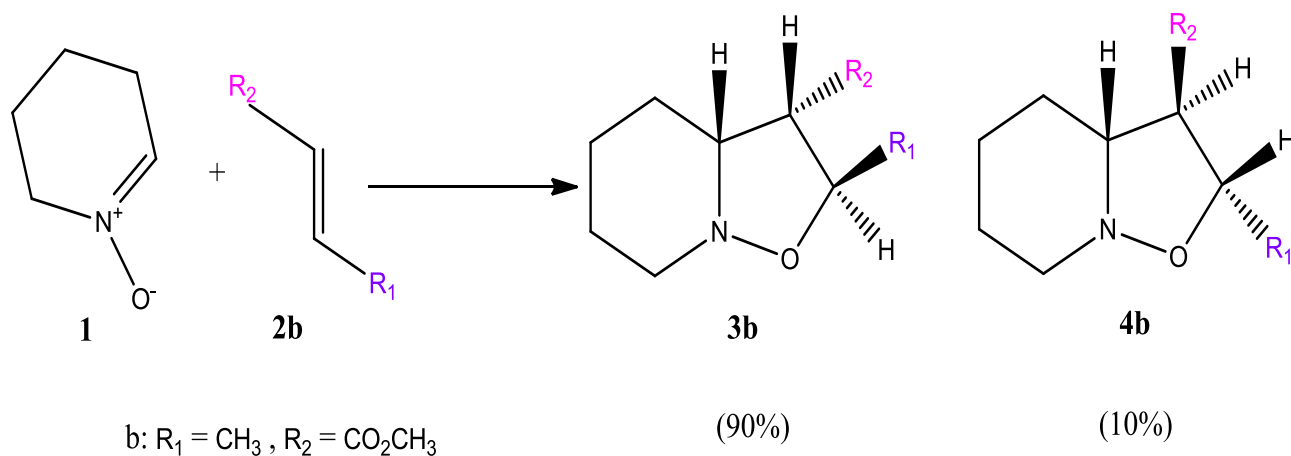
have been devoted to the study of regio- and stereoselectivities of 1,3DC reactions of nitrene with alkenes [11].

Recently, *Liu et al* [12]. have performed DFT calculations at B3LYP/6-31G* level on the 1,3-dipolar cycloaddition reaction of the simplest nitrene to dipolarophiles containing electronreleasing substituents. Here, the endo approach is kinetically favoured because of stabilization of the secondary orbital interactions.

In another study, *Domingo et al* [13,14]. have studied the 1,3-dipolar cycloaddition reaction of nitrenes with several dipolarophiles using DFT methods at the B3LYP/6-31G* and B3LYP/6-31+G* levels. Their calculations predict an asynchronous concerted mechanism and both stereo and regioselectivity were found dependent on the computational model and computational level.

More recently, *Marakchi et al* [15]. have studied the mechanism of the 1,3-dipolar cycloaddition reaction between nitrene and sulfonyl ethene chloride using ab initio and DFT methods. HF and DFT calculations predict meta path, in agreement with the experimental results.

The experimental investigations of *Asrof et al* [16]. Have determined the 1,3-DC reaction between 2,3,4,5-tetrahydropyridine-1 -oxide **1** and methyl crotonate **2b** to giving a mixture of substituted isoxazolidines (see Scheme II.1) **3b** and **4b** with a ratio of 90:10, respectively. The stereochemistry of the major adduct is depicted in **4b** having endo oriented carbomethoxy group which is known to manifest favourable secondary orbital interaction. In what follows, we present a theoretical study on the cycloaddition reaction involving cyclic 5-membered nitrene with methyl crotonate.



Scheme II.1: 1,3-Dipolar cycloaddition of 2,3,4,5 tetrahydropyridine 1-oxide with methylcrotonate

II.2. The 1,3-Dipolar cycloaddition reactions

II.2.1. The 1,3-dipole

A 1,3-dipole is defined as an *a-b-c* structure that undergoes 1,3-DC reactions with multiple bonds [17-19]. Basically, 1,3-dipoles can be divided into two different types: the allyl anion type and the propargyl/allenyl anion type as in Figure II.1.

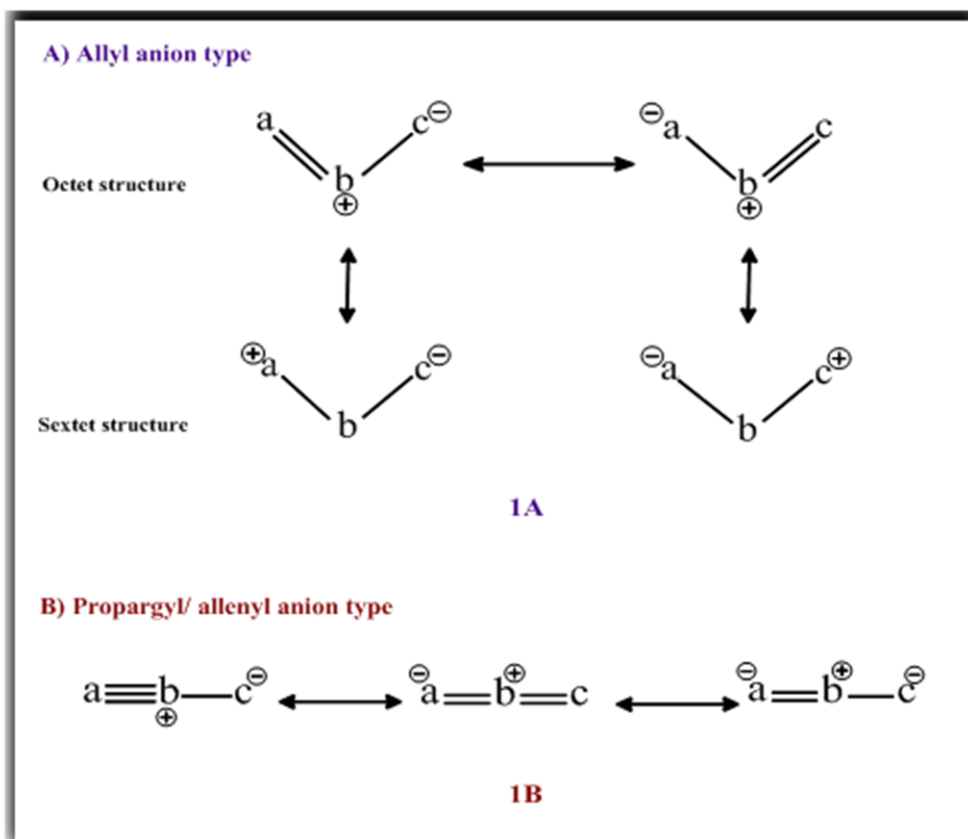


Figure II.1: Classification of the parent 1,3-dipoles

- **The allyl anion type:** is characterized by four electrons in three parallel p_z orbitals that are perpendicular to the plane of the dipole and so that the 1,3-dipole is bent. Two resonance structures in which the three centers have an electron octet, and two in which a or c has an electron sextet, can be drawn. The central atom b can be nitrogen, oxygen, or sulfur (**1A**).
- **The propargyl/allenyl anion type:** has an extra π orbital that is located in the plane orthogonal to the allenyl anion type molecular orbital (MO). The propargyl/allenyl anion type is linear and the central atom b is limited to nitrogen (**1B**). The classification of the parent 1,3-dipoles is depicted in Table II.1 [17].

Table II.1: Classification of the parent 1,3-dipoles

<u>Allyl anion type</u>			
Nitrogen in the middle		Oxygen in the middle	
	Nitrones		Carbonyl Ylides
	Azomethine Imines		Carbonyl Imines
	Azomethine Ylides		Carbonyl Oxides
	Azimines		Nitrosimines
	Azoxy Compounds		Nitrosoxides
	Nitro Compounds		Ozone
<u>Propargyl/ allenyl anion type</u>			
Nitrillium Betaines		Diazonium Betaines	
	Nitrile Oxides		Diazoalkanes
	Nitrile Imines		Azides
	Nitrile Ylides		Nitrous Oxides

II.2.2. The dipolarophile

The dipolarophile in a 1,3-dipolar cycloaddition is a reactive alkene moiety containing 2π electrons. Thus, depending on which dipole that is present, α,β -Unsaturated aldehydes, ketones, and esters, allylic alcohols, allylic halides, vinylic ethers *etc.* are examples of dipolarophiles that can readily react with 1,3-dipoles (Figure II.2). In addition to these compounds, molecules with a double bond such as carbonyls, thiocarbonyls, imines *etc.* can also undergo cycloaddition with 1,3-dipoles [20].

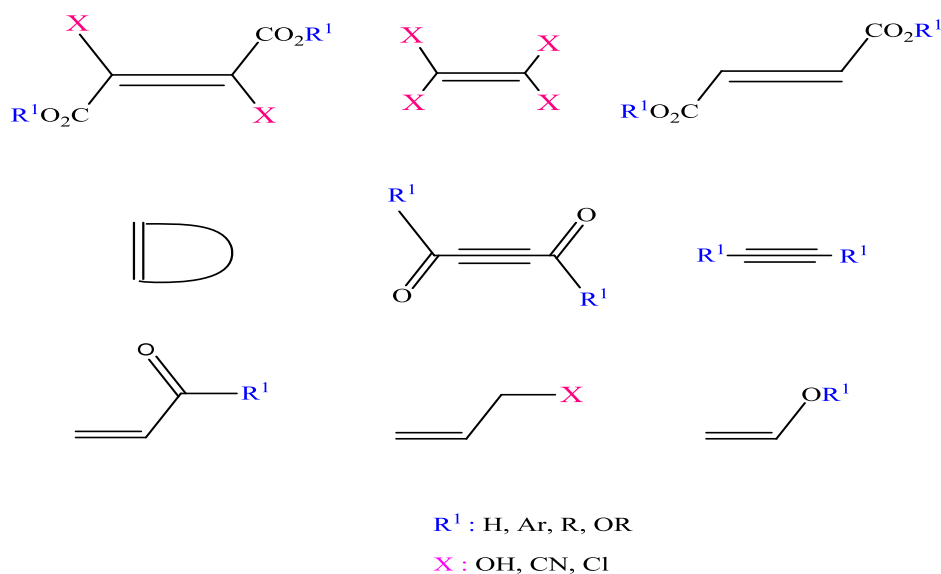
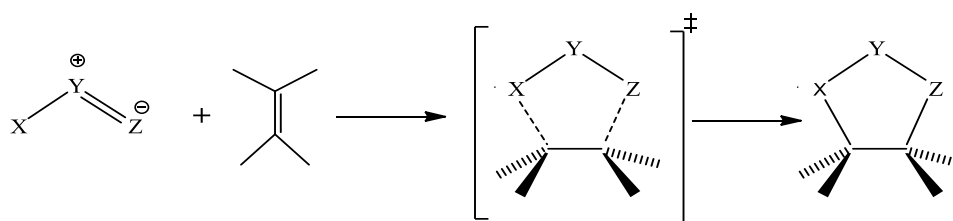


Figure II.2: Examples of dipolarophiles in 1,3-dipolar cycloaddition reactions

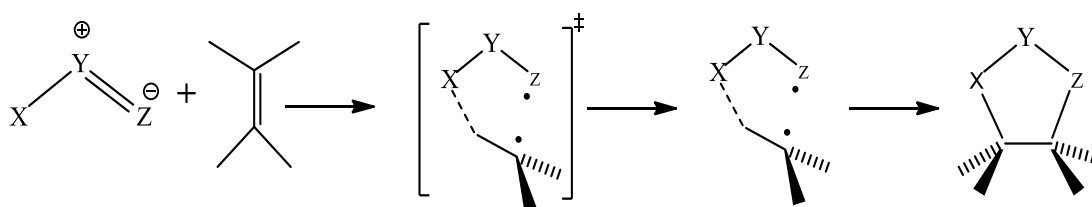
II.3. Mechanistic aspects

The mechanism of the 1,3-DC reaction is, the subject of a great controversy in the experimental as well as theoretical aspects. Two extreme theories were proposed for this reaction:

- One is the concerted synchronous mechanism with simultaneous formation of the two new sigma bonds in the transition state, suggested by Huisgen [21,22]; (Scheme II.2).
- Two is the stepwise mechanism containing two continuous processes, through which a diradical transition state is present, forward by Firestone [23–25]; (Scheme II.3).



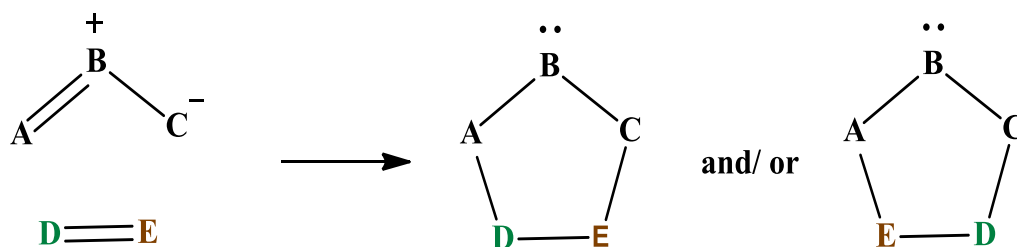
Scheme II.2. Concerted mechanism for DCR



Scheme II.3: Stepwise diradical mechanism for DCR

II.4. Regioselectivity [26]

A 1,3-dipolar cycloaddition can potentially produce a mixture of products. Regiochemistry is concerned with the way in which the dipolarophile adds onto the 1,3-dipole. In the case of an unsymmetrical dipolarophile, there may be two regioisomers.

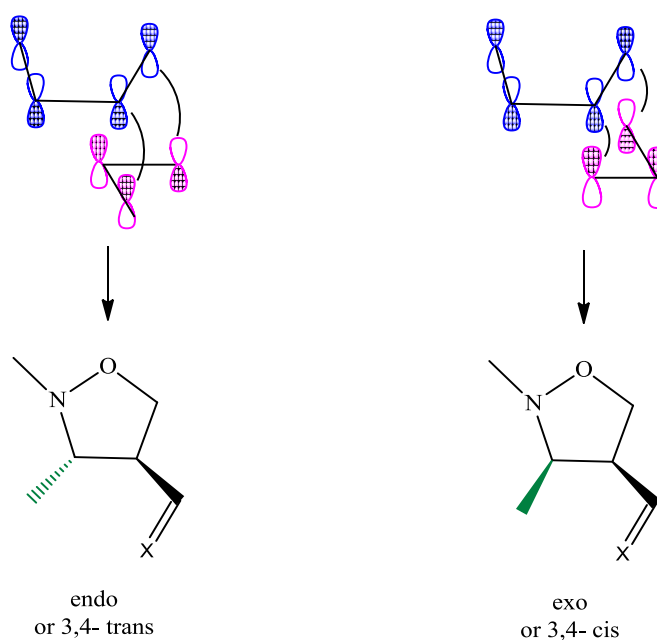


However, most 1,3-dipolar cycloadditions are highly selective and give only one of the two possible regioisomers. Two key pieces of information are required in order to determine the preferred product of a 1,3-dipolar cycloaddition.

- The dominant HOMO-LUMO interaction between the two reactants
- The orbital coefficients of the reactants

II.5. Stereoselectivity

Most of the 1,3-dipolar cycloaddition reactions are highly stereoselective and this may be due to the concerted nature of the reaction. When two chiral centres are formed, one arising from the dipole and the other from the dipolarophile, diastereomeric products (cis- and trans-) may be produced via endo and exo transition states. The predominant formation of each diastereomer depends on attractive p-orbital overlap of unsaturated substituents (favouring an endo transition state) and repulsive vanderWaals steric interactions (favouring an exo transition state) [27].



This we demonstrate in section below. The reaction of 2,3,4,5-tetrahydropyridine-1-oxide with methyl crotonate.

II.6. Theoretical analysis of 1,3-dipolar cycloaddition reactions

II.6.1. Frontier molecular orbital theory and Sustmann classification

Frontier molecular orbital theory is a powerful tool for understanding the different reactivities of 1,3-dipoles with different dipolarophiles. In FMO theory [28], only the highest occupied molecular orbitals (HOMO) and the lowest unoccupied molecular orbitals (LUMO) of both reactants are considered during bond formation. The theory can only be applied to a reaction with a concerted mechanism which proceeds through a single transition state. No intermediates are formed and the symmetry of the frontier orbitals is conserved in the change to products. Any other orbital interactions between the occupied and unoccupied orbitals of the reactants are ignored. Generally, the HOMO of one reactant reacts with the LUMO of the other. The favoured interaction occurs between those orbitals with the lowest HOMO - LUMO energy gap.

The actual energy of the frontier orbitals is determined by both the skeletal atoms and the substituents of the molecule in question. Electron withdrawing groups such as ketones and esters lower the energies of molecular orbitals. Electron donating groups, on the other hand, raise the energies of the molecule's orbitals. Different substituents can therefore play a significant role in the outcome of a reaction.

Sustmann [29, 30] has classified 1,3-dipolar cycloadditions into three categories based on the relative FMO energies of the dipole and the dipolarophile (Figure II.3).

Type I: HOMOdipole-LUMOdipolarophile interaction is dominant.

(normal electron demand)

Type II: Both HOMO-LUMO gaps are approximately equal.

(neutral electron demand)

Type III: LUMOdipole-HOMOdipolarophile interaction is dominant.

(inverse electron demand)

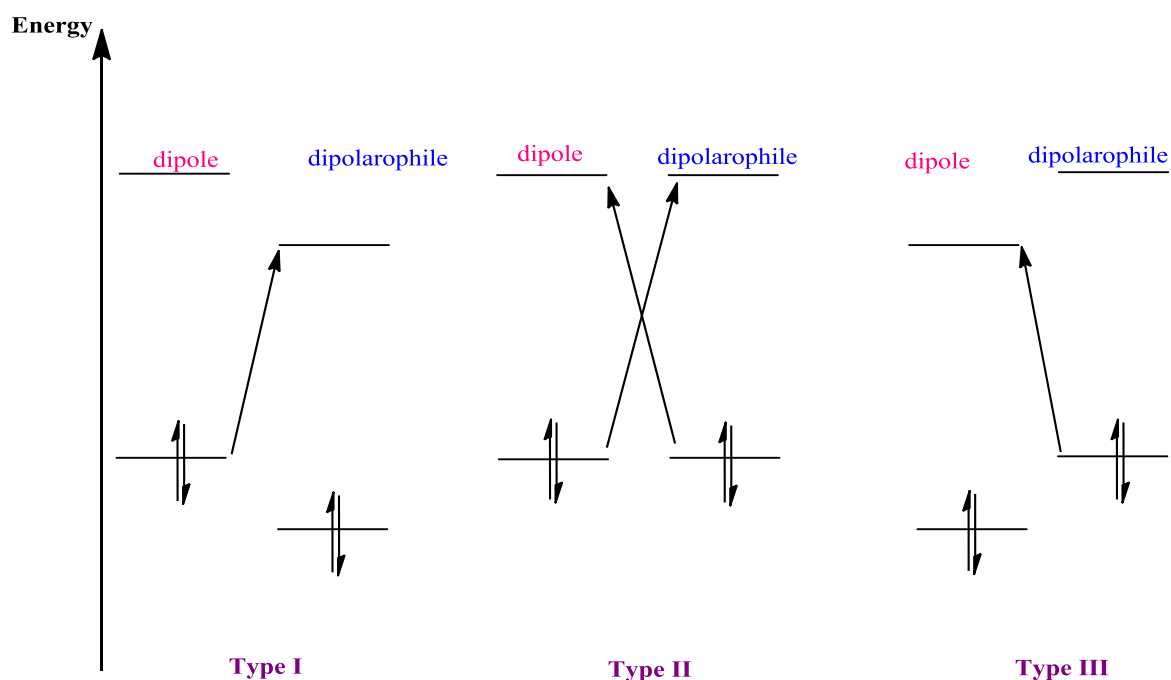


Figure II.3: Sustmann classification of 1,3-dipolar cycloadditions

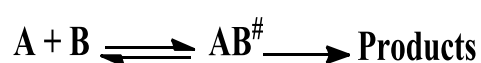
II.6.2. Transition state theory (TST)

At the beginning of the past century, some important theories based on experiments, such as the TST [31], were developed in kinetic chemistry, which permitted to establish fundamental concepts used in the study of molecular mechanisms. Within this theory, the concept of the activation complex or TS enabled the establishment of a relationship between the experimental activation energy [32] and the energy of the TS associated to an organic reaction. The goal of TST providing a computational tool for predicting rate constants from a knowledge of the potential energy surface (PES) controlling the reaction. However, this goal was not attained until much later because theoretical chemists were unable to calculate the PES with sufficient accuracy. Thus, for many years, TST was used primarily for correlating and interpreting rate constants, including kinetic isotope effects, and that aspect of its use continues [33].

The Eyring equation [34] of transition state theory has been applied for determination of rate constants for elementary bimolecular reactions in the gas phase. Following the theory of reactions rates, the rate constant for the reaction between A and B is given by

$$K = \frac{k_B T}{h} \frac{Q_{\#}}{Q_A Q_B} e^{-\frac{E_a}{RT}} \quad (1)$$

in which Q represents the total partition function, for the concerned species.



The bimolecular reaction is depicted in the activation energy diagram in (Figure II.4).

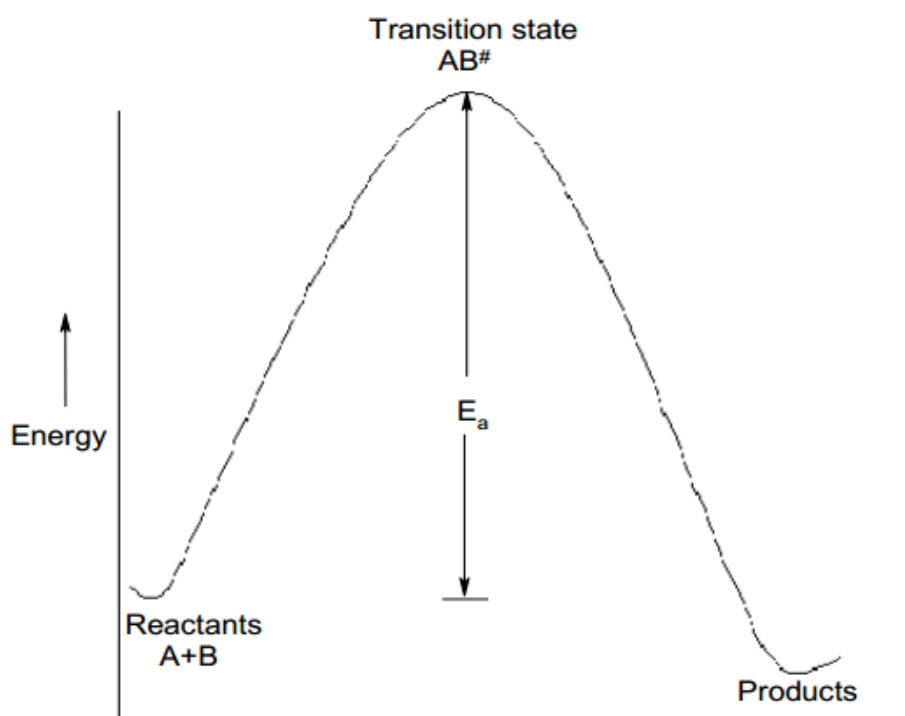


Figure II.4: Reaction co-ordinate diagram

II.6.3. Chemical Reactivity Indexes

Parallel to the development of quantum chemical models to approach the Hohenberg-Kohn, Parr developed the so-called “CDFT” in the late 1970s and early 1980s [35]. CDFT is a DFT- subfield in which relevant concepts and principles are extracted from electron density, making it possible to understand and predict the chemical behaviour of a molecule. Parr and co-workers, and later a large community of theoretical chemists, were able to give precise definitions for chemical concepts which had already been known and used for many years in various branches of chemistry, thus providing their calculations with a quantitative use.

II.6.3.1. Global properties

The main global reactivity indices defined within CDFT are the electronic chemical potential μ and the chemical hardness η , from which the electrophilicity ω index was further obtained by combination of the two former. Then, several proposals for a nucleophilicity N index were developed apart

II.6.3.1.1. Electronic chemical potential

In 1983, Parr defined the electronic chemical potential μ as the energy changes of the system with respect to the electron number N at a fixed external potential $v(\mathbf{r})$, i.e. the potential created by the nuclei (see Eq. [2]) [36]. The electronic chemical potential μ is associated with the feasibility of a system to exchange electron density with the environment at the ground state.

$$\mu = \left(\frac{\partial E}{\partial N} \right)_{v(r)} \quad (2)$$

Applying the finite difference approximation, the following simple expression is obtained:

$$\mu \approx - \frac{I+A}{2} \quad (3)$$

where I and A are the ionisation potential and the electron affinity of an atom or molecule, respectively. Although a large number of experimental I values for organic molecules can be obtained, a very small number of experimental A values can be found. Using Koopmans' theorem [37] and Kohn–Sham's formalism [38] within DFT, these energies can be approached by the frontier HOMO and LUMO energies as I by $-E_{\text{HOMO}}$ and A by E_{LUMO} . Consequently, the electronic chemical potential μ can be expressed as:

$$\mu \approx \frac{E_{\text{HOMO}} + E_{\text{LUMO}}}{2} \quad (4)$$

Note that this relationship comprises only the FMO energies but not additional physical significance and, therefore, the use of the HOMO and LUMO energies of a molecule to approach the I and A values in the CDFT is completely different to the analysis of the HOMO–LUMO interactions between two molecules used in the FMO theory.

II.6.3.1.2. Chemical hardness

In 1963, Pearson established a classification of Lewis acids and bases into hard and soft [39]. He proposed that in an acid/base reaction, the most favourable interactions take place between hard/hard or soft/soft pairs, the HSAB principle. Within CDFT, Parr defined, in 1983, a quantitative expression for the chemical hardness η , which can be expressed as the changes of the electronic chemical potential μ of the system with respect to the electron number N at a fixed external potential $v(\mathbf{r})$ (see Eq. [5]) [36]. Chemical hardness η can be thought as the resistance of a molecule to exchange electron density with the environment.

$$\eta = \left(\frac{\partial \mu}{\partial N} \right)_{v(r)} = \left(\frac{\partial^2 E}{\partial N^2} \right)_{v(r)} \quad (5)$$

Applying the finite difference approximation, the following simple expression is obtained:

$$\eta \approx \frac{I - A}{2} \quad (6)$$

which by substitution of I by $-E_{HOMO}$ and A by $-E_{LUMO}$ can be expressed as:

$$\eta \approx \frac{E_{LUMO} - E_{HOMO}}{2} \quad (7)$$

Usually, the term $1/2$ is neglected, so that chemical hardness η is expressed as:

$$\eta \approx E_{LUMO} - E_{HOMO} \quad (8)$$

On the other hand, chemical softness S was introduced as the inverse of chemical hardness η (see Eq. [9])

$$S = \frac{1}{\eta} \quad (9)$$

II.6.3.1.3. Electrophilicity

In 1999, Parr defined the electrophilicity ω index [40], which gives a measure of the energy stabilisation of a molecule when it acquires an additional amount of electron density from the environment. Thus, a good electrophile can be characterized by a high value of μ and a low value of η [41]. (see Eq. [10]).

$$\omega = \frac{\mu^2}{2\eta} \quad (10)$$

The electrophilicity ω index has become a powerful tool for the study of the reactivity of organic molecules participating in polar reactions [42].

II.6.3.1.4. Nucleophilicity

The global nucleophilicity (N) model has been recently introduced, based on the relationship $N = -IP$, where IP is the gas phase (intrinsic) ionization potential and can be straightforwardly extended to describe the local nucleophilicity. Following those

derivations concerning the local electrophilicity and philicities, it is taken that the global nucleophilicity index (N) can be written as,

$$N = \sum_k N_k \quad (11)$$

The nucleophilicity index N for a given system was therefore defined as,

$$N = E_{\text{HOMO}(\text{nucleophile})} - E_{\text{HOMO}(\text{TCE})} \quad (12)$$

Where, E_{HOMO} is the HOMO energy the nucleophile molecule in eV units for the reactants we want to study and $E_{\text{HOMO}}(\text{TCE})$ is the HOMO energy of the reference tetracyanoethylene molecule in eV unit. The reference TCE is taken due to the lowest HOMO energy allowing us to have a positive global nucleophilicity scale. With the simplest approximation to nucleophilicity, the IP values can be approximated in terms of the HOMO energy in a molecule within a given molecular orbital (MO) scheme.

II.6.3.2. Local properties

Another index that directly comes from CDFT is the electronic Fukui function $f(r)$. In 1984, Parr proposed the $f(r)$ function [43], named frontier function or Fukui function, for a molecule, which represents the changes in electron density at a point r with respect to the variation of the number of electrons N at a fixed external potential $v(r)$ (see Eq. [13]).

$$f(r) = \left(\frac{\partial \rho(r)}{\partial N} \right)_{v(r)} \quad (13)$$

The function $f(r)$ reflects the ability of a molecular site to accept or donate electrons. High values of $f(r)$ are related to a high reactivity at point r [44]. By applying a finite difference approximation to Equation (13), two definitions of Fukui functions depending on total electronic densities are obtained:

$$f(\mathbf{r})^+ = \rho_{N+1}(\mathbf{r}) - \rho_N(\mathbf{r}) \quad \text{for nucleophilic attacks} \quad (14)$$

and

$$f(\mathbf{r})^- = \rho_N(\mathbf{r}) - \rho_{N-1}(\mathbf{r}) \quad \text{for electrophilic attacks} \quad (15)$$

where $\rho_{N+1}(\mathbf{r})$, $\rho_N(\mathbf{r})$, and $\rho_{N-1}(\mathbf{r})$ are the electronic densities at point \mathbf{r} for the system with $N + 1$, N , and $N - 1$ electrons, respectively.

In this sense, the regioselectivity analysed within the FMO theory and that using Fukui functions based on FMOs is numerically identical, but conceptually completely different; while FMO theory establishes the most favourable MO overlap, CDFT establishes the most favourable nucleophilic/electrophilic interaction in a polar reaction, which depends on the total molecular electron density but not on any specific MO. However, these mathematical expressions did not exactly match with the theoretical concept of the electronic Fukui functions as derived from the CDFT.

A great deal of theoretical work devoted to the study of molecular mechanisms of polar reactions involving non-symmetric reagents has shown that the most favourable reactive channel is that associated with the two-center interaction between the most electrophilic center of the electrophile and the most nucleophilic center of the nucleophile [45]. Thus, the local electrophilicity ω_k and local nucleophilicity N_k (see Eq. [16] and [17]), which permit the distribution of the global electrophilicity ω and nucleophilicity N indices at the atomic sites k through the electrophilic P_K^+ and nucleophilic P_K^- Parr functions, respectively, enable the characterisation of the most electrophilic and nucleophilic centers in the molecule, and thus, the prediction of the regio- and chemo- selectivities in polar reactions.

$$\omega_k = \omega \cdot P_K^+ \quad (16)$$

$$N_k = N \cdot P_K^- \quad (17)$$

II.7. Materials and methods

The reported quantum chemical calculations were performed at the B3LYP/6-31G(d) level of theory using GAUSSIAN 09 suite of programs [46]. Full geometry optimizations followed by frequency calculations at the same level of theory were carried out for all stationary points. The intrinsic reactions coordinate (IRC) [47] path was calculated in order to check that each TS connects well to the two corresponding minima in the energy profiles of the proposed mechanism. Atomic electronic populations were computed using the natural bond orbital (NBO) method [48]. Bulk solvent effects of Dichloromethane were considered implicitly by performing single point energy calculations on the gas phase stationary structures using the polarisable continuum model (PCM) as developed by Tomasi's group [49] on the basis of the self-consistent reaction field (SCRF) background [50]. Values of enthalpy and free energy were obtained by frequency calculations over B3LYP/6-31G(d) geometries.

II.8. Results and Discussion

II.8.1. Regiochemistry study based on FMOs and reactivity indices

The determination of NED (Normal Electron Demand) character or IED (Inverse Electron Demand) character of the cycloaddition reaction is necessary for the prediction of regioselectivity. This characterization can be performed by using electronic chemical potential (μ), global electrophilicity (ω), global nucleophilicity (N) and HOMO–LUMO gap. This latter is calculated by considering the directly involved orbitals in the reaction. Frontier molecular orbitals (Figure II.5) analysis shows that HOMO and LUMO of the nitron are π molecular orbitals (MOs). However, in the methyl crotonate, the HOMO is a nonbonding MO localized essentially on the oxygen of the carbonyl group and, in consequence, it is expected that it will not be directly involved in the 1,3-DC process. The HOMO-1 of dipolarophile is a bonding π molecular orbital with a large contribution than the HOMO ones. Thus the relative reactivity can be explained with analysis of this MO.

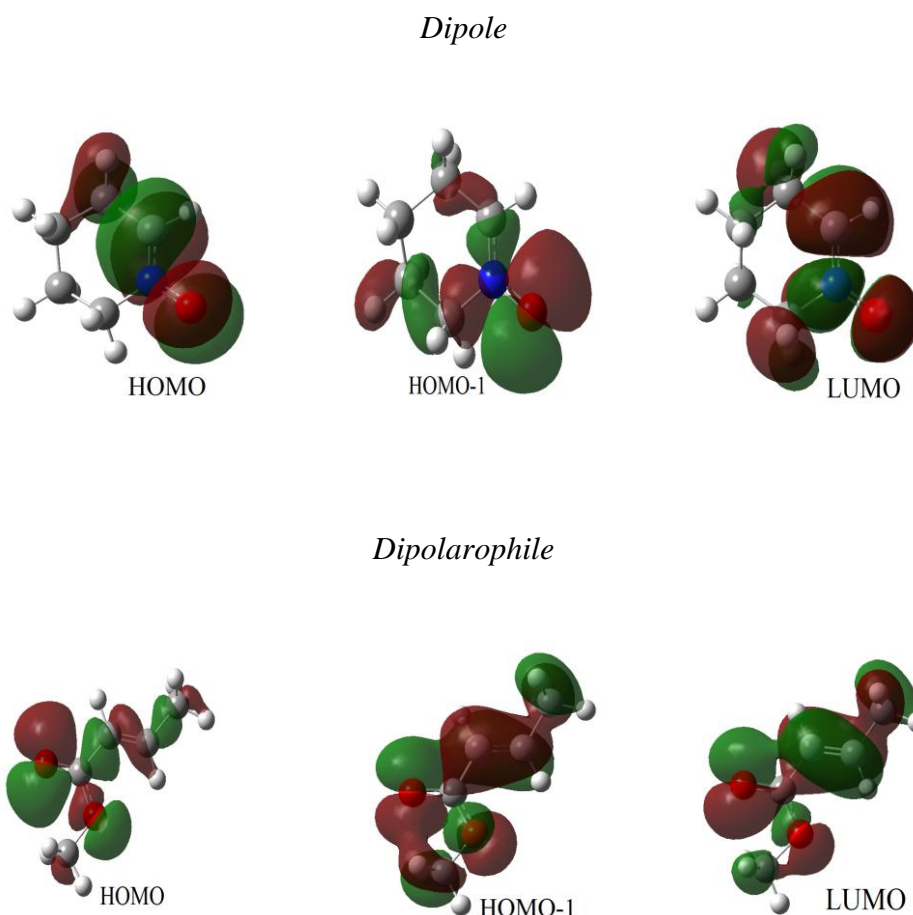


Figure II.5: Optimized geometries and visualized FMOs for the reactants

For a better visualization of the FMO approach, we have presented the two possible interactions:

- interaction between the highest occupied molecular orbital of the 1, 3-dipole (HOMO_{dipole}) and the lowest unoccupied molecular orbital of the dipolarophile (LUMO_{dipolarophile}) that is corresponding the Normal Electron Demand (NED) : $\Delta E(I) = \text{HOMO}_{\text{dipole}} - \text{LUMO}_{\text{dipolarophile}}$ (4.406 eV) a large number of 1,3-DC reactions is classified in.

- Interaction between the LUMO of dipole and the HOMO-1 of dipolarophile. This type is named as the inverse-electron demand (IED): ΔE (II) = HOMO-1 dipolarophile-LUMOdipole (7.072 eV). These types of interactions are illustrated in (Figure II. 6).

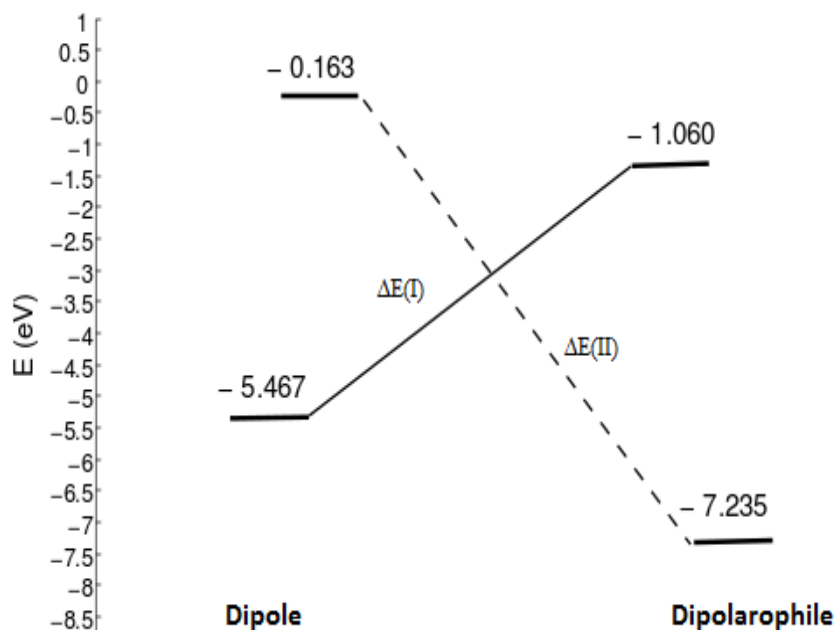


Figure II. 6: The interactions between HOMO and LUMO orbitals of a 1, 3-dipole/dipolarophile

The values of ΔE (HOMO-LUMO gap) for NED character indicating lower than that corresponding to IED. This agrees with Sustmann's type I reaction [30]. The regioselectivity of 1,3-DC reactions can be explained on the basis of large-large and small-small FMO interactions that are more favoured than large-small and small-large ones. The coefficients of frontier molecular orbitals of the reactants, given in Table II.2, show that the most favoured interactions are between O7 of the dipole with C1 of the dipolarophile and C1 of dipole interact with C2 of dipolarophile leading to the formation of meta regioisomer as the major product.

Table II.2: FMOs Molecular Coefficients of the dipole and dipolarophile

Dipole				Dipolarophile			
HOMO		LUMO		HOMO-1		LUMO	
C1	O7	C1	O7	C1	C2	C1	C2
-0.3898	0.4677	0.3877	0.3021	0.3449	0.3850	0.4025	-0.2447

In Table II.3 are displayed the HOMO and LUMO energies, electronic chemical potential μ , chemical hardness η , global electrophilicity ω , and global nucleophilicity N of the dipole and dipolarophile.

Table II.3: FMO energies (a.u), electronic chemical potential (a.u), chemical hardness (a.u), electrophilicity index (eV) and nucleophilicity index (eV).

Reactant	HOMO	LUMO	μ	η	ω	N^a
Dipole	-0.201	-0.006	-0.103	0.195	0.740	3.649
Dipolarophile	-0.266	-0.039	-0.152	0.227	1.384	1.881

-The HOMO energy of tetracyanoethylene is -0.3351 a.u at the same level of theory.

-Chemical potential, hardness, electrophilicity and nucleophilicity values are associated to the HOMO-1 of dipolarophile

The electronic chemical potential (μ) of dipole (-0.103 eV) is greater than that of dipolarophile (-0.152 eV). Consequently, the charge transfer will take place from the 1, 3-dipole to dipolarophile in agreement with the FMO energy predictions. The Values of electrophilicity indices (ω) of reactants are 1.384 and 0.740 eV for methyl crotonate (dipolarophile) and nitron (dipole), respectively. Thus, nitron will act as a nucleophile whereas methyl crotonate act as an electrophile. Many studies dealing with cycloaddition reactions are based on the electrophilic (f_k^+) and nucleophilic (f_k^-), Parr functions attacks. Negative values of Fukui functions can be obtained from various population analyses. However, Hirshfeld's population [51] guarantees positive Fukui functions values, thus it is a good means to predict the right regioselectivity [52]. Therefore, Fukui functions based on the Hirshfeld's population were calculated for the studied reactants and the corresponding results are given in Table II.4. Prediction of regioselectivity can also be performed by using the Chattaraj's polar model where local philicity indices are used [53]. Values of local electrophilicity ω_k and local nucleophilicity N_k are listed in Table II.4 and the corresponding most favourable two-centre interaction is shown in Figure II.7.

In dipolar cycloaddition, the most favourable attack takes place between C1 atom of the dipolarophile (the preferred position for a nucleophilic attack) and O7 of the dipole, leading to the formation of the meta-regioisomer. Which is in good agreement with the experimental data.

Table II.4: Local properties of dipole and dipolarophiles calculated at B3LYP/6-31G (d) level of theory

Reactant	Site	NPA				Chelpg			
		f^+	f^-	ω	N	f^+	f^-	ω	N
Dipole	O7	0.189	0.423	0.140	1.543	0.211	0.355	0.156	1.295
	C1	0.298	0.312	0.220	1.138	0.435	0.268	0.180	0.977
Dipolarophile	C1	0.225	0.246	0.300	0.469	0.234	0.268	0.311	0.511
	C2	0.116	0.392	0.154	0.747	0.114	0.397	0.151	0.288

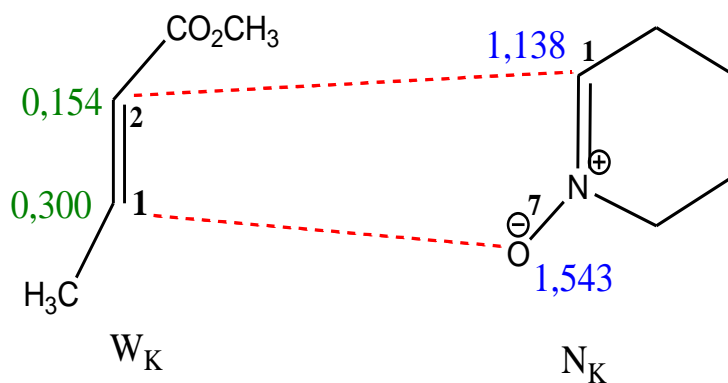
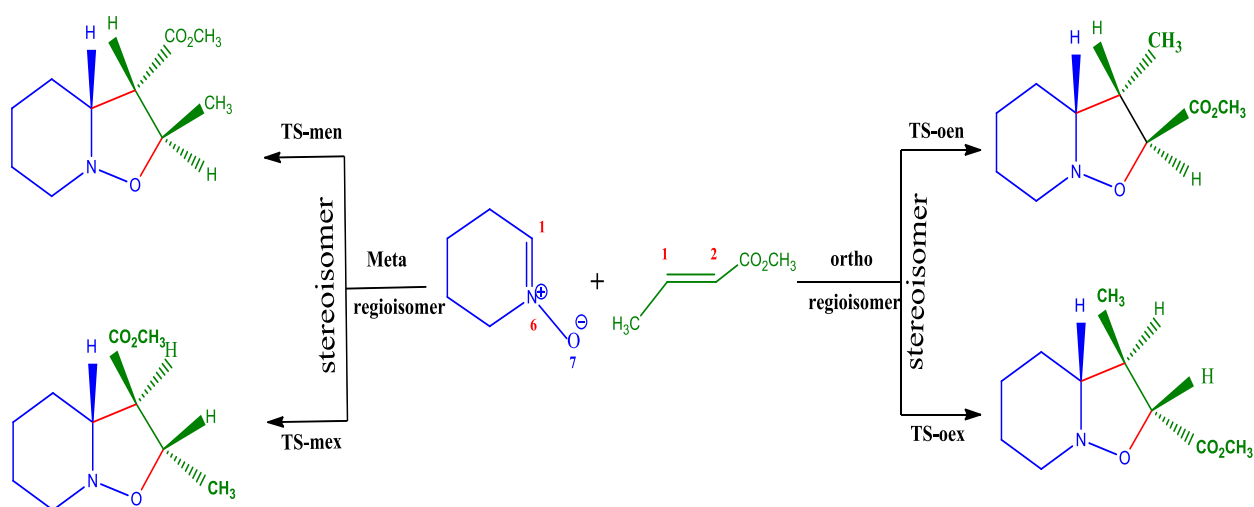


Figure II.7: Prediction of the favoured interactions between dipole and dipolarophile using DFT based indices

II.8.2. Mechanistic study of the cycloaddition reaction based on activation energy

II.8.2.1. Energies of the Transition Structures

The 1-3 DC reactions of nitron with dipolarophile can happen along four possible reactive channels corresponding to the endo and exo approach modes in two different regioisomeric channels; the meta and ortho reactions (see Scheme II.4). The right pathway corresponds to the O7-C2 and C1-C1 forming bond processes, while the left pathway corresponds to the O7-C1 and C1-C2 ones. Four transition states designated as: TS-men, TS-mex, TS-oen, and TS-oex have been confirmed by frequency calculations. The geometries of the four TSs are represented in Figure II.8. The studied energies and relative energies of reactants in both the gas phase and DCM solvent are given in Table II.5. The PES schemes, corresponding to the four reactive channels, are given in Figure (II.9 and II.10).



Scheme II.4: The exo and endo approaches of tetrahydropyridine -1 -oxide to methyl crotonate

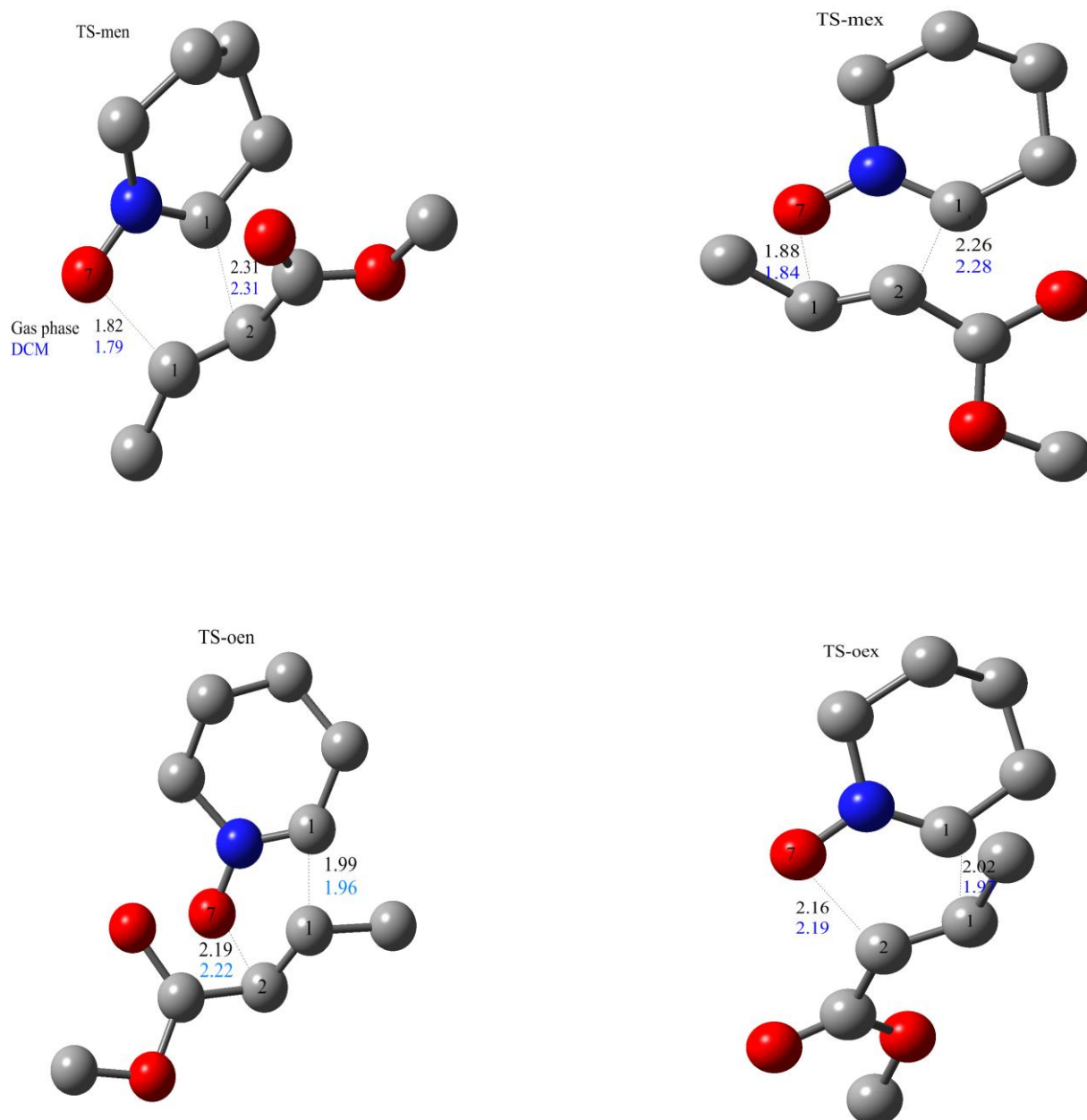


Figure II.8: Optimized transition structures of the 1-3 DC reaction between tetrahydropyridine-1-oxide and methyl crotonate

The calculated activation energies and relative electronic energies of the stationary points involved in the 1,3-DC of 2,3,4,5-tetrahydropyridine-1-oxide with methylcrotonate in both

cases the gas-phase and dichloromethane solvent are shown in Table II.5. The activation energy **TS-men** (6.45 kcal/mol) is lower compared to that of the **TS-mex** (10.91 kcal/mol), **TS-oen** (11.25 kcal/mol) and **TS-oex** (15.47 kcal/mol), In dichloromethane, the reactants are slightly more stabilized than the TSs and this results the activation energies increase respectively **TS-men** (10.45), **TS-mex** (13.80), **TS-oen** (14.73) and **TS-oex** (17.43) kcal/mol, that is to say, the formation of pathways **P-men** in the gas phase is easier than that in the solvent. We have to note that these 1,3-DC reactions are strongly exothermic. In addition, solvent effects decreased their exothermicity while their corresponding activation energies increased as a consequence of the greater solvation of polar nitron [54]. The comparison of relative energies with relative free energies given in Table II.5 shows that the trends in region- and stereoselectivity are essentially coincident. Finally, we can conclude that the energy results indicate that the **P-men** shows a very high reactivity, both kinetically and thermodynamically.

Table II.5: Total energies (a.u), Relative energies ΔE (in kcal/mol), relative free energies ΔG (in kcal/mol) and enthalpies ΔH (in kcal/mol) in gas phase and in DCM, of the stationary points involved in the 1, 3-DC reaction between 2,3,4,5-tetrahydropyridine-1-oxide and methyl crotonate

Stationary point	Gas phase				Dichloromethane solvent			
	ΔG	ΔH	ΔE^*	E_T	ΔG	ΔH	ΔE^*	E_T
Dipole				-325.8570278				-325.8656038
Dipolarophile				-345.7852858				-345.7903845
TS-men	21.75	8.05	6.45	-671.6320213	24.98	11.88	10.45	-671.6393249
TS-mex	25.71	12.38	10.91	-671.6249129	28.56	15.27	13.80	-671.6339965
P-men	-6.16	-20.20	-23.62	-671.6742828	-1.80	-15.70	-19.04	-671.6813970
P-mex	-2.72	-16.07	-19.56	-671.6734902	1.17	-12.20	-15.75	-671.6810906
TS-oex	30.52	16.97	15.47	-671.6176606	32.56	19.00	17.43	-671.6282077
TS-oen	26.25	12.76	11.25	-671.6243785	29.81	16.28	14.73	-671.6325104
P-oen	-3.87	-17.80	-15.56	-671.6671208	0.47	-13.70	-12.07	-671.6752383
P-oex	-2.69	-16.60	-13.77	-671.6642700	1.10	-12.70	-10.90	-671.6733701

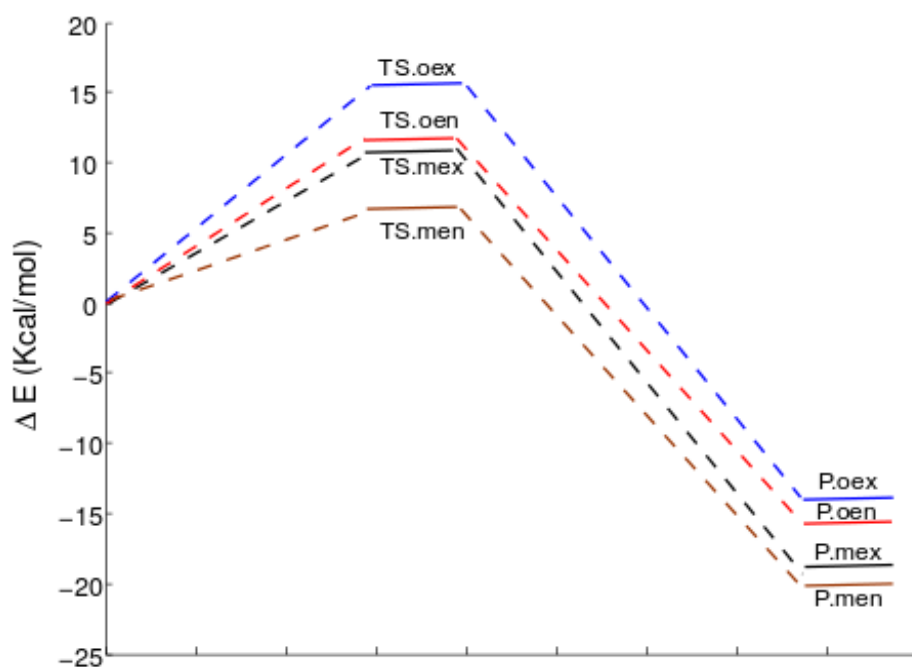


Figure II.9: Energy profiles, in kcal/mol, for the 1,3-DC reactions of stationary point in gas phase

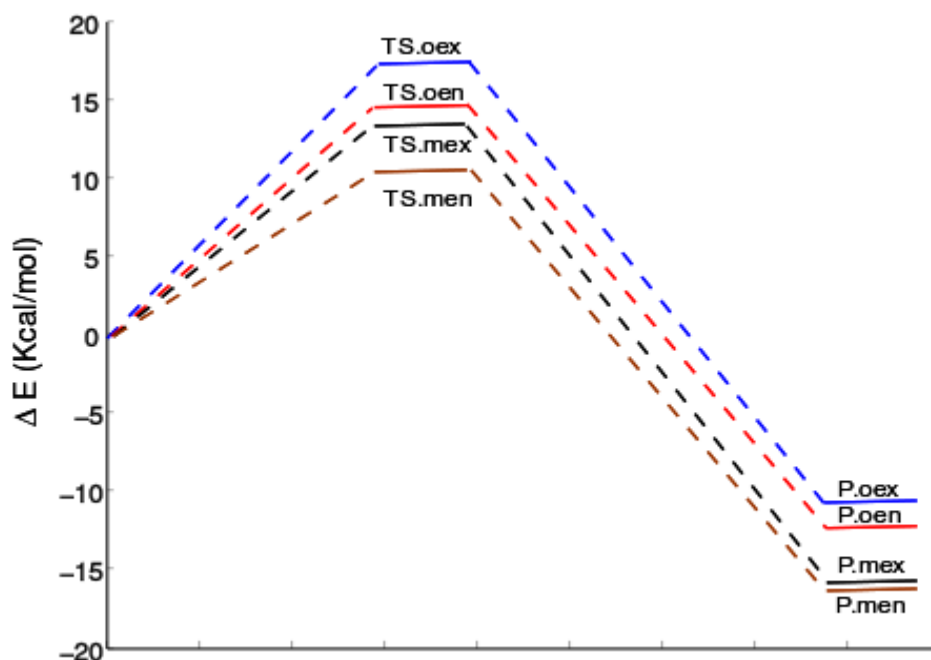


Figure II.10. Energy profiles, in kcal/mol, for the 1,3-DC reactions of stationary point in solvent phase

II.8.2.2. Geometry analysis

Table II.6: Values of $|\Delta d|$ in TS-men, TS-mex, TS-oen and TS-oex of the 1, 3-DC reaction of dipolarophile with dipole in the gas phase and solvent DCM

	Gas phase			DCM		
	Meta channels			Meta channels		
	d(O ₇ -C ₁)	d(C ₁ -C ₂)	Δd	d(O ₇ -C ₁)	d(C ₁ -C ₂)	Δd
TS-men	1.82	2.31	0.48	1.79	2.31	0.52
TS-mex	1.88	2.26	0.37	1.84	2.28	0.44
	Ortho channels			Ortho channels		
	d(O ₇ -C ₂)	d(C ₁ -C ₁)	Δd	d(O ₇ -C ₂)	d(C ₁ -C ₁)	Δd
TS-oen	2.19	1.99	0.20	2.22	1.96	0.26
TS-oex	2.16	2.02	0.14	2.19	1.97	0.40

Comparison of the most relevant geometrical parameters of the four TSs involved in gas phase and in solvent 13DCs of the 2,3,4,5-tetrahydropyridine-1-oxide with methyl crotonate is presented in Figure II.8. The corresponding selected geometric parameters are given in Table II.6. The lengths of the bonds of the C₁-C₂ are: (2.31 – 2.62) Å and O₇-C₁ (1.82 – 1.88) Å at the **TS-men** and **TS-mex**, respectively, while at the **TS-men** and **TS-mex** involved in the DCM are: C₁-C₂ (2.31-2.28) Å and O₇-C₁ (1.79 – 1.84) Å. These values indicate that they correspond to asynchronous bond formation processes where the lengths of the O-C are shorter than the C-C bonds. However, in the **TS-oen** and **TS-oex**, O-C forming bonds (2.19 and 2.16 Å) are longer than C-C forming bonds (1.99 and 2.02 Å); the lengths of the bonds in the solvent are the same as in the gas phase i.e. O-C is longer than C-C. This shows a change of the dissymmetry on the bond formation process for the two regioisomeric pathways. The degree of asynchronicity of bond formation at the TSs is determined by considering the difference between the lengths of the two new σ forming bonds such as $\Delta d = [d_1 - d_2]$ for four pathways. Values of Δd for the **TS-men** are high asynchronous and more favorable stereoisomeric with respect to the other channels.

II.9. Conclusion

In this work, the 1,3-DC reaction of 2,3,4,5-tetrahydropyridine-1-oxide with methyl crotonate, the regio- and stereoselectivities have been thoroughly probed using DFT methods at the B3LYP/6-31G(d) theoretical level. The ortho/meta regioisomeric pathways along with the endo and exo stereoisomeric channels have been studied on the basis of both kinetic and thermodynamic controls. The regioselectivity has been analyzed and confirmed through DFT based indices. The calculated electrophilic, f_k^+ , and nucleophilic, f_k^- , indices prove clearly that the meta channels are the major and most favourable regioisomeric paths. In all the studied cases, the reaction pathways leading to the endo/meta are the most favourable. These 1,3-DCs have basically asynchronous concerted mechanisms and not one presents a stepwise mechanism as a consequence there was no intermediate localization along the present study. The solvent slightly affects the activation energies due to the better solvation of the reactants. All in all, the obtained results using the present theoretical approaches are in good agreement with experimental data.

II.10. References

- [1] W. Carruthers. In *Some Modern Methods of organic Synthesis*; 2nd ed.; Cambridge University Press: Cambridge (1978); W. Carruthers. In *Cycloaddition Reactions in Organic Synthesis*; ed.; J.E. Baldwin and P.D. Magnus; Pergamon: Oxford (1990).
- [2] T. K. Das, A Quantum theoretical study of 1,3- dipolar cycloaddition reactions ; Doctoral thesis: Univ. of Burdwan; (2012).
- [3] S. Bouacha, A. H. Djerourou, *Int. J. Chem. Model.*, **4**, 515 (2013).
- [4] R. Huisgen, R. Grashey, J. Sauer, *The Chemistry of Alkenes*, Interscience: New York (1964).
- [5] K. A. Kumar, *Int.J.ChemTech Res.*, **6**, 3033 (2013).
- [6] T. K. Das, S. Salampuria, M. Banerjee, *J. Mol. Struct. Theochem.*, **959**, 22 (2010).
- [7] A. Padwa, In *Comprehensive Organic Chemistry*; Pergamon Press: Oxford, 1069 (1991).
- [8] P. Ding, M. Miller, Y. Chen, P. Helquist, A. J. Oliver, O. Wiest, *Org. Lett.*, **6**, 1805 (2004).
- [9] K. Marakchi, R. Ghailane, O. Kabbaj, N. Komaha, *J. Chem. Sci.*, **126**, 283 (2014).
- [10] S. Bouacha, A. H. Djerourou, I. Chemo, *J. Sci.*, **4**, 1941 (2013).
- [11] J. Liu, S. Niwayama, Y. You, K. N. Houk, *J. Org. Chem.*, **63**, 1064 (1998).
- [12] V. Cadra, R. Portoles, J. Murga, S. Uriel, J. A. Marco, L. R. Domingo, R. J. Zaragoza, *J. Org. Chem.*, **65**, 7000 (2000).
- [13] L. R. Domingo, R. J. Zaragoza, *J. Org. Chem.*, **65**, 7000 (2000).
- [14] L. R. Domingo, *Eur. J. Org. Chem.*, 2265 (2000).
- [15] K. Marakchi, O. Kabbaj, N. Komaha, R. Jalal, M. Esseffar, *J. Mol. Struct. Theochem.*, **620**, 271 (2003).
- [16] Sk. A. Ali, J. H. Khan, M. I. M. Wazeer, P. H. Perzanowski. *Tetrahedron.*, **45**, 5979 (1989).
- [17] (a) R. Huisgen, In *1,3-Dipolar Cycloaddition Chemistry*; A. Padwa, Ed.; Wiley: New York, **1**, 1 (1984); (b) R. Huisgen, "Steric Course and Mechanism of 1,3-Dipolar cycloadditions", in *Advances in Cycloaddition*, Ed., D. P. Curran, JAI Press Inc, Greenwich, **1**, 1 (1988).

- [18] A. Padwa, In *Comprehensive Organic Synthesis*; B. M. Trost, I. Fleming, Eds.; Pergamon Press: Oxford, **4**, 1069 (1991).
- [19] P. A. Wade, In *Comprehensive Organic Synthesis*; B. M. Trost, I. Fleming, Eds.; Pergamon Press: Oxford, **4**, 1111 (1991).
- [20] S. Karlsson, Asymmetric 1,3-Dipolar Cycloaddition Reactions of Azomethine Ylides, Thiocarbonyl Ylides, and Nitrones; Doctoral thesis: Univ. of Denmark; (2003).
- [21] R. Huisgen, *Angew. Chem. Int. Ed. Engl.*, **2**, 565 (1963).
- [22] R. Huisgen, *J. Org. Chem.*, **33**, 2291 (1968).
- [23] R.A. Firestone, *J. Org. Chem.*, **33**, 2285 (1968).
- [24] R.A. Firestone, *J. Org. Chem.*, **37**, 2181 (1972).
- [25] R.A. Firestone, *Tetrahedron.*, **33**, 3009 (1977).
- [26] E. M. Moloney, Cycloaddition reactions of imidazolium and phthalazinium dicyanomethanide 1,3-dipoles: synthesis, mechanism and the effect of water; Doctoral thesis: Univ. of Ireland; (2017).
- [27] R. Sandhya, Studies on developing a facile route for the synthesis of highly substituted quinoline and indole derivatives; Doctoral thesis: Univ. of India; (2013).
- [28] R. B. Woodward, R. Hoffmann, *Angew. Chem. Int. Ed.*, **8**, 781 (1969).
- [29] R. Sustmann, *Tetrahedron Lett.*, 2717 (1971).
- [30] R. Sustmann, *Tetrahedron Lett.*, 2721 (1971).
- [31] M. Trautz, *Anorg. Allg. Chem.*, **96**, 28 (1916).
- [32] S. A. Arrhenius, *Phys. Chem.*, **4**, 96 (1889).
- [33] C. E. Dykstra, F. Gernot, S. K. Kwang, E. S. Gustavo, Theory and Applications of Computational Chemistry: The First Forty Years. Elsevier (2005).

- [34] D.A. McQuarrie, J.D. Simon, *Physical Chemistry*, Viva Books Private Limited, First South Asian Edition (1998).
- [35] R. G.Parr, W.Yang, *Annu. Rev. Phys. Chem.*, **46**, 701 (1995).
- [36] (a) R. G. Parr, R. G. Pearson, *J. Am. Chem. Soc.*, **105**, 7512 (1983); (b) R. G.Parr, W. Yang, *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, NY, USA, (1989).
- [37] T. Koopmans, *Physica.*, **1**, 104 (1933).
- [38] W. Kohn, L. J. Sham, *Phys. Rev. B* **140**, 1133 (1965).
- [39] (a) R. G. Pearson, *J. Am. Chem. Soc.*, **85**, 3533 (1963); (b) R. G. Pearson, Acids and bases. *Science* **151**, 172 (1966); (c) R. G. Pearson, J. Songstad, *J. Am. Chem. Soc.*, **89**, 1827 (1967).
- [40] R. G. Parr, L. v. Szentpály, S. Liu, Electrophilicity index. *J. Am. Chem. Soc.*, **121**, 1922 (1999).
- [41] R. B. Woodward, R. Hoffmann, *Acc. Chem.Res.*, **1**, 17 (1968).
- [42] P. Pérez, L. R.Domingo, A. Aizman, R. Contreras, *The Electrophilicity Index in Organic Chemistry*, in *Theoretical Aspects of Chemical Reactivity*; Elsevier: New York, NY, USA, **19** (2007).
- [43] R. G.Parr, W. Yang, *J. Am. Chem. Soc.*, **106**, 4049 (1984).
- [44] P. Geerlings, F. de Proft, W. Langenaeker, *Chem. Rev.*, **103**, 1793 (2003).
- [45] L. R. Domingo, M. J. Aurell, P. Pérez, R. Contreras, *Tetrahedron.*, **58**, 4417 (2002).
- [46] Gaussian 09, Revision B.01, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery, J.E. Peralta, F.

Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, T. Keith, R. Kobayashi, R. Normand, J. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, O. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, (2010) Gaussian, Inc., Wallingford CT.

[47] K. Fukui, *J. Phys. Chem.*, **74**, 4161 (1970).

[48] A. E. Reed, F. Weinhold, *J. Chem. Phys.*, **78**, 4066 (1983).

[49] J. Tomasi, M. Persico, *Chem. Rev.*, **94**, 2017 (1994).

[50] E. Cancès, B. Mennucci, J. Tomasi, *J. Chem. Phys.*, **107**, 3032 (1997).

[51] F.L. Hirshfeld, *Theor. Chem. Acc.*, **44**, 38 (1977).

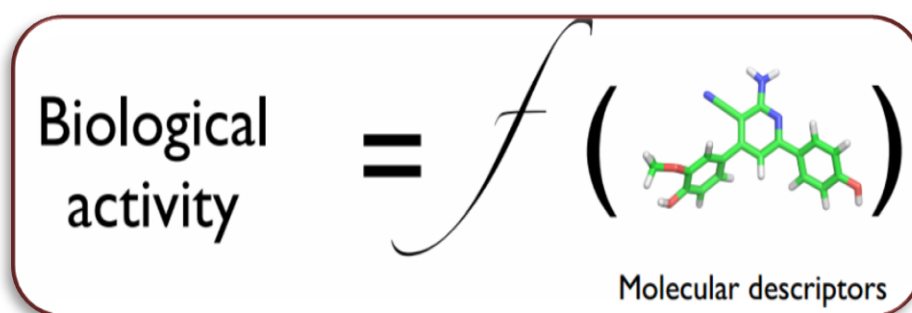
[52] J. Padmanabhan, R. Parthasarathi, U. Sarkar, V. Subramanian, P.K. Chattaraj, *Chem. Phys. Lett.*, **383**, 122 (2004).

[53] R.G. Parr, L. Von Szentpaly, S. Liu, *J. Am. Chem. Soc.*, **121**, 1922 (1999).

[54] K. Marakchi, R. Ghailane, O. Kabbaj, N. Komaha, *J. Chem. Sci.*, **126**, 283 (2014).

CHAPTER III

DFT- Based Reactivity And QSAR Modeling of 1,2,4,5-Tetrazine Inhibitors



III.1.Introduction

Lung Cancer is a malignant tumor, which threatening the human life and health. It is the most common type of all human cancers all over the world [1,2]. The main cause of death for lung cancer is smoking, the early diagnosis which leads to a rise in tumor tissues and cell proliferation [3,4]. Although treatments include surgery, chemotherapy, radiation therapy [5]. The mortality rate reached 158080 deaths in 2016 (US) [6]. Therefore, it is necessary to find a cure for this disease.

National Cancer Institute is now investigating 1,2,4,5-tetrazine molecule to explore their effectiveness against cancer [7]. 1,2,4,5-tetrazine represent a significant sort of heterocyclic compounds that find many practical and synthetic applications also provide a broad of natural products and bioactive compounds [8,9]. Natural products demonstrate a wide range spectrum of biological activities with a high potential for medicinal applications [10]. 1,2,4,5-tetrazine and its derivatives were found to have a high potential of biological properties including as anti-inflammatory, anti-malarial, anti-viral, anti- mite, herbicidal, antibacterial activities, also possess antitumor activity [11-17] and have been used as pesticides and herbicides [18].

In the recent years, the density functional theory (DFT) has become the most popular quantum chemical method were used to compute several molecular properties such as chemical, physical and biological systems [19]. The chemical descriptors such as atomic charges, molecular orbital energies, frontier orbital densities, etc. Which have been used in the development of quantitative structure activity relationships (QSAR) for predicting reactivity in terms of the structure and physicochemical properties of molecules [20].

Quantitative Structure Activity Relationship (QSAR) is a way of finding a simple equation that can be used to calculate some property from the molecular structure of a compound. QSAR attempt to correlate structural molecular features (descriptors) with physicochemical properties such as biological activities for a set of compounds, by means of statistical methods. As a result, a simple mathematical relationship is established [21].

It has been nearly 45 years since the QSAR modeling was first introduced into the practice of agrochemistry, drug design, toxicology, and industrial and environmental chemistry. Its growing power in the following years may be mainly attributed to the rapid and extensive development in methodologies and computational techniques that have

allowed to delineate and refine the many variables and approaches used to model molecular properties [22]. Furthermore, the interest in QSAR is more and more growing because nowadays these tools are used not only for research purposes but also to produce data on chemicals in the interest of time and cost effectiveness [23].

Now a day in modern QSAR (also termed as 2D or 3D-QASR), QSAR relations are determined by using statistical correlation of structural descriptor and biological activity. Physiochemical properties are correlated to biological activity using statistical method [24]. There are several modules available in commercial tools [25] (SYBYL, MOE and Schrodinger suite) that make QSAR studies simpler than ever.

In general, QSAR modeling involves a systematic process with multiple steps, including dataset preparation, molecular descriptors selection and generation, mathematical or statistical models derivation, model training and validation using a training dataset and model testing on a testing dataset [26].

III.2. Objectives of QSAR [27]

Mostly all the QSAR methods focus on the following goals:

1. Quantitative relationship between the structure and physiochemical properties of substances and their biological activity are being used as the foundation stone in search of new medicines. The mathematical and statistical analysis helps us to predict the drug activity.
2. QSAR makes it easy now to reach the conclusion for any of the congener that still not in process, in way that whether it will optimal and profitable or not.
3. To quantitatively correlate and recapitulate the relationships between trends in chemical structure alterations and respective changes in biological endpoint for comprehending which chemical properties are most likely determinants for their biological activities.
4. To optimize the existing leads so as to improve their biological activities.
5. To predict the biological activities of untested and sometimes yet unavailable compounds.

III.3. Molecular Descriptors [21]

A common question in QSAR is how to describe molecules and their physicochemical properties (descriptors). The nature of the descriptors used and the extent to which they instruct the structural properties related to the biological activity is a critical part of a QSAR study [28]. The various descriptors in use can be largely categorized as being constitutional, topological, electrostatic, geometrical, or quantum chemical [29].

III.3. 1. Constitutional descriptors

Give a simple description of what is in the molecule. For example, the number of heteroatoms, the number of rings, the number of double bonds, etc. Constitutional descriptors often appear in a QSAR equation when the property being predicted varies with the size of the molecule.

III.3. 2. Topological descriptors

Topological descriptors are numbers that give information about the bonding collection in a molecule. They are derived from graph representation of chemical structures; they attempt to encode the size, shape, or branching in the compound by handling of graph-theoretical aspects of the structures. Some examples are Randic indices, Kier and Hall indices, Weiner index (sum of the chemical bonds existing between all pairs of heavy atoms in the molecule), the connectivity index and others [30].

III.3. 3. Electrostatic descriptors

Electrostatic descriptors are single values that give information about the molecular charge division. Some examples are polarity indices and polarizability. One of the most commonly used electrostatic descriptors is the topological polar surface area (TPSA), which gives an indication of the portion of the molecular surface composed of polar groups against nonpolar groups. Another deeply used descriptor is the octanol–water partition coefficient, which is designated by a specific prediction scheme such as ClogP or MlogP.

III.3. 4. Geometrical descriptors

Geometrical descriptors are single values that describe the molecule's size and shape as well as the degree of complementarity of a ligand and the receptor. They are developed from three-dimensional models of molecules, and derived from molecular surface area calculations. Some examples are moments of inertia, molecular volume molecular surface area, and other parameters that describe length, height, and width.

III.3. 5. Quantum chemical descriptors

Give information about the electronic structure of the molecule. They are obtained by molecular orbital calculations and they mainly describe electronic interaction. These includes, the energy of the highest occupied molecular orbital, E_{HOMO} , which is a quantitative measure for the chemical reactivity of the compound ionization potential of a molecule, the energy of the lowest unoccupied molecular orbital, E_{LUMO} , which accounts for the electron affinity, refractivity, and total energy. The $E_{\text{HOMO}} - E_{\text{LUMO}}$ gap or ionization potential can be important descriptors for predicting how molecules will react [31].

III.4. Biological Parameters

In QSAR analysis, it is vital important that the biological data be both accurate and precise to develop a meaningful model. The equilibrium constants and rate constants that are used extensively in physical organic chemistry and medicinal chemistry are related to free energy values ΔG . Thus for use in QSAR, standard biological equilibrium constants such as K_i or K_m should be used in QSAR studies. Percentage activities (e.g., % inhibition of growth at certain concentrations) are not appropriate biological endpoints because of the nonlinear characteristic of dose-response relationships. These types of endpoints may be transformed to equieffective molar doses. Only equilibrium and rate constants pass muster in terms of the free-energy relationships or influence on QSAR studies. Biological data are usually expressed on a logarithmic scale because of the linear relationship between response and log dose in the midregion of the log dose-response curve. Inverse logarithms for activity ($\log 1/C$) are used so that higher values are obtained for more effective analogs. Various types of biological data have been used in QSAR analysis [32].

Types of Biological Data Utilized in QSAR Analysis

Source of Activity	Biological Parameters
1. Isolated receptors	
Rate constants	Log k
Michaelis-Menten	Log 1/K _m
Constants	
Inhibition constants	Log 1/K _i
Affinity data	pA ₂ ; pA ₁
2. Cellular systems	
Inhibition constants	Log 1/IC ₅₀
Cross resistance	Log CR
In vitro biological data	Log 1/C
Mutagenicity states	Log TA ₉₈
3. In vivo systems	
Biocentration factor	Log BCF
In vivo reaction rates	Log I (Induction)
Pharmacodynamic Rates	Log T (total clearance)

III.5. Statistical methods

Statistical methods are the mathematical basis for the development of QSAR models. Chemometric methods [33] are used to extract information from QSAR data using tools of statistics and mathematics. The applications of these methods are combined with the important goal of explanation and prediction of non-synthesised test compounds. Many different statistical methods are available such as: Multiple Linear Regression (MLR), Principal Component Analysis (PCA), Partial Least Squares (PLS), Principal Components

Regression (PCR), Artificial Neural Networks (ANN) and Support Vector Machine (SVM) [34]. The majority methods have been used are regression methods, of which multiple linear regression was the first used [35].

III.5.1. Multiple linear regression

Multiple linear regression or MLR [36] is one of the most popular methods of QSAR due to its simplicity, transparency, reproducibility, and easy interpretability. The generalized expression of an MLR equation will be like the following:

$$Y = a_0 + a_1 \times X_1 + a_2 \times X_2 + a_3 \times X_3 + \dots + a_n \times X_n$$

In the above expression, Y is the response or dependent variable (biological data), X_1 , X_2 , ..., X_n are descriptors (features or independent variables) present in the model with the corresponding regression coefficients a_1 , a_2 , ..., a_n , respectively, and a_0 is the constant term of the model. The interpretation of contribution of individual descriptors X_1 , X_2 , ..., X_n is straightforward depending on the corresponding coefficient value and its algebraic sign. Each regression coefficient should be significant at $p < 0.05$ which can be checked from a 't' test. The descriptors present in an MLR model should not be much intercorrelated.

III.6. Chemometric Tools [37]

III.6. 1. Determination coefficient (R^2)

One can define the determination coefficient (R^2) in the following manner:

$$R^2 = 1 - \frac{\sum(Y_{\text{obs}} - Y_{\text{calc}})^2}{\sum(Y_{\text{obs}} - \bar{Y}_{\text{obs}})^2} \quad (1)$$

In the above equation, Y_{obs} stands for the observed response value, while Y_{calc} is the model-derived calculated response and \bar{Y}_{obs} is the average of the observed response values. For the ideal model, the sum of squared residuals being 0, the value of R^2 is 1.

As the value of R^2 deviates from 1, the fitting quality of the model deteriorates. The square root or R^2 is the multiple correlation coefficient (R).

III.6. 2. Adjusted R^2 (R_a^2)

If one goes on increasing the number of descriptors in a model for a fixed number of observations, R^2 values will always increase, but this will lead to a decrease in the degree of freedom and low statistical reliability. Thus, a high value of R^2 is not necessarily an indication of a good statistical model that fits well the available data. To reflect the explained variance (the fraction of the data variance explained by the model) in a better way, adjusted R^2 which has been defined in the following manner:

$$R_a^2 = \frac{(N-1) \times R^2 - p}{N-1-p} \quad (2)$$

In the above expression, p is the number of predictor variables used in the model development.

III.6. 3. Variance ratio (F)

To judge the overall significance of the regression coefficients, the variance ratio (the ratio of regression mean square to deviations mean square) can be defined as follows:

$$F = \frac{\frac{\sum (Y_{\text{calc}} - \bar{Y})^2}{p}}{\frac{\sum (Y_{\text{obs}} - Y_{\text{calc}})^2}{N-p-1}} \quad (3)$$

The F value has two degrees of freedom: p , $N - p - 1$. The computed F value of a model should be significant at $p < 0.05$. For overall significance of the regression coefficients, the F value should be high.

III.6. 4. Standard error of estimate (s)

For a good model, the standard error of estimate of Y should be low and this is defined as follows:

$$s = \sqrt{\frac{(Y_{\text{obs}} - Y_{\text{calc}})^2}{N - p - 1}} \quad (4)$$

It has a degree of freedom of $N - p - 1$.

III.7. Validation of QSAR Models

Cross-validation is important statistical validation technique that avoids over-fitting models on training data, as over-fitting will give low accuracy on validation. It also helps to right set of descriptors, an appropriate algorithm and associated parameter for a given dataset [38]. In the leave-one-out (LOO) method of CV, the training set is primarily modified by eliminating one compound from the set. The QSAR model is then rebuilt based on the remaining molecules of the training set using the descriptor combination originally selected, and the activity of the deleted compound is computed based the resulting QSAR equation. This cycle is repeated until all the molecules of the training set have been deleted once, and the predicted activity data obtained for all the training set compounds are used for the calculation of various internal validation parameters. Finally, the model predictivity is judged using the predicted residual sum of squares (PRESS) and cross-validated R^2 (Q^2) for the model [39].

$$\text{PRESS} = \sum (Y_{\text{obs}} - Y_{\text{pred}})^2 \quad (5)$$

$$Q^2 = 1 - \frac{\sum (Y_{\text{obs}(\text{train})} - Y_{\text{pred}(\text{train})})^2}{\sum (Y_{\text{obs}(\text{train})} - \bar{Y}_{\text{training}})^2} = 1 - \frac{\text{PRESS}}{\sum (Y_{\text{obs}(\text{train})} - \bar{Y}_{\text{training}})^2} \quad (6)$$

In Eqs. (5)–(6), Y_{obs} and Y_{pred} correspond to the observed and LOO predicted activity values, n refers to the number of observations, $Y_{\text{obs}(\text{train})}$ is the observed activity,

$Y_{\text{pred}}(\text{train})$ is the predicted activity of the training set molecules based on the LOO technique.

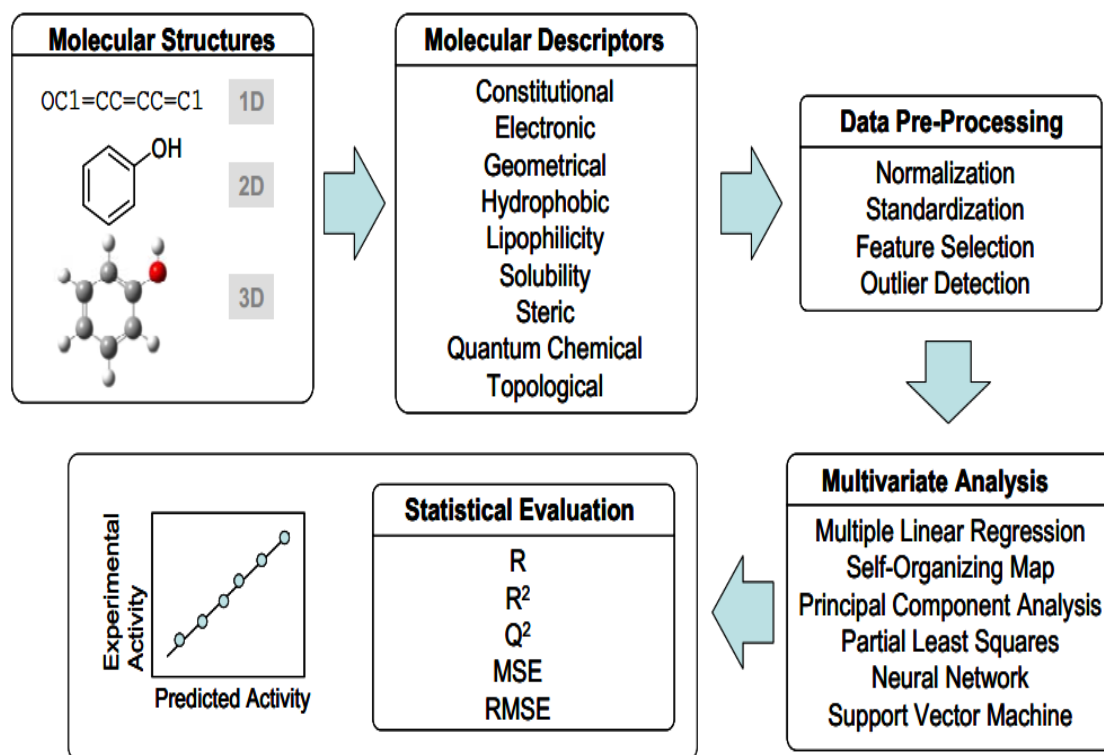


Figure III.1: Schematic overview of the QSAR process [40]

❖ **When can we accept the developed QSAR model as reliable and predictive one?**

A developed QSAR model can be accepted generally in QSAR (MLR) studies when it can satisfy the following criterion (The following values are the minimum recommended values for significant QSAR model meanwhile these evaluation measures are depend on the response measure scale or measure unit [41, 24]):

Statistical measures	Minimum recommended values
N	Number of molecules (>20 molecules)
K	Number of descriptor in a model (Statistically $n/5$ descriptor in a model).
Df	Degree of freedom (n-K-1) (higher is better)
R²	Coefficient of determination (≥ 0.6)
R	Correlation coefficient (≥ 0.8)
S	Standard deviation is not much larger than standard deviation of the biological data.
Q²	Cross- validated (> 0.5)
SEE	Standard error of estimate (smaller is better)

III.8. Material and methods

III.8.1. Computational methods

The reported quantum chemical calculations were performed DFT methods at the B3LYP/6-311++G(d,p) level of theory for all 1,2,4,5-tetrazines. The geometry optimization in the gas phases was carried out using Gaussian 09 suite of programs [42]. The optimized geometry in the gas phases of molecules was further reoptimized in the solvent effect (water) using the polarisable continuum model PCM [43]. Atomic electronic populations were computed using the natural bond orbital (NBO) method in the both gas and solvent phases [44].

Utilizing the check point file obtained from these calculations, the Gaussian 09 cubegen utility can generate the density cubes necessary to obtain each Fukui function. The cubegen utility in Gaussian 09 calculates the ρ_{N+1} , ρ_N , and ρ_{N-1} grids for use in generating electron-density mapped surfaces in GaussView [45]. GaussView then calculates the electron-density mapped surfaces from arithmetic operations on the cube files, namely $f(-) = \rho_N - \rho_{N-1}$ and $f(+) = \rho_{N+1} - \rho_N$, where $f(-)$ and $f(+)$ show the sites that are most susceptible to electrophilic or nucleophilic attack, respectively. All Fukui functions

mapped surfaces were parameterized in exactly the same manner using GaussView, where Isovalue = 0.020, Density = 0.040. For this work, these surface-mapping parameters were obtained by matching previously reported $f(-)$ and $f(+)$ Fukui functions for 1,2,4,5-tetrazine.

III.8.2. QSAR modeling

A total of 18 of 1,2,4,5-tetrazine derivatives has been studied and analyzed in order to find quantitative structure activity relationship between the antitumor lung cancer activity and the structure of these molecules. The biological parameters used in this study were collected from literature [46] and listed in Table III.3.

The multiple linear regression (MLR) analysis was employed to derive the QSAR models for some 1,2,4,5-tetrazine derivatives. MLR and correlation analysis were carried out by using statistical software SPSS version 19 for Windows [47].

The QSAR study was performed by choosing some of electronic descriptors such as: HOMO-LUMO energy gap (GAP), electrophilicity(ω), the electronic chemical potential (μ), dipole moment (DM), atomic net charge (qN1, qC3, qN4, qC6), nucleophilic frontier electron density (f_{N1}^N, f_{N2}^N) and electrophilic frontier electron density f_{C3}^E .

Frontier orbital electron densities also involve the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), providing useful measures of donor–acceptor interactions in the molecular space [48].

Nucleophilic atomic frontier electron density is defined as: $f_i^N = \sum(C_{LUMO})^2 \times 100$.

Electrophilic atomic frontier electron density is defined as: $f_i^E = \sum(C_{HOMO})^2 \times 100$.

III.9. Results and discussion

III.9.1. Analysis of the DFT reactivity indices of 1,2,4,5-tetrazine

The DFT reactivity indices of the 1,2,4,5-tetrazine were analyzed by using the global indices: chemical hardness (η), electronic chemical potential (μ), global electrophilicity (ω) and other parameters such dipole moment (DM) and total energy (Et) are presented in

Table III.1, while the local properties are displayed in Table III.2, have been analyzed in both gas and solvent phases.

Table III.1: Reactivity descriptors for 1,2,4,5-tetrazine at the B3LYP/6-311++G(d,p) level

	E_t (u.a)	HOMO (u.a)	LUMO (u.a)	η (u.a)	μ (u.a)	ω (eV)	DM (Debye)
Gas phase	-297.631	-0.215	-0.031	0.184	-0.123	1.115	0.846
Aqueous phase	-297.643	-0.220	-0.033	0.187	-0.126	1.154	1.085

The electronic chemical potential and the electrophilicity values of the 1,2,4,5-tetrazine in the gas phase are higher than in the aqueous phase, this indicates that the charge transfer is better in the gas than in the aqueous phase. Also, the electrophilic indices of 1,2,4,5 tetrazine may change in the presence in a solution.

The total energy vary slightly between the structures in the gas and aqueous phases. Which corresponds that the structure of 1,2,4,5-tetrazine is more stable in the gas than the aqueous phase.

The dipole moment is greater in water (DM=1.085 Debye) than in the gas phase (DM=0.846 Debye). This, suggests the dipole moment of 1,2,4,5-tetrazine increases with increasing the polarity of the solvent.

Table III.2: Fukui function values of 1,2,4,5-tetrazine in gas and aqueous phases

Atoms	Gas phase		Aqueous phase	
	$f(-)$	$f(+)$	$f(-)$	$f(+)$
N1	0.1967	0.0261	0.2103	0.0366
N2	0.1481	0.0350	0.1524	0.1627
C3	0.0405	0.0363	0.0449	0.2429
N4	0.1967	0.0261	0.2103	0.0366
N5	0.1481	0.0350	0.1524	0.1627
C6	0.0405	0.0363	0.0449	0.2429
H7	0.0578	0.2158	0.0445	0.0352
H8	0.0569	0.1871	0.0480	0.0222
H9	0.0578	0.2151	0.0442	0.0352
H10	0.0570	0.1864	0.0480	0.0222

In Table III.2, we have reported the values of Fukui function calculated by NBO charge in the both gas and aqueous effect. Where, the electrophile attack characterized a largest value of $f(-)$. The results indicate the more favorable site for electrophilic attack in both gas and aqueous phases at the atom N1. While, the nucleophile attack favor a largest value of $f(+)$. Therefore, the more reactive sites in the gas phase at the atom H7, but in the aqueous phase predict reactivity site on C3. In addition, a visualization of Fukui indices of the 1,2,4,5-tetrazine is shown in figure III.2 to demonstrate the reactivity centers of the studied molecule.

We note that the reactivity of 1,2,4,5-tetrazine in the aqueous phase is more reactive in the electrophilic and nucleophilic cases than in the gas phase.

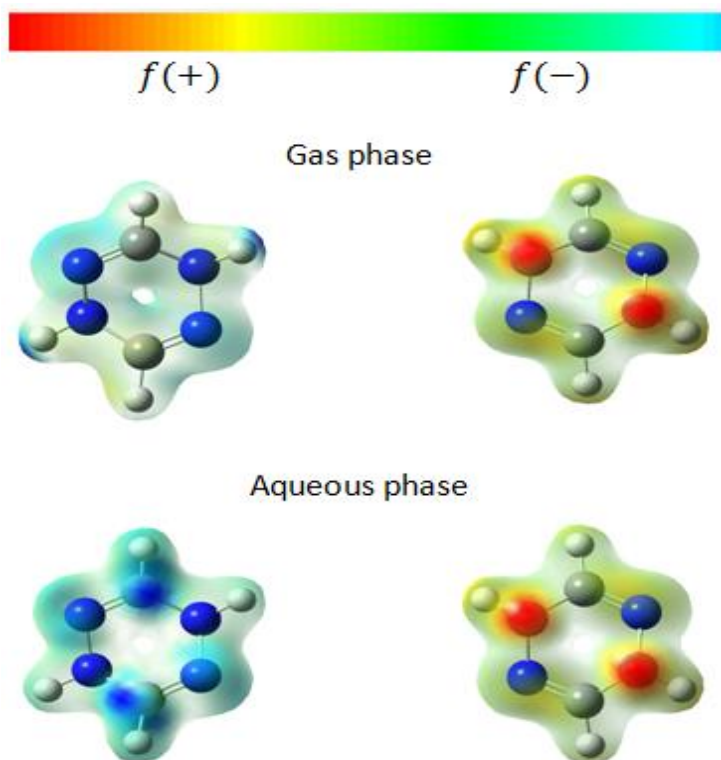
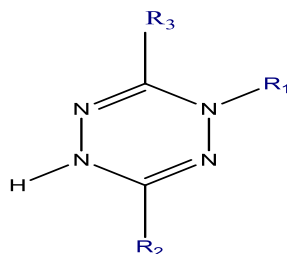


Figure III.2: Electron-density mapped $f(+)$ and $f(-)$ Fukui function for 1,2,4,5-tetrazine in both gas and aqueous phases (the blue regions show the areas of the molecules most susceptible to nucleophilic attacks and the red regions show the areas of the molecules most susceptible to electrophilic attacks)

III.9.2. Study of Quantitative structure-activity relationship (QSAR) for 1,2,4,5 - tetrazine derivatives

The analysis of QSAR was performed using pIC_{50} of 18 molecules have been evaluated in vitro antitumor activity against lung cancer cell lines (A-549); these compounds are listed in Table III.3. In order to identify a quantitative relationship between the structure and antitumor activity. The values of the eight electronic descriptors in both gas and aqueous phases are listed in Table III.4.

Table III.3: Chemical structures and experimental activity of the 1,2,4,5-tetrazine derivatives under study



Comp	R ₂ = R ₃	R ₁	pIC ₅₀ exp
1	Ph	COOCH ₃	4.403
2	Ph	COO(CH ₂) ₂ CH ₃	4.389
3	Ph	COO(CH ₂) ₂ CH(CH ₃) ₂	4.355
4	4-CF ₃ C ₆ H ₄	COOCH ₂ CH ₃	6.240
5	4-ClC ₆ H ₄	COOCH ₃	4.306
6	Ph	H	4.701
7	4-CF ₃ C ₆ H ₄	H	4.398
8	4-ClC ₆ H ₄ CH ₂	H	4.286
9	4-ClC ₆ H ₄	H	4.286
10	2-OH-5-ClC ₆ H ₃	H	4.879
11	Ph	COCH ₃	4.350
12	Ph	COCH ₂ CH ₃	4.316
13	Ph	COCH(CH ₃) ₂	5.654
14	4-CF ₃ C ₆ H ₄	COCH ₃	5.011
15	4-CF ₃ C ₆ H ₄	COCH ₂ CH ₃	5.020
16	4-CF ₃ C ₆ H ₄	COCH(CH ₃) ₂	4.500
17	4-CF ₃ C ₆ H ₄	COCH ₂ Cl	5.069
18	Ph	CONH-(2-methoxyphenyl)	4.348

Table III.4: Quantum chemical descriptors of 1,2,4,5-tetrazine derivatives in both gas and aqueous phases

Comp.	Gas phase								Aqueous phase							
	f_{N1}^N	f_{N2}^N	f_{C3}^E	GAP	DM	qN1	qC3	qC6	f_{N1}^N	f_{N2}^N	f_{C3}^E	GAP	DM	qN1	qC3	qC6
1	0.594	0.712	1.508	0.151	2.280	-0.387	0.413	0.409	0.417	0.321	1.284	0.155	3.691	-0.384	0.428	0.412
2	0.389	0.308	0.174	0.128	3.862	-0.308	0.363	0.402	0.385	0.329	0.214	0.131	5.429	-0.297	0.369	0.424
3	0.547	0.624	0.185	0.151	1.912	-0.389	0.413	0.410	0.269	0.292	1.466	0.155	3.285	-0.386	0.428	0.412
4	0.612	0.694	0.230	0.142	2.557	-0.389	0.408	0.406	0.381	0.404	0.192	0.146	3.868	-0.387	0.424	0.407
5	0.490	0.559	0.233	0.148	2.296	-0.388	0.412	0.408	0.284	0.306	0.217	0.151	3.695	-0.385	0.427	0.410
6	0.195	0.114	0.095	0.105	0.000	-0.476	0.406	0.406	0.194	0.122	0.107	0.106	0.000	-0.477	0.413	0.413
7	0.176	0.175	0.002	0.128	0.000	-0.201	0.347	0.346	0.174	0.173	0.002	0.130	0.000	-0.208	0.359	0.359
8	0.328	0.170	0.072	0.168	1.595	-0.490	0.419	0.419	0.259	0.130	0.394	0.176	2.519	-0.50	0.425	0.425
9	0.194	0.114	0.090	0.104	0.000	-0.477	0.404	0.423	0.194	0.120	0.098	0.103	0.000	-0.476	0.429	0.410
10	0.200	0.122	0.074	0.09	0.000	-0.480	0.409	0.409	0.196	0.132	0.091	0.096	0.000	-0.482	0.417	0.417
11	0.379	0.391	1.373	0.150	3.826	-0.382	0.416	0.396	0.214	0.209	1.307	0.155	5.901	-0.372	0.432	0.404
12	0.412	0.382	1.413	0.150	3.621	-0.382	0.414	0.396	0.243	0.209	1.382	0.155	5.667	-0.374	0.429	0.404
13	0.456	0.396	0.355	0.151	3.637	-0.379	0.414	0.399	0.256	0.213	0.316	0.156	5.768	-0.370	0.429	0.406
14	0.413	0.432	0.142	0.142	3.732	-0.390	0.320	0.388	0.258	0.257	0.132	0.147	5.859	-0.374	0.397	0.400
15	0.475	0.466	0.150	0.142	3.366	-0.390	0.318	0.388	0.305	0.279	0.146	0.146	5.438	-0.381	0.334	0.395
16	0.489	0.429	0.147	0.142	3.620	-0.390	0.317	0.391	0.307	0.260	0.132	0.147	5.729	-0.380	0.333	0.397
17	0.455	0.461	0.163	0.143	4.826	-0.389	0.318	0.381	0.290	0.296	0.149	0.147	7.314	-0.375	0.339	0.388
18	0.347	0.200	0.106	0.105	2.220	-0.359	0.421	0.414	0.645	0.477	0.050	0.102	5.132	-0.362	0.376	0.417

Our work is based on the development of the best QSAR models to explain the correlation between the different electronic descriptors and the biological activity of the 1,2,4,5-tetrazine derivatives in the both gas and aqueous phases.

The use of the nineteen compounds does not give any model satisfied statically. The compounds 5, 13 and 14 are three outliers, therefore, is necessary to delete these compounds for improving the quality of the regression models. After removal of compounds 5, 13 and 14, QSAR models were obtained and presented by the following mathematical equations:

❖ **In gas phase:**

$$pIC_{50} = 10.294 + 0.778f_{N2}^N - 0.169f_{C3}^E - 8.258 \text{ GAP} + 0.081\text{DM} - 4.229qN1 - 16.761qC6 \quad (1)$$

$$n = 15; R = 0.909; R^2 = 0.827; S = 0.164; F = 6.359$$

❖ **In aqueous phase:**

$$pIC_{50} = 2.400 - 7.371f_{N1}^N + 10.510f_{N2}^N - 0.821f_{C3}^E + 4.992\text{GAP} - 2.840qN1 + 6.043qC3 \quad (2)$$

$$n = 15; R = 0.914; R^2 = 0.835; S = 0.278; F = 6.757$$

QSAR models having $R^2 > 0.6$ will only be considered for validation. The value $R^2 = 0.827$ (model 1) and $R^2 = 0.835$ (model 2) allowed us to indicate firmly the correlation between different parameters (independent variables) with the antitumor activity.

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

In model (1), the antitumor activity in the gas phase increases by increasing values of molecular descriptors f_{N2}^N and DM. The positive coefficient of the f_{N2}^N shows that the antitumor activity needs a higher nucleophilic frontier electron density of 2- position Azote atom, indicting the ability of Azote atom to accept electrons. In the other hand antitumor activity in the solvent phase increases by increasing values of f_{N2}^N , qC3 and GAP (model 2). The positive coefficient of qC3 indicates the more

positive charge of the carbon at position 3, the higher activity, which suggest that the positions 3 of 1,2,4,5-tetrazine ring should be occupied by electro-withdrawing substituent.

In order to test the validity of the predictive power of selected MLR models (1 and 2), the leave-one out technique (LOO technique) was used [49-51]. The developed models were validated by calculation of the following statistical parameters: predicted residual sum of squares (PRESS), total sum of squares deviation (SSY) and cross validated correlation coefficient (r^2_{adj}) (Table III.5).

PRESS is an important cross-validation parameter as it is a good approximation of the real predictive error of the models. Its value being less than SSY points out that model predicts better than chance and can be considered statically significant. The smaller PRESS value means the better of the model predictability. From the results depicted in Table III.5, model 1 and 2 are statistically significant. Furthermore, for reasonable QSAR model, the PREES/SSY ratio should be lower than 0.4 [52]. The data presented in Table III.5 indicate that for the developed models this ratio is 0.173 for the first model and 0.164 for the second one. The high value of r^2_{cv} and r^2_{adj} are essential criteria for the best qualification of the QSAR models in the both gas and aqueous phases. However, the only way to estimate the true predictive power of developed model is to predict the by calculation of pIC_{50} values of the investigated 1,2,4,5-tetrazine using model 1 and 2, respectively (Table III.6).

Table III.5: Cross-validation parameters in both gas and aqueous phase

Model	PRESS	SSY	PRESS/SSY	S_{PRESS}	r^2_{adj}	r^2_{cv}
(1)	0.215	1.241	0.173	0.119	0.697	0.827
(2)	0.618	3.753	0.164	0.202	0.712	0.835

However, the only way to estimate the true predictive power of developed models is to predict the by calculation of pIC_{50} values of the investigated 1,2,4,5-tetrazines using model 1 and 2 (Table III.6). Whereas, the gas phase results are showing an average of the squared values of residues equal to 0.012 while the aqueous phase results are showing an average of the squared values of residues equal to 0.035. Thus, gas phase

calculations are more accurate for studying a quantitative structure activity relationship of our studied molecules as tumor inhibitors.

Table III.6: Experimental, predicted and residual activity of 1,2,4,5-tetrazine derivatives in gas and aqueous phases

Comp.	pIC ₅₀ Exp	Gas phase		Aqueous phase	
		pIC ₅₀ Pred	Residue	pIC ₅₀ Pred	Residue
1	4.403	4.298	0.104	4.111	0.292
2	4.389	4.303	0.085	4.578	-0.189
3	4.306	4.441	-0.135	4.745	-0.390
4	4.701	4.694	0.006	6.074	0.165
6	4.398	4.408	-0.010	4.549	0.151
7	4.286	4.198	0.087	4.355	0.042
8	4.286	4.429	-0.143	4.411	-0.125
9	4.879	4.789	0.089	4.622	-0.362
10	4.350	4.399	-0.049	4.652	0.226
11	4.316	4.368	-0.052	4.390	-0.040
12	5.011	4.863	0.147	4.103	0.212
15	5.020	4.866	0.153	4.804	0.215
16	4.500	4.800	-0.300	4.592	-0.092
17	5.069	5.082	-0.013	5.109	-0.040
18	4.348	4.316	0.031	4.440	-0.092

The correlation plots of the calculated antitumor activity values in gas and solvent phases are very significant and shows a good deal of correspondence with experimentally reported data (Figure III.3). Thus, our QSAR models in gas and aqueous phases can be successfully applied to predict the antitumor activities in these molecules generations.

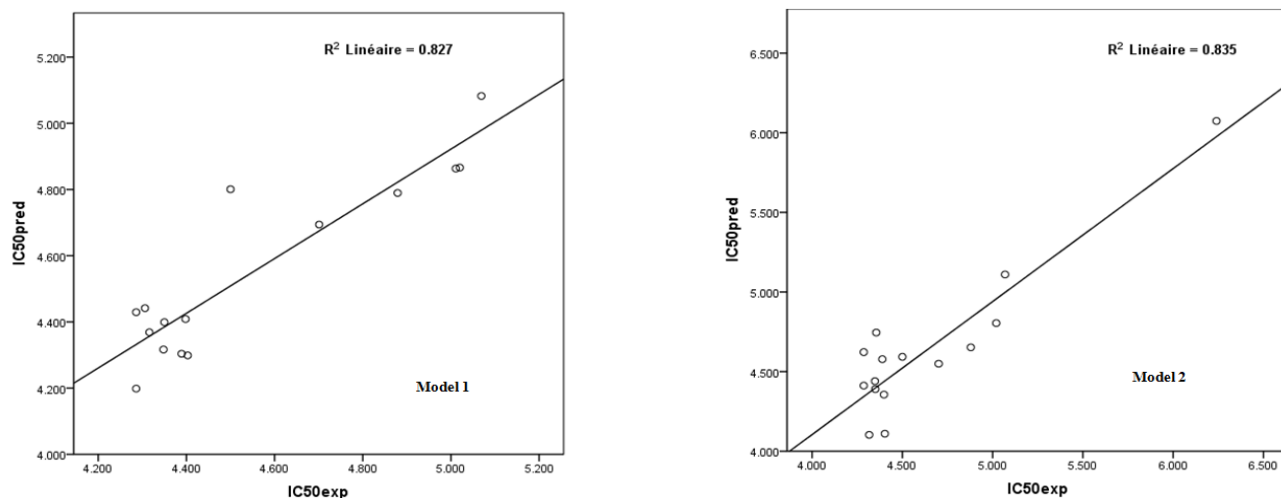


Figure III.3: Predicted plots versus experimental observed antitumor activity for models in both gas and aqueous phase

To investigate the presence of a systematic error in developing the QSAR models in both gas and solvent phases, the residuals of predicted values of the biological activity ($\log 1/IC_{50}$) was plotted against the experimental values, as shown in Figure III.4.

The propagation of the residuals on both sides of zero indicates that no systemic error exists, as suggested by Jalali-Heravi and Kyani [53]. It indicates that these models can be successfully applied to predict the antitumor activity of this class of molecules.

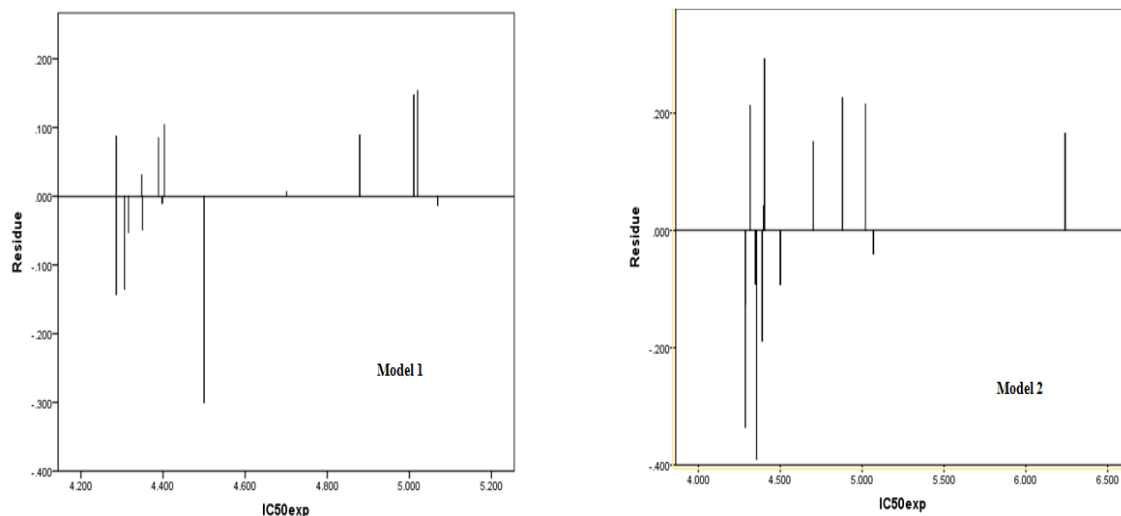


Figure III.4: Plots of residual against experimental observed in gas and aqueous phase

III.10. Conclusion

The reactivity of 1,2,4,5-tetrazine have been thoroughly probed using a DFT method at the B3LYP/6-311++G(d,p) level of theory. The results presented in gas and solvent phases indicated that the application of a solvation effect on the studied molecule affects its reactivity. Whereas, Fukui function values showed the most susceptible sites for electrophile and nucleophile attack in the gas phase are N1 and H7 respectively, while it showed the major sites are N1 and C3 for electrophile and nucleophile attack in the aqueous phase.

In order to determine the influence of the electronic descriptors of 1,2,4,5-tetrazine on the studied antitumor activity, we applied a DFT-based QSAR modeling in gas and aqueous phases.

Our QSAR models, which have been validated using cross validation parameters, showed a common series of important descriptors in both phases gas and aqueous for antitumor activity enhancement, which are: f_{N2}^N , f_{C3}^E , Gap, qN1. Thus, we can say that the antitumor activity of novel 1,2,4,5-tetrazine derivatives can be predicted based on their electronic behavior.

III.11. References

- [1] D.M. Parkin, F. Bray, J. Ferlay, P. Pisani, *CA Cancer J Clin.*, **55**, 74 (2005).
- [2] C.G. Kang, H.J. Lee, S.H. Kim, E.O. Lee, *J Nat Product.*, **79**, 156 (2016).
- [3] L. Jun, Z. Xiaoli, S. Tingting, M. Chao, W. Jun, K. Hualei, T. Jing, S. Zhifeng, Z. Xiaodong, X. Ling, *Evid Based Complement Alternat Med.*, **1**, 13 (2016).
- [4] J.Y. Young Yoon, L. Jung-Dong, S.W. Joo, D.R. Kang, *Ann Occup Environ Med.*, **28**, 15 (2016).
- [5] X. Qifei, J. Xiaoding, Z. Weixing, C. Chuo, H. Gaoyun, L. Qianbin, *Arab. J. Chem.*, **103**, 91 (2015).
- [6] American Cancer Society, Inc.No. 500816, 2016.
- [7] M. Gopalakrishnan, P. Sureshkumar, V. Kanagarajan, J. Thanusu, *J. Sulfur Chem.*, **28**, 383 (2007).
- [8] H. Neunhoeffer, Tetrazines–In Comprehensive Heterocyclic Chemistry, A.R. Katritzky, C.W. Rees (Eds.) Pergamon Press, London.**1**, 531(1984).
- [9] Q.W. Li, S.S. Zang, J. Gao, L. Wei-hua, X. Liang-Zang, Y. Zhi-Gang, *Bioorg. Med. Chem.*, **14**, 7146 (2006).
- [10] B.N. Berad, S.M. Bhiwagade, A.G. Ulhe, *Der Pharma Chemica.*, **4**, 1730 (2012).
- [11] K. Kottke, H. Kuehmstedt, H. Landmann, H. Wehlan, *Chem. Abstr.*, **100**, 103388. East Ger. Patent, **203**, 545 (1984).
- [12] D. Nhu, S. Duffy, V.M. Avery, A. Hughes, J.B. Baell, *Bioorg. Med. Chem. Lett.*, **20**, 4496 (2010).
- [13] R. Guo-Wu, H. Wei-Xiao, *Bioorg. Med. Chem. Lett.*, **15**, 3174 (2005).
- [14] H.I. Falfushynska, L.L. natyshyna, O.B. Stoliar, *Comp. Biochem. Physiol., Part C: Toxicol. Pharmacol.*, **155**, 329 (2012).

- [15] N. Malecki, P. Caroto, B. Rigo, J.H. Goosens, J.P. Henchart, *Bioorg. Med. Chem.*, **12**, 641 (2004).
- [16] V. Kumar, C. Mohan, M. Gupta, M.P. Mahajan, *Tetrahedron.*, **61**, 3533 (2005).
- [17] W.X. Hu, G.W. Rao, Y.Q. Sun, *Bioorg. Med. Chem. Lett.*, **14**, 1177 (2004).
- [18] P. Bhardwaj, N. Gupta, *Iran. J. Org. Chem.*, **8**, 1827 (2016).
- [19] S. Belaidi, T. Salah, N. Melkemi, L. Sinha, O.J. Prasad, *Comput. Theor. Nanosci.*, **12**, 2421 (2015).
- [20] D. R. Roy, U. Sarkar, P. K. Chattaraj, et al, *Mol. Divers.*, **10**, 119 (2006).
- [21] O. Deeb, *J. med. chem. drug discov.*, **3**, 57 (2012).
- [22] M. Charton, B. I. Charton, *Advances in Quantitative Structure–Property Relationships*; JAI Press: Amsterdam, The Netherlands, 228 (2002).
- [23] T. I. Oprea, 3D QSAR Modeling in Drug Design. In *Computational Medicinal Chemistry for Drug Discovery*; P. Bultinck, H. De Winter, W. Langenaeker, J. P. Tollenaere, Eds.; Marcel Dekker: New York, NY, 571–616 (2004).
- [24] N.K. Mahobia, R.D. Patel, N.W. Sheikh, S. K. Singh, A. Mishra, R. Dhardubey, *Der Pharma Chemica.*, **5**, 260 (2010).
- [25] J. Shamsara, *Open Med Chem J.*, **11**, 212 (2017).
- [26] K. Z. Myint, X. Qun Xie, *Int. J. Mol. Sci.*, **11**, 3847 (2010).
- [27] B. Jhanwar, V. Sharma, R. K. Singla, B. Shrivastava, *Pharmacologyonline.*, **1**, 313 (2011).
- [28] G.M. Downs, *Molecular Descriptors*. In *Computational Medicinal Chemistry for Drug Discovery*. P. Bultinck, H. De Winter, W. Langenaeker, J. P. Tollenaere, (Eds.). Marcel Dekker; New York, 515-538 (2004).
- [29] H. Hentabli, S. Naomie, F. Saeed, *Jurnal Teknologi (Sciences & Engineering).*, **78**, 101 (2016).

- [30] K. Roy, R.N. Das, A review on principles, theory and practices of 2D-QSAR. *Current Drug Metabol*, **15**, 346 (2014).
- [31] C.Silipo, A. Vittoria, Three-Dimensional Structure of Drugs. In *Comprehensive Medicinal Chemistry*. Vol 4. Quantitive Drug Design. C. Hansch, P.G. Sammes, J.B. Taylor, eds. Pergamon Press, New York, 154-204 (1990).
- [32] C. D. SELASSIE, *Burger's Medicinal Chemistry and Drug Discovery*, John Wiley & Sons, Inc (2003).
- [33] L. Eriksson, E. Johansson, N. Kettaneh-Wold, S. Wold, *Multi- and Megavariate Data Analysis. Principles and Applications*. Umetrics: Umea (2001).
- [34] L. Xu, W.J. Zhang, *Anal.Chim. Acta.*, **446**, 477 (2001).
- [35] C. Hansch, T. Fujita, *J. Amer. Chem. Soc.*, **86**, 1616 (1964).
- [36] G. W. Snedecor, W. G. Cochran, *Statistical methods*. Oxford and IBH, New Delhi (1967).
- [37] I. Mitra, P.P. Roy, S. Kar, P. Ojha, K. Roy, *J. Chemometr.*, **24**, 33 (2010).
- [38] N. K. Mahobia, R. D. Patel, N. W. Sheikh, S.K. Singh, A. Mishra, R. Dhardubey, *Der Pharma Chemica.*, **2**, 263 (2010).
- [39] K. Roy, I. Mitra, *Comb Chem High Throughput Screen.*, **14**, 450 (2011).
- [40] Ch. Nantasenamat, Ch. I. Na-Ayudhya, T. Naenna, V. Prachayasittikul, *EXCLI J.*, **8**, 76 (2009).
- [41] R. Veerasamy, H. Rajak, A. Jain, S. Sivadasan, C. P. Varghese, R. K. Agrawal, *Drug Des Discov.*, **2**, 516 (2011).
- [42] Gaussian 09, Revision B.01, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M.

Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery, J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, T. Keith, R. Kobayashi, R. Normand, J. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, O. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, (2010) Gaussian, Inc., Wallingford CT.

[43] J. Tomasi, M. Persico, *Chem. Rev.*, **94**, 2027 (1994).

[44] A. E Reed, F. Weinhold, *J. Chem. Phys.*, **78**, 4066 (1983).

[45] R. Todeschini, V. Consonni, *Molecular Descriptors for Chemoinformatics*; Wiley-Vch Verlag GmbH & Co. KGaA, Weinheim, 1257 (2009).

[46] R. Guo-Wu, W. C. Wang, J. Wang, Z. Zhen-Guo, H. Wei-Xiao, *Bioorg. Med. Chem. Lett.*, **23**, 6474 (2013).

[47] SPSS for windows (2010) computer program. Version 19.0. IBM Corp.

[48] H.D. Watts, M.N.A. Mohamed, J.D. Kubicki, *Phys. Chem. Chem. Phys.*, **13**, 20974 (2011).

[49] Y. Dai, Xu. Zhang, X. Zhang, H. Wang, Z. Lu, *J. Mol. Model.*, **14**, 807 (2008).

[50] B. Debnath, S. Gayen, A. Basu, K. Srikanth, T. Jha, *J. Mol. Model.*, **10**, 328 (2004).

[51] A. Sarkar, T. R. Middy, A.D. Jana, *J. Mol. Model.*, **18**, 2621 (2012).

[52] S.O. Podunavac-Kuzmanovic, D.D. Cvetkovic, D.J. Barna, *Int. J. Mol. Sci.*, **10**, 1670 (2009).

[53] M. Jalali-Heravi, A.J. Kyani, *J Chem Inf Comput Sci.*, **44**, 1328 (2004).

GENERAL CONCLUSION

General conclusion

This work was done on the study of two different applications: in the one hand, the first application contained chemical reactivity of 1,3-dipolar cycloaddition of 2,3,4,5-tetrahydropyridine-1-oxide with methyl crotonate. In the other hand, the second application performed biological activity of 1,2,4,5-tetrazine which included the reactivity study of this molecule also Quantitative structure–activity relationships (QSAR) of some 1,2,4,5-tetrazine derivatives.

- In the first application:

We have checked in the present theoretical study, the regioselectivity and stereoselectivity of 1,3-dipolar cycloaddition reaction of the cyclic nitrene: 2,3,4,5-tetrahydropyridine-1-oxide and methyl crotonate using the frontier molecular orbital analysis, CDFT and TST.

The analysis of DFT global reactivity indices, charge transfer analysis in the TSs, and FMO interactions reveal that all cyclizations under investigation had a normal electron demand character for the reaction. The regioselectivity has been analyzed and confirmed through DFT based indices, proved clearly that the Meta channels are the major. In addition we note that the solvent is slightly affects in the activation energies.

All this theoretical study shows a clear preference for the meta-regioselectivity of the cycloaddition process in conformity with the experimental data.

- In the second application:

We have started with the reactivity of 1,2,4,5-tetrazine in the gas and aqueous phases, where we have found that the results presented in both cases indicated that the application of the salvation modifies the values of the reactivity descriptors. Whereas, Fukui function values showed the most susceptible sites for electrophile and nucleophile attack in the gas phase are N1 and H7 respectively, while it showed the major sites are N1 and C3 for electrophile and nucleophile attack in the aqueous phase. Afterward, we have performed the study of Quantitative structure activity relationship (QSAR) on a series of 1,2,4,5-tetrazine derivatives as antitumor activity against lung cancer were established in gas and

General conclusion

aqueous phases. A multiple linear regression analysis was performed to derive quantitative structure activity relationship models which were further evaluated internally for the prediction of activity.

The regression models indicated that the higher values of descriptors in gas phase: f_{N2}^N , DM, and in aqueous phase: f_{N2}^N , GAP, qC3 increase the antitumor activity of the 1,2,4,5-tetrazine.

The QSAR models show that the descriptors derived from DFT can successfully be utilized to predict the antitumor activity of the 1,2,4,5-tetrazine molecules.

As plausible perspectives to this work, we suggest:

- To study the other selectivities (chemoselectivity, diastereoselectivity,...) of 1,3 dipole cycloaddition reactions.
- To study the effects influencing 1,3-dipole cycloaddition reactions such as catalysts, solvents, temperature, pressure, etc.
- To study 1,3-dipolar cycloaddition reactions in competition with Diels-Alder reactions .
- To study the method 3D-QSAR (Quantitative Structure Activity Relationship) for models development.

APPENDIX

Theoretical study of the regio- and stereoselectivity of the 1,3-DC reaction of 2,3,4,5-tetrahydropyridine-1-oxide with methyl crotonate

Halima. Hazhazi^(a), Youcef. Boumedjane^(a), Boulanouar. Messaoudi^{(b,c)*}

(a) Laboratoire de Chimie Moléculaire et Environnement (LCME), Equipe de Chimie Informatique et Pharmaceutique (ECIP), Faculté des Sciences et Sciences Exacte - Département des sciences de la matière, Université de Biskra, BP 145 RP, 07000 Biskra, Algérie

(b) Laboratoire de Thermodynamique Appliquée et Modélisation Moléculaire, Département de Chimie, Faculté des Sciences, Université A. Belkaid, BP 119, Tlemcen, 13000, Algérie

(c) Ecole Préparatoire en Sciences et Techniques Tlemcen, BP 165 RP Bel horizon, 13000 Tlemcen, Algérie

Abstract

A theoretical study of the regio- and stereoselectivities of the 1,3-dipolar cycloaddition reaction between methyl crotonate and 2,3,4,5-tetrahydropyridine-1-oxide has been carried out using density functional theory (DFT) calculations at B3LYP/6-31G(d) level of theory. Analysis of the global reactivity and local electrophilicity indices has been used to explaining the regioselectivity of the titled reaction. Overall, our results show that the studied 1, 3-dipolar cycloaddition reactions favor the formation of the meta-endo cycloadduct in both cases. The bond order and charge transfer at the transition states and activation energies indicate that these reactions proceed via an asynchronous concerted mechanism. Thermodynamic and kinetic quantities for the possible stereoisomeric and regioisomeric pathways have been calculated at gas and solvent phase. Solvent effects do not modify the gas-phase selectivities but slightly increases the reactivity of the reagents. A good concordance is found between the obtained results and the experimental outcomes.

* Corresponding author:

messaoudiboulanouar@gmail.com

Received 20 Sept 2016,

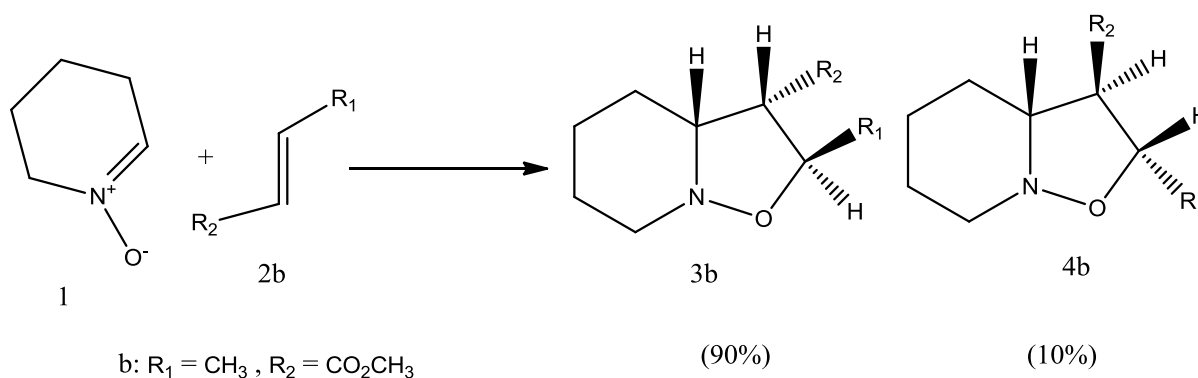
Revised 16 Dec 2016,

Accepted 16 Jan 2017

Keywords: 1,3-Dipolar cycloaddition; 2,3,4,5-tetrahydropyridine-1-oxide; Selectivity; DFT calculations.

1. Introduction

Cycloaddition reactions are among the most powerful tools for synthesis and mechanistic interest in organic chemistry area because of their capacity to construct in region- and /or stereoselectively method[1]. The current understanding of the underlying principles of 1,3- dipolar cycloaddition (1,3-DC) reactions has grown from the interplay between theory and experiment[2]. The 1,3-DC are a versatile manner for obtaining five-membered heterocycles[3]. Several experimental and theoretical studies continued to populate the literature over different manners of the 1,3 dipolar cycloaddition[4]. Reactions between nitrones and alkenes to obtaining isoxazolidines are well-known due to their great importance in construction processes[2-5]. Substituted isoxazolidines are interesting biological active compounds[6] that could be used as enzyme inhibitors[7-8], and applied as synthetic intermediates of a variety of compounds[9]. Many theoretical investigations have been devoted to the study of regio- and stereoselectivities of 1,3-DC reactions of nitron with alkenes[10]. Recently, reactivity descriptors based on the density functional theory (DFT) have been widely used for the prediction of the regioselectivity[11]. Kumar Das et al[4]. have studied the 1,3-dipolar cycloadditions of both 1-phenylethyl-trans-2-methyl nitron with styrene and 1-phenylethyl nitron with allyl alcohol. An analysis of frontier orbitals interaction, electrophilicity difference, Pauling's bond order and Wiberg bond index in the transition state was found to be in good agreement with the experimental data. A few years ago, Cossó et al[12]. have used B3LYP/6-31G(d) level of theory to study the 1,3-dipolar cycloaddition reaction of unsubstituted nitron with nitroethene. Asynchronicity in the bond formation process for the two regioisomeric approaches was found to be electron-deficient dipolarophile controlled. Liu et al[13]. have performed DFT calculations at B3LYP/6-31G(d) on the 1,3-dipolar cycloaddition reaction of the simplest nitron to dipolarophiles containing electron-releasing substituents. Another time again, the endo approach is kinetically favoured because of the stabilising secondary orbital interactions. Moreover, Nacerddine et al[14]. have studied the region and stereoselectivities of the 1,3- dipolar cycloaddition of C-diethoxyphosphoryl-N- methylnitron with substituted alkenes. The given analysis of potential energy surface shows that these 1,3-dipolar cycloaddition reactions favour the formation of the ortho-trans cycloadduct in both cases to be in good agreement with experimental findings. Very recently, Marakchi et al[15]. have studied the mechanism of the 1,3-dipolar cycloaddition reaction between nitron and sulfonyl ethene chloride using ab initio and DFT methods. HF and DFT calculations predict meta path, in agreement with the experimental results, while MP2 calculations provided the ortho regioselectivity. In another work, Khorief et al[16]. have studied the region and stereoselectivities of the 1,3-dipolar cycloaddition reaction between C-phenyl-N-methylnitron and Ethylvinylether. The ortho/ endo were produced to be more favorable kinetically and thermodynamically. Later, Chafaa et al[17]. have reported a DFT calculations at B3LYP/6-31+G(d) level of the 1,3-dipolar cycloaddition of C-diethoxyphosphoryl-N-methylnitron and N-(2-florophenyl)acrylamide. Analysis of the bond order and charge transfer at the transition states indicates a one -step asynchronous mechanism. Experimentally, Asrof et al[18]. have determined the 1,3-DC reaction between 2,3,4,5-tetrahydropyridine-1-oxide **1** and methyl crotonate **2b** to giving a mixture of substituted isoxazolidines (see Scheme 1) **3b** and **4b** with a ratio of 90:10, respectively. The stereochemistry of the major adduct is depicted in **4b** having endo oriented carbomethoxy group which is known to manifest favourable secondary orbital interaction. Our interest in this study is focused on the investigation and interpretation of the regio- and stereoselectivities of the 1,3-DC reactions of 2,3,4,5 tetrahydropyridine-1-oxide(dipole) with methyl crotonate (dipolarophile) by using several theoretical frontier molecular orbital (FMO) interactions, conceptual DFT, and the analysis of stationary points.



Scheme 1: 1,3-Dipolar cycloaddition of 2,3,4,5-tetrahydropyridine-1-oxide with methylcrotonate

2/Theory and computational details

The reported quantum chemical calculations were performed at the B3LYP/6-31G(d) level of theory using GAUSSIAN 09 suite of programs[19]. Full geometry optimizations followed by frequency calculations at the same level of theory were carried out for all stationary points. The intrinsic reactions coordinate (IRC)[20] path was calculated in order to check that each TS connects well to the two corresponding minima in the energy profiles of the proposed mechanism. Atomic electronic populations were computed using the natural bond orbital (NBO) method[21]. Bulk solvent effects of Dichloromethane were considered implicitly by performing single point energy calculations on the gas phase stationary structures using the polarisable continuum model (PCM) as developed by Tomasi's group[22] on the basis of the self-consistent reaction field (SCRf) background[23-24]. Values of enthalpy and free energy were obtained by frequency calculations over B3LYP/6-31G(d) geometries.

The global electrophilicity index ω [25], is given by

$$\omega = (\mu^2/2\eta) \quad (1)$$

where “ μ ” is the electronic chemical potential and “ η ” is the chemical hardness[26]. These quantities could be expressed as a function of the frontier molecular orbitals HOMO and LUMO, ϵ_H and ϵ_L , as:

$$\mu \approx (\epsilon_{\text{HOMO}} + \epsilon_{\text{LUMO}})/2 \quad (2)$$

$$\eta \approx (\epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}}) \quad (3)$$

A new empirical (relative) nucleophilicity index N [27] has been recently introduced on the basis of the HOMO energies[28]:

$$N = E_{\text{HOMO(nucleophile)}} - E_{\text{HOMO(TCE)}} \quad (4)$$

The tetracyanoethylene (TCE) is taken as a reference because of its lower HOMO energy in a large series of molecules [29].

The local electrophilicity index, ω_k [30] condensed to atom k is easily obtained by projecting the global quantity onto any atomic center k in the molecule by using the electrophilic Fukui index (i.e. the Fukui function for the nucleophilic attack f_k^+). This gives the following equation:

$$\omega_k = \omega f_k^+ \quad (5)$$

The local nucleophilicity condensed to atom k (N_k)[31] was then calculated as:

$$N_k = N f_k^- \quad (6)$$

For an atom k in a molecule, three different types of condensed Fukui function could be considered,[32]

$$f^+ = [\rho_k(N+1) - \rho_k(N)] \quad (\text{for nucleophilic attack}), \quad (7a)$$

$$f^- = [\rho_k(N) - \rho_k(N-1)] \quad (\text{for electrophilic attack}), \quad (7b)$$

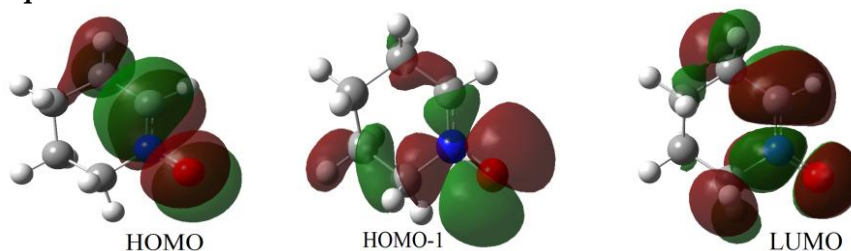
where $\rho_k(N)$, $\rho_k(N - 1)$ and $\rho_k(N + 1)$ are the gross electronic populations of the site k in neutral, cationic, and anionic systems, respectively.

3/Results and discussion

3.1 Regioselectivity study based on DFT reactivity indices

The determination of NED/IED reaction character is necessary for the prediction of regioselectivity. This characterization can be performed by using electronic chemical potential (μ), global electrophilicity (ω), global nucleophilicity N and HOMO–LUMO gap. This latter is calculated by considering the directly involved orbitals in the reaction. Frontier molecular orbitals Fig.1 analysis shows that HOMO and LUMO of the nitron are π molecular orbitals (MOs).

Dipole



Dipolarophile

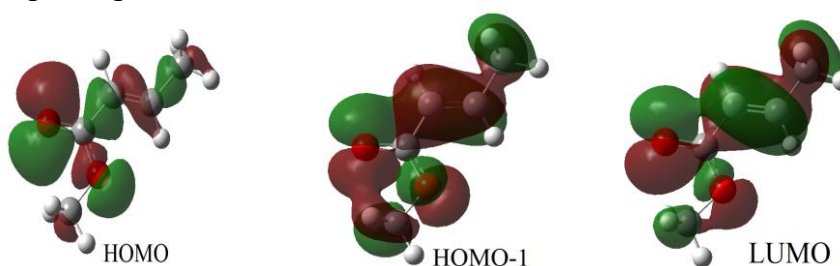


Fig1. Optimized geometries and visualized FMOs for the reactants

On one hand, the HOMO of the methyl crotonate is a nonbonding MO localized essentially on the carbonyl group's oxygen. Consequently, it will not be directly involved in the 1, 3-DC process. On the other hand, the dipolarophile's HOMO⁻¹ is a bonding π molecular orbital showing an important contribution of the active sites' atoms Fig.1 The HOMO-1 density (dipolarophile) is more pronounced than the HOMO ones. Thus, the cycloaddition reaction will take place between the LUMO of the dipole and the HOMO-1 of the dipolarophile. In the light of FMO theory, Sustmann[33] classified various types. Type (I) is the FMO interaction between the highest occupied molecular orbital of the 1, 3-dipole (HOMO_{dipole}) and the lowest unoccupied molecular orbital of the dipolarophile (LUMO_{dipolarophile}) that is corresponding the normal-electron demand (NED) $\Delta E(I) = \text{HOMO}_{\text{dipole}} - \text{LUMO}_{\text{dipolarophile}}$ (4.406 eV) a large number of 1,3-DC reactions is classified in. Type (II) interaction between the LUMO of dipole and the HOMO₋₁ of dipolarophile. This type is named as the inverse-electron demand (IED) $\Delta E(II) = \text{HOMO}_{-1 \text{ dipolarophile}} - \text{LUMO}_{\text{dipole}}$ (7.072 eV). These types of interactions are illustrated in Fig 2.

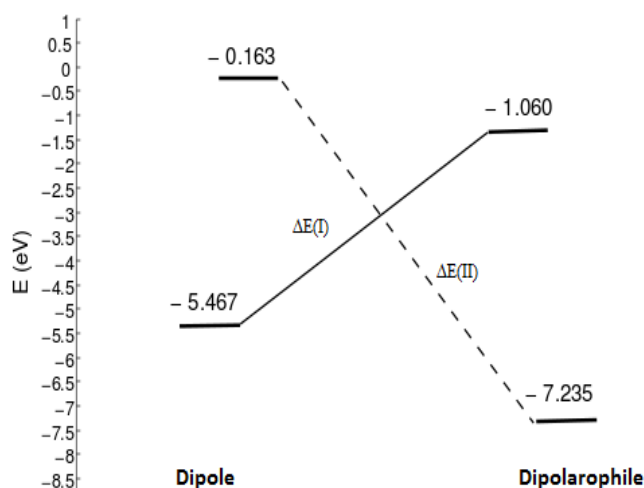


Fig 2. The interactions between HOMO and LUMO orbitals of a 1, 3-dipole/dipolarophile

According to the Houk rule[34], the regioselectivity of 1,3-DC reactions can be explained on the basis of large- large and small- small FMO interactions that are more favoured than large- small and small- large ones. The coefficients of frontier molecular orbitals (FMOs) of the reactants, given in Table 1, show that the most favoured interactions are between O_7 of the dipole and C_1 of the dipolarophile and C_1 of dipole interact with C_2 of dipolarophile leading to the formation of meta regioisomer as the major product. This fact is in agreement with the study of K. Marakchi et al[15] of 1,3-DC between pyrrolidine-1-oxide and methyl crotonate.

Table 1. FMOs Molecular Coefficients of the dipole and dipolarophile

Dipole		Dipolarophile					
HOMO	LUMO	HOMO-1		LUMO			
C1	O7	C1	O7	C1	C2	C1	C2
-0.3898	0.4677	0.3877	0.3021	0.3449	0.3850	0.4025	-0.2447

In Table 2 are displayed the HOMO and LUMO energies, electronic chemical potential μ , chemical hardness η , global electrophilicity ω , and global nucleophilicity N^a of the dipole and dipolarophile.

Table 2. FMO energies (a.u), electronic chemical potential (a.u), chemical hardness (a.u), electrophilicity index (eV) and nucleophilicity index (eV)

	HOMO	LUMO	μ	η	ω	N^a
Dipole	-0.201	-0.006	-0.103	0.195	0.740	3.649
Dipolarophile	-0.266	-0.039	-0.152	0.227	1.384	1.881

HOMO energy of tetracyanoethylene is -0.3351 a.u at the same level of theory. Chemical potential, hardness, electrophilicity and nucleophilicity values are associated to the HOMO-1 of dipolarophile.

The electronic chemical potential (μ) of dipole (-0.103 eV) is greater than that of dipolarophile (-0.152 eV). Consequently, the charge transfer will take place from the 1, 3-dipole to dipolarophile. The Values of electrophilicity indices (ω) of reactants are 1.384 and 0.740 eV for methyl crotonate (dipolarophile) and nitron (dipole), respectively. As it can be seen, the values of ΔE (HOMO-LUMO gap) for NED character are predicted to be lower than that corresponding to IED. Many studies dealing with cycloaddition reactions are based on the electrophilic (f_k^-) and nucleophilic (f_k^+), Parr functions attacks. Negative values of Fukui functions can be obtained from various population analyses. However, Hirshfeld's population [35] guarantees positive Fukui functions values, thus it is a good means to predict the right regioselectivity[36,37]. Therefore, Fukui functions based on the Hirshfeld's population were calculated for the studied reactants and the corresponding results are given in Table 3. Prediction of regioselectivity can also be performed by using the Chattaraj's polar model where local philicity indices are used[38]. Values of local electrophilicity ω_k and local nucleophilicity N_k are listed in Table 3 and the corresponding most favourable two-centre interaction is shown in Fig 3. In dipolar cycloaddition, the most favourable attack takes place between C1 atom of the dipolarophile (the preferred position for a nucleophilic attack) and O7 of the dipole, leading to the formation of the meta-regioisomer.

Table 3. Local properties of dipole and dipolarophiles calculated at B3LYP/6-31G (d) level of theory

Reactant	Site	NPA				Chelpg			
		f^+	f^-	ω	N	f^+	f^-	ω	N
Dipole	O ₇	0.189	0.423	0.140	1.543	0.211	0.355	0.156	1.295
	C ₁	0.298	0.312	0.220	1.138	0.435	0.268	0.180	0.977
Dipolarophile	C ₁	0.225	0.246	0.300	0.469	0.234	0.268	0.311	0.511
	C ₂	0.116	0.392	0.154	0.747	0.114	0.397	0.151	0.288

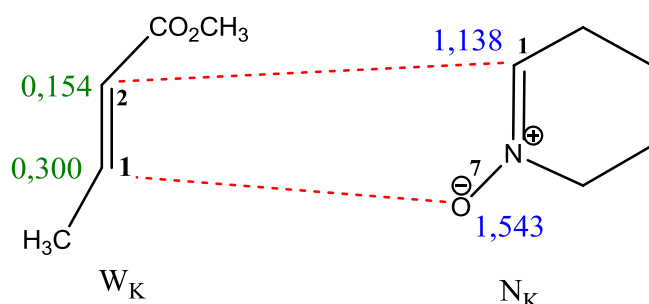
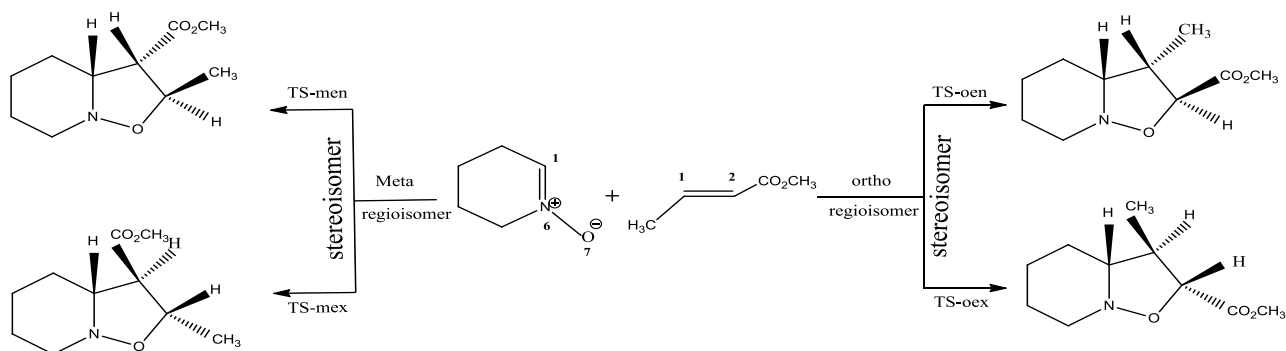


Fig 3. Prediction of the favoured interactions between dipole and dipolarophile using DFT based indices

3.2. Mechanistic study of the cycloaddition reaction based on activation energy

3.2.1. Energies of the Transition Structures

The 1-3 DC reactions of nitron with dipolarophile can happen along four possible reactive channels corresponding to the endo and exo approach modes in two different regioisomeric channels; the meta and ortho reactions (see Scheme 2). The right pathway corresponds to the O7-C2 and C1-C1 forming bond processes, while the left pathway corresponds to the O7-C1 and C1-C2 ones. Four transition states designated as: TS-men, TS-mex, TS-oen, and TS-oex have been confirmed by frequency calculations. The geometries of the four TSs are represented in Fig 4. The studied energies and relative energies of reactants in both the gas phase and DCM solvent are given in Table 4. The PES schemes, corresponding to the four reactive channels, are given in Fig 5 and 6.



Scheme 2. The exo and endo approaches of tetrahydropyridine -1-oxide to methyl crotonate.

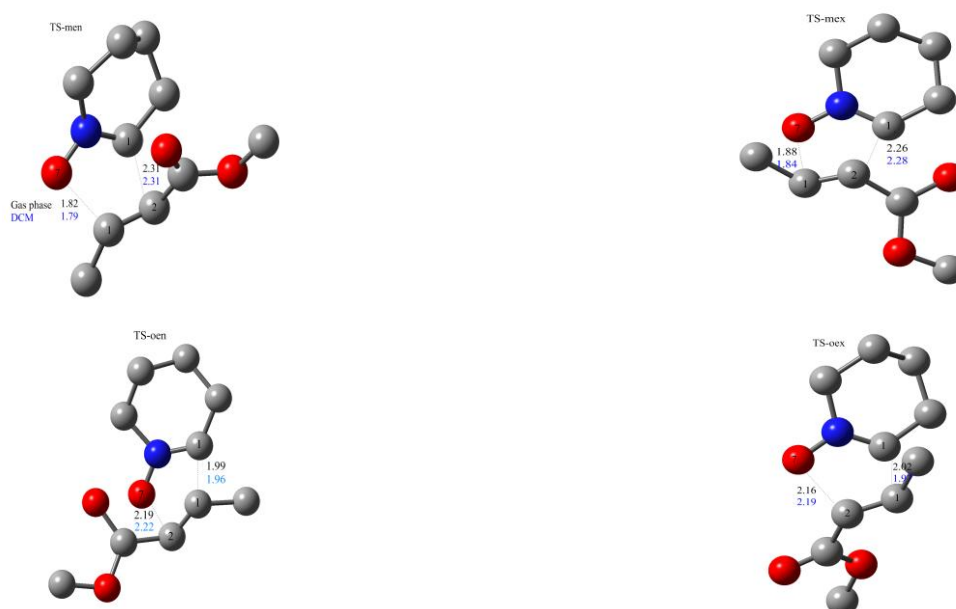


Fig 4. Optimized transition structures of the 1-3 DC reaction between tetrahydropyridine-1-oxide and methyl crotonate

The calculated activation energies and relative electronic energies of the stationary points involved in the 1,3-DC of 2,3,4,5-tetrahydropyridine-1-oxide with methylcrotonate in both cases the gas-phase and dichloromethane solvent are shown in Table 4. The activation energy **TS-men** (6.45 kcal/mol) is lower compared to that of the **TS-mex** (10.91 kcal/mol), **TS-oen** (11.25 kcal/mol) and **TS-oex** (15.47 kcal/mol). In dichloromethane, the reactants are slightly more stabilized than the TSs and this results the activation energies increase respectively **TS-men** (10.45), **TS-mex** (13.80), **TS-oen** (14.73) and **TS-oex** (17.43) kcal/mol, that is to say, the formation of pathways **P-men** in the gas phase is easier than that in the solvent. We have to note that these 1,3-DC reactions are strongly exothermic. In addition, solvent effects decreased their exothermicity while their corresponding activation energies increased as a consequence of the greater solvation of polar nitrene[39]. The comparison of relative energies with relative free energies given in Table 4 shows that the trends in region- and stereoselectivity are essentially coincident. Finally, we can conclude that the energy results indicate that the **P-men** shows a very high reactivity, both kinetically and thermodynamically.

Table 4. Total energies (a.u), Relative energies ΔE (in kcal/mol), relative free energies ΔG (in kcal/mol) and enthalpies ΔH (in kcal/mol) in gas phase and in DCM, of the stationary points involved in the 1, 3-DC reaction between 2,3,4,5-tetrahydropyridine-1-oxide and methyl crotonate

Stationary point	Gas phase				Dichloromethane solvent			
	ΔG	ΔH	ΔE^*	E_T	ΔG	ΔH	ΔE^*	E_T
Dipole				-325.8570278				-325.8656038
Dipolarophile				-345.7852858				-345.7903845
TS-men	21.75	8.05	6.45	-671.6320213	24.98	11.88	10.45	-671.6393249
TS-mex	25.71	12.38	10.91	-671.6249129	28.56	15.27	13.80	-671.6339965
P-men	-6.16	-20.20	-23.62	-671.6742828	-1.80	-15.70	-19.04	-671.6813970
P-mex	-2.72	-16.07	-19.56	-671.6734902	1.17	-12.20	-15.75	-671.6810906
TS-oex	30.52	16.97	15.47	-671.6176606	32.56	19.00	17.43	-671.6282077
TS-oen	26.25	12.76	11.25	-671.6243785	29.81	16.28	14.73	-671.6325104
P-oen	-3.87	-17.80	-15.56	-671.6671208	0.47	-13.70	-12.07	-671.6752383
P-oex	-2.69	-16.60	-13.77	-671.6642700	1.10	-12.70	-10.90	-671.6733701

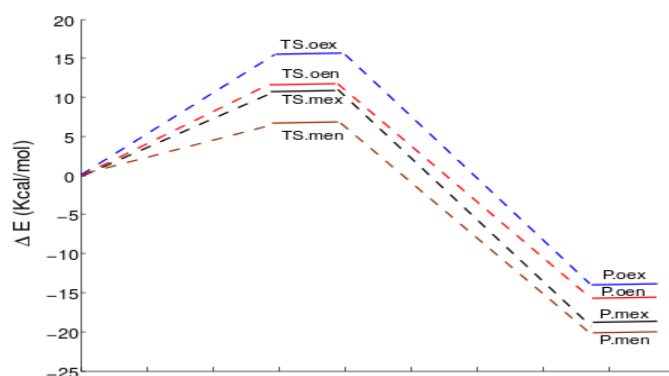


Fig 5. Energy profiles, in kcal/mol, for the 1,3-DC reactions of stationary point in gas phase

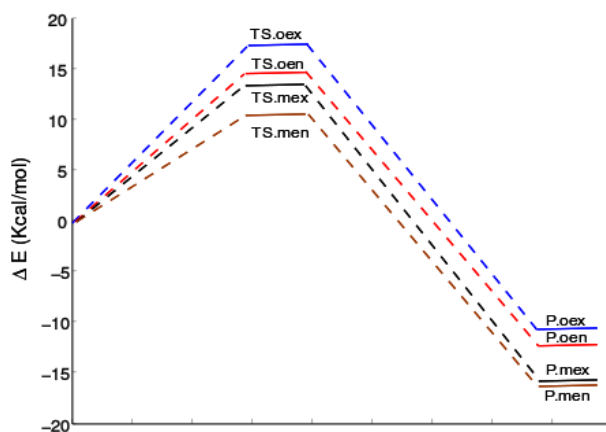


Fig 6. Energy profiles, in kcal/mol, for the 1,3-DC reactions of stationary point in solvent phase

3.2.2. Geometry analysis

Table 5. Values of $|\Delta d|$ in TS-men, TS-mex, TS-oen and TS-oex of the 1, 3-DC reaction of dipolarophile with dipole in the gas phase and solvent DCM

	Gas phase			DCM		
	Meta channels			Meta channels		
	d(O ₇ -C ₁)	d(C ₁ -C ₂)	Δd	d(O ₇ -C ₁)	d(C ₁ -C ₂)	Δd
TS-men	1.82	2.31	0.48	1.79	2.31	0.52
TS-mex	1.88	2.26	0.37	1.84	2.28	0.44
	Ortho channels			Ortho channels		
	d(O ₇ -C ₂)	d(C ₁ -C ₁)	Δd	d(O ₇ -C ₂)	d(C ₁ -C ₁)	Δd
TS-oen	2.19	1.99	0.20	2.22	1.96	0.26
TS-oex	2.16	2.02	0.14	2.19	1.97	0.40

Comparison of the most relevant geometrical parameters of the four TSs involved in gas phase and in solvent 13DCs of the 2,3,4,5-tetrahydropyridine-1-oxide with methyl crotonate is presented in Figure 4. The corresponding selected geometric parameters are given in Table 5. The lengths of the bonds of the C1-C2 are: (2.31 – 2.62) Å and O7-C1 (1.82 – 1.88) Å at the **TS-men** and **TS-mex**, respectively, while at the **TS-men** and **TS-mex** involved in the DCM are: C1-C2 (2.31-2.28) Å and O7-C1 (1.79 – 1.84) Å. These values indicate that they correspond to asynchronous bond formation processes where the lengths of the O-C are shorter than the C-C bonds. However, in the **TS-oen** and **TS-oex**, O-C forming bonds (2.19 and 2.16 Å) are longer than C-C forming bonds (1.99 and 2.02 Å); the lengths of the bonds in the solvent are the same as in the gas phase i.e. O-C is longer than C-C. This shows a change of the dissymmetry on the bond formation process for the two regioisomeric pathways.

The degree of asynchronicity of bond formation at the TSs is determined by considering the difference between the lengths of the two new σ forming bonds such as $\Delta d = [d_1 - d_2]$ for four pathways. Values of Δd for the **TS-men** are high asynchronous and more favorable stereoisomeric with respect to the other channels.

Conclusion:

In this work, the 13DC reaction of 2,3,4,5-tetrahydropyridine-1-oxide with methyl crotonate, the regio- and stereoselectivities have been thoroughly probed using DFT methods at the B3LYP/6-31G(d) theoretical level. The ortho/meta regioisomeric pathways along with the endo and exo stereoisomeric channels have been studied on the basis of both kinetic and thermodynamic controls. The regioselectivity has been analyzed and confirmed through DFT based indices. The calculated electrophilic, f_k^+ , and nucleophilic, f_k^- , indices prove clearly that the meta channels are the major and most favourable regioisomeric paths. In all the studied cases, the reaction pathways leading to the endo/meta are the most favourable. These 13DCs have basically asynchronous concerted mechanisms and not one presents a stepwise mechanism as a consequence there was no intermediate localization along the present study. The solvent slightly affects the activation energies due to the better solvation of the reactants. All in all, the obtained results using the present theoretical approaches are in good agreement with experimental data.

Acknowledgements

The Ministry of Higher Education and Scientific Research of the Algerian Government (project CNEPRU F01420130050) is gratefully acknowledged for its support.

Supplementary material

The Cartesian coordinates of the important optimized stationary geometries (reactants, intermediates, TSs and products) along the 13DC reaction of 2,3,4,5-tetrahydropyridine-1-oxide with methyl crotonate at B3LYP/6-31 G(d) level of theory are given in supplementary material.

References

- [1]. W. Carruthers. In *Some Modern Methods of organic Synthesis*; 2nd ed.; Cambridge University Press: Cambridge, 1978; W. Carruthers. In *Cycloaddition Reactions in Organic Synthesis*; ed.; J.E. Baldwin and P.D. Magnus; Pergamon: Oxford, 1990.
- [2]. A. Padwa, In *Comprehensive Organic Chemistry*; Pergamon Press: Oxford, 1991; pp 1069-1109.
- [3]. (a) A. Padwa, In *1,3-Dipolar Cycloaddition Chemistry*; Wiley interscience: New York, (1984) pp 1-2; (b) K.V. Gothelf, K. A. Jorgenson, *Chem Rev.*, 98 (1998) 863-910.
- [4]. T. K. Das, S. Salampuria, M. Banerjee, *J. Mol. Struct. Theochem.*, 959 (2010) 22-29.
- [5]. (a) A. Padwa, Y. Tomioka, M. K. Venkatramanan, *Tetrahedron Lett.*, 28 (1987) 755-758; (b) E. Breuer, H. G. Aurich, In *Nielsen Nitrones, nitronates and nitroxides*; New York: Wiley, 1989; (c) K. B. G. Torssell, In *Nitrile oxides, nitrones and nitronates in organic synthesis*; New York: VCH, 1998; (d) K. V. Gothelf, K. A. Jørgensen, *Chem. Rev.*, 98 (1998) 863-910; (e) S. Majumder, P. Bhuyan, *J. Tetrahedron Lett.*, 53 (2012) 762-764.
- [6]. A. R. Minter, B. B. Brennan, A. K. Mapp, *J. Am. Chem. Soc.*, 126 (2004) 10504-10505.
- [7]. P. Ding, M. Miller, Y. Chen, P. Helquist, A. J. Oliver, O. Wiest, *Org. Lett.*, 6 (2004) 1805-1808.
- [8]. G. Wess, W. Kramer, G. Schuber, A. Enhsen, K. H. Baringhaus, H. Globmik, S. Müller, K. Bock, H. Klein, M. John, G. Neckermann, A. Hoffmann, *Tetrahedron Lett.*, 34 (1993) 819.
- [9]. K. Marakchi, R. Ghailane, O. Kabbaj, N. Komaha, *J. Chem. Sci.*, 126 (2014) 283-292.
- [10]. S. Bouacha, A. H. Djerourou, I. Chemo, *J. Sci.*, 4 (2013) 1941-3955.
- [11]. R. G. Parr, R.G. Pearson, *J. Am. Chem. Soc.*, 105 (1983) 7512-7516.
- [12]. F. P. Cossío, I. Marao, H. Jiao, P. V. R. Schleyer, *J. Am. Chem. Soc.*, 121 (1999) 6737-6746.
- [13]. J. Liu, S. Niwayama, Y. You, K. N. Houk, *J. Org. Chem.*, 63 (1998) 1064-1073.
- [14]. A. K. Nacereddine, W. Yahia, S. Bouacha, A. Djerourou, *Tetrahedron Lett.*, 51 (2010) 2617- 2621.
- [15]. K. Marakchi, O. Kabbaj, N. Komaha, R. Jalal, M. Esseffar, *J. Mol. Struct. Theochem.*, 620 (2003) 271-281.
- [16]. A. K. Nacerreddi, W. Yahia, C. Sobhi, H. Layeb, Z. Lechtar, A. Djerourou, *J. Appl. Biopharm. Pharmacokinet.*, 1 (2013) 18-23.
- [17]. F. Chafaa, D. Hellel, A. K. Nacereddine, A. Djerourou, *Mol. Phy.*, 114 (2015) 663-670.
- [18]. Sk. A. Ali, J. H. Khan, M. I. M. Wazeer, P. H. Perzanowski. *Tetrahedron.*, 45 (1989) 5979-5986.
- [19]. Gaussian 09, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, F. J. B. Farkas, J. V. Ortiz, J. Cioslowski, D. Fox, J. Wallingford, CT, 2009.

- [20]. K. Fukui, *J. Phys. Chem.*, 74 (1970) 4161–4163.
- [21]. A. E. Reed, F. Weinhold, *J. Chem. Phys.*, 78 (1983) 4066-4073.
- [22]. J. Tomasi, M. Persico, *Chem. Rev.*, 94 (1994) 2017–2094.
- [23]. E. Cancès, B. Mennucci, J. Tomasi, *J. Chem. Phys.*, 107 (1997) 3032–3041.
- [24]. M. Cossi, V. Barone, R. Cammi, J. Tomasi, *Chem. Phys. Lett.*, 255 (1996) 327–335.
- [25]. R. G. Parr, L. von Szentpaly, S. Liu, *J. Am. Chem. Soc.*, 121 (1999) 1922-1924.
- [26]. (a) R. G. Parr, R. G. Pearson, *J. Am. Chem. Soc.*, 105 (1983) 7512-7514; (b) R. G. Parr, W. Yang, In *Density Functional Theory of Atoms and Molecules*; Oxford University: New York., (1989).
- [27]. (a) L. R. Domingo, E. Chamorro, P. Pérez, *J. Org. Chem.*, 73 (2008) 4615-4624; (b) L. R. Domingo, P. Pérez, *Org. Biomol. Chem.*, 9 (2011) 7168-7175.
- [28]. W. Kohn, L. Sham, *J. Phys. Rev.*, 140 (1965) 1133-1138.
- [29]. (a) L. R. Domingo, E. Chamorro, P. Pérez, *J. Org. Chem.*, 73(12) (2008) 4615-4624. (b) P. Jaramillo, L. R. Domingo, E. Chamorro, P. Pérez, *J. Mol. Struct. Theochem.*, 865(2008) 68-72.
- [30]. L. R. Domingo, M. J. Aurell, P. Pérez, R. Contreras, *J. Phys. Chem. A.*, 106 (2002) 6871.
- [31]. P. Pérez, L. R. Domingo, M. Duque-Norna, E. Chamorro, *J. Mol. Struct. Theochem.*, 895 (2009) 86.
- [32]. W. Yang, W. J. Mortier, *J. Am. Chem. Soc.*, 108 (1986) 5708.
- [33]. R. Sustmann, R. Shubert, *Tetrahedron Lett.*, 13 (1972) 4271.
- [34]. K.N. Houk, *Accounts Chem Res.*, 8(11) (1975) 361-69.
- [35]. F.L. Hirshfeld, *Theor. Chem. Acc.*, 44 (1977) 129-38.
- [36]. J. Padmanabhan, R. Parthasarathi, U. Sarkar, V. Subramanian, P.K. Chattaraj, *Chem. Phys. Lett.*, 383 (2004) 122.
- [37]. R. K. Roy, S. Pal, K. Hirao, *J. Chem. Phys.*, 110 (1999) 8236.
- [38]. R.G. Parr, L. Von Szentpaly, S. Liu, *J. Am. Chem. Soc.*, 121 (1999) 1922.
- [39]. K. Marakchi, R. Ghailane, O. Kabbaj, N. Komiha, *J. Chem. Sci.*, 126 (2014) 283–292.

Abstract

The objective of the present thesis is the theoretical study of some organic compounds, which is divided into two different applications:

- i) 1,3-dipolar cycloaddition reaction of methyl crotonate with 2,3,4,5 tetrahydropyridine 1-oxide for interpretation and prediction the regio and stereoselectivity by means of several theoretical approaches, namely, the transition state theory state (TST) - Theory of frontier molecular orbital (FMO) - Indices derived from conceptual DFT. The findings obtained in this work are in agreement with experimental data
- ii) QSAR studies have been carried out on eighteen molecules of 1,2,4,5-tetrazine derivatives, firstly we have analyzed the molecular reactivity of 1,2,4,5-tetrazine by conceptual DFT. Also, multiple linear regression (MLR) procedure was used to find the correlation between chemical descriptors and antitumor activity of 1,2,4,5-tetrazine derivatives and the leave-one-out (LOO) method to estimate the predictivity of our models. High correlation between experimental and predicted of the activity values was observed, indicating the validation and the good quality of the derived QSAR models.

Keywords: 1,3-Dipolar reaction, Regioselectivity, Stereoselectivity, Effect of solvent, Conceptual DFT, 1,2,4,5-tetrazine, QSAR, MLR.

Résumé

Le travail présenté dans cette thèse, qui a pour objectif l'étude théorique de quelques composés organiques, est divisé en deux applications différentes:

- i) Réaction de cycloaddition 1,3-dipolaire du crotonate de méthyle avec le 2,3,4,5 tétrahydropyridine 1-oxyde pour l'interprétation, la prédiction de la régio et la stéréosélectivité au moyen de différentes approches quantiques, en l'occurrence, la Théorie de l'état de transition (TST)-Théorie des orbitales moléculaires frontières (FMO)- Indices dérivant de la DFT conceptuelle. Les résultats obtenus dans ce travail sont en accord avec les constatations expérimentales
- ii) Une étude QSAR a été effectuée sur dix-huit molécules de dérivés de 1,2,4,5-tétrazine. Nous avons tout d'abord analysé la réactivité moléculaire de 1,2,4,5-tétrazine par DFT conceptuelle. De plus, la régression linéaire multiple a été utilisée pour déterminer la corrélation entre les descripteurs chimiques et l'activité antitumorale des dérivés du 1,2,4,5-tétrazine et la méthode de validation croisée (LOO) utilisé pour la prédiction des modèles obtenus. Une forte corrélation a été observée entre les valeurs expérimentales et les valeurs prédites de l'activité antitumoral, ce qui indique la validité et la qualité des modèles QSAR obtenus.

Mots-clés : Réaction 1,3-Dipolaire, Régiosélectivité, Stéréosélectivité, Effet de solvant, DFT conceptuelle, 1,2,4,5-tétrazine, QSAR, MLR.