

MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA  
RECHERCHE SCIENTIFIQUE

UNIVERSITE MOHAMED KHIDER BISKRA

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

Département des Sciences de la matière

**THESE**

Présentée par

**YOUCEF Oumhani**

En vue de l'obtention du diplôme de :

**DOCTORAT EN SCIENCES**

**Option :**

Chimie théorique et pharmaceutique

*Intitulée :*

*Etude par des calculs quantiques et empiriques, des  
propriétés moléculaires dans des hétérocycles à intérêt  
pharmaceutique*

Soutenue le : 18/06/2016

Devant la commission d'Examen

M. BARKAT Djamel	Professeur	Univ. Mohamed Khider - Biskra	Président
M. BELAIDI Salah	Professeur	Univ. Mohamed Khider - Biskra	Directeur de thèse
M. DIBI Ammar	Professeur	Univ. Hadj Lakhdar - Batna	Examineur
M. LANEZ Touhami	Professeur	Univ. Hamma Lakhdar -El Oued	Examineur
M. TERKI Belgacem	Professeur	Univ. Kasdi Merbah - Ouargla	Examineur

*Study by quantum and empirical calculations of  
molecular properties in heterocyclic pharmaceutical  
interest*

*To my parents*

*YOUCEF Messaoud*

*and*

*HASSANI Zohra*

## *Acknowledgements*

First of all, i thank allah for all his gifts of life. Wich allah grant us the best.

Secondly, I would thank my supervisor Mr. BELAIDI Saleh, professor at Biskra University for his efforts, and his critical spirit in my work.

I also have the honor to thank the members of my jury: professor DIBI Ammar from the university of Batna, professor BARKAT Djamel from the university of Biskra, professor LANEZ Touhami from the university of El-Oued, and professor TERKI Belgacem from the university of Ouargla, for judging my working thesis.

I also deeply thank my colleagues in Group of Computational and Pharmaceutical Chemistry, LMCE Laboratory, University of Biskra, and my friends for their spirit of collaboration.

My gratitude goes to my family, my parents, my brothers Imad and Ashraf and darling Saida for their help, their patience and encouragement which embosomed my work.

# Table of contents

Acknowledgements.....	iv
Table of contents.....	v
List of Figures .....	ix
List of Tables.....	xi

<b>General Introduction .....</b>	<b>1</b>
References .....	3

## **CHAPTER I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides**

I.1. Physicochemical properties and chemical reactivity of DNA .....	6
I.1. Introduction .....	6
I.1.2. Structure of Deoxyribonucleic Acid.....	7
I.1.2.1. Primary Structure .....	7
I.1.2.2. Secondary Structure .....	9
I.1.2.2.1. Sugar-Phosphate Backbone and the Glycosyl Bond .....	9
I.1.2.2.2. Sugar Pucker .....	11
I.1.2.3. Hydrogen Bonding and Base Stacking .....	13
I.1.3. From DNA to proteins .....	15
I.1.4. DNA damage .....	16
I.1.4.1. Types of DNA damage .....	16
I.1.4.2. Simple adducts .....	17
I.1.4.2.1. Oxidation .....	17
I.1.4.2.2. Alkylation .....	18
I.1.4.2.3. Hydrolysis .....	18
I.1.4.3. The biological effect of radiation .....	19
I.1.4.4. Environmentally caused DNA double strand breaks .....	20
I.1.4.5 DNA double-strand repair pathways - A short overview.....	21
I.2. Classification and pharmacological properties of macrolides .....	23

I.2.1. Macrolide structures .....	23
I.2.2. Chemical structure of macrolide antibiotics .....	25
I.2.3. The need for novel antibacterial agents .....	26
I.2.3.1. New antibiotics .....	27
I.2.4. Macrolide-binding site in the ribosome .....	27
I.2.5. Modes of inhibition of protein synthesis by macrolides .....	29
I.2.6. Antibiotic Resistance .....	30
I.2.6.1. Macrolide action and drug resistance .....	31
I.2.7. Erythromycin .....	32
I.2.7.1. Metabolism of erythromycin A and clarithromycin .....	34
I.2.8. 12 Membered $\alpha,\beta$ -unsaturated macrolides .....	36
I.2.8.1. Patulolides .....	36
I.2.8.2. Cladospolides .....	37
I.2.8.3. Mycolactone .....	38
References .....	39

## CHAPTER II: Molecular Modeling

II.1. Introduction. ....	44
II.2. Quantum mechanics .....	45
II.3. Quantum mechanics Calculations .....	46
II.3.1. Schrödinger equation .....	47
II.3.2. Born-Oppenheimer approximation .....	47
II.3.3. Hartree-Fock calculation (HF) .....	48
II.3.4. Basis sets .....	48
II.3.5. Density Functional Theory (DFT) .....	49
II.4. Semi-empirical Calculations .....	50
II.5. Molecular Dynamics .....	51
II.6. Molecular mechanics .....	51
II.6.1. Bonded interactions .....	52
II.6.2. Bond stretching .....	52
II.6.3. Angle bending .....	53
II.6.4. Torsional terms .....	53

II.6.5. Non-bonded interactions .....	54
II.6.6. Electrostatic interactions .....	54
II.6.7. Van der Waals interactions .....	55
II.7. The force field .....	56
II.7.1. Type of force field .....	57
II.7.1.1. AMBER .....	57
II.8. Minimization Methods .....	58
II.9. Computer Simulations Methods .....	60
II.9.1. Energy Minimization Method .....	60
II.9.2 Steric energy .....	61
II.10. Monte Carlo (MC) Method .....	64
II.11. Structure-Activity Relationships (SAR) .....	66
II.11.1. QSAR .....	66
References .....	67

### **CHAPTER III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA**

III.1. Introduction. ....	71
III.2. Computational Methods.....	72
III.3. Results and Discussion .....	73
III.3.1. Geometric and Electronic structure of thymine, thymidine and thymidylic acid.....	73
III.3.2. Molecular geometry of thymidine 1-3a .....	82
III.3.3. Molecular geometry of thymidylic acid 1-3b .....	87
III.4. Conclusion .....	89
References .....	89

## **CHAPTER IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides**

IV.1. Introduction .....	93
IV.2. Computational methods .....	95
IV.3. Results & discussion .....	96
IV.3.1. Conformational analysis of 12-membered $\alpha$ , $\beta$ -unsaturated macrolides .....	96
IV.3.2. Geometric and Electronic Structure of 12-membered Macrolide 12d (1-3) (Type 3) .....	100
IV.3.3. Structure - activity relationships and rules of five for 12-membered macrolides .....	103
IV.4. Conclusion .....	109
References .....	109
<b>General Conclusion .....</b>	<b>114</b>

## *List of Figures*

<b>Figure 1.1:</b> Structure of nucleosides.....	8
<b>Figure 1.2:</b> Primary structure for the tetradeoxynucleotide adenylyl-3',5'-thymidylyl-3',5'-guanylyl-3',5'-cytidine (d(ATGC)).....	9
<b>Figure 1.3:</b> Diagram of torsion angles in a nucleotide along the sugar-phosphate backbone and the glycosyl bond .....	11
<b>Figure 1.4:</b> Different sterically permissible sugar pucker conformations including 3'-endo, 2'-endo and 2'-exo-3'-endo .....	13
<b>Figure 1.5:</b> Central dogma of biology schematically illustrated .....	15
<b>Figure 1.6:</b> A representation of the different types of DNA damage that can occur after exposure to chemicals or radiation .....	17
<b>Figure 1.7:</b> Examples for antimicrobials macrolides .....	25
<b>Figure 1.8:</b> Chemical structures of the macrolide antibiotics. ....	26
<b>Figure 1.9:</b> The 2nd generation erythromycins and the synthetic pathways to them .....	33
<b>Figure 1.10:</b> The degradation of erythromycin A under acidic condition .....	34
<b>Figure 1.11:</b> Chemical structures of representative macrolide antibiotics. The numbering of the macrolactone atoms is illustrated on the erythromycin structure .....	36
<b>Figure 1.12:</b> Cladospolide family .....	38
<b>Figure 1.13:</b> Mycolactone B .....	39
<b>Figure 2.1:</b> Accuracy vs. time of the different molecular modeling methods.....	45
<b>Figure 2.2:</b> Bond Stretching .....	62
<b>Figure 2.3:</b> Bond Bending .....	63

<b>Figure 2.4:</b> Stretch-Bend Interaction .....	63
<b>Figure 2.5:</b> Out of Plane Bending .....	64
<b>Figure 3.1:</b> 3D structures of thymine, thymidine and thymidylic acid .....	74
<b>Figure 3.2:</b> Sigmatropic hydrogen migration in thymine (a – f) .....	75
<b>Figure 3.3:</b> Sigmatropic hydrogen migration in thymidine (1a-3a) and thymidylic acid (1b-3b) .....	78
<b>Figure 3.4:</b> Molecular electrostatic potential surface of thymidine 1a and thymidylic acid 1b by DFT Method .....	81
<b>Figure 4.1:</b> $\alpha$ , $\beta$ -unsaturated macrolactone .....	97
<b>Figure 4.2:</b> Main conformational types of macrolides .....	97
<b>Figure 4.3:</b> 3D structure of the most favored conformer in 12 macrolide 12d, 1-3 (HyperChem) .....	98
<b>Figure 4.4:</b> 3D MESP contour map for 12d (1-3) macrolide/ type 3.....	103
<b>Figure 4.5:</b> 3D Conformation of cladospolide B (Gauss View 3.09). .....	103
<b>Figure 4.6:</b> Chemical structure of the studied macrolides.....	104

## *List of Tables*

<b>Table 3.1:</b> Heats of formation $\Delta H^\circ_f$ (in kcal/mole), dipole moments (in Debye), $\Delta E$ (HOMO/LUMO) and atomic charges for thymine a to f by PM3 Method .....	76
<b>Table 3.2-a:</b> Heats of formation $\Delta H^\circ_f$ (in kcal/mole), dipole moments (in Debye), $\Delta E$ (HOMO/LUMO) and atomic charges for thymidine 1 to 3 by PM3 Method .....	79
<b>Table 3.2-b:</b> Heats of formation $\Delta H^\circ_f$ (in kcal/mole), dipole moments (in Debye), $\Delta E$ (HOMO/LUMO) and the atomic charges for thymidylic acid 1 to 3 by PM3 Method .....	80
<b>Table 3.3:</b> Bond lengths ( $\text{Å}^\circ$ ) of thymidine 1-3a and thymidylic acid 1-3b by PM3 Method .....	84
<b>Table 3.4:</b> Bond angles ( $^\circ$ ) of thymidine 1-3a and thymidylic acid 1-3b by PM3 Method. ....	85
<b>Table 3.5:</b> Dihedral angles ( $^\circ$ ) of thymidine tautomers 1-3a by PM3 Method. ....	86
<b>Table 3.6:</b> Dihedral angles ( $^\circ$ ) of thymidylic acid tautomers 1-3b by PM3 Method .....	88
<b>Table 4.1:</b> Energetic difference and Boltzmann population for different macrolide types .....	98
<b>Table 4.2:</b> Calculated values of dihedral angles of 12 symmetric macrolide / T3.... .....	99

<b>Table 4.3:</b> Calculated values of dihedral angles of 12 dissymmetric macrolide /T3 for ester and diene system .....	100
<b>Table 4.4:</b> Calculated values of bond lengths of 12d (1-3) macrolide /T3 .....	101
<b>Table 4.5:</b> Calculated values of angles of 12d (1-3) macrolide /T3.....	101
<b>Table 4.6:</b> QSAR proprieties for 12-membered macrolides .....	105

# General Introduction

Computational chemistry is supporting research of new chemical compound. It uses sophisticated software to assist in the identification of new chemical compound. The theoretical chemist must be able to predict and reinterpret the experience using molecular modeling.

Molecular modeling (quantum mechanics, molecular mechanics and molecular dynamics), is the sum of theoretical methods and computational techniques that is used to predict molecular behaviors specifically interactions between molecules [1].

Quantitative structure – activity relationships (QSARs), as one of the most important areas in chemometrics [2]. QSARs are a suite of tools used to link chemical activities with molecular structure and composition [3], and is actively used in drug design [4, 5]. The concept of using QSARs to link structure and activity was introduced over 100 years ago and subsequently widely used in medical and biological research [6, 7]. To develop a QSAR model, several statistic methods can be used. [8].

Molecular descriptors define the molecular structure and physicochemical properties of molecules by a single number. A wide variety of descriptors have been reported for using in QSAR analysis [2].

## General Introduction.

---

This research is located in the context of a fundamental and original research on two types of molecules, macrolides to 12 ring members and containing molecules into DNA, namely thymine, thymidine and thymidylic acid.

The main objective of this work is the application of different methods of molecular modeling to predict the chemical reactivity and the expected biological activity in new bioactive molecules for macrolides and study for derivatives of DNA; sigmatropic hydrogen migration bonding.

This work comprises four chapters:

The first chapter is divided into two parts: in the first part, we present generalities on the physicochemical properties and the chemical reactivity of DNA. In the second part, we will show, on generalities regarding classification and the pharmacological properties of macrolides.

In the second chapter, we present a theoretical overview on the various methods of molecular modeling used in our work (quantum mechanics (QM): DFT and ab-initio and molecular mechanics (MM).

The third chapter is a study by molecular modeling sigmatropic hydrogen migration in the formation of Tautomeric Forms for the following molecules: thymine, thymidine and thymidylic acid in DNA , ( work published in Journal : Quantum Matter, Volume 5, Number 5, pp. 1-7 ; 2016).

The fourth Chapter is divided into two parts: the first part includes a structural, electronic and energy study of macrolides derivatives to 12 membered. In this part we

present the results of a comparative study of two methods used in the calculation, so the substitution effect on energy and electronics parameters of this molecules.

The second part is conformational analysis and stereoselectivity 12 membered macrolides. In this part we realized a study by molecular mechanics of the conformations of these macrolides. At the end of chapter we present a qualitative study of properties and structure activity relationships in a selected series of macrolides, ( work published in: *Journal of Computational and Theoretical Nanoscience*, Volume 12, Number 11, , pp. 1-7 ; December 2015).

### References:

- [1] V. Jahed, A. Zarrabi, A. Bordbar, M.S. Hafezi. *Food Chemistry* 165, 241–246, (2014).
- [2] R. Miri, M. Akhond, M. Shamsipur. B. Hemmateenejad. *Chemometrics and Intelligent Laboratory Systems* 64. 91–99. (2002).
- [3] B. Chen, T. Zhang, T. Bond, Y. Gan. *Journal of Hazardous, Materials* 299. 260–279. (2015).
- [4] H.M Patel, M.N Noolvi, P. Sharma, V. Jaiswal, S Bansal, S. Lohan, S. S. Kumar, V. Abbot, S. Dhiman, V. Bhardwaj . *Medicinal Chemistry Research*. 23 (12), 4991-5007. (2014).
- [5] C. Nantasenamat , C. Isarankura-Na-Ayudhya, T. Naenna, V. Prachayasittikul. *Excli Journal*. 8.74-88. (2009).
- [6] M. Dehmer, K. Varmuza, D. Bonchev, F. Emmert-Streib, *Statistical Modelling of Molecular Descriptors in QSAR/QSPR*, John Wiley & Sons, (2012).

## General Introduction.

---

[7] C. Hansch, A. Leo, Exploring QSAR: Fundamentals and Applications in Chemistry and Biology, American Chemical Society, (1995).

[8] V. Guillén-Casla, N. Rosales-Conrado, M. E. León-González, L. V. Pérez- Arribas, and L. María Polo-Díez, *Journal of Food Composition and Analysis*. 24, 456, (2011).

**Chapter I:**

**Physicochemical properties  
and chemical reactivity of DNA  
and classification and pharmacological  
properties of macrolides.**

## **I.1. Physicochemical properties and chemical reactivity of DNA**

### **I.1.1. Introduction**

Nucleotides are the most versatile of the major molecular living cell constituents.

They are notable for their involvement in reactions essential to the propagation and maintenance of life. More specifically, nucleotides are necessary for energy transfer, storage and the decoding of genetic information as well as having a structural and catalytic role in the cell. Deoxyribonucleic acid (DNA) is composed of a polymer of nucleotides. In order to fully comprehend the function of DNA and exploit it to its fullest, it is important to have a complete understanding of its primary and secondary structure. Having an appreciation for DNA structure will not only help us understand its biological role, but will aid in the attempt to use DNA as a constituent in nanotechnology and genetic engineering. The DNA molecule has many appealing features for use in nanotechnology, including its miniscule size and consistent diameter, its short structural repeat and its stiffness [1; 2]. The better we understand DNA's three dimensional structure, the better we will understand its biological role and its potential as a tool for genetic engineering and nanotechnology. [2]

In 1953, the first structure of DNA was elucidated by Watson and Crick in the journal *Nature* [3; 4] and given the designation B-DNA. This particular duplex is characterized by a 3.4°A rise of the stacked base pairs and 10 base pairs per helical turn. This DNA duplex conformation is considered the "normal" or "standard" form. In conjunction with Watson and Crick, Franklin and Gosling were working on the first alternative form of DNA, A-DNA (Franklin and Gosling 1953). This conformation

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

arose from the dehydration of B-DNA fibers; and this change was found to be reversible upon rehydration of the DNA duplexes. The A-DNA conformation is characterized by a 2.7 °A rise and 11 base pairs per helical turn [4].

The structure was solved in the Cavendish Laboratory, Cambridge by Francis Crick and James Watson, using X-ray diffraction data from fibers of DNA obtained by Rosalind Franklin at King's College, London [5].

Since the discoveries of 1953, the biological function of DNA has completely opened up, and a new era of research has begun. The entire human genome has now been sequenced as well as the genomes of several other species. Using X-ray diffraction it is now possible to study in more detail the structure of short oligonucleotides as well as DNA complexes with proteins, drugs or metals among other molecules. Besides the enormous amount of information we have gained about the biological function of DNA, this understanding has also opened the doors for genetic engineering and nanotechnology [2].

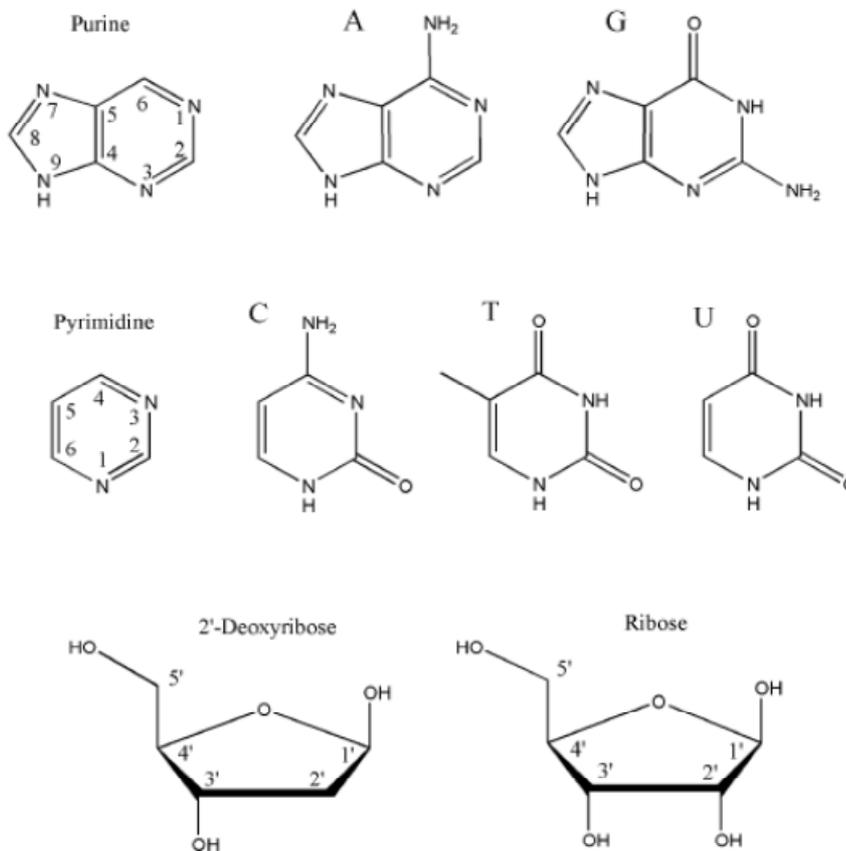
### **I.1.2. Structure of Deoxyribonucleic Acid**

#### **I.1.2.1 Primary Structure**

DNA and RNA molecules (also known as nucleic acids) are comprised of a long chain polymer of nucleotides. Nucleotides consist of a nitrogenous base: Adenine (A), Cytosine (C), Guanine (G) and Thymine (T) for DNA or Uracil (U) for RNA. (Figure 1.1) [6; 7].

## Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.

---



**Figure 1.1:** structure of nucleosides.

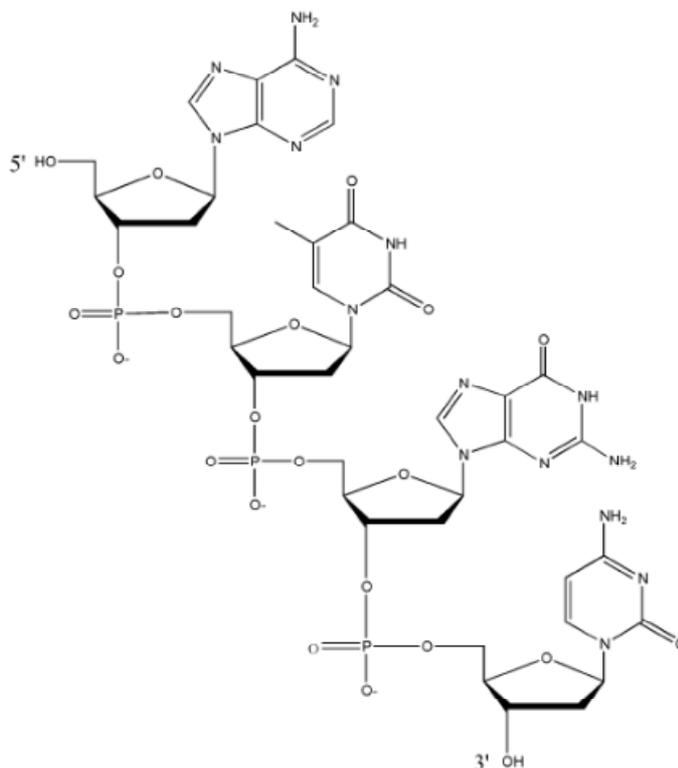
If the base is attached to a  $\beta$ -D-ribose, the molecule is termed a nucleoside. Therefore the five nucleosides are then called guanosine, adenosine, cytidine, uridine and thymidine. If they are attached to  $\beta$ -D-2'-deoxyribose, as in DNA, they are called deoxyguanosine (dG), deoxyadenosine (dA), deoxycytidine (dC), deoxyuridine (dU) and deoxythymidine (dT). If the base is attached to a  $\beta$ -D-ribose which is in turn attached to at least one phosphate, that is a complete nucleotide and they are labelled, guanylic acid, adenylic acid, cytidylic acid, uridylic acid and thymidylic acid. If the pentose is  $\beta$ -D-2'-deoxyribose, then the prefix deoxy is put before the names as it is for

## Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.

---

the nucleosides. A phosphate group can be bound to either the C3' or the C5' of the pentose to form the 3'-nucleotide or the 5'-nucleotide respectively (Figure 1.2) [2].

In organisms, the genetic material is usually double-stranded DNA and the RNA is single-stranded. For this reason, RNA is more flexible and can form a much greater variety of complex three-dimensional structures than double-stranded DNA (dsDNA). However, single-stranded DNA (ssDNA), used in in vitro experiments or in DNA computing, can also form complex structures. [6]



**Figure 1.2:** Primary structure for the tetradenucleotide adenylyl-3',5'-thymidylyl-3',5'-guanylyl-3',5'-cytidine (d(ATGC)).

### I.1.2.2 Secondary Structure

#### I.1.2.2.1 Sugar-Phosphate Backbone and the Glycosyl Bond

The conformational flexibility of the nucleotide unit is limited by a variety of internal constraints that greatly restrict rotational freedom. The conformation of a molecule is

## Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.

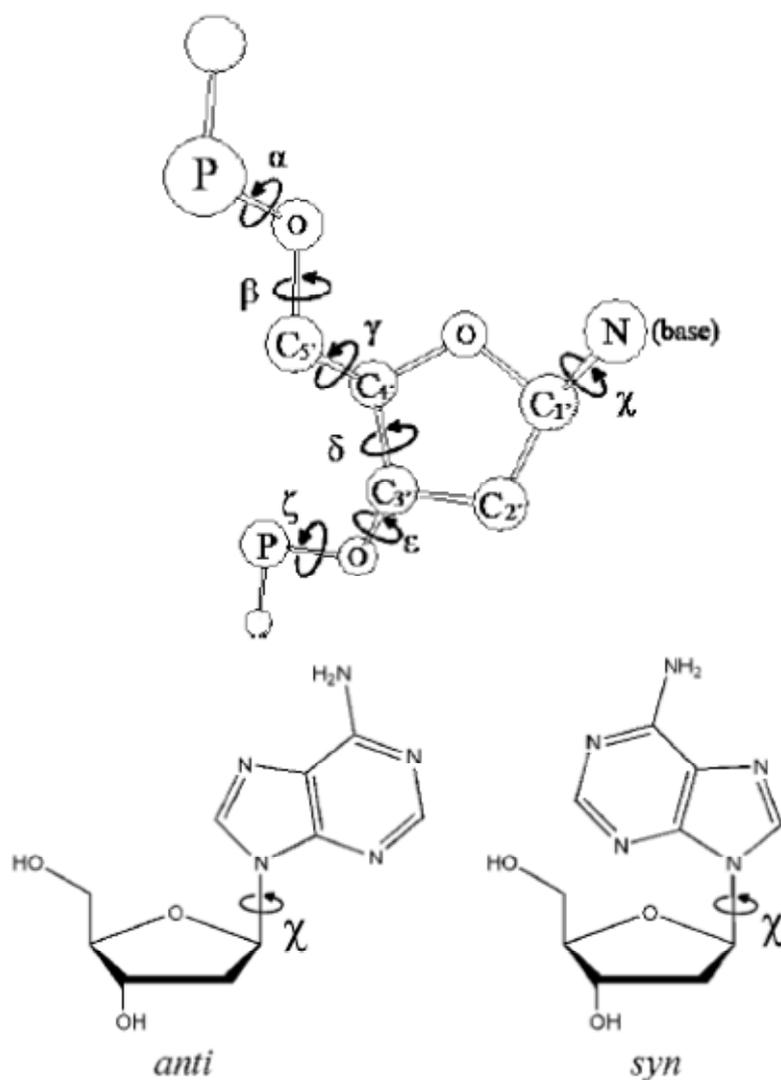
---

usually described by bond lengths, bond angles and rotations of groups of atoms about bonds by torsion angles [8]. Specifically, there are seven torsion angles that are important for describing the secondary structure of a nucleotide,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\zeta$  and  $\chi$  (Figure 1.3). The first six describe the torsion angles of the sugar phosphate backbone, the last torsion angle,  $\chi$ , describes the rotation around the glycosyl bond. The rotation of a base around its glycosyl bond is greatly hindered by steric interactions, especially for the pyrimidine bases due the 2-keto group. Generally, the glycosyl bond is only found in one of two sterically permissible positions: *syn* and *anti*.

For example, the two structures shown in Figure 1.3 are called *syn*-adenosine and *anti*adenosine [2].

In the *anti* conformation, most of the base is pointing away from the sugar, whereas in the *syn* conformation, most of the base is over toward the sugar giving rise to close interatomic contacts [9, 10]. The amount of steric hindrance that exists in the *syn* conformation is affected by the conformation of the sugar or the sugar pucker [2].

The glycosyl bonds can be either in the *syn* or *anti* conformations (bottom), *syn* being the conformation with the most steric hindrance.



**Figure 1.3:** Diagram of torsion angles in a nucleotide along the sugar-phosphate backbone and the glycosyl bond.

#### I.1.2.2.2 Sugar Pucker

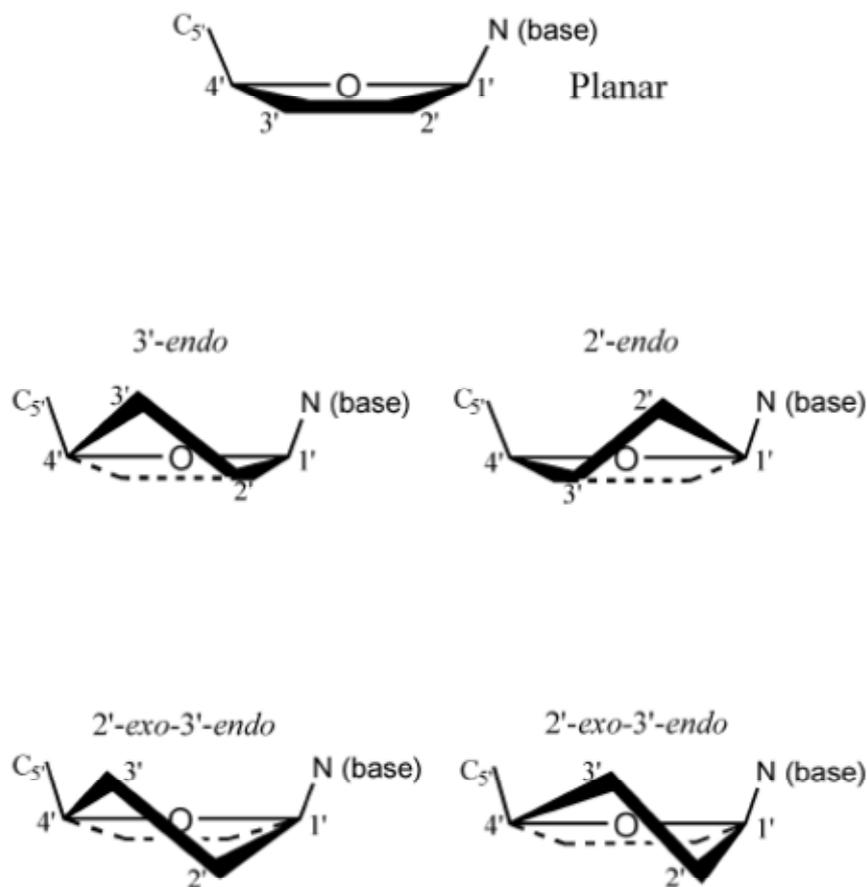
The sugar pucker describes the flexibility of the ribose ring which is generally non planar. The sugar may be puckered in an envelope form with four of the five atoms coplanar and the fifth atom being out of this plane by 0.5 Å; or in a twist form with two adjacent atoms displaced on opposite sides of the plane formed by the other three atoms [8; 11] (Figure 1.4). In the 2'-endo and 3'-endo conformations, it is the C2' or the C3',

## Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.

---

respectively, that is on the same side of the plane as the C5'. In the 2'-exo-3'-endo conformations, the C2' is on the opposite side of the plane as C5' and the C3' is on the same side of the plane. In conformations with two atoms out of the plane, it is rare that the two will be displaced equally out of the plane. Therefore, the atom farthest from the plane has major puckering, whereas the other atom has minor puckering. Generally, in DNA, the 3.-*endo* and 2.-*endo* conformations are favored with a larger portion of nucleotides being in the 2.-*endo* conformation. Now that we know about the sugar-phosphate backbone and the sugar pucker, it is important to understand how the polynucleotide chains bind in relation to one another. Here is where the helix and the studies of Watson and Crick come into play [2].

The planar conformation is not sterically permissible, but rather an illustration to show the plane (top). Of the two 2'-exo-3'-endo conformations, the one on the left shows major 2'-exo with minor 3'-endo and the conformation on the right illustrates the opposite, minor 2'-exo with major 3'-endo. In DNA, the 2'-endo conformation is the most common [2].



**Figure 1.4:** Different sterically permissible sugar pucker conformations including 3-endo, 2-endo and 2-exo-3-endo.

### I.1.2.3 Hydrogen Bonding and Base Stacking

There are two different types of non-covalent interactions that play a vital role in maintaining duplex stability. The first interaction is within the plane of the bases and is known as hydrogen bonding. Hydrogen bonding occurs when a hydrogen atom bridges two electronegative atoms; it is mostly governed by electrostatics. The second interaction, known as base stacking interactions or  $\pi$ - $\pi$  interactions, is perpendicular to the plane of the bases and is stabilized by London dispersion forces and hydrophobic effects [8]. The hydrophobic effect refers to the tendency to minimize the surface area

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

of the bases exposed to water. London dispersion involves induced dipole moments and electrostatic interactions. Base pair hydrogen bonding depends on composition while base stacking is dependent on composition and sequence. In non polar solvents, hydrogen bonding is more prominent and base stacking effects are suppressed [2].

In water, base stacking effects are dominant and hydrogen bonding interactions are negligible due to the competition for binding sites with the water molecules. Base stacking and hydrogen bonding have been recently discovered to influence each other.

For example, the hydrogen bonding ability of a stacked base depends on the hardness and the orientation of the stacking molecule [2; 12].

The A and T bases pair via two hydrogen bonds, while G and C pair with hydrogen bonds. The hydrogen bond is a short non-covalent bond. A non-covalent bond, in contrast to a covalent bond, does not involve the sharing of pairs of electrons and is therefore a weaker bond.

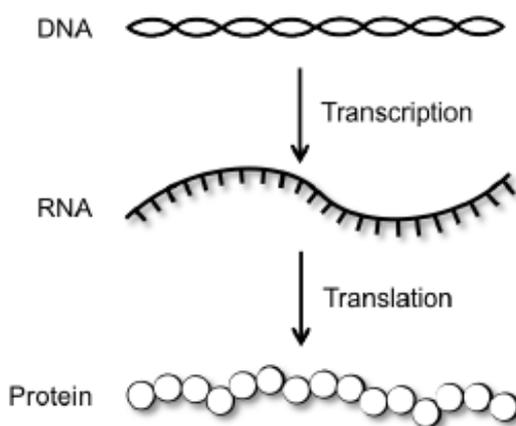
These bonds can also be deformed by stretching and bending. Additionally, the hydrogen bond in DNA is weaker than most hydrogen bonds because of geometric constraints of the helix structure. The hydrogen bonds and stacking of bases provide similar amounts of stability to the DNA structure. The individual stabilization interactions are weak, but taken together, form a stable helix.

In each of these a different surface (the Hoogsteen bonding surface) of the nucleotide molecule is used for the bond. These different base pair schemes serve to alter the shape of the molecules, which allows different parts of the molecules to be accessed by other molecules for different chemical reactions. This change of shape alters the chemical reactivity of the DNA molecule depending on the exact shape and conformation of the molecules [13].

### **I.1.3 From DNA to proteins**

The information stored along the DNA molecules is used to create RNA (ribonucleic acid) and proteins molecules (Figure 1.5) [14].

Proteins are long unbranched polymers just as DNA molecules, although the monomer units are not nucleotides but amino acids. In comparison to the four different nucleotides that are the building blocks of DNA, there is in total 20 different amino-acids available for protein fabrication [15]. The large number of different building blocks opens up for a wide range of physical and chemical properties for different proteins. Due to the large variations in physical and chemical properties proteins are involved in processes ranging from oxygen transportation to the copying of DNA [14].



**Figure 1.5:** Central dogma of biology schematically illustrated.

When a certain protein is needed in the cell, the gene encoding that particular protein is first copied into a RNA in a process called transcription. The sequence of nucleic acids along the RNA chain is complementary to the nucleotide sequence along the DNA strand being used as a template. The resulting RNA chain can then be used as a template

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

for protein production in a process called translation, where three consecutive bases along the RNA chain encodes for one amino-acid [14].

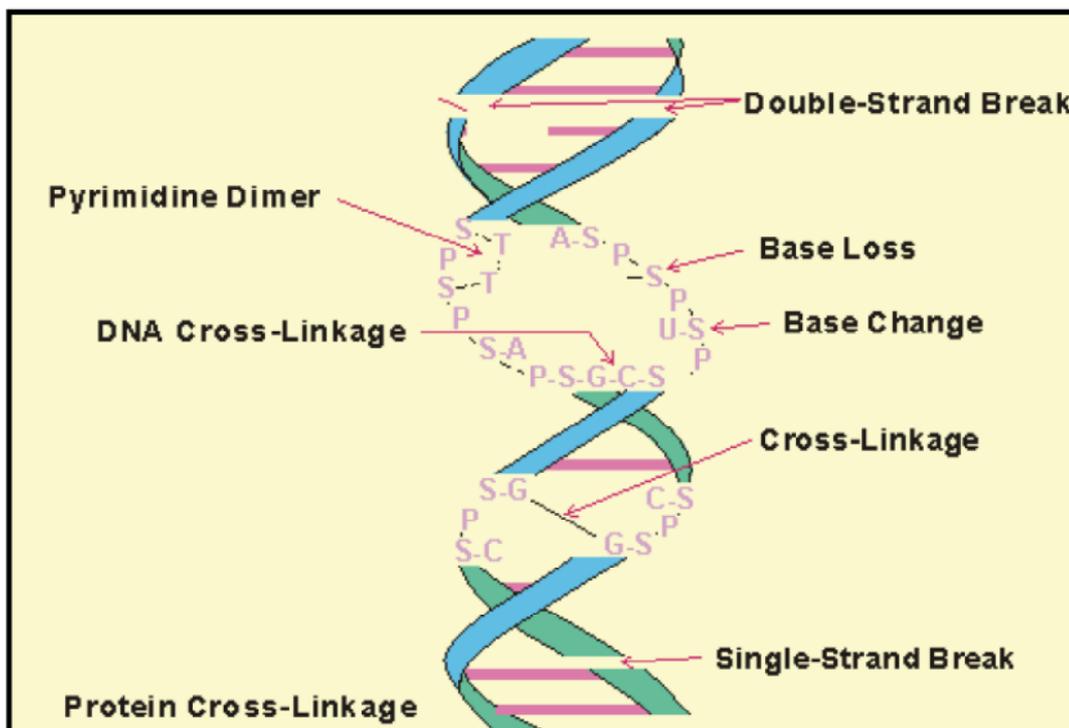
### **I.1.4 DNA damage**

DNA damage is normally divided into two major classes, namely endogenous DNA damage and environmental DNA damage. Endogenous DNA damage includes the DNA damage caused by reactive species generated during cellular metabolism, hydrolytic damage and DNA replication errors. Environmental DNA damage is referred to that caused by physical and chemical agents generated outside cells.

Any damage to the molecular structure of DNA has the potential to cause genomic instability, mutagenesis or even cell death. Unfortunately, DNA damage is unavoidable: DNA is continually exposed to insults resulting from exogenous and endogenous DNA damaging agents as well as by challenges posed by DNA replication [16].

#### **I.1.4.1 Types of DNA damage**

DNA can be damaged in numerous ways. Spontaneous damage due to replication errors, deamination, depurination and oxidation is compounded in the real world by the additional effects of radiation and environmental chemicals. The number of ways that DNA molecules can be damaged is very large. Examples of the types of DNA damage are shown in figure 1.6. [17]



**Figure 1.6:** A representation of the different types of DNA damage that can occur after exposure to chemicals or radiation.

### I.1.4.2 Simple adducts

#### I.1.4.2.1 Oxidation

Oxidative damage involves ROS damaging DNA and can produce a variety of lesions, including base adducts, single strand breaks (SSB) and double strand breaks (DSB) [18]. For example the most reactive ROS is the hydroxyl radical. The hydroxyl radical can cause the oxidation of DNA bases, the most common adduct formed is 8-oxodeoxyguanosine (8-oxodG). The hydroxyl radical is also able to react with adenine in a similar manner to guanine but this type of lesion is far less prevalent in DNA damage than 8-oxodG. [17]

If the 8-oxodG lesion is not repaired, it results in guanine no longer pairing with cytosine but instead pairs with adenosine during DNA replication. This is due to the fact

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

that the oxidized guanine bases no longer have three hydrogens available for hydrogen binding in cytosine and can only form two bonds and therefore binds preferably to adenosine. This results in guanine to thymine transition mutations if not repaired [17].

### **I.1.4.2.2 Alkylation**

Alkylation is the transfer of an alkyl group from one molecule to another, leading to a base adduct, with the simplest type of modification being methylation. Alkylation is accomplished by an alkyl electrophile, alkyl nucleophile or sometimes as alkyl radical. Undesired alkylation of DNA can lead to a possible mutation if unrepaired. An alkyl radical such as a methyl radical has only three out of four valency electrons of carbon bonded to hydrogen. This leaves one free unpaired electron which is free to interact with nucleophile groups on DNA. For example, alkylation of the O6 on a guanine can completely alter its ability to bind with cytosine resulting in it being read as deoxy-Adenine and therefore binding to thymine [19]. This is due to the molecule's ability to now only form two hydrogen bonds instead of three [17].

### **I.1.4.2.3 Hydrolysis**

The hydroxyl radical is also able to remove hydrogen from the deoxyribose-phosphate backbone. This results in DNA cleavage between the deoxyribose sugar and DNA base, leading to depurination and depyrimidation to form an apurinic or apyrimidinic site or AP site. If left unrepaired AP sites can lead to mutation during semi conservative replication as a random nucleotide base will be inserted into the strand synthesized opposite them [17].

### **I.1.4.3 The biological effect of radiation**

The biological effect of radiation can be mainly related to damages of the DNA. An ionizing radiation has a potential to directly interact with structures of the target to cause ionization, thus initiating the chain of events to lead to biological changes [20; 21; 22]

The inherent instability of DNA constitutes both an opportunity and a threat. DNA lesions can block important cellular processes such as DNA replication and transcription, cause genome instability and impair gene expression. Lesions can also be mutagenic and change the coding capacity of the genome, which can lead to devastating diseases and conditions associated with genome instability, including cancer, neurodegenerative disorders and biological ageing. Furthermore, mutagenic chemicals and radiation can also have a healing effect; they can for instance be used to treat cancer, by introducing DNA lesions that halt cell proliferation and stimulate programmed cell death [23].

The reactivity of DNA with endogenous or environmental agents results in chemical modifications, which can have mutagenic or lethal effect upon replication [24].

Damage to DNA is unavoidable and arises in many ways. It is estimated that in a single human cell the number of DNA damage events range from  $10^4$  to  $10^6$  per day [25]. DNA damage can be caused by spontaneous cleavage of chemical bonds in DNA, by environmental agents such as ultraviolet radiation from sunlight and cigarette smoke [26] or exposure to therapeutic agents used in the treatment of cancers including chemotherapeutic drugs and ionizing radiotherapy. DNA damage can be also caused by reaction with genotoxic chemicals, like reactive oxygen species (ROS) that are by-products of normal cellular metabolism or occur in the environment. ROS can cause

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

oxidative damage to DNA leading to single- and double-strand breaks (SSBs and DSBs) [27].

A change in the normal DNA sequence, called a mutation, can occur during replication when a DNA polymerase inserts a wrong nucleotide as it reads a damaged template. Mutations also occur at a low frequency as the result of copying errors introduced by DNA polymerases when they replicate an undamaged template. If mutations were left uncorrected, cells might accumulate too many mutations that can have detrimental effects on genetic stability. Failure to repair DNA lesions may result in blockages of transcription and replication, mutagenesis, and/or cytotoxicity [27].

### **I.1.4.4 Environmentally caused DNA double strand breaks**

The genome is not only exposed to endogenous mutagens like oxidative byproducts of cellular respiration, but also environmental agents like ionizing radiation, UV-light or genotoxic chemicals can cause various DNA damages. These include directly or indirectly introduced DSBs [28].

Ionizing radiation (IR) occurs naturally e.g. by radioactive decay of instable atomic nuclei or by cosmic radiation. Besides, IR is used in medical procedures like X-ray inspections or radiation therapy in cancer treatment [29]. IR produces a broad spectrum of different DNA damages, which are introduced via the production of reactive oxygen species. Most often IR leads to DNA base damages or introduction of DNA single-strand breaks (SSBs), which are repaired by base excision repair (BER) or single strand repair pathways. IR-caused DSBs occur when two SSBs are introduced in close proximity on opposite DNA strands. Therefore, IR caused DSBs often possess single strand overhangs. In addition IR produces DNA breaks with 3' termini carrying

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

phosphate or phosphoglycolate groups, which need to be removed before ligation of the breaks [28].

UV light on the other hand can indirectly provoke DSB formation by introducing 6-4 photoproducts and cyclobutane pyrimidine dimers into the DNA. These bulky lesions may induce replication-fork collapse and thereby DSBs if not repaired properly by the nucleotide excision repair (NER) machinery [30]. Similar effects are induced by different genotoxic chemicals, which also create replication blocking lesions like *e.g.* different alkylating agents, the intrastrand crosslinking anti-cancer-drug cisplatin or the interstrand crosslinking agent mitomycin [28].

In addition, chemicals, which poison the topoisomerase enzymes, can promote DSB formation by stabilizing the cleavage complex in which the topoisomerase is covalently attached to the cleaved DNA [31]. The Topoisomerase I (TopI) inhibitor camptothecin (CPT) triggers the accumulation of TopI-bound SSBs, which may be converted to DSBs when a replication fork collides with the cleavage complex. Topoisomerase II (TopII) enzymes introduce DSBs in the DNA during their catalytic cycle. Top II inhibitors like etoposide increase the concentration of cleavage complexes, which can be converted to permanent DSBs by collision with polymerases or helicases [28].

### **I.1.4.5 DNA double-strand repair pathways - A short overview**

To protect the genome and thus ensure the integrity of its coded information, cells have evolved different sophisticated mechanisms to repair DSBs: The two major pathways here are homologous recombination (HR) and non-homologous end joining (NHEJ) [28].

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

Traditionally, DNA melting has been performed in bulk and monitored by observing the UV-absorption of light in the 260 nm wavelength region. Monitoring a single wavelength is possible due to phenomena termed hyperchromicity, which describes the effect of a constant absorbance wavelength during an increasing absorbance. The increased absorbance arises from a decrease in hydrophobic interactions between the bases in the single stranded parts of the DNA chain. Within a ds DNA molecule the stacked bases interact with each other through their  $\pi$ -electron clouds, which are also responsible for the UV-absorption. The base-base interaction hinders the  $\pi$ -electron clouds from absorbing light in a dsDNA, but when the bases are separated due to melting the bases interaction with each other decreases and the  $\pi$ -electrons are free to absorb light [14; 32].

During DNA melting, the weaker AT-rich areas melt first since the bases A and T are only connected by two hydrogen bonds, whereas the bases C and G are connected by three hydrogen bonds. This sequence dependence in melting behavior can be used to get sequence information as discussed later. When DNA melts, bubbles of ssDNA are within seconds created along the dsDNA chain. These bubbles then slowly grows and more bubbles are created until the two chains are completely separated [14].

The melting probability for a base-pair is dependent on its environment. A base-pair adjacent to melted bases has a higher probability of melting than a base-pair within a double-stranded part of the chain, giving a cooperativity of the melting behavior [33].

Inside the cell nucleus the genome is constantly under attack from a range of chemical species as well as UV-radiation [34]. Breaks in the DNA threaten the genome integrity and can potentially lead to cell cycle arrest or even apoptosis. Furthermore, many types

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

of cancer arises due to failure to correctly repair DNA breaks. Both single-stranded and double stranded breaks occurs but the double-stranded breaks are more critical to the cell. Due to the high risks accompanying DNA breaks, cells have developed pathways to accurately detect and repair damaged DNA [14].

The formation of pyrimidine dimers upon irradiation with UV light results in an abnormal covalent bond between adjacent pyrimidine bases. The photoreactivation process reverses this damage by the action of a photolyase. This enzyme is activated by the energy absorbed from blue/UV light (300-500nm wavelenght). Photolyase is a phylogenetically old enzyme present and functional in many species from bacteria to animals. However, in humans and other placental mammals this photolyase activity is missing and its function is replaced by the less efficient nucleotide excision repair mechanism [27].

### **I.2. Classification and pharmacological properties of macrolides**

#### **I.2.1. Macrolide structures**

The term *macrocycle* refers to medium- and large-ring compounds, with, respectively, 8–11 and 12 or more atoms in the ring. Macrocyclic structures that have one or more ester linkages are generally referred to as *macrolides* or *macrocylic ring lactones*. In some cases, macrocyclic lactams have also been described as macrolides. Originally macrolides denoted a class of antibiotics derived from species of *Streptomyces* and containing a highly substituted macrocyclic lactone ring aglycone with few double bonds and one or more sugars, which may be aminosugars, non-nitrogen sugars or both. To our knowledge the largest naturally occurring macrolides are the 60-membered

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

quinolidomycin<sup>6</sup> and the largest constructed macrolide is the 44-membered swinholidide [35].

Macrolides are an important class of antibiotics effective against a number of pathogenic bacteria. These mostly bacteriostatic drugs are commonly used to treat respiratory tract infections including community acquired pneumonia, pharyngitis, and tonsillitis, along with skin and soft tissue infections, urogenital infections, and orodental infections.

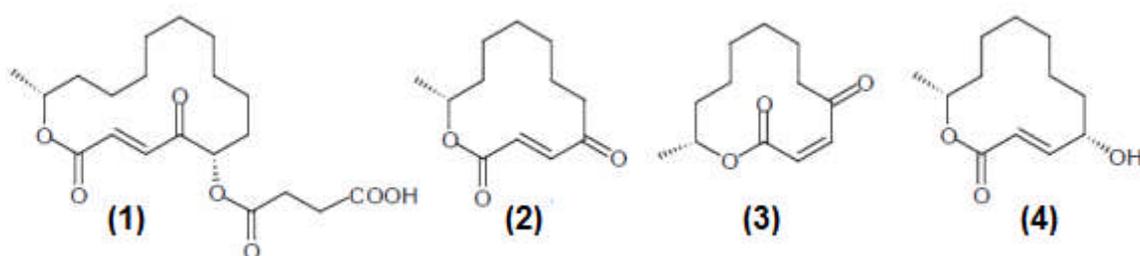
Macrolides were introduced into medical practice almost 60 years ago and continue to be viewed as excellent antibiotics with high potency and low toxicity [36].

Macrolides have been known for 15 decades, and, since the introduction of erythromycin into therapy, a number of these molecules have been developed for clinical use. For years, these antibiotics have represented a major alternative to the use of penicillins and cephalosporins for the treatment of infections due to gram-positive microorganisms (mostly  $\beta$ -hemolytic streptococci and pneumococci); however, the worldwide development of macrolide resistance with wide variations, according to both the country and the bacterial species, has sometimes constrained to limit the use of these antibiotics to certain indications. Although the evolution of the macrolide class has been marked, in the 1990s, by the development of semisynthetic macrolides with improved pharmacokinetics and tolerability, these new derivatives have proved unable to overcome erythromycin resistance. By contrast, our increasing knowledge of the molecular mechanisms of resistance to macrolides has led to the pharmaceutical industry's design of derivatives, such as the ketolides, that have activity against certain types of erythromycin-resistant organisms [37].

## Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.

---

Most of the pharmacologically active macrolides have highly substituted structures as can be seen from the few examples above. However, complexity of the structure is not essential for antibiotic activity. Relatively simple macrolides such as macrolide A26771B (1) and patulolides A (2), B (3), and C (4) are antimicrobial compounds (figure 1.7) [35].



**Figure 1.7:** Examples for antimicrobials macrolides.

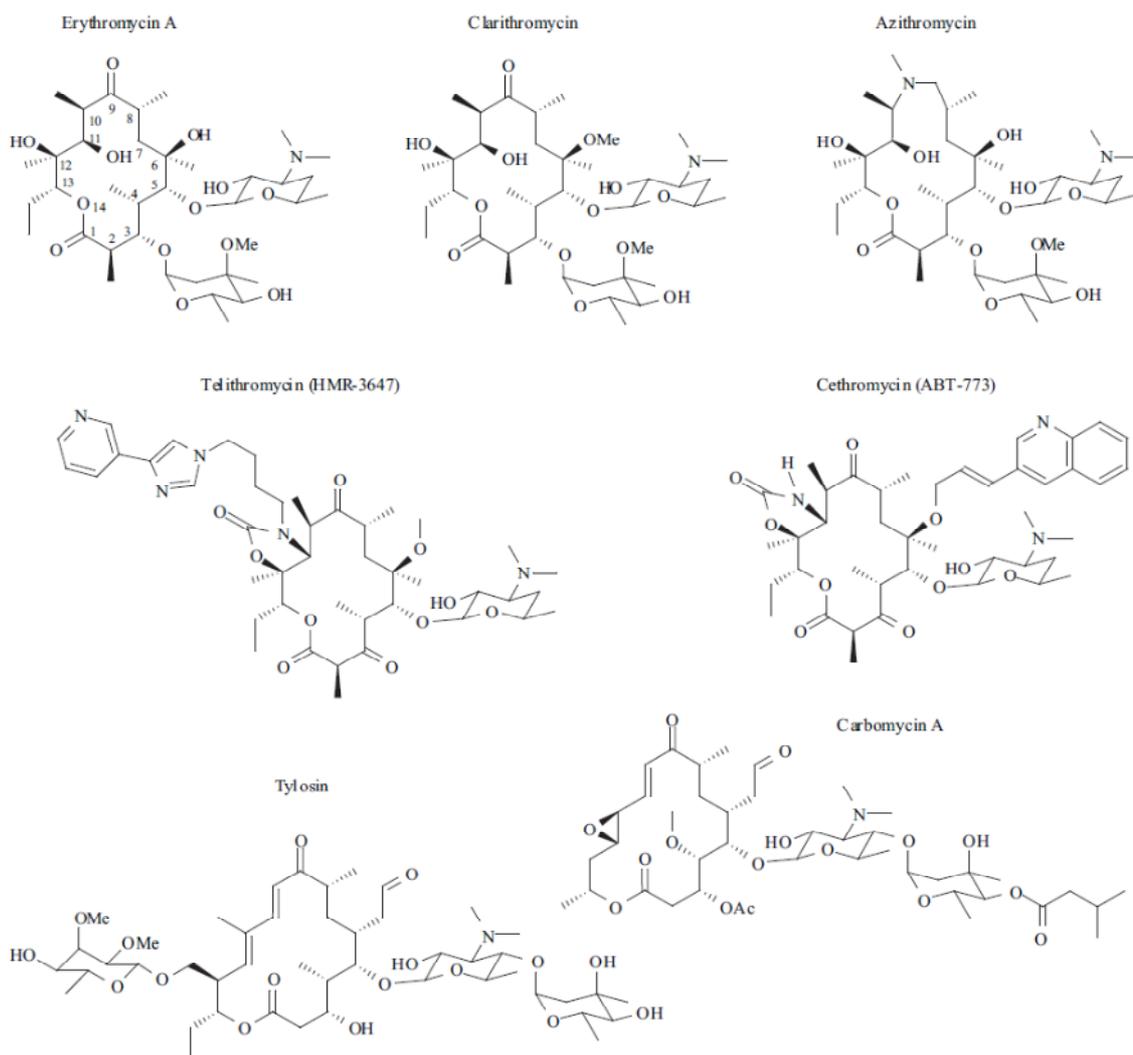
### I.2.2. Chemical structure of macrolide antibiotics

Antibacterial agents, or antibiotics, are a class of a much larger group of compounds called antimicrobial agents. Antibiotics used to refer to only naturally occurring molecules produced by a variety of microorganisms [38].

Macrolides belong to one of the most commonly used families of clinically important antibiotics used to treat infections caused by Gram-positive bacteria such as *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Streptococcus pyogenes*. [39; 40] Chemically, macrolides are represented by a 14-, 15- or 16-membered lactone ring carrying one or more sugar moieties and additional substitutions linked to various atoms of the lactone ring, (Figure 1.8) [39].

## Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.

---



**Figure 1.8:** Chemical structures of the macrolide antibiotics.

First row: drugs of the first generation (erythromycin) and second generation (14-membered ring clarithromycin and 15-membered ring azithromycin). Second row: ketolides, the macrolides of the third generation. Third row: examples of 16-membered ring macrolides.

### I.2.3. The need for novel antibacterial agents

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

The prevalence of drug resistant bacteria is growing at an alarming rate in both developing and developed countries [41]. From this statement alone, it should be clear that the need for the development of novel antibacterial agents is of utmost importance. In the current antibacterial drug pipeline, there is only a miniscule glimmer of hope. This rapid increase in resistant bacteria coupled with the slow development of novel agents has led some experts to call this time the “dawn of the post-antibiotic era.” [38]

There exists a perpetual need for new antibiotics. Most drugs will be just as effective in the future as they are today, but that is not the case with antibiotics. Eventually, the inevitable rise of resistance will erode the utility of today’s antibiotics. There are three factors that intensify this supply problem by discouraging antibiotic development [42].

First, antibiotics are used in smaller quantities than other drugs. The standard antibiotic course lasts only weeks compared to treatment for chronic illness which can last a lifetime. Therefore, antibiotics yield lower revenues than most drugs. Second, the use of newly approved antibiotics is often limited to serious bacterial infections. The third reason is an increase in regulatory requirements to get a drug licensed, which makes clinical trials cost prohibitive. However, most newly approved drugs can be prescribed to all who may benefit from their use. These factors ultimately result in this quandary: Resistance is on the rise while antibiotic discovery and development are on the decline [38].

### **I.2.3.4 New antibiotics**

Having new antibiotics approved for use is a great thing, but those antibiotics which utilize the same mechanism of action as previously approved drugs always run the risk of increasing the rate of resistance. [38]

### **I.2.4 Macrolide-binding site in the ribosome**

The discovery of the significance of (hetero) aryl anchored at the terminus of a suitable linker enable us design the novel macrolides capable of interacting with a new binding site of bacterial ribosome [43].

The general location of the macrolide binding site on the large ribosomal subunit has been initially mapped using a combination of biochemical and genetic methods [39].

Nonetheless, the details of the molecular interactions of the various classes of macrolides with the ribosome have just started to emerge with the release of several crystallographic structures of the archaeal and bacterial large ribosomal subunits and their complexes with antibiotics [39].

Because the polar groups locate on the frontal face of the aglycone and most of the alkyl groups install on the other face, the two sides of the macrolides present hydrophilic and hydrophobic properties, respectively. The hydrophobic side is believed to make van der Waals contacts with hydrophobic areas on the peptide tunnel wall. Recently, comparison of reduced efficacy of 4,8,10-tridemethyl telithromycin and 4,10-didemethyl telithromycin with parent telithromycin undoubtedly gave new sight into the important function of these methyl groups [43].

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

A macrolide molecule is coordinated in its binding site by multiple hydrophobic and hydrogen bonds (and possibly, a covalent bond in case of some 16-membered ring macrolides) between its functional groups and 23S rRNA [39].

These interactions with RNA account for most of the free energy of drug binding. In addition, some macrolides can reach ribosomal proteins L4 and/or L22. The mucinose sugar of tylosin interacts with L22 and the furosamine residue of spiramycin interacts with L4. The proximity of these proteins to the macrolide binding site explains why mutations in L22 and L4 protein genes can render cells resistant to macrolides [39].

### **I.2.5 Modes of inhibition of protein synthesis by macrolides**

The precise mechanism of protein synthesis inhibition by macrolides depends on the specific chemical structure of the drug molecule. This affects its interaction with the ribosome as well as the mode of the inhibitory action. Four modes of inhibition of protein synthesis have been ascribed to macrolides [39]:

- 1) Inhibition of the progression of the nascent peptide chain during early rounds of translation [44];
- 2) Promotion of peptidyl tRNA dissociation from the ribosome [45];
- 3) Inhibition of peptide bond formation;
- 4) Interference with 50S subunit assembly [46].

All of these mechanisms have some correlation with the location of the macrolide binding site on the ribosome [39].

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

Most macrolides are weak bases and are unstable in acids. The action of macrolides can be bactericidal or bacteriostatic, the effect depending on the concentration and the type of microorganism targeted by the drug. Macrolides bind to the 50S subunit of the bacterial ribosome and inhibit the transpeptidation and translocation process, causing premature detachment of incomplete polypeptide chains. The antimicrobial spectrum of erythromycin is very similar to penicillin; it has proved to be a safe and effective alternative for penicillin-sensitive patients.

Macrolides are effective against Gram-positive bacteria and spirochetes but not against most Gram-negative organisms, the exceptions being *Neisseria gonorrhoeae* and *Haemophilus influenzae*. *Mycoplasma pneumoniae*, *Legionella* species and some chlamydial organisms are also susceptible. The macrolides are administered orally and intravenously.

They diffuse readily into most tissues, but do not cross the blood-brain barrier; additionally there is poor penetration into synovial fluid. The majority of side effects associated with macrolides are mild and transient. As a class they are generally well-tolerated. The most common complaints involve the gastrointestinal tract and include diarrhea, nausea, vomiting and abdominal pain. Patients may complain of an abnormal or metallic aftertaste. Hepatotoxicity is a very rare but serious side effect associated with the estolate salt of erythromycin [47].

### **I.2.6 Antibiotic Resistance**

Resistance is defined as the relative insusceptibility of a microorganism to a particular treatment under a particular set of circumstances [48]. Some researchers believe that

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

resistance is an ecological phenomenon stemming from the response of bacteria to the widespread use of antibiotics and their presence in the environment. They believe that the rise in the frequency of antibiotic resistance among pathogens should be a “cause of great concern” and suggest “a commitment to act responsibly” [47].

### **I.2.6.1 Macrolide action and drug resistance:**

Macrolides inhibit protein synthesis by interacting with bacterial ribosomal RNA. Two distinct mechanisms have been responsible for the majority of macrolide resistance observed in clinical isolates. The first mechanism involves the production of a ribosomal methylase. This enzyme methylates a specific adenine residue on the ribosomal RNA thus preventing binding of the macrolide, lincosamide and streptogramin B (MLSB) and conferring MLSB resistance. The methylase is the product of a family of genes called *erm* that can be inducibly regulated or constitutively expressed. The second mechanism involves a family of genes called *mef* and is commonly referred to as efflux resistance [49].

Bacteria resist macrolide and lincosamide antibiotics in 3 ways:

- (1) Through target-site modification by methylation or mutation that prevents the binding of the antibiotic to its ribosomal target,
- (2) Through efflux of the antibiotic, and
- (3) By drug inactivation.

These mechanisms have been found in the macrolide and lincosamide producers, which often combine several approaches to protect themselves against the antimicrobial that

## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

they produce. In pathogenic microorganisms, the impact of the 3 mechanisms is unequal in terms of incidence and of clinical implications [37].

The extensive use of these antibiotics has led inevitably to the spread of resistant strains. Expression of some of the resistance determinants is inducible by macrolides. Of particular interest are Erm methyltransferases, which specifically methylate a unique nucleotide within the macrolide binding site. The mechanism of Erm induction depends on ribosome stall within the translated regulatory open reading frame preceding the Erm cistron, and is apparently closely related to the general mode of macrolide action on protein synthesis. However, the details of the mechanism of Erm induction are not known [39].

Equally important, the study of arising macrolide resistance shed light on how to fight against mutation of pathogens [43].

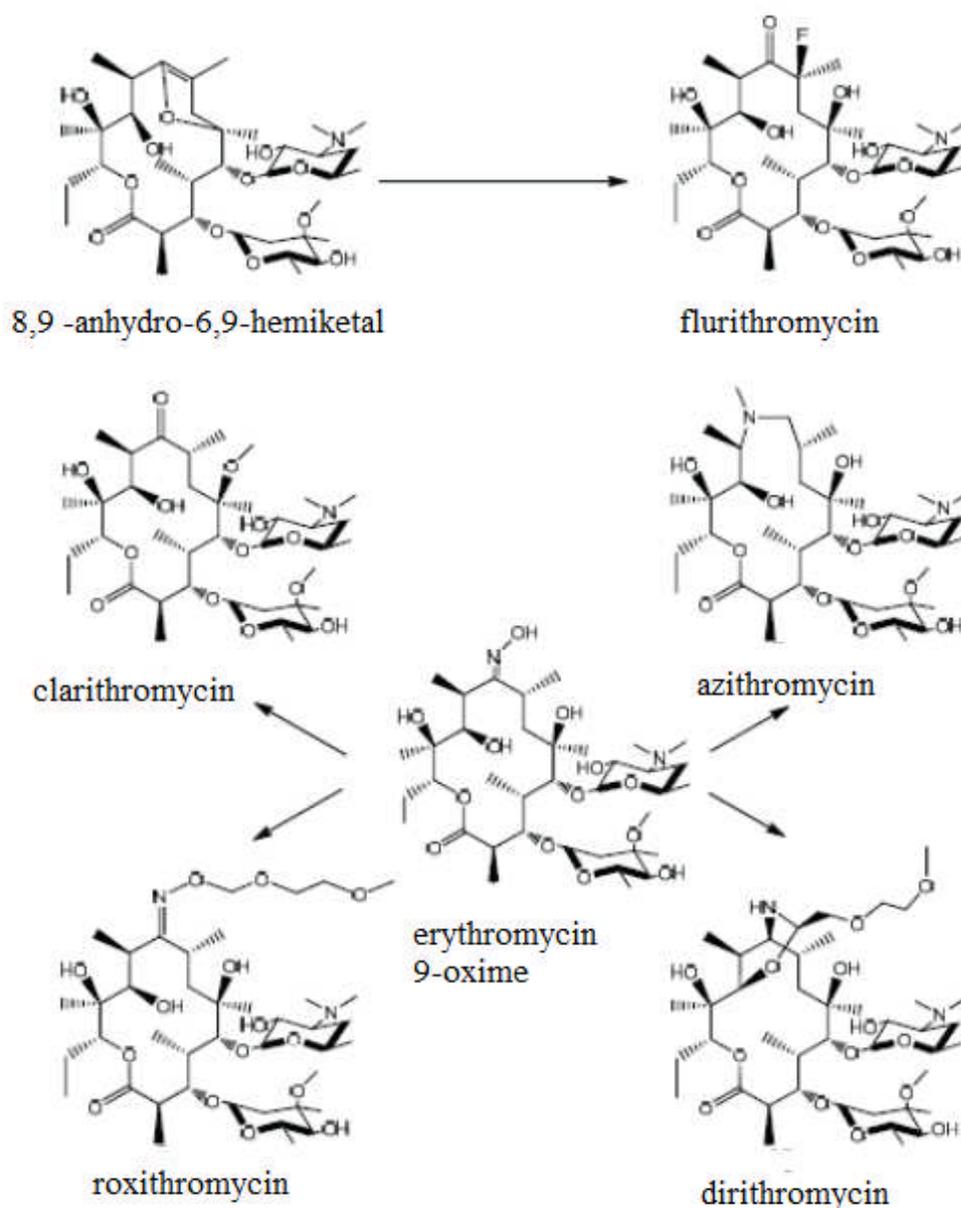
### **I.2.7 Erythromycin**

Erythromycin A, a natural product of *Saccharopolyspora erythraea*, was the first macrolide to be advanced to medical use in the early 1950s for the treatment of bacterial infections. Erythromycin (Figure 1.9) is effective against many Gram-positive pathogenic bacteria including *Staphylococcus aureus*, *Streptococcus pneumoniae*, *S. pyogenes*, and *Enterococcus* sp.; some Gram-negative bacteria such as *Neisseria gonorrhoeae*, *Bordetella pertussis*, and *Haemophilus influenzae*; and intracellular pathogens such as *Mycoplasma* sp., *Legionella* sp., and *Chlamydia* sp. Semisynthetic derivatives of erythromycin, including drugs such as clarithromycin, roxithromycin, and azithromycin

**Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

(Figure 1.9), which belong to the second generation of macrolides, exhibit increased acid stability, better oral bioavailability, improved pharmacodynamics [36].



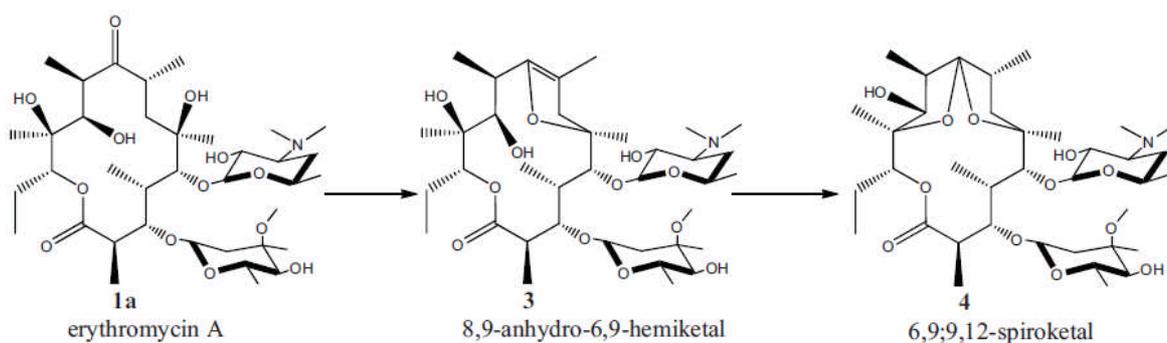
**Figure 1.9:** The 2<sup>nd</sup> generation erythromycins and the synthetic pathways to them.

The 2<sup>nd</sup> generation erythromycins successfully block the acidic degradation process which involves the 9-ketone, 6-hydroxy, and 8-hydrogen. For this reason,

## Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.

---

clarithromycin, azithromycin and the others possess better bioavailability, reduced gastrointestinal side effects and a superior spectrum of coverage (figure 1.10) [43].



**Figure 1.10:** The degradation of erythromycin A under acidic condition.

In the 1960s and 1970s, the structure-activity relationships were extensively explored, which resulted in erythromycin 9-oxime, erythromycin 9-amine, erythromycin 11,12-cyclic carbonate, 3-descladinosyl derivatives and 4"- modified compounds. Although most of the resulting structures do not have attractive antibacterial activity, these modifications are nonetheless valuable because some of the functionalities have been presented in the currently designed structures [43].

### I.2.7.1 Metabolism of erythromycin A and clarithromycin

Both erythromycin A and clarithromycin are metabolised through CYP450 3A4.3 However, there are differences in their abilities to bind to and inhibit the cytochrome P-450 isoform CYP 3A4. On the basis of these differences, macrolides (in general) are classified into three groups on the basis of data provided by in vitro experiments:

1) Group 1 include erythromycin A and troleandomycin. Both drugs bind strongly to and markedly inhibit CYP 3A4.

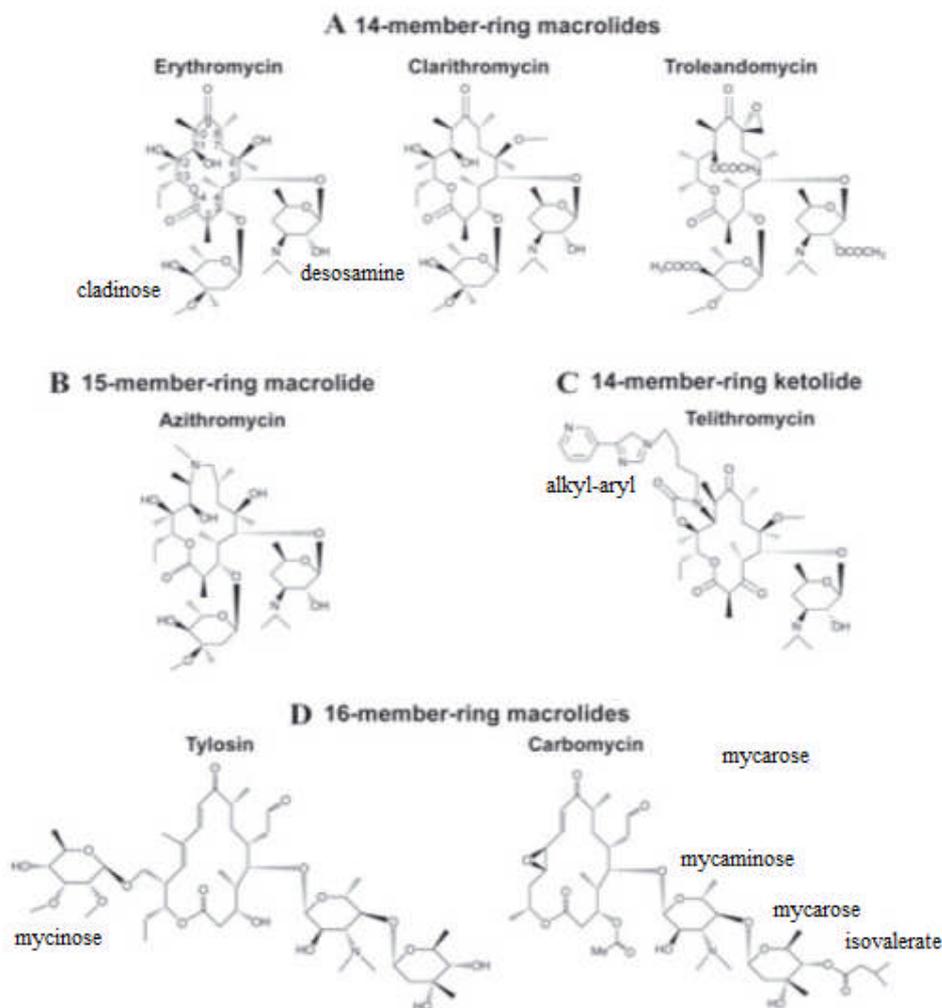
## **Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

2) Clarithromycin belongs to Group 2 agents. This drug exhibits a lower affinity for CYP 3A4 compared to erythromycin A, and form complexes to a lesser extent.

3) Group 3 include azithromycin and dirithromycin. These compounds have been shown to interact poorly with the cytochrome P-450 system in vitro. However, results obtained from some clinical studies showed that clarithromycin is similar to erythromycin A in some drug interactions (e.g. with psychotropic agents) [50].

The emergence and broad spread of resistance prompted the development of the new generation of macrolides. Ketolides (Figure 1.11), representing the third generation, show improved potency against many sensitive and some resistant strains and are often associated with bactericidal activity. Macrolide antibiotics with an extended macrolactone ring, such as 16-membered macrolides (Figure 1.11), find extensive use in veterinary medicine and are sometimes also used in humans [36].



**Figure 1.11:** Chemical structures of representative macrolide antibiotics.

The numbering of the macrolactone atoms is illustrated on the erythromycin structure.

### I.2.8 12 Membered $\alpha,\beta$ -unsaturated macrolides

Twelve membered macrocycles are the smallest to be generally considered as large ring compounds. According to CA there are 319 known compounds with a 12-membered lactone ring structure, of which nearly 60 are  $\alpha,\beta$ -unsaturated. Many 12-membered macrolides, such as the antibiotics methymycin 6 and neomethymycin are

physiologically active. Cladospolides are plant growth regulators, and structurally even simpler patulolides have antimicrobial properties [35].

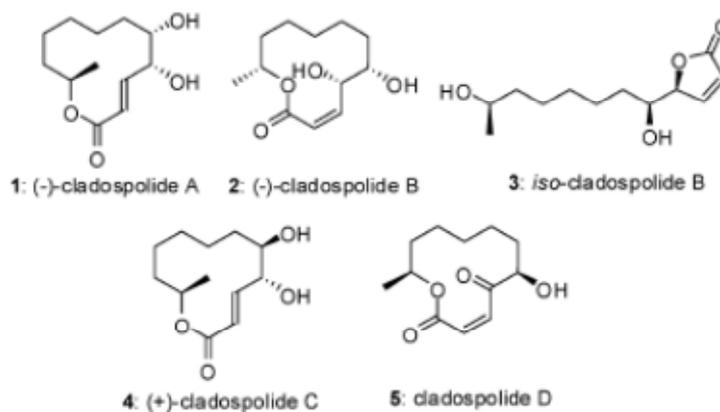
### **I.2.8.1 Patulolides**

Patulolide A was isolated from the culture broth of *Penicillium urticae* S11R59 in 1985 and later three more 12-membered macrolides, patulolides B and C, were isolated. *Penicillium urticae* S11R59 is a blocked mutant that produces patulolides instead of the mycotoxic patulin (figure 1.7) [51].

Patulolides have been shown to be pure acetogenic hexaketides derived from the head-to-tail condensation of six acetate units. Patulolides A and B inhibit various strains of fungi and yeasts and they show some activity against grampositive, gram-negative as well as enteric bacteria [35].

### **I.2.8.2 Cladospolides**

Marine fungi, being a potential source of new biologically active secondary metabolites, are a topic of growing interest. Cladospolides (A–D) (Figure 1.12) are such secondary metabolites, responsible for the host plant's growth, and were isolated 1–4 from different *Cladosporium* sp. [52; 53] Structurally this class of natural products differ in the position, number and stereochemistry of hydroxyl groups as well as the double bond and these trivial differences in functionality affect their biological profiles significantly [52].

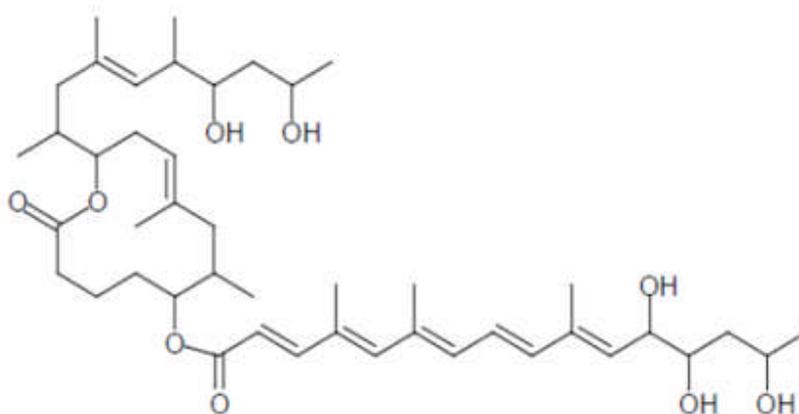


**Figure 1.12:** cladospolide family.

cladospolide A 1 is found to inhibit root growth of lettuce seedlings, while cladospolide B 2 promotes the growth. Likewise, cladospolide C 4 was also found to inhibit shoot elongation of rice seedlings. Unlike other congeners, cladospolide D 5, the recently isolated species from *Cladosporium* sp. FT-0012 by Omura and co-workers, shows antimicrobial activity with IC<sub>50</sub> values of 0.1 and 29 mg mL<sup>-1</sup> against *Mucor racemosus* and *Pyricularia oryzae* respectively [52].

### I.2.8.3 Mycolactone

Recent findings in the field of macrolides include the first isolated mycobacterial toxins, mycolactone A and its isomer B (figure 1.13), containing a 12-membered lactone ring.<sup>21</sup> Mycolactones are shown to be directly responsible for the necrosis and immunosuppression associated with the skin disease Buruli ulcer [35].



**Figure 1.13:** mycolactone B

**References:**

- [1] N. C. Seeman, *Nature* 421, 427-431 (2003).
- [2] Angela G. Hoffort, *Thesis of Master*, Saskatchewan Saskatoon University (2006).
- [3] Robert B. Macgregor Jr., Gregory M.K. Poon, *Computational Biology and Chemistry* 27 461–467 (2003).
- [4] Jeffrey M. Vargason, *Thesis of doctorate*, Oregon State University,( 2002)
- [5] Aaron Klug, *J. Mol. Biol.* 335, 3–26 (2004).
- [6] M. S. Andronescu, *Thesis of Master*, Academy of Economic Studies, Romania, (2000).
- [7] Serena F. Abbondante, *Thesis of doctorate*, University of Canberra, (2009).
- [8] Saenger, W. *Principles of Nucleic Acid Structure* (Springer Verlag, New York), (1984).

**Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

- [9] J. Donohue, K. N. Trueblood, *J. Mol. Biol.*, 2, 363-371(1960).
- [10] A. E. V. Haschemeyer, A. Rich, *J. Mol. Biol.*, 27, 369-384 (1967).
- [11] L. D. Hall, *Chem. Ind.*, 950-951 (1963).
- [12] P., Mignon, S., Loverix, J. Steyaert, and P. Geerlings, *Nucleic Acids Res.*, 33, 1779-1789 (2005).
- [13] Erin Lynn Knapp, *Thesis of Master*, University of Tennessee, Knoxville, (2013).
- [14] Camilla Freitag, *Licentiate Thesis*, Lund University, (2013).
- [15] B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Garland Science Pub.*, 1 (2004).
- [16] Yu Fu, M.Sc, *Thesis of doctorate*, University of Saskatchewan Saskatoon, (2008).
- [17] Seth Eli Remington, *Thesis of doctorate*, Newcastle University, (2010).
- [18] M. D. Evans, M. Dizdaroglu and M. S. Cooke, *Reviews in Mutation Research* 567, 1-61(2004).
- [19] A.Dudas, M. Chovanec, *Mutat Res.*, 566, 131-167 (2004).
- [20] J. Valentin, Low-dose Extrapolation of Radiation-Related Cancer Risk, *Elsevier*, 2006.
- [21] S. Lehnert, Biomolecular Action of Ionizing Radiation, *Taylor and Francis Group, LLC*, 2007.
- [22] Wei Han, K. N. Yu, Nova Science Publishers, 4, 1-13 (2010).
- [23] H.J. Muller, Artificial Transmutation of the Gene. *Science*, 66(1699), 7-84 (1927).
- [24] S. Broomfield, T. Hryciw, and W. Xiao, *Mutat Res*, 486(3), 167-84 (2001).

**Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

- [25] E.C. Friedberg, *J. Bioessays*, 13(6), 295-302 (1991).
- [26] K.H. Kraemer, *Proc. Natl. Acad. Sci. U S A*, 94(1), 11-4 (1997).
- [27] Andreea Daraba, *Doctoral thesis*, University of Szeged, (2014).
- [28] Christian Bernd Schiller, *thesis of doctorate*, munich university, (2011).
- [29] A. Ciccia, S. J. Elledge, *J. Mol. Cell.*, 40(2), 179-204 (2010).
- [30] C. L. Limoli, E. Giedzinski, W. M. Bonner and J. E. Cleaver, *Proc. Natl. Acad. Sci. USA*, 99(1): 233-238 (2002).
- [31] F. Degrassi, M. Fiore and F. Palitti, *Curr Med Chem Anticancer Agents* 4(4): 317-325 (2004).
- [32] R. H. Garrett, C. M. Grisham, *Biochemistry*. Available Titles Cengage NOW Series, Brooks/Cole Publishing Company, (2010).
- [33] R. Metzler, T. Ambjörnsson, A. Hanke, and H. C. Fogedby, *Journal of physics. Condensed matter*, 21, 034111 (2009).
- [34] J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry*. W. H. Freeman, 2010.
- [35] Leena Kaisalo, *thesis of doctorate*, University of Helsinki, (2002).
- [36] K. Kannan, A. S. Mankin, *Ann. N.Y. Acad. Sci.*, 1241, 33–47 (2011).
- [37] Roland Leclercq, *Clinical Infectious Diseases*, 34, 482–92 (2002).
- [38] Joshua Randal Brown, *thesis of doctorate*, University of Tennessee (2010).
- [39] M. Gaynor, A. S. Mankin, *Current Topics in Medicinal Chemistry*, 3, 949-961 (2003).
- [40] W. Scott Champney, R. Burdine, *Antimicrob. Agents Chemother.*, 39(9), 2141–2144 (1995).

**Chapter I: Physicochemical properties and chemical reactivity of DNA and classification and pharmacological properties of macrolides.**

---

- [41] S. J. Projan, P. A. Bradford, *Curr. Opin. Microbiol.*, 10, 441-446 (2007).
- [42] C. Nathan, F. M. Goldberg, *Nat. Rev. Drug Discov.*, 4 (11), 887-891 (2005)
- [43] Jian-Hua Liang, Xu Han, *Current Topics in Medicinal Chemistry*, 13, 3131-3164 (2013).
- [44] D. Vazquez, J.W. Corcoran, F.E. Hahn, *Springer-Verlag: New York, Heidelberg, Berlin*, 459-479 (1975).
- [45] J.R. Menninger, D.P. Otto, *Antimicrob. Agents Chemother.*, 21, 810-818 (1982).
- [46] W.S. Champney, R. Burdine, *Antimicrob. Agents Chemother.*, 39, 2141-2144 (1995).
- [47] Ginger M. Shipp, *thesis of doctorate*, Iowa State University, (2009).
- [48] X. Zhao, P. Lacasse, *J. Anim. Sci.* 86, 57-65 (2008).
- [49] Zhenkun Ma, Peter A. Nemoto, *Curr. Med. Chem.*, 1, 15-34 (2002).
- [50] Biljana Arsic, *thesis of doctorate*, University of Manchester, (2012).
- [51] D. Rodphaya, T. Nihira, S. Sakuda, Y. Yamada, *J. Antibiot.*, 41, 1649 (1988).

## **Chapter II:**

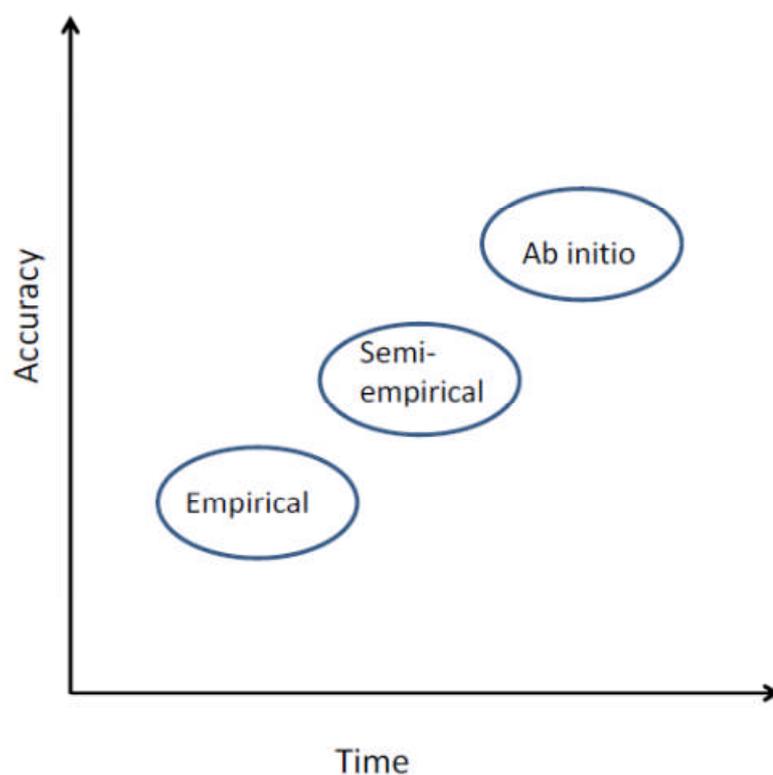
# **Molecular Modeling**

### II.1.Introduction

Computational chemistry may be defined as the applications of mathematical methods, molecular mechanics and quantum mechanics to chemical systems. Molecular modeling can be used to calculate the structure and properties of the molecular systems such as dipole moments, vibrational frequencies, spectroscopic and reactivity [1].

Molecular modeling or computational chemistry is one of the best methods to calculate the characteristics of materials when experimental data is not available. The main idea of computational chemistry is to study materials at an atomic and molecular scale. Therefore, the macroscopic data are obtained by manipulating atomic-scale parameters, such as bond length, bond angles, etc. Computational chemistry includes different approaches such as ab initio methods or quantum mechanics, semi-empirical, and empirical methods or molecular mechanics. [2]

Depending on the complexity of the system, any of these approaches might be used. A main goal of a study is to use the fastest calculation with the least computer power to get the most accurate result. In general, a more accurate the calculation is, the more time is required. In Figure 2.1, the accuracy and time consumption of these different methods are roughly compared. [2]



**Figure 2.1:** Accuracy vs. time of the different molecular modeling methods.

## II.2. Quantum mechanics

Quantum mechanics is a computational method, which models molecular systems using the quantum theory. Quantum theory was developed, because classical Newtonian mechanics can not describe some aspects of the microworld, specifically: quantization of microparticles energy, Heisenberg's uncertainty principle, wave-particle duality etc.. [3].

Two main approaches can be used [4, 5]: Ab-initio methods and semi-empirical methods.

Ab-initio methods are computational methods, derived directly from theory and including no experimental data. On the other hand, semi-empirical methods use empirical parameters and/or ignore some terms in the Hamiltonian.

Ab-initio methods are derived directly from physical and chemical principles, they use mathematical approximations. Specifically, the Hamiltonian is expressed by different approximations called theory levels [5, 6]. And the wavefunction is replaced by a set of simple functions called basis set [3, 6]

Semi-empirical methods are the approximate methods in which parameters involved in the equations are taken from experiment, some are neglected, and some others are estimated by fitting to experimental data. [7]

### II.3. Quantum mechanics calculations

The quantum mechanics (QM) approach postulates the fundamental principles and then uses these postulates to deduce experimental results. For the definition of the state of a system in QM, the function of the coordinates of particles is referred as the wave function or state function  $\Psi$ . In general, the state changes with time, thus for one-particle, one-dimensional system,  $\Psi=\Psi(x, t)$ . The wave function contains all possible information about a system [7, 8].

In order to find the future state of a system from the knowledge of its first state, an equation is needed that tells how the wave function changes with time ( $t$ ) [9]. Schrödinger's time dependent equation describes the particle by a wave function  $\Psi(r, t)$  [7, 9].

**II.3.1. Schrödinger equation**

In the QM methods, both nuclei and electrons dynamics can be computed by solving the time dependent Schrödinger equation, which is given as follows, [10]

$$i\hbar \frac{\partial \psi(\mathbf{R}, \mathbf{r}, t)}{\partial t} = \mathcal{H}\psi(\mathbf{R}, \mathbf{r}, t) \quad (1)$$

where,  $\psi$  is the time dependent wave function,  $\mathbf{R}$  and  $\mathbf{r}$  are positions of the nuclei and electrons in space, respectively and  $t$  is time. The Hamilton operator  $\mathcal{H}$  is related to the electrons and nuclei, and it is the sum of the operators for the kinetic and coulomb electrostatic energy. The Hamilton operator  $\mathcal{H}$  further can be separated as follows: [10]

$$\mathcal{H} = T_e + V_{ee} + V_{eN} + V_{NN} + T_N \quad (2)$$

where  $T_e$  and  $T_N$  is kinetic energy operator for electrons and nuclei.  $V_{ee}$ ,  $V_{eN}$  and  $V_{NN}$  is columbic energy operator for the electrons only, between the electrons and the nuclei, and the nuclei only, respectively. [10]

**II.3.2. Born-Oppenheimer approximation**

Solving above Schrödinger equation (eq. 1) is computationally expensive and not practically feasible for small molecules with more than few atoms. The Born-Oppenheimer approximation alleviates this problem; mass of the electron is much less than mass of the nuclei and consequently, the wave function can be solved separately in steps for electronic motions, nuclear vibrations, and molecular rotations. Using this approximation, one can compute electron and nuclei dynamics separately for a molecule [10].

### II.3.3. Hartree-Fock calculation (HF)

Hartree-Fock (HF) is the most common *ab-initio* calculation to determine the ground state wave function of molecular systems  $\Psi(r, R)$ . The HF method is generally used to solve the time independent Schrödinger equation  $H \Psi(r, R) = E \Psi(r, R)$ .

This method considers several approximations for many-body systems, which includes the Born- Oppenheimer approximation. Another important approximation is the single particle approximation. Standard electronic structure methods approximate the total wave functions of the molecular system into the product of the single-electron wave functions. The effect of HF theory describes that the each electron in a molecule as moving in the average electric field generated by the other electron and nuclei. That means the Columbic repulsions among electrons are not explicitly taken into consideration; however, their average effect is included in the calculation.

Variational calculation of HF theory implies that the approximate energies calculated are all equal or greater than the exact energy; and, based on the size of the basis set, we can determine the accuracy of the calculation [1].

### II.3.4. Basis sets

There are a lot of specific basis sets [6], which differs in the type and number of basis functions. The most common basis sets are the following:

**STO-3G:** This basis set approximates the wavefunction of each orbital by a set of three Gaussians and it is the simplest possible basis set. An analogous case is STO-nG.

**6-31G:** This notation means that each valence orbital is expressed by two sets of Gaussians (one with 3 Gaussians and second with one Gaussian) and other orbitals by one set of six Gaussians. Analogous cases are 3-21G, 6-311G, etc..

**6-31G\*:** This basis set is an extension of 6-31G, which has one more Gaussian (called polarization function) for all atoms other than hydrogens (analogously all basis sets with \*).

These basis sets belong to the most advanced available bases. [3]

### II.3.5. Density Functional Theory (DFT)

Typically, there are two approaches to calculation of total energy of molecules. One of them is the wave function based methods that are more accurate but require higher computer power. The goal of these methods is to calculate the molecular wave function. The other one is density functional theory (DFT) with the goal of calculation of the molecular electron probability density  $\rho$  in order to get the molecular electronic energy from  $\rho$ . This latter approach needs less computer power and is capable of modeling systems with large number of electrons. In general, DFT is an energy minimization method. In this method, the atom nuclei are assumed in fixed positions and the electrons around the nuclei are described as an electron density. In DFT calculation, the energy of the system and the electron density is calculated by solving a set of one-electron Schrödinger equations (the Kohn-Sham equations) instead of solving a many-electron Schrödinger equation. In density functional theory method, the attempt is to solve the Schrodinger equation in terms of electron density, which is defined by

$$\rho[\vec{r}] = \int \psi * \psi d\vec{r} \quad (3)$$

The electron density gives the number of electrons at a given position denoted by  $\vec{r}$ .

Based on Hohenberg-Kohn Theorem in DFT, the energy of the system becomes

$$E_0[\rho_0] = \bar{T}[\rho_0] + \bar{V}_{Ne}[\rho_0] + \bar{V}_{ee}[\rho_0] \quad (4)$$

where  $E_0$  is the ground state energy,  $\rho_0$  is the ground state electron density,  $T$  is the kinetic energy,  $\bar{V}_{Ne}$  is the potential energy from nuclei-electron interactions, and  $\bar{V}_{ee}$  is the potential energy from electron-electron interaction.

In this equation  $E_0$ ,  $\bar{T}$ ,  $\bar{V}_{Ne}$ , and  $\bar{V}_{ee}$  are mathematically defined as functionals. A functional is similar to a function (which takes a number and returns a number), but takes a function and returns a number. Therefore, if the functionals in equation 4 are known, we can calculate the ground state energy by finding the ground state wave function.

Kohn and Sham suggested to take a reference system of non-interacting electrons in an unknown external potential which leads to the same electron density of the ground state density,  $\rho_0$  [2].

### II.4. Semi-empirical Calculations

Semi-empirical methods consider only valance electrons of the system and the core electrons are subsumed into the nuclear core. The key point in semi-empirical methods is the overlap matrix  $S$  (in Roothaan-Hall equation FC=SCE), which is approximated by the identity matrix  $I$  [11]. Therefore, all diagonal elements of the overlap matrix are equal to one and all off-diagonal elements are zero. Thus, the Roothaan-Hall equation FC=SCE becomes FC=CE (F represents the Fock matrix, is a sum of one- and two-

electron contributions,  $C$  is the molecular orbital coefficients and  $E$  is the energy levels) [7, 11].

### II.5. Molecular Dynamics

By using statistical mechanics, microscopic information calculated via molecular dynamics can be used to define macroscopic quantities such as internal energy, temperature, pressure, and enthalpy. Using this link, one can define particular statistical mechanical ensembles for a simulation. These ensembles include constant number, volume, and energy (NVE); constant number, pressure, and temperature (NPT); constant number, pressure, and enthalpy (NPH); and constant number, volume, and temperature (NVT). Associated with each constant property is a controller algorithm used to maintain the property value [12].

### II.6. Molecular mechanics

Today, many of the MM force fields use relatively simple four components of the intra- and inter-molecular forces within the system. Energetic penalties are associated with the deviation of bond lengths, bond angles and torsion angles from their reference or equilibrium values. [7]

Molecular mechanics is a classical mechanical technique for computing energies and selected properties of molecular systems. In an all-atomic molecular representation, the electronic motion is ignored and each atom is defined as a single particle. The potential energy of the system is calculated as a function of the nuclear position using a force field. Force fields are empirical models, designed to reproduce primarily the

structural properties of molecules. They define the functional form and parameters describing the interactions between the atoms. Most atomic force fields have a similar functional form that is a compromise between accuracy and computational efficiency.

The basic molecular mechanics functional form describes the interactions between the atoms as a summation of bonded and non-bonded terms (eq. 5). The bonded terms are computed for the atoms covalently linked and the non-bonded terms are computed for all pairs of atoms that are at least separated by three bonds. The potential  $U(r^N)$  depends on all the positions  $r$  of the  $N$  atoms composing the system (eq. 5), [13]

$$U(r^N) = \sum_i^{bonds} U_{bond}(r_i) + \sum_i^{angles} U_{angle}(\theta_i) + \sum_i^{torsions} U_{torsion}(\phi_i) + \sum_i^N \sum_{j=i}^N U_{coulomb}(r_{ij}) + U_{vdw}(r_{ij}) \quad (5)$$

The bonded terms are computed for all bonds, angles and torsional angles present in the molecule. The non-bonded terms model the electrostatic and van der Waals interactions between all pairs of atoms, i.e.,  $i$  and  $j$ . [13]

### II.6.1 Bonded interactions

The bonded interactions are used to model all the intra-molecular bonded interactions. They account for the interaction of pairs to quartets of bonded atoms [13].

### II.6.2. Bond stretching

The  $U_{bond}$  models the potential energy for the bond stretching between covalently bonded atoms. Since the forces between bonded atoms is very strong, it is rare that the distance between the atoms deviate significantly from its equilibrium position.

In molecular mechanics,  $U_{bond}$  is often approximated by a harmonic oscillator (eq. 6),

$$U_{bond}(r) = \frac{k_i}{2}(r - r_0)^2 \quad (6)$$

In this case, the potential energy varies with the square of the displacement from the reference bond length  $r_0$  (Hooke's law) [13].

### II.6.3. Angle bending

The  $U_{angle}$  models the interaction between valence angles, *i.e.* three consecutive bonded atoms. It has also been described using Hooke's law (eq.7),

$$U_{angle}(\theta) = \frac{k_i}{2}(\theta - \theta_0)^2 \quad (7)$$

Each angle is defined by the spring constant  $k_i$  and a reference angle  $\theta_0$ . Less energy is required for distorting than for the bond-stretching term. [13]

### II.6.4. Torsional terms

Since substantial energy is required to cause deformation for the bond-stretching and the angle-bending terms, most of the conformational changes in a molecule are due to the interactions between the torsional and non-bonded contributions.  $U_{torsion}$  models the interaction between a bonded quartet of atoms. The torsional potential is almost always a periodic potential cosine series expansion (eq. 8),

$$U_{torsion}(\phi) = k_d(1 + \cos(n\phi - \phi_0)) \quad (8)$$

The value of  $k_d$  is a quantitative indication of the barrier height to rotation. The phase factor  $\phi_0$  defines where the energy minimum are. The multiplicity term  $n$  determines the number of minima [13].

### II.6.5. Non-bonded interactions

The molecular structures are not only determined by the intra-molecular bonded interactions, but are also strongly dependent on the non-bonded forces between the atoms. In most force fields, the non-bonded terms are divided into two types: (i) the electrostatic and (ii) the van der Waals interactions. [13]

### II.6.6. Electrostatic interactions

An unequal distribution of charge in a molecule arises due to the difference in the electronegativity of its different constituting atoms. The electronegative atoms attract more the electrons and acquire a small negative charge. In an all-atom force field, each particle will have a specific partial atomic charge, designed to reproduce the quantum mechanic electrostatic properties of the molecules. Coulomb's law is used to compute the electrostatic interaction as a sum of interactions between pairs of atoms (eq. 9), [13]

$$U_{coulomb}(r) = \frac{q_i q_j}{4\pi\epsilon_0 r} \quad (9)$$

$q_i$  and  $q_j$  are the partial charges of two atoms and  $\epsilon_0$  is the permittivity of free space.

The electrostatic contribution is commonly ignored in the force fields of the bonded pair and triplet of atoms, since the stretching (1 – 2) and bending (1 – 3) terms already account for the interactions among atoms. For the atoms composing the torsional angle (1 – 4), the electrostatic interactions are either scaled down or ignored. The computation

of electrostatic contributions between all the atoms of a system is expensive, *i.e.* it scales as  $N^2$ , where  $N$  is the total number of atoms. [13]

### II.6.7. Van der Waals interactions

The Dutch scientist Johannes Diderik van der Waals observed and quantified deviations in the behavior of rare gas compared to ideal gas behavior. His findings suggested the presence of interactions different from electrostatic ones, since the multipole moments of these atoms are zero. The van der Waals interactions arise from a balance between attractive forces at long-range and repulsive forces at short distances and passes through a minimum at around  $2.5 - 4 \text{ \AA}$  (*e.g.*  $3.8 \text{ \AA}$  for argon).

These forces are due to two quantum effects:

(i) attractive London forces.

(ii) repulsive forces derived from the Pauli principle.

The computation of these forces are complicated at the quantum level. Therefore in order to model this effect at the classic level, the best known van der Waals potential is the Lennard-Jones 12 – 6 function (eq. 10), [13]

$$U_{IJ}(r) = 4\varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r} \right)^{12} - \left( \frac{\sigma_{ij}}{r} \right)^6 \right] \quad (10)$$

The Lennard-Jones potential is characterized by a single minimum of depth  $\varepsilon$ , and converges to zero at infinity. The parameter  $\sigma$  defines the collision diameter, *i.e.* the separation at which the energy is zero. The determination of the van der Waals parameters  $\sigma$  and  $\varepsilon$  for all pair of atom-types would be very time consuming. To overcome this problem, they are commonly obtained from the parameters of

monoatomic species using the Lorentz-Berthelot mixing rule. The collision diameter  $\sigma_{ij}$  is computed as the arithmetic mean of each atom (eq. 11). The well depth is equal to the geometric mean (eq.12), [13]

$$\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}) \quad (11)$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \quad (12)$$

### II.7. The force field

Force field can be regarded as a mathematical function of the conformation of the system.

There are several principles of designing the force field:

- 1) Electrons and nuclei are regarded as atom like particles.
- 2) The shape of atom-like particles are spherical and their charges can be obtained from the theory.
- 3) Interactions are based on springs and classical potentials.
- 4) Interactions must be specified for particular sets of atoms.
- 5) Interactions are determinations of atom like particles energies.

The parameters in the force field are obtained usually either from *ab initio* or semi-empirical quantum mechanical calculations, or by fitting experimental data: neutron, Xray and electron diffraction, NMR, infrared, Raman and neutron spectroscopy [14].

### II.7.1. Type of force field

The energy equation mentioned previously is valuable for most force fields, including CHARMM, AMBER, GROMOS and OPLS [15, 16].

#### II.7.1.1. AMBER

One of the molecular mechanics software packages that is widely used is AMBER.

AMBER was developed by American scientists using Fortran 77, and it can be used in different environments. The force field of the package is based on harmonic bonded interactions with 6-12 non-bonded interactions. The molecular dynamics simulations are implemented with constant temperature and pressure with the possibility of using different constraints and restraints. Before AMBER, there were several available software packages, but unfortunately they were made only for a particular class of compounds [17].

The principal use of AMBER force fields is in the area of proteins and nucleic acids. Because of the limitations, it has limited use in drug design [18].

The effect of the polarisability of molecules is incorporated in AMBER calculation of the total energy, where the induced moments are in the interaction with fixed charges and not with each other (bond, angle, dihedral, van der Waals, electrostatic, polarization components in total energy calculation):

$$E_{\text{total}} = \sum_{\text{bonds}} K_r (r-r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta (\theta-\theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{1}{2} V_n [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \left( \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{R_{ij}^c} \right) + \sum_i^{\text{atoms}} (\mu \cdot E_i), \text{ where:} \quad (13)$$

$$\mu_i = \alpha_i \mathbf{E}_i, \quad \mathbf{E}_i = E_i^0 + \sum_{j=1, j \neq i} T_{ij} \mu_j, \quad (14)$$

$$E_i^0 = \sum_{j=1, j \neq i} q_j \frac{r_{ij}}{r_{ij}^2}, \quad (15)$$

$$T_{ij} = \frac{1}{r_{ij}} \left[ 3r_{ij} \left( \frac{r_{ij}}{r_{ij}^2} \right) - 1 \right]. \quad (16)$$

Internal restraints can be regarded as “time-averaged”, which can be calculated using the following equation:

$$\bar{r} = (1/C) \left( \int_0^t e^{-(t'-t)/\tau} r(t')^{-i} dt' \right)^{-1/i} \quad (17)$$

Where  $\bar{r}$  is the time-averaged value of the internal coordinate,  $\tau$  is the exponential decay constant,  $r(t')$  is the value at time  $t'$ ,  $i$  is the average over internals to the inverse of  $i$ ,  $C$  is the normalization integral.

The free energy can be calculated according to the equation: [14]

$$\Delta G_i = G(\lambda(i+1)) - G(\lambda(i)) = \langle \partial V / \partial \lambda \rangle_{\lambda(i+1)} - \langle \partial V / \partial \lambda \rangle_{\lambda(i)} (\lambda(i+1) - \lambda(i)) / 2. \quad (18)$$

## II.8. Minimization Methods

Energy is a function of the atomic coordinates of the system and software programs attempt to generate the xyz coordinates of atoms which correspond to a minimum energy.

This is accomplished by minimization procedures and these techniques are iterative in which the atomic coordinates are altered from one iteration to next in order to minimize the energy.

Commonly used optimization methods are: [7]

(i) **The Steepest Descent (SD) Method:** The SD method uses a first-order minimization algorithm and changes the coordinates of the atoms in the system gradually in order to reach closer and closer to minimum energy. In the SD method,  $r$  is defined as a vector of all coordinates of  $N$  atoms in the system. Then, the net force on each atom  $F$  and the potential energy at each iteration step  $t$  are calculated. New positions are calculated by [19]:

$$r_i^{t+1} = r_i^t + F(r_i) \quad i = 1, 2, 3, \dots, N. \quad (19)$$

(ii) **The Conjugate Gradients (CG) Method:** In the SD method, both gradients and direction of successive steps are orthogonal, however in the CG method, the gradients at each point are orthogonal but the directions are conjugate. New positions are calculated by: [7]

$$r_i^{t+1} = r_i^t + h_i^{t+1} \quad i = 1, 2, 3, \dots, N. \quad (20)$$

where,

$$h_i^{t+1} = F(r_i^{t+1}) + \gamma_i^t h_i^t \quad \text{and, a scalar constant}$$

$$\gamma_i^t = \frac{F(r_i^{t+1}) \cdot F(r_i^{t+1})}{F(r_i^t) \cdot F(r_i^t)}$$

(iii) *The Powell Method:* Another first-order minimization algorithm is the Powell method and belongs to the CG family of optimization method. It applies a more efficient CG method to determine the descent direction and it is more tolerant to inexact line searches [19, 20].

(iv) *The Newton-Raphson Method:* This method does not use only first derivatives, but also the second derivatives in the minimization procedures. Second derivatives give information about the curvature of the function [7, 19, 20].

## II.9. Computer Simulations Methods

### II.9.1 Energy Minimization Method

The aim of energy minimization is to find a set of coordinates representing a molecular conformation such that the potential energy of the system is at a minimum. This can be formally stated as follows: given a function  $f$  which depends on one or more independent variables  $x_1, x_2, x_3, \dots, x_i$ , find the values of those variables where  $f$  has a minimum value [19] [21]. At the minimum, the first derivative of the function with respect to each variable is zero and second derivatives are all positive:

$$\left(\frac{\partial f}{\partial x_i} = 0; \text{ and } \frac{\partial^2 f}{\partial x_i^2} > 0\right).$$

Information provided by energy minimization calculations in some cases can be sufficient to predict accurately the properties of a system. If low energy conformations of a system on an energy surface can be identified then statistical mechanics techniques can be used to derive a partition function from which thermodynamic properties can be calculated. However, this is possible only for small molecular assemblies in the gas phase. Computer simulation methods assist to study large systems and predict their

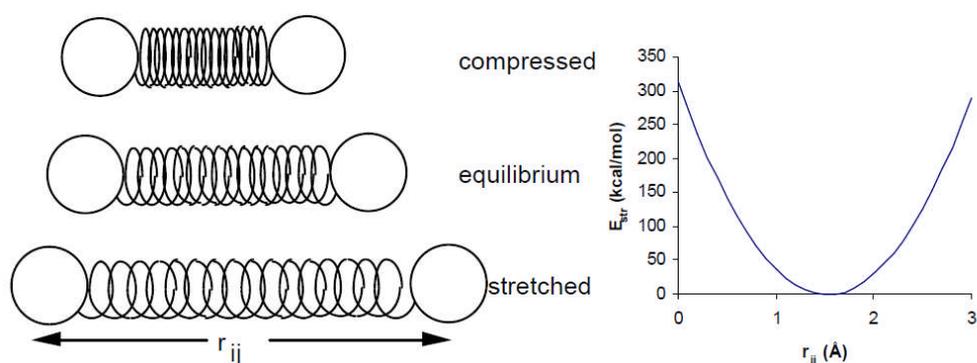
properties through the use of techniques that consider small replications of a macroscopic system with manageable numbers of atoms [19]. Simulations generate a time dependent behavior of these small replications in such a way that accurate values of structural and thermodynamic properties can be obtained with a feasible computation time [7].

Molecular mechanics assumes the steric energy of a molecule to arise from a few, specific interactions within a molecule. These interactions include the stretching or compressing of bonds beyond their equilibrium lengths and angles, torsional effects of twisting about single bonds, the Van der Waals attractions or repulsions of atoms that come close together, and the electrostatic interactions between partial charges in a molecule due to polar bonds. To quantify the contribution of each, these interactions can be modeled by a potential function that gives the energy of the interaction as a function of distance, angle, or charge [22, 23]. The total steric energy of a molecule can be written as a sum of the energies of the interactions:

$$E_{\text{steric energy}} = E_{\text{str}} + E_{\text{bend}} + E_{\text{str-bend}} + E_{\text{oop}} + E_{\text{tor}} + E_{\text{vdW}} + E_{\text{qq}} \quad (21)$$

### II.9.2 Steric energy

Steric energy ( $E_{\text{str}}$ ) represents the energy required to stretch or compress a bond between two atoms, (Figure 2.2).



**Figure 2.2:** Bond Stretching.

A bond can be thought of as a spring having its own equilibrium length,  $r_0$ , and the energy required to stretch or compress it can be approximated by the Hookian potential for an ideal spring:

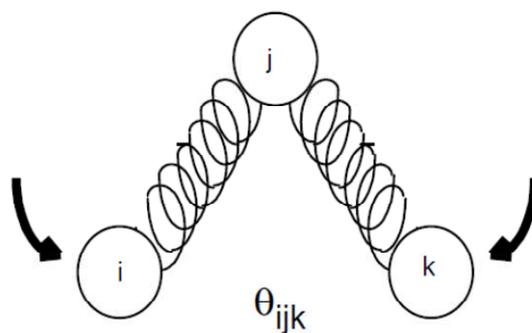
$$E_{\text{str}} = 1/2 k_{s,ij} (r_{ij} - r_0)^2 \quad (22)$$

where  $K_{s,ij}$  is the stretching force constant for the bond and  $r_{ij}$  is the distance between the two atoms, (Figure 2.2).

$E_{\text{bend}}$ : is the energy required to bend a bond from its equilibrium angle  $\theta_0$ . Again this system can be modeled by a spring, and the energy is given by the Hookian potential with respect to angle:

$$E_{\text{bend}} = 1/2 k_{b,ijk} (\theta_{ijk} - \theta_0)^2 \quad (23)$$

where  $k_{b,ijk}$  is the bending force constant and  $\theta_{ijk}$  is the instantaneous bond angle (Figure 2.3).

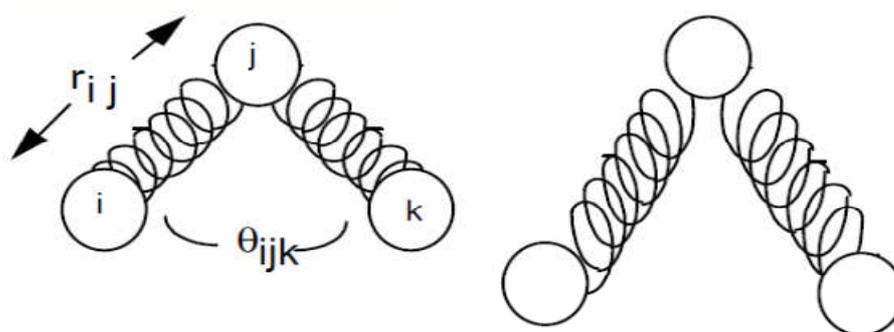


**Figure 2.3:** Bond Bending.

$E_{str-bend}$  : is the stretch-bend interaction energy that takes into account the observation that when a bond is bent, the two associated bond lengths increase (Figure 2.4). The potential function that can model this interaction is:

$$E_{str-bend} = 1/2 k_{sb,ijk} ( r_{ij} - r_o ) ( \theta_{ijk} - \theta_o ) \quad (24)$$

where  $k_{sb,ijk}$  is the stretch-bend force constant for the bond between atoms i and j with the bend between atoms i, j, and k.

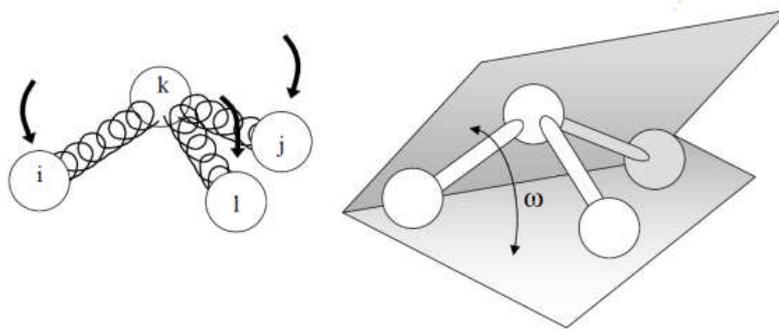


**Figure 2.4:** Stretch-Bend Interaction

$E_{oop}$ : is the energy required to deform a planar group of atoms from its equilibrium angle,  $\omega_o$ , usually equal to zero [24]. This force field term is useful for sp<sup>2</sup> hybridized atoms such as doubly bonded carbon atoms, and some small ring systems. Again this system can be modeled by a spring, and the energy is given by the Hookian potential with respect to planar angle:

$$E_{oop} = 1/2 k_{o,ijkl} (\omega_{ijkl} - \omega_0)^2 \quad (25)$$

where  $k_{o,ijkl}$  is the bending force constant and  $\theta_{ijk}$  is the instantaneous bond angle (Figure 2.5).



**Figure 2.5:** Out of Plane Bending

*Torsional Interactions:*  $E_{tor}$  is the energy of torsion needed to rotate about bonds.

Torsional energies are usually important only for single bonds because double and triple bonds are too rigid to permit rotation. Torsional interactions are modeled by the potential:

$$E_{tor} = 1/2 k_{tor,1} (1 + \cos \phi) + 1/2 k_{tor,2} (1 + \cos 2 \phi) + 1/2 k_{tor,3} (1 + \cos 3 \phi) \quad (26)$$

The angle  $\phi$  is the dihedral angle about the bond. The constants  $k_{tor,1}$ ,  $k_{tor,2}$  and  $k_{tor,3}$  are the torsional constants for one-fold, two-fold and three-fold rotational barriers, respectively [24].

## II.10. Monte Carlo (MC) Method

MC is another computer simulation method, while the low energy conformations of a system are connected to the time, in a MD simulation; in a MC simulation, each

conformation depends only to the predecessor and not upon any other conformations previously visited [20]. The MC technique derives conformations randomly and uses a special set of criteria to decide whether or not to accept each new conformation. These criteria ensure that the derived conformation is equal to its Boltzman factor

$$\exp\{-\mathcal{G}(\mathbf{r}^N)/k_B T\}$$

Where  $\mathcal{G}(\mathbf{r}^N)$  is calculated using the potential energy function [19].

In MC method, each new conformation of the system is generated by randomly rotating the bonds. The energy of the new system is then calculated using the potential energy function, and if the energy of the new system is lower than the energy of its predecessor then the new conformation is accepted. If the energy of the new system is higher than its predecessor, then the Boltzmann factor of the energy difference is calculated with

$$\exp\{-(\mathcal{G}_{new}(\mathbf{r}^N) - (\mathcal{G}_{old}(\mathbf{r}^N)))/k_B T\}$$

and compared with a generated random number between 0 and 1 [19]. If the Boltzmann factor is greater than the random number then the new conformation is accepted, if not, then initial conformation is retained for the next move [19]. Indeed, numbers generated from random number generator are not truly random because the same sequence of numbers should always be generated when the program is run with the same initial conditions [7].

There are some difficulties applying the MC simulations to flexible molecules. It has been found that, even small movements away from an equilibrium bond length, cause a large increase in the energy. One of the widely used methods to overcome this problem is to freeze out some of the internal degrees of freedom, such as bond lengths and bond angles [19, 25, 20].

### II.11. Structure-Activity Relationships (SAR)

Molecular properties are coded by molecular structure. Compounds with similar structures often tend to have similar pharmacological activity. However, they usually exhibit differences in potency, undesirable side effects and in some cases different binding affinities. These structurally related differences are commonly referred to as SAR.

SAR can be defined as “the relationship between chemical structure and pharmacological activity for a series of compounds” [7, 26, 27].

#### II.11.1. QSAR

The success of the SAR approach to drug design depends not only on the knowledge and experience of the drug design group but also may be related with luck. QSAR is an attempt to remove the luck factor from drug design by establishing a relationship in the form of a mathematical equation between biological activity and the measurable physicochemical parameters of a drug that represents its properties such as lipophilicity, shape and electron distribution, which have major effects on the activity [26, 27]. Quantitative structure-property relationships (QSPR) are also used, particularly when a specific property other than biological activity is considered. If an equation is formed, then a medicinal chemist could determine from the equation the value of parameter, and hence the structures, that would optimize the activity. These predictions allow medicinal chemists to make a more informed choice as to what analogues to design and synthesize. Obviously, this could considerably cut down the cost of drug development.

The relationship between these numerical properties and the activity is described by a general equation;  $v = f(p)$ , where  $v$  is the activity (usually defined as  $\log(1/C)$ , where  $C$  is the concentration of the compound required to produce standard response in a given time), and  $p$  is the molecular descriptor (i.e., structure-derived properties of the molecule). These properties that influence the activity of a drug are quite diverse, the major ones being lipophilicity, steric effects and electronic effects.

The parameters commonly used to represent these properties are partition coefficients for lipophilicity, Taft steric constants for steric effects, and Hammett  $\sigma$  constants for electronic effects [7].

### References:

- [1] Praveena Battu, University of Hyderabad, India (2011).
- [2] Elaheh Kamaloo, *thesis of master*, Worcester Polytechnic Institute (2013).
- [3] Durdagi, S., *thesis of doctorate*, Berlin University (2009).
- [4] Leach, A.R.: *Molecular modeling*. Longman, (1996).
- [5] Jensen, F.: *Computational chemistry*. Wiley, (1999).
- [6] Atkins, P.W. *Physical chemistry*. Oxford university press, (1992).
- [7] Durdagi, S.; Supuran, C. T.; Strom, A. T.; Doostdar, N.; Kumar, M.K.; Barron, A. R.; Mavromoustakos, T.; Papadopoulos, M.G. *J. Chem. Inf. Model.* 49, 1139 -1143 (2009).
- [8] Levine, I.N. *Quantum Chemistry*, 5th edition, Prentice Hall,( 2000).

- [9] Hinchliffe, A. *Computational Quantum Chemistry*, 1st edition, John Wiley and Sons Ltd. 1989.
- [10] Rajendra Kumarn, *thesis of doctorate*, Georg-August University Göttingen (2013).
- [11] Lipkowitz, K.B., Boyd, D.B. *Reviews in Computational Chemistry*, VCH Publishers, 1990.
- [12] Derrick B. Callander, *thesis of doctorate*, Georgia Institute of Technology (2005).
- [13] Thomas Lemmin, *thesis of doctorate*, polytechnic federal school lausanne (2013).
- [14] Biljana Arsic, *thesis of doctorate*, University of Manchester (2012).
- [15] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. Mackerell Jr., *J. Comput. Chem.*, 31, 671, (2010).
- [16] W. L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.*, 110,1657, (1988).
- [17] P. K. Weiner and P. A. Kollman, *J.Comput. Chem.*, 2, 287, (1981).
- [18] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.*, 25, 1157, (2004).
- [19] Leach, A.R. *Molecular Modelling: Principles and Applications*, 2<sup>nd</sup> edition, Pearson Educ. Ltd., England, (2001).
- [20] Jensen, F. *Introduction to Computational Chemistry*, John Wiley & Sons Ltd.(1999).
- [21] Hinchliff, A. *Molecular Modelling for Beginners*. John Wiley & Sons Ltd.(2003).
- [22] Hehre, W. J., *A Guide to Molecular Mechanics and Quantum Chemical Calculations*, Wavefunction, Inc., Irvine, CA, , 436-441 (2003).

- [23] Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.*, *117*, 5179-5197, (1995).
- [24] Gasteiger, J., M. Marsili, *Tetrahedron*, *36*, 3219-3228, (1980).
- [25] Cramer, C.J. *Essentials of Computational Chemistry: Theories and Models*, 1<sup>st</sup> edition, John Wiley & Sons Ltd. (2002).
- [26] Patrick, G.L. *An Introduction to Medicinal Chemistry*, 2nd edition, Oxford University Press Inc., New York, (2001).
- [27] Thomas, G. *Medicinal Chemistry: An Introduction*, John Wiley & Sons Ltd (2000).

**Chapter III:**

**Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA**

### **III.1. Introduction**

The influence of ultra-violet and visible light in pyrimidine bases were reported to modify the biological activity [1]. When the nature of base is changed by chemical, photochemical or thermal reaction, its size and hydrogen bonding ability are altered [2, 3]. Photochemical and thermal sigmatropic hydrogen migration in the formation of tautomers of pyrimidines (cytosine, thymine and uracil) and purines (adenine and guanine) play a vital role in changing the base sequence to cause chromosomal mutations, sometimes useful and occasionally harmful [4]. These tautomeric shifts alter the distribution of electrons and protons in the molecule with the change of functional group as  $-\text{NH}_2$  (amine)  $\leftrightarrow$   $=\text{NH}$  (imino) and  $>\text{C}=\text{O}$  (keto)  $\leftrightarrow$   $=\text{C}-\text{O}-\text{H}$  (enol). Then it cannot form a hydrogen bonded pair within a DNA molecule with its usual partner. Thus, it is expected to influence selective hydrogen bonding in nucleic acids and alter the sequences of bases for mutation during DNA duplication [5].

This damage can lead a change in the sequence of bases with mutations and increase likelihood of the development of cancerous cells [6]. There is an increasing tendency observed to simplify the descriptions of natural systems, relying on the creating of models being some approximation of the systems [7]. Many side effects of chemical therapy prompted medical and biological scientists to lean towards biological therapy [8-13]. In this connection, the advanced technology of genetic engineering for the synthesis of genes and discovery of miracle enzymes are required to cure many diseases [14].

Quantum chemistry methods play an important role in obtaining molecular geometries and predicting various properties [15]. To obtain highly accurate geometries and physical

## **Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.**

---

properties for molecules that are built from electronegative elements, expensive DFT electron correlation methods are required [16-20].

The molecular electrostatic potential surface MESP which is a plot of electrostatic potential mapped onto the iso-electron density simultaneously displays molecular shape, size and electrostatic potential values and has been plotted for the molecules. The molecular electrostatic potential (MESP) mapping is very useful in the investigation of the molecular structure with its physiochemical property relationships [9, 14].

A portion of the molecule that has a negative electrostatic potential is susceptible to electrophilic attack. The red and blue regions in the MESP map refer to the regions of negative and positive potentials and correspond to the electron rich and electron-deficient regions respectively whereas the green color signifies the neutral electrostatic potential [21, 22].

The mechanistic investigation of sigmatropic rearrangement in thymine affected in the conformational changes [23]. In view of these observations, we applied the PM3 and DFT methods for conformational analyses [24-29], for that our aim is to analyze the geometry and electronic structures of thymine, thymidine and thymidylic acid.

### **III.2. Computational Methods**

Semi-empirical molecular orbital calculations were performed using PM3 and DFT methods included in Gaussian 09, this work also includes calculation of 3D MESP surface map to reveal the information regarding charge transfer within the molecule [30].

### **III.3. Results and Discussion**

#### **III.3.1. Geometric and Electronic structure of thymine, thymidine and thymidylic acid**

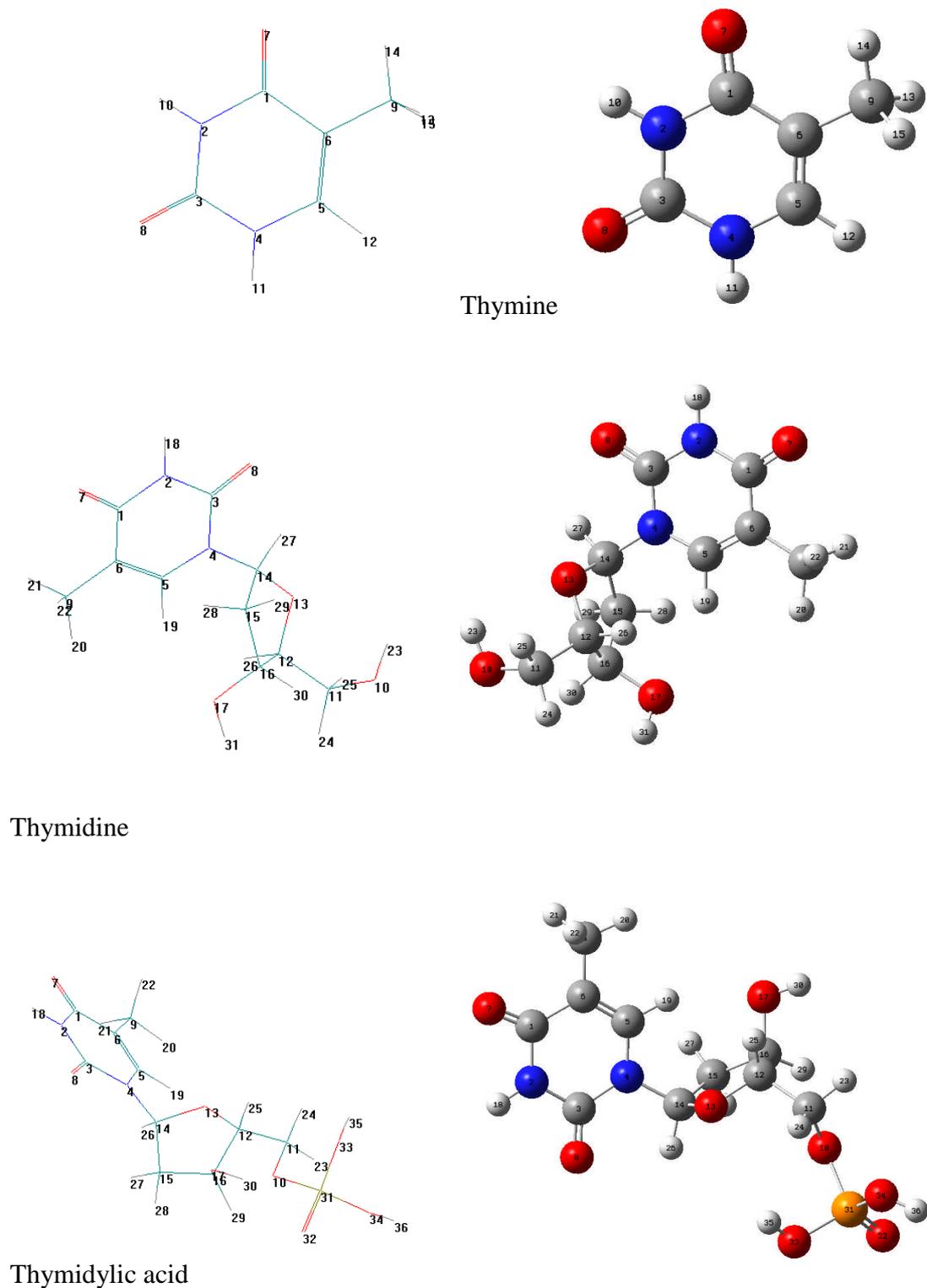
The electronic structures of thymine, thymidine and thymidylic acid along with the numbering of the system in this context are shown in figure 3.1. The calculated heats of formation, dipole moments, frontier orbitals energies and net charges on typical atoms of thymine tautomers (a-f), thymidine 1a-3a and thymidylic acid 1b-3b are presented in table 3.1 tables 3.2-a and table 3.2-b respectively.

Photochemical or thermal (1; 3) sigmatropic rearrangement [23] will take place by shifting of a hydrogen atom with its  $\sigma$ -bond within a  $\pi$ -bond frame-work to a new site in two pathways: i.e. the hydrogen moves along the top or bottom face of the  $\pi$ -system or across the  $\pi$ -system from top to bottom or vice versa.

## Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

### Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

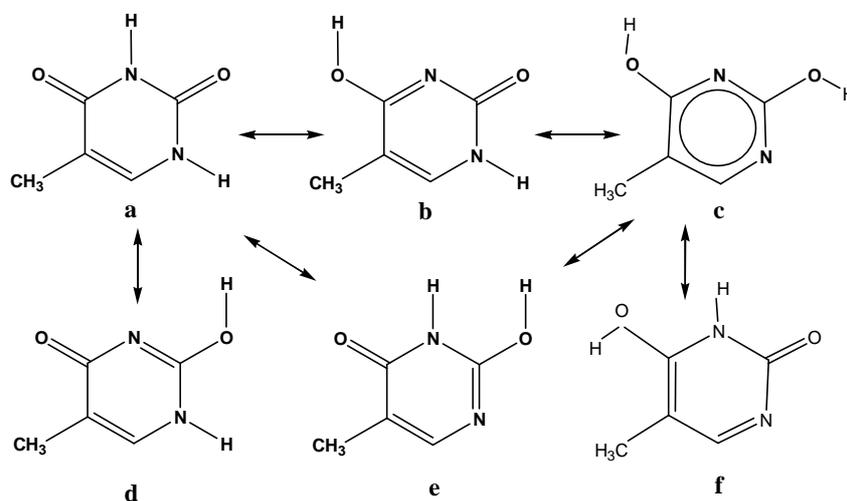
---



**Figure 3.1:** 3D structures of thymine, thymidine and thymidylic acid.

### Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

The former is called suprafacial and later antarafacial migration respectively. But in a photochemical reaction, promotion of an electron from HOMO to LUMO, the suprafacial pathway is allowed in the case of a to f tautomers (figure 3.1 and figure 3.2). (1;3) sigmatropic hydrogen migration in tautomers of b, d and f tautomers is easily converted to c tautomer, because of an aromatic sextet. The net charges of O<sub>7</sub> are more in the case of a, d and O<sub>8</sub> is more in the case of a, b then shifting of proton is more susceptible to form tautomers.



**Figure 3.2:** Sigmatropic hydrogen migration in thymine (a – f).

According to the heat of formation data, the stability of tautomers follow the order  $a < d < c < f < b < e$  (table 3.1). The value of dipole moment is in the order  $e > b > f > a > d > c$ .

### Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen

#### Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

**Table 3.1:** Heats of formation  $\Delta H^{\circ}f$  (in kcal/mole ), dipole moments (in Debye),  $\Delta E$  (HOMO/LUMO) and atomic charges for thymine a to f by PM3 Method.

Property	Thymine					
	A	B	C	D	E	F
$\Delta H^{\circ}f$ (kcal/mole )	-76.0293	-64.8901	-68.4575	-69.4981	-53.5455	-66.7318
moment dipolaire( $\mu$ ),	3.9923	7.0065	1.3102	2.1567	7.9113	5.9485
$\Delta E$ (HOMO/LUMO)	0,32879	0,31879	0,3373	0,31953	0,33201	0,31437
HOMO (ev)	0.34682	-0.34221	-0.35167	-0.33887	-0.34000	-0.33565
LUMO (ev)	-0.01803	-0.02342	-0.01437	-0.01934	-0.00799	-0.02128
atomic charge						
C1	0.305450	0.206352	0.202081	0.293143	0.335380	0.124938
C3	0.205484	0.258795	0.180000	0.118432	0.086241	0.253843
C5	-0.086006	-0.052831	0.051202	0.039899	-0.125014	0.090023
C6	-0.234529	-0.295913	-0.248760	-0.264279	-0.215506	-0.365454
C9	-0.045275	-0.028641	-0.035461	-0.034992	-0.052889	-0.012466
N2	-0.010573	-0.197260	-0.229411	0.078267	-0.218457	0.078616
N4	0.098537	0.083737	-0.232940	-0.261362	0.114704	-0.181030
O7	-0.353438	-0.189341	-0.214971	-0.364286	-0.308897	-0.338991
O8	-0.389501	-0.344582	-0.189643	-0.231717	-0.179612	-0.012466

The strong kind of dipole-dipole interactions make hydrogen bonding in the formation of the double helix of DNA, and it permits the self duplication of the molecule, which is the basis of heredity.

Polar bonds allow molecules to interact in various ways with each other. Such secondary interactions are relatively weak and play a vital role in the chemistry of life. When the 2-

### **Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.**

---

deoxyribose is united with N<sub>4</sub>-atom of thymine through a glycosidic bond, the tautomeric forms of thymine are strongly favored in the solid and solution. The tautomeric forms are important in determining the nature of hydrogen bonding in nucleic acids.

The thymidine tautomers 1-3a are formed by attachment of 2-deoxy (-) ribose at N9-position of thymine, through a glycosidic bond (figure 3.1 and figure 3.3). But the promotion of an electron from HOMO to LUMO, a photochemical reaction is allowed in the suprafacial pathway in case of 3a tautomer and all others allowed antarafacial thermal (1;3) sigmatropic rearrangement (figure 3.1 and figure 3.3). Stereochemically, migration of hydrogen can be suprafacial or antarafacial in the biological systems and natural reactions. In the transition state, a three-center bond is required to involve overlap between the s orbital of the hydrogen and lobes of p orbitals of the two terminal atoms. But the s orbital of the hydrogen and lobes of p orbitals in opposite phase cannot be overlapped or bonded simultaneously to both terminal orbitals and therefore, (1; 3) shift of hydrogen is symmetry-forbidden.



### Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen

#### Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

**Table 3.2-a:** Heats of formation  $\Delta H^{\circ}f$  (in kcal/mole), dipole moments (in Debye),  $\Delta E$  (HOMO/LUMO) and atomic charges for thymidine 1 to 3 by PM3 Method.

Property	Thymidine		
	1a	2a	3a
$\Delta H^{\circ}f$ (kcal/mole )	-205.8958	-184.8589	-193.9647
moment dipolaire( $\mu$ ),	4.8923	8.3023	7.4777
$\Delta E$ (HOMO/LUMO)	0,33297	0,33243	0,32015
HOMO (ev)	-0.34547	-0.34362	-0.33835
LUMO (ev)	-0.01250	-0.01119	-0.01820
atomic charge			
C1	0.304492	0.335123	0.207751
C3	0.234994	0.121901	0.286220
C5	-0.066713	-0.115506	-0.026046
C6	-0.229749	-0.203697	-0.289640
C9	-0.039066	-0.048609	-0.036729
C11	0.059745	0.058919	0.059125
C12	-0.016096	-0.010536	-0.016105
C14	0.100899	0.102244	0.098196
C15	-0.146479	-0.144295	-0.148990
C16	0.056981	0.058431	0.057183
N2	-0.017807	-0.219789	-0.206344
N4	0.036623	0.035968	0.018703
O7	-0.354081	-0.302605	-0.187068
O8	-0.392075	-0.184599	-0.347547
O10	-0.307926	-0.305436	-0.308036
O13	-0.267686	-0.276896	-0.264144
O17	-0.305710	-0.301006	-0.307537
H18	0.121079	0.195372	0.199404

### Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen

#### Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

**Table 3.2-b:** Heats of formation  $\Delta H^{\circ}f$  (in kcal/mole), dipole moments (in Debye),  $\Delta E$  (HOMO/LUMO) and the atomic charges for thymidylic acid 1 to 3 by PM3 Method.

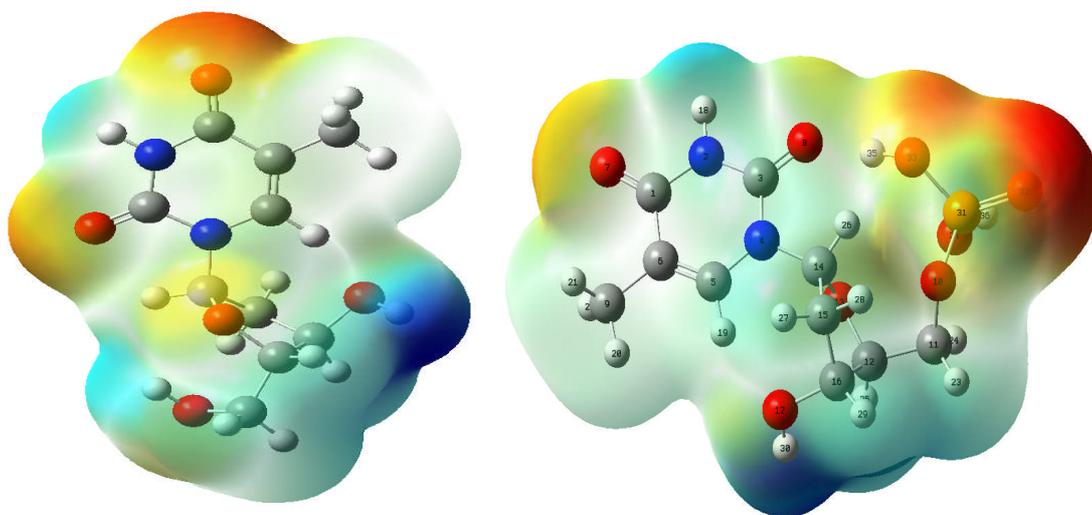
Property	acide thymidylique		
	1b	2b	3b
$\Delta H^{\circ}f$ (kcal/mole )	-409.6794	-389.6958	-193.9647
moment dipolaire( $\mu$ )	5.8877	9.6902	7.9133
HOMO (ev)	-0.34337	-0.33861	-0.33765
LUMO (ev)	-0.01075	-0.00619	-0.01755
$\Delta E$ (HOMO/LUMO)	0.33262	0.33242	7.9133
Atomic charge			
C1	0.305202	0.336806	0.208106
C3	0.235094	0.124323	0.286616
C5	-0.064264	-0.109246	-0.024430
C6	-0.234036	-0.209837	-0.292456
C9	-0.037483	-0.046276	-0.035573
C11	0.140831	0.137183	0.141143
C12	0.006390	0.017428	0.009204
C14	0.106880	0.106003	0.103853
C15	-0.143363	-0.139653	-0.145147
C16	0.062194	0.066097	0.062018
N2	-0.019102	-0.224876	-0.208035
N4	0.034677	0.040116	0.018510
O7	-0.356561	-0.307851	-0.187784
O8	-0.392101	-0.187485	-0.350637
O10	-0.640883	-0.636470	-0.650699
O13	-0.273616	-0.278633	-0.268307
O17	-0.311650	-0.307542	-0.313533
H18	0.120627	0.199451	0.199287

### Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

---

According to the heats of formation data, the stability of thymidine tautomers is in the order of  $2a < 3a < 1a$ . These tautomers are formed with the difference in the heats of formation of 10.1058 k.cal/mol and 21.0369 k.cal/mol respectively in the case of 2a to 3a and 2a to 1a.

The value of dipole moment is in the order of  $2a > 3a > 1a$ . The hydrogen bonding is a strong kind of dipole-dipole attraction, in which a hydrogen atom stands as a bridge between the electronegative atoms by holding one with a covalent bond and the other by purely electrostatic forces. Tautomer 2a has more polarity than the other tautomers (1a and 3a). The position of equilibrium depends upon the polarity of the solution and also the presence of tautomers.



**Figure 3.4:** Molecular electrostatic potential surface of thymidine 1a and thymidylic acid 1b by DFT Method .

Tautomers of thymidylic acid 1-3b are formed by attachment of phosphate at O<sub>10</sub>-atom position of thymidine (figure 3.1 and figure 3.3). But, a photochemical reaction allows

### **Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.**

---

the promotion of an electron from HOMO to LUMO, in suprafacial pathway in the case of 1b and 3b tautomers.

Thymidylic acid tautomer 2b is allowed antarafacial thermal (1;3) sigmatropic rearrangement. The net charge of O<sub>10</sub> is more in the case of 1-3b tautomers than respective thymidine tautomers 1-3a. The proton shifting from N<sub>2</sub>- to O<sub>8</sub> in the case of 1b to 2b is considered by increasing the atomic charges at N<sub>2</sub>, N<sub>4</sub>, and C<sub>1</sub> and decreasing at O<sub>7</sub>, O<sub>8</sub>, O<sub>10</sub>, C<sub>3</sub>, C<sub>11</sub> and C<sub>15</sub> atoms (table 3.2-b).

The proton shifting from N<sub>2</sub> to O<sub>7</sub> in the case of tautomers 1b to 3b is considered by decreasing net charges at O<sub>7</sub>, O<sub>8</sub>, N<sub>4</sub>, C<sub>1</sub> and C<sub>9</sub> atoms and increasing at O<sub>10</sub>, N<sub>2</sub>, C<sub>3</sub>, C<sub>6</sub>, C<sub>11</sub> and C<sub>15</sub> atoms. The net charges are highest at oxygen atoms (i.e., O<sub>32</sub>, O<sub>33</sub>; and O<sub>34</sub>), when they are attached to phosphorus atom (figure 3.4).

According to the heats of formation data, the stability of thymidylic acid tautomers is in the order of 1b > 2b > 3b. These tautomers are formed with the difference in heats of formation of 19.9836 kcal/mol and 215.7147 kcal/mol respectively in the case of 1b to 2b and 1b to 3b. From this data, it can be predicted that the conversion of 1b to 2b is a lower energy process than the conversion of 1b to 3b.

The value of dipole moment is more in the case of tautomer 2b, and it may exist in more polar medium.

#### **III.3.2. Molecular geometry of thymidine 1-3a**

### Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

---

The (1; 3) sigmatropic hydrogen migration in thymine of thymidine can be shown in the figure 3.3 as per the results of PM3 calculations. The molecular deformations will depend on the changes in the parameters of bond lengths, bond angles, dihedral angles and the change in energy content of the molecule. The spatial arrangement of atoms in a molecule is considered for the (1;3) sigmatropic hydrogen migration in thymine of thymidine and its conformations are shown in figure 3.3.

The main data of bond lengths (table 3.3), bond angles (table 3.4), and dihedral angles (table 3.5) of thymidine tautomers 1-3a and 1-3b (table 3.6) of tautomers are only listed for the sake of simplicity from full-optimized PM3 calculations.

When (1;3) sigmatropic hydrogen migration is involved from N<sub>2</sub> to O<sub>8</sub> and N<sub>2</sub> to O<sub>7</sub> then respective tautomers 2a and 3a will be formed. From table 3.3, it is found that the bond length of N<sub>2</sub>-H in 1a (0.9998A) is larger than O<sub>8</sub>-H bond in 2a (0.9535A) and O<sub>7</sub>-H bond in 3a (0.9509A). The proton shifting from N<sub>2</sub> to O<sub>8</sub> is in the formation of 2a from 1a, the bond lengths are decreased in the case of C<sub>14</sub>-N<sub>4</sub> and C<sub>10</sub>-N<sub>2</sub> by 0.0014 A and 0.1114 A and increased in the case of N<sub>4</sub>-C<sub>3</sub>, N<sub>4</sub>-C<sub>5</sub>, N<sub>2</sub>-C<sub>1</sub>, C<sub>3</sub>-O<sub>8</sub> and O<sub>13</sub>-C<sub>14</sub> by 0.0036 A, 0.0106 A, 0.0104 A, 0.1259A and 0.0100A.

According to the table 3.4, bond angles of C<sub>11</sub>C<sub>12</sub>O<sub>13</sub>, N<sub>2</sub>C<sub>3</sub>N<sub>4</sub>, N<sub>4</sub>C<sub>3</sub>O<sub>8</sub> and N<sub>2</sub>C<sub>1</sub>O<sub>7</sub>, are increased by 0.0226° , 6.746° , 0.213° and 1.4025° and decreased in case of bond angle of C<sub>1</sub>N<sub>2</sub>C<sub>3</sub> and C<sub>3</sub>N<sub>4</sub>C<sub>14</sub> by 2.4343° and 0.2257° .

### Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic

#### Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

As per the table 3.5, dihedral angle is changed from -ac to -sc, in the conformation of tautomers 1a to 2a in the case of  $O_{13}C_{14}N_4C_3$ , +ap conformation is observed in the case of  $N_4C_3N_2H_{18}$  of 1a and +sp conformation in the case of  $N_4C_3O_8H_{18}$ , of tautomer 2a.

**Table 3.3:** Bond lengths ( $\text{\AA}$ ) of thymidine 1-3a and thymidylic acid 1-3b by PM3 Method.

Bond lengths ( $\text{\AA}$ )	Thymidine			acide thymidylique		
	1a	2a	3a	1b	2b	3b
$O_{10} - C_{11}$	1.3991	1.3992	1.3991	1.3954	1.3957	1.3935
$C_{12} - O_{13}$	1.4307	1.4307	1.4303	1.4304	1.4271	1.4293
$O_{13} - C_{14}$	1.4207	1.4307	1.4209	1.42	1.4271	1.4205
$C_{14} - N_4$	1.4989	1.4975	1.4967	1.5005	1.4964	1.4977
$N_4 - C_3$	1.4396	1.4432	1.4606	1.439	1.4425	1.4601
$N_4 - C_5$	1.4171	1.4277	1.3992	1.4165	1.4269	1.3989
$C_3 - N_2$	1.4189	1.3075	1.4226	1.4193	1.3083	1.4227
$N_2 - C_1$	1.43	1.4404	1.3198	1.4297	1.4399	1.3199
$C_1 - O_7$	1.2202	1.2138	1.3461	1.2205	1.2143	1.3462
$C_3 - O_8$	1.2252	1.3511	1.2169	1.2252	1.3508	1.2172
$N_2 - H_{18}$	0.9998			0.9998		
$O_8 - H_{18}$		0.9535			0.9538	
$O_7 - H_{18}$			0.9509			0.9509
$O_{10} - P_{31}$				1.7013	1.6996	1.7048
$O_{32} - P_{31}$				1.4449	1.4441	1.441
$O_{33} - P_{31}$				1.6595	1.6616	1.6736
$O_{34} - P_{31}$				1.6806	1.6801	1.6684
$O_{33} - H_{35}$				0.944	0.9442	0.9429
$O_{34} - H_{36}$				0.943	0.9429	0.9444

(1;3) Sigmatropic hydrogen migration from  $N_2$  to  $O_7$  is involved in the formation of 3a from 1a. The bond lengths of  $N_4 - C_5$ ,  $N_2 - C_1$ , and  $C_3 - O_8$  are decreased by 0.0179

### Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic

#### Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

$A^\circ$ ,  $0.1102 A^\circ$ , and  $0.0083A^\circ$  and increased in the case of  $N_4 - C_3$ ,  $C_3 - N_2$  and  $C_1 - O_7$  by  $0.021 A^\circ$ ,  $0.0037 A^\circ$  and  $0.1259A^\circ$  (table 3.3).

Table 3. 4 shows that the bond angles are decreased by  $2.9314^\circ$ ,  $2.5087^\circ$  and  $5.4848^\circ$  respectively in the case of  $C_1N_2C_3$ ,  $N_4C_3O_8$  and  $N_2C_1O_7$  and increased in case of bond angle of  $C_3N_4C_{14}$  by  $1.4178^\circ$ .

**Table 3.4:** Bond angles ( $^\circ$ ) of thymidine 1-3a and thymidylic acid 1-3b by PM3 Method.

Bond angles ( $^\circ$ )	Thymidine			acide thymidylique		
	1a	2a	3a	1b	2b	3b
$C_{11}C_{12}O_{13}$	108.5386	108.5612	108.6317	109.3223	109.9955	109.7279
$C_{11}C_{12}O_{16}$	113.4151	113.6213	113.3586	114.4873	114.0347	113.9384
$C_3N_4C_{14}$	119.0686	118.8429	120.4864	119.1675	118.8359	120.4921
$C_5N_4C_{14}$	119.227	118.8207	119.6733	119.2857	118.9057	119.6379
$N_2C_3N_4$	118.2454	124.9914	118.215	118.2575	124.9433	118.2528
$C_1N_2C_3$	122.2453	119.811	119.3139	122.2616	119.7763	119.2897
$N_4C_3O_8$	122.3561	122.5691	119.8474	122.4587	122.615	119.9002
$C_3N_2H_{18}$	117.7839			117.7873		
$N_2C_1O_7$	115.7781	117.1806	110.2933	115.7859	117.1651	110.2958
$C_3O_8H_{18}$		109.9585			110.0289	
$C_1O_7H_{18}$			108.7763			108.7678
$O_{10}P_{31}O_{32}$				114.5377	113.6915	112.176
$O_{32}P_{31}O_{33}$				117.745	117.9183	119.0658
$O_{32}P_{31}O_{34}$				115.6723	116.6627	117.8424
$P_{31}O_{33}H_{35}$				118.1268	117.8621	118.8045
$P_{31}O_{34}H_{36}$				116.3001	116.9071	117.2213

### Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen

#### Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

The bond angle of  $C_3N_2H_{18}$  in 1a ( $117.7839^\circ$ ) is changed to  $C_1O_7H_{18}$  in tautomer 3a ( $108.7763^\circ$ ). The dihedral angle is +ap conformation in the case of  $N_4C_3N_2H_{18}$  tautomer 1a and  $N_2C_1O_7H_{18}$  of tautomer 3a (table 3.5).

**Table 3. 5:** Dihedral angles ( $^\circ$ ) of thymidine tautomers 1-3a by PM3 Method.

Dihedral angles ( $^\circ$ )	Thymidine					
	1a	*	2a	*	3a	*
$O_{10}C_{11}C_{12}O_{13}$	65.0089	+sc	64.2857	+sc	65.299	+sc
$O_{10}C_{11}C_{12}C_{16}$	-55.5669	-sc	-55.729	-sc	-55.4579	-sc
$C_{11}C_{12}C_{16}O_{17}$	-118.6508	-ac	-109.9122	-ac	-120.8097	-ac
$C_{12}O_{13}C_{14}N_4$	-113.6525	-ac	-115.9736	-ac	-115.3012	-ac
$O_{13}C_{14}N_4C_3$	-98.5239	-ac	-69.7314	-sc	-109.6707	-ac
$O_{13}C_{14}N_4C_{x5}$	62.9715	+sc	80.3047	+sc	63.4635	+sc
$C_{14}N_4C_3N_2$	170.1536	+ap	158.221	+ap	174.4139	+ap
$N_4C_3N_2C_1$	-6.6722	-sp	-3.9002	-sp	-0.1877	-sp
$C_{12}C_{11}O_{10}H_{23}$	-63.7632	-sc	-66.6397	-sc	-63.6901	-sc
$C_{12}C_{16}O_{17}H_{31}$	73.9803	+sc	72.0881	+sc	75.5018	+sc
$C_{14}N_4C_3O_8$	-13.1899	-sp	-27.1574	-sp	-6.9293	-sp
$C_{14}N_4C_5H_{19}$	13.3127	+sp	25.5952	+sp	5.4328	+sp
$C_3N_2C_1O_7$	-177.4533	-ap	-179.7599	-ap	179.9868	-ap
$N_4C_3N_2H_{18}$	179.3647	+ap				
$N_4C_3O_8H_{18}$			7.7834	+sp		
$N_2C_1O_7H_{18}$					179.8494	+ap

\* Conformation analyses using prefixes a = anti, s = syn, p = peri-planar, c = clinal, and + & - signs.

It can be predicted that the sequence of (1;3) sigmatropic hydrogen migration in thymine of thymidine from 1a to 2a or 3a, changes the conformation and dipole moment. The polar bonds may interact with other molecules in various media.

## Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

---

Further, it is relatively weak in the secondary interactions and play vital role in the replication, translation, transcription, recombination and mutation will take place among the genes.

### III.3.3. Molecular geometry of thymidylic acid 1-3b

Table 3.3 represents that the bond length of N<sub>2</sub>-H<sub>18</sub> in 1b (0.9998A°) is decreased to O<sub>8</sub>-H<sub>18</sub> bond in 2b (0.9535 A°) and O<sub>7</sub> - H<sub>18</sub> bond in 3b (0.9509A°), if the (1;3) sigmatropic hydrogen migration is involved from N<sub>2</sub> to O<sub>8</sub> and N<sub>2</sub> to O<sub>7</sub> in the formation of respective tautomers 2b and 3b. When the (1;3) migration of proton from N<sub>2</sub> to O<sub>8</sub> in thymine of thymidylic acid with the formation of 2b from 1b, the bond lengths of N<sub>4</sub> - C<sub>5</sub>, N<sub>2</sub> -C<sub>1</sub> , C<sub>3</sub> - O<sub>8</sub> and N<sub>4</sub> - C<sub>3</sub> are increased by 0.0104A°, 0.0102A° , 0.1256 A° and 0.0035A° and decreased in the case of C<sub>14</sub> - N<sub>4</sub> and C<sub>3</sub> - N<sub>2</sub> by 0.0041A° and 0.111 A°.

Table 3.4 shows that the bond angles of N<sub>2</sub>C<sub>3</sub>N<sub>4</sub>, N<sub>2</sub>C<sub>1</sub>O<sub>7</sub>, N<sub>4</sub>C<sub>3</sub>O<sub>8</sub>, and O<sub>32</sub>P<sub>31</sub>O<sub>33</sub> , are increased by 6.6858° , 1.3792° , 0.1563° and 0.1733° and decreased in case of bond angle of C<sub>1</sub>N<sub>2</sub>C<sub>3</sub> and O<sub>10</sub>P<sub>31</sub>O<sub>32</sub> by 2.4853° and 0.8462° respectively. The bond angle of C<sub>3</sub>N<sub>2</sub>H<sub>18</sub> in 1b (117.7873°) is changed to C<sub>3</sub>O<sub>8</sub>H<sub>18</sub> with the bond angle of 110.0289° in 2b.

As per the table 3.6, dihedral angle is changed from -ap to +ap in the conformation of tautomers from 1b to 2b in the case of C<sub>3</sub>N<sub>2</sub>C<sub>1</sub>O<sub>7</sub> and O<sub>32</sub>P<sub>31</sub>O<sub>33</sub>H<sub>35</sub> respectively. Dihedral angle is +ap in the conformation of N<sub>4</sub>C<sub>3</sub>N<sub>2</sub>H<sub>18</sub>, of 1b and +sp in the conformation of N<sub>4</sub>C<sub>3</sub>O<sub>8</sub>H<sub>18</sub>, of tautomer 2b.

When the (1;3) migration of proton N<sub>2</sub> to O<sub>7</sub> in the formation of thymidylic acid tautomer 3b from 1b, the bond lengths of N<sub>4</sub> - C<sub>3</sub>, C<sub>1</sub> - O<sub>7</sub> , N<sub>9</sub>-C<sub>10</sub>, and C<sub>14</sub>-O<sub>28</sub> are

### Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen

#### Migration in the formation of Tautomeric forms of some Biomolecules in DNA.

increased by  $0.0211 \text{ A}^\circ$  and  $0.1257 \text{ A}^\circ$  and decreased in the case of  $\text{C}_{12} - \text{O}_{13}$ ,  $\text{C}_{14} - \text{N}_4$ ,  $\text{N}_4 - \text{C}_5$  and  $\text{N}_2 - \text{C}_1$  by  $0.0011 \text{ A}^\circ$ ,  $0.0028 \text{ A}^\circ$ ,  $0.0176 \text{ A}^\circ$  and  $0.1098 \text{ A}^\circ$  (table 3.3).

**Table 3. 6:** Dihedral angles ( $^\circ$ ) of thymidylic acid tautomers 1-3b by PM3 Method.

Dihedral angles ( $^\circ$ )	acide thymidylique					
	1b	*	2b	*	3b	*
$\text{O}_{10}\text{C}_{11}\text{C}_{12}\text{O}_{13}$	75.0802	+sc	72.8498	+sc	70.8506	+sc
$\text{O}_{10}\text{C}_{11}\text{C}_{12}\text{C}_{16}$	-46.9406	-sc	-49.4154	-sc	-51.2424	-sc
$\text{C}_{11}\text{C}_{12}\text{C}_{16}\text{O}_{17}$	-119.0743	-ac	-122.0017	-ac	-121.9429	-ac
$\text{C}_{12}\text{O}_{13}\text{C}_{14}\text{N}_4$	-115.0188	-ac	-128.8488	-ac	-118.2642	-ac
$\text{O}_{13}\text{C}_{14}\text{N}_4\text{C}_3$	-98.6903	-ac	-70.642	-sc	-107.7864	-ac
$\text{O}_{13}\text{C}_{14}\text{N}_4\text{C}_5$	63.2472	+sc	79.6738	+sc	64.7129	+sc
$\text{C}_{14}\text{N}_4\text{C}_3\text{N}_2$	170.5075	+ap	158.9619	+ap	174.616	+ap
$\text{N}_4\text{C}_3\text{N}_2\text{C}_1$	-6.577	-sp	-3.8942	-sp	-0.6682	-sp
$\text{C}_{11}\text{C}_{12}\text{C}_{16}\text{H}_{29}$	3.3259	+sp	0.5795	+sp	0.4745	+sp
$\text{C}_{12}\text{O}_{13}\text{C}_{14}\text{H}_{26}$	129.1243	+ac	116.9398	+ac	126.3425	+ac
$\text{C}_{12}\text{C}_{16}\text{O}_{17}\text{H}_{30}$	75.9224	+sc	72.5164	+sc	77.1087	+sc
$\text{C}_{14}\text{N}_4\text{C}_3\text{O}_8$	-12.6907	-sp	-26.3635	-sp	-6.6707	-sp
$\text{C}_{14}\text{N}_4\text{C}_5\text{H}_{19}$	12.8739	+sp	24.7223	+sp	5.4258	+sp
$\text{C}_3\text{N}_2\text{C}_1\text{O}_7$	-177.4643	-ap	179.8371	+ap	179.9773	+ap
$\text{N}_4\text{C}_3\text{N}_2\text{H}_{18}$	179.2208	+ap				
$\text{N}_4\text{C}_3\text{O}_8\text{H}_{18}$			8.3756	+sp		
$\text{N}_2\text{C}_1\text{O}_7\text{H}_{18}$					179.9451	+ap
$\text{P}_{31}\text{O}_{32}\text{O}_{10}\text{C}_{11}$	-156.3673	-ap	-162.5793	-ap	-160.6822	-ap
$\text{O}_{10}\text{O}_{11}\text{P}_{31}\text{O}_{34}$	-30.288	-sc	-35.6338	-sc	-31.7116	-sc
$\text{O}_{32}\text{P}_{31}\text{O}_{34}\text{H}_{36}$	-9.8211	-sp	-1.9299	-sp	-128.0938	-ac
$\text{O}_{32}\text{P}_{31}\text{O}_{33}\text{H}_{35}$	-168.7755	-ap	163.012	+ap	-86.1135	-sc

\* Conformation analyses using prefixes a = anti, s = syn, p = peri-planar, c = clinal, and + & - signs.

### **Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.**

---

From table 3.4, it is revealed that the bond angle of  $O_{32}P_{31}O_{33}$  is increased by  $1.3208^\circ$  and decreased in the case of  $N_4C_3O_8$  and  $N_2C_1O_7$  by  $2.5585^\circ$  and  $5.4901^\circ$ .

The bond angle of  $C_3N_2H_{18}$  in 1b ( $117.7873^\circ$ ) is changed to  $C_1O_7H_{18}$  in 3b ( $108.7678^\circ$ ) tautomer. Dihedral angle is +ap in the conformation of  $N_4C_3N_2H_{18}$ , of 1b and in the conformation of  $N_2C_1O_7H_{18}$ , of tautomer 2b (table 3.6).

#### **III.4. Conclusion**

The present study provides a discussion of several qualitative approximations using PM3 and DFT calculations to predict weak interactions.

Our investigation about Photochemical or thermal (1: 3) sigmatropic hydrogen migration in thymine that effects on the conformations of thymidine and thymidylic acid showed that hydrogen migration among hetero-atoms on these studied tautomers may exist within the energy range of 22.4838kcal/mol. Dipole moment and total charge in tautomers portray dipole-dipole interactions for hydrogen bonding in nucleic acids.

Further, the utility of theoretical prediction is important for evaluating the nature of hydrogen bonding in nucleic acids, storage, retrieval and transfer of genetic information.

This study reveals that the stability of conformations is highly dependent upon the polarity of the medium.

#### **References**

- [1] S. Ignacimuthu, Plant Biotechnology, Oxford and IBH Publishing Co Ltd, New Delhi, (1977).

### **Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.**

---

- [2] DE. Hatchway, G. F. Kolar, Chem Soc Revs, 9, 241(1980).
- [3]. T. Ke, B. Xiao-Zhi, L. Xiao-Qiang, Z. Yue, S. Ji-Hong, T. Chao-Wu, H. Da-Hai, J. Comput. Theor. Nanosci. 11, 1785 (2014).
- [4] AD. Maclaren, D. Shugar, Photochemistry of proteins and nucleic acids, Pergamon Press Ltd, Oxford, (1964).
- [5] EW. Sinnott, LC. Dunn, and T. Dobzhansky, Principles of Genetics, Tata-McGraw-Hill Book Company Ltd, New Delhi, (1992).
- [6] S. T. Reid, Adv Heterocycl Chem, 11, 1 (1970).
- [7] Z. Pluta, T. Hryniewicz, International Letters of Chemistry Physics and Astronomy, 4, 73 (2013).
- [8] J. D. Watson, Molecular biology of the Genes, 3rd Edn (WA Benjamin, New York) (1976).
- [9] P. Sagar Gami, V. Kalpesh Vilapara, R. Hasmukh Khunt, S. Jayesh Babariya, T. Yogesh Naliapara, International Letters of Chemistry, Physics and Astronomy, 11, 127, (2014).
- [10] Z. Wang, F. Wang, C. Su, and Y. Zhang, J. Comput. Theor. Nanosci.10, 2323 (2013).
- [11] A. Eghdami and M. Monajjemi, Quantum Matter 2, 324 (2013).
- [12] Y. Chen, D. Xu, and M. Yang J. Comput. Theor. Nanosci. 10, 2916 (2013).
- [13] B. Wu, X. Kong, Z. Cao, Y. Pan, Y. Ren, Y. Li, Q. Yang, and F. Lv, J. Comput. Theor. Nanosci. 10, 2403 (2013).
- [14] R. Weiss, Gene therapy at a crossroads, Washington Post Health, Washington (1994).
- [15] R. Mazri, S. Belaidi, A. Kerassa, T. Lanez, International Letters of Chemistry Physics and Astronomy, 14, 146 (2014).

### **Chapter III: Molecular Modeling of Photochemical and Thermal Sigmatropic Hydrogen Migration in the formation of Tautomeric forms of some Biomolecules in DNA.**

---

- [16] Z. Almi, S. Belaidi, T. Lanez, N. Tchouar, International Letters of Chemistry Physics and Astronomy, 18, 113 (2014).
- [17] S.R. Ggadre, R.K. Pathak, J. Chem. Phys, 93, 1770 (1990).
- [18] F. J. Luque, M. Orozco, P.K. Bhadane, and S. R. Gadre, J. Phys. Chem, 97, 9380 (1993).
- [19] G.Thomas, Rev. Theor. Sci. 2, 289 (2014).
- [15] D. S. Cleydson Breno Rodrigues, V. Josinete Braga, D. S. F. Adriane, et al., J. Comput. Theor. Nanosci. [11], 553 (2014).
- [20] P. Rengin, O. Hüseyin, E. Şakir, Rev. Theor. Sci. 2, 301 (2014).
- [21] R. K. Srivastava, V. Narayan, A. Kumar, O. Prasad and L. Sinha, Res. J. Recent Sci, 1, 11 (2012).
- [22] A. El Guerdaoui, Y. El Kahoui, M. Bourjila, R. Tijar A. El gridani, International Letters of Chemistry, Physics and Astronomy, 19, 26 (2014).
- [23] Bojja Rajeshwar Rao, Indian Journal of Chemistry, 48B, 1411 (2009).
- [24] A. Srivastava, S. Jain, and A. K. Nagawat, Quantum Matter 2, 469 (2013).
- [25] A. Srivastava, N. Saraf, and A. K. Nagawat, Quantum Matter 2, 401 (2013).
- [26] A. Srivastava, N. Jain, and A. K. Nagawat, Quantum Matter 2, 307 (2013).
- [27] M. Narayanan and A. J. Peter, Quantum Matter 1, 53 (2012).
- [28] G. H. Cocolletzi and N. Takeuchi, Quantum Matter 2, 382 (2013).
- [29] M. Ibrahim and H. Elhaes, Rev. Theor. Sci. 1, 368 (2013).
- [30] T.Salah, S. Belaidi, N. Melkemi, N. Tchouar, Rev. Theo. Sci. 3, 355-364 (2014).

**Chapter IV:**

**In silico Approach for Conformational  
Analysis, drug-likeness properties and  
Structure activity Relationships of 12-  
Membered Macrolides**

## **IV.1. Introduction**

Quantum chemistry methods play an important role in obtaining molecular geometries and predicting various properties [1]. To obtain highly accurate geometries and physicochemical properties for molecules that are built from electronegative elements, expensive ab initio/HF electron correlation methods are required [2-4]. Density functional theory methods offer an alternative use of inexpensive computational methods which could handle relatively large molecules [5-11].

Conformational stereoselectivity of macrolactone compounds has been studied since 1981, when Still and Galynker [12] have published their investigation on chemical consequences of conformation in the macrolactone compounds.

The theoretical studies are moving towards the rational design which means that knowledge of the relationship between the physicochemical properties and molecular structure allows the scientist to develop new active molecules, with a fairly good anticipation [13].

The volumic representation of a molecule: shape, volume and surface area accessible to the environment allows an approach to the complementary of drug-receptor interactions that ensured when there is an adjustment optimal of the contact surface of the two partners.

The term of antibiotic subsequently has been used to mean a chemical substance which has the ability to inhibit or even destroy the bacterial growth [14].

## **Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

Macrolide antibiotics have been the focus of widespread research due to increasing bacterial resistance. There have been significant synthetic and theoretical efforts to generate new core structures to address this challenge.

Structure elucidation of a large number of obtained molecules, shows the existence of two parts. The first one is a macrocyclic system from 12 to 40 links with several asymmetric centers and lactone function; the second is a sugar part. The two main classes of these macrolides are presented by two molecules; the first is erythromycin A, which is an active antibiotic against a large number of bacteria, and the second is amphotericin B which presents a strong antifungal [15].

The process of drug development is time-consuming and cost-intensive. Several years are required for lead identification, optimization, in vitro and in vivo testing before starting the first clinical trials.

Drug-likeness is a qualitative concept used in drug design, which is estimated from the molecular structure before the substance is even synthesized and tested.

The calculation of drug-like properties can give us better assumption of biological activity of certain molecule. The theoretical calculation and maintain of certain properties of a molecule can fulfill the parameters which are essential to show certain biological activity [16].

Quantitative and qualitative Structure-Activity Relationships (QSAR) are attempts to correlate molecular structure, or properties derived from molecular structure [17-18] with a particular kind of chemical or biochemical activity.

The paper deals with a specific organization form of molecular matter. Other forms are given for example in the references [19-25].

## **Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

Our present research destined to study molecular geometry and electronic properties of macrocyclic compounds using MM/MM<sup>+</sup>, PM3 and ab-initio methods, Afterward, we extended a dataset of antibiotic macrolides in order to identify compounds with a good balance of properties, we use structure activity/property relationships (SAR/SPR) and Lipinski's rules respectively to their antibiotic activity.

### **IV.2. Computational methods**

The molecular modeling calculations for the studied macrolides are performed by HyperChem version 8.08 [26] and Gaussian 09 [27].

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

In the next step, geometries were fully re-optimized by using Ab initio/HF (STO-3G). This work also includes calculation of 3D MESP surface map and 2D MESP contour map to reveal the information regarding charge transfer within the molecule [28].

Finally, macrolide compounds with antibiotic activity characteristics were used to calculate the following physicochemical properties: (SAG) with a grid calculation (solvent-accessible or Van der Waals) much slower than the approximate calculation, but is more accurate for a given set of atomic radii. It is recommended that this method can be used as a benchmark for the approximate surface area calculation. The grid method used is that described by Bodor and al. Molecular volume (V) calculation is very similar to the Surface Area (Grid) calculation; it employs a grid method described

## **Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

by Bodor et al. [29]. The hydration energy (HE) is a key factor determining the stability of different molecular conformations.

The calculation is based on exposed surface area. Calculation of octanol-water partition coefficient (log P) is carried out using atomic parameters derived by Ghose, Pritchett and Crippen and later extended by Ghose and coworkers [30].

The polarizability (Pol) is estimated from an additivity scheme presented by Miller, different increments are associated with different atom types. For a variety of organic molecules, the estimates are accurate to within 1 to 3 % [31]. The molecular weight (MW) of a system calculation is based on a general applicability method.

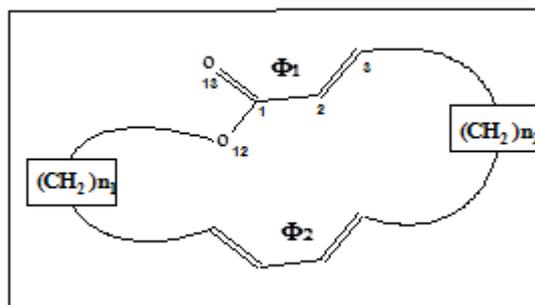
### **IV.3. Results & discussion**

#### **IV.3.1. Conformational analysis of 12-membered $\alpha$ , $\beta$ -unsaturated macrolides**

The most stable structures can be characterized by three structural characters: the diene group, the  $\alpha$ ,  $\beta$ -unsaturated ester group, and the two saturated chains (figure 4. 1). Thus, we have obtained eight types of conformations which are present in the majority of cases in an 11 kcal/mol energy range above the global minimum. The conformation types are classed from 1 to 8 [32].

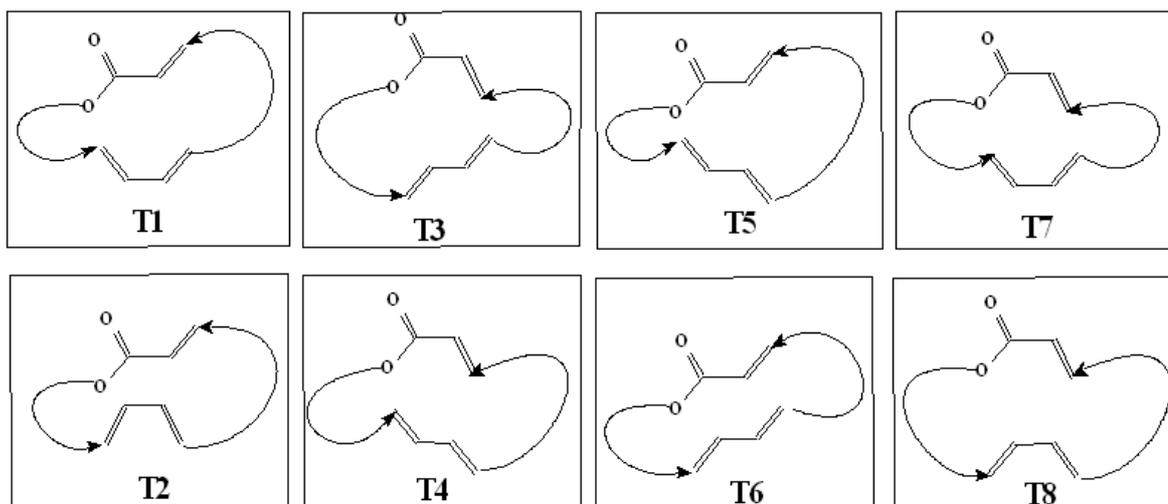
**Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---



**Figure 4.1:**  $\alpha, \beta$ -unsaturated macrolactone.

For types (2, 4, 6, 8), the two planes of two conformational sites, diene and  $\alpha, \beta$ -unsaturated ester group are pseudo parallels; but for types (1, 3, 5, 7), the two planes of the two sites are pseudo antiparallels (figure 4. 2).



**Figure 4.2:** Main conformational types of macrolides.

In 2 kcal/mol difference, the macrocycle 12s is characterized by the first conformer type 3, which is the most favored with 29.5 % rate followed by a type 4 with 28.1 % (table 4. 1).

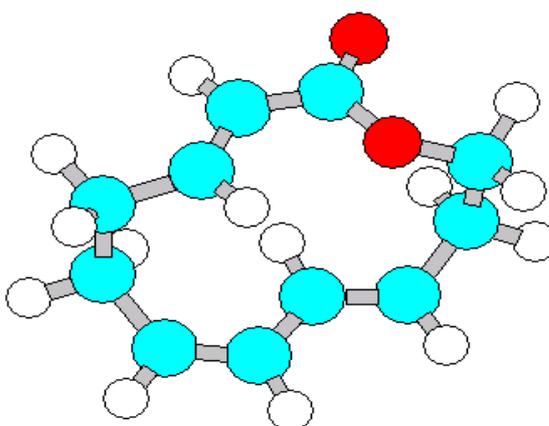
**Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

**Table 4.1:** Energetic difference and Boltzmann population for different macrolide types.

Macrolides	12 symmetric ( $n_1 = n_2 = 2$ )			12 dissymmetric ( $n_1 = 1, n_2 = 3$ )			12 dissymmetric ( $n_1 = 3, n_2 = 1$ )		
	Type	$\Delta E$	%	Type	$\Delta E$	%	Type	$\Delta E$	%
to kcal/mol	3	00.00	29.5	3	00.00	35.0	3	00.00	24.9
	4	00.20	28.1	4	00.34	32.2	8	00.71	21.0
							4	01.26	18.4
							7	01.68	16.6
	5	02.16	17.4	7	05.07	10.2	5	05.13	07.2
	6	03.46	12.7	8	06.15	07.8	6	06.96	04.6
Sup to 2 Kcal/mol	8	07.02	05.3	5	07.45	05.7	2	07.48	04.0
	7	08.80	03.5	6	07.49	05.7	1	08.31	03.3
	1	11.32	01.9	1	11.66	02.1			
	2	11.87	01.6	2	13.31	01.4			

$\Delta E$ : Energetic difference to the absolute minimum, %: Boltzmann population

Then, the macrocycle 12d (1-3) is presented preferably in the type T3 (figure 4.3) with 35.0 % followed by type 4 with 32.2%.



**Figure 4.3:** 3D structure of the most favored conformer in 12 macrolide 12d, 1-3

(HyperChem).

## Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.

---

Finally, the macrocycle 12d (3-1) is presented preferably in T3 with 24.9% followed by T8 with 21.0% and T4 with 18.4%. The percentages of other conformers are listed in Table 4. 1.

The geometric study allow to see that ester  $\alpha$ ,  $\beta$  –unsaturated system for the 12S macrolide type 3 has an S-TRANS form with a dihedral angle O13-C1-C2-C3 = 173.972° (Table 4.2) using molecular mechanic calculation; 165.348° using PM3 method and 167.056° via ab-initio/HF(STO-3G) method. Also it allows to see that the diene system, it has an S-TRANS form, with a dihedral angle of C6-C7-C8-C9 = 166.749° using MM calculation, 130.034° via PM3 method and 149.076° using ab-initio method.

**Table 4.2:** Calculated values of dihedral angles of 12 symmetric macrolide /T3

Angle	MM+	PM3	Ab initio
C1-C2-C3-C4	166.394	167.874	164.256
C2-C3-C4-C5	089.819	106.255	100.541
C3-C4-C5-C6	061.621	086.687	066.264
C4-C5-C6-C7	081.571	088.645	086.829
C5-C6-C7-C8	002.570	002.360	002.099
C6-C7-C8-C9	166.749	130.034	149.076
C7-C8-C9-C10	168.485	173.073	168.192
C8-C9-C10-C11	091.054	121.898	100.502
C9-C10-C11-O12	051.628	048.146	051.000
C10-C11-O12-C1	061.621	071.792	078.874
C11-O12 -C1-C2	139.136	132.736	147.687
C11-O12-C1-O13	030.217	047.952	028.387
O12-C1-C2-C3	004.325	015.406	008.878
O13-C1-C2-C3	173.972	165.348	167.056

In the other side the ester  $\alpha$ ,  $\beta$ –unsaturated system for 12d (1-3) macrolide type 3 contain an S-TRANS form with a dihedral angle O13-C1-C2-C3 = 174.296° as found from MM results and 170.456° via PM3 method. For the diene system, it has an S-

## Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.

---

TRANS form with a dihedral angle of C6-C7-C8-C9 = 171.454° via MM and 178.515° using PM3 method (Table 4. 3).

**Table 4.3:** Calculated values of dihedral angles of 12 dissymmetric macrolide /T3 for ester and diene system.

	MM		PM3	
	Ester	Diene	Ester	Diene
12 d (1-3)	174.296	171.454	170.456	178.515
12 d (3-1)	175.798	169.525	172.998	122.812

Otherwise, The ester  $\alpha$ ,  $\beta$ -unsaturated system for 12d (3-1) macrolide type 3, contain an S-TRANS form with a dihedral angle O13-C1-C2-C3 = 175.798° from MM and 172.998° via PM3 method, in the same way we found a dihedral angles C5-C6-C7-C8 = 169.525° using MM and 122.812° using PM3 method for the diene system with an S-TRANS form.

Finally, we conclude that ester  $\alpha$ ,  $\beta$ -unsaturated and diene systems for the two macrocycles are perpendicular on medium plans of cycles.

### IV.3.2. Geometric and Electronic Structure of 12-membered Macrolide 12d (1-3) (Type 3)

The optimized geometrical parameters of the studied conformers are obtained using MM/MM+, PM3 and ab-initio/HF (STO-3G) methods, listed in Table 4.4 and Table 4.5 with an observed similarity in the obtained results, regarding bond length. Therefore we can say that there is a similarity of calculations results between methods.

**Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

**Table 4.4:** Calculated values of bond lengths of 12d (1-3) macrolide /T3

Distance	MM/MM <sup>+</sup>	PM3	ab initio/HF
C1-O12	1.3497	1.3846	1.3625
C1-O13	1.2115	1.2136	1.2061
C1-C2	1.3602	1.4809	1.4709
C2-C3	1.3437	1.3362	1.3203
C3-C4	1.5098	1.4858	1.5003
C4-C5	1.5451	1.5306	1.5679
C5-C6	1.5148	1.4900	1.5168
C6-C7	1.3465	1.3376	1.3254
C7-C8	1.3434	1.4579	1.4709
C8-C9	1.3433	1.3372	1.3210
C9-C10	1.5130	1.4881	1.5090
C10-C11	1.5418	1.5339	1.5442
C11-C12	1.4121	1.4174	1.4536

**Table 4.5:** Calculated values of angles of 12d (1-3) macrolide /T3

Angle	MM/MM <sup>+</sup>	PM3	ab initio/HF)
C1-C2-C3	120.815	121.908	120.971
C2-C3-C4	123.408	124.002	125.272
C3-C4-C5	109.708	109.419	108.582
C4-C5-C6	112.512	113.111	112.919
C5-C6-C7	126.760	125.732	126.929
C6-C7-C8	124.500	123.946	125.414
C7-C8-C9	124.243	121.708	123.848
C8-C9-C10	121.888	122.689	123.356
C9-C10-C11	110.093	111.880	110.049
C10-C11-O12	110.525	113.206	109.404

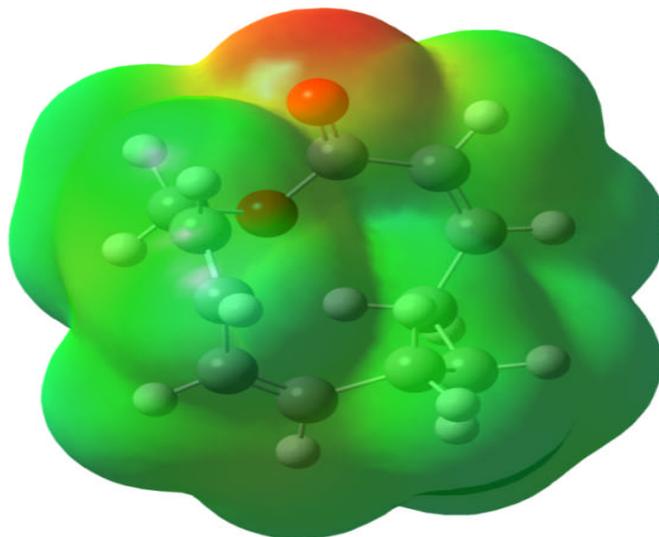
**Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

C11-O12 -C1	118.520	118.204	119.327
O12-C1-O13	120.374	117.759	122.213
C2-C1-O13	117.517	126.326	112.453

The molecular electrostatic potential MESP surface which is a plot of electrostatic potential mapped onto the iso-electron density surface [33], the importance of the MESP lies in the fact that it simultaneously displays the molecular size and shape as well as positive, negative and neutral electrostatic potential regions in terms of the electrostatic surface, which explain the investigation of the molecular structure with its physiochemical property relationships [34]. The MESP surface map and contour map of 12d (1-3) macrolide/ type 3 (Figure 4. 4) show the two regions characterized by red color (negative electrostatic potential) around the  $\alpha$ ,  $\beta$ -unsaturated ester group which explain the ability for an electrophilic attack on these positions, also for the green and blue regions explain the neutral and the positive electrostatic potential surface.

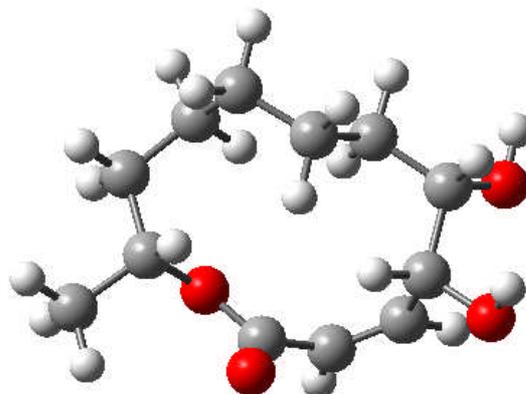
Finally, we can say the variation in electrostatic potential produced by a molecule is largely responsible for binding of a drug to its active sites (receptor), as the binding site in general is expected to have opposite areas of electrostatic potential.



**Figure 4.4:** 3D MESP contour map for 12d (1-3) macrolide/ type 3.

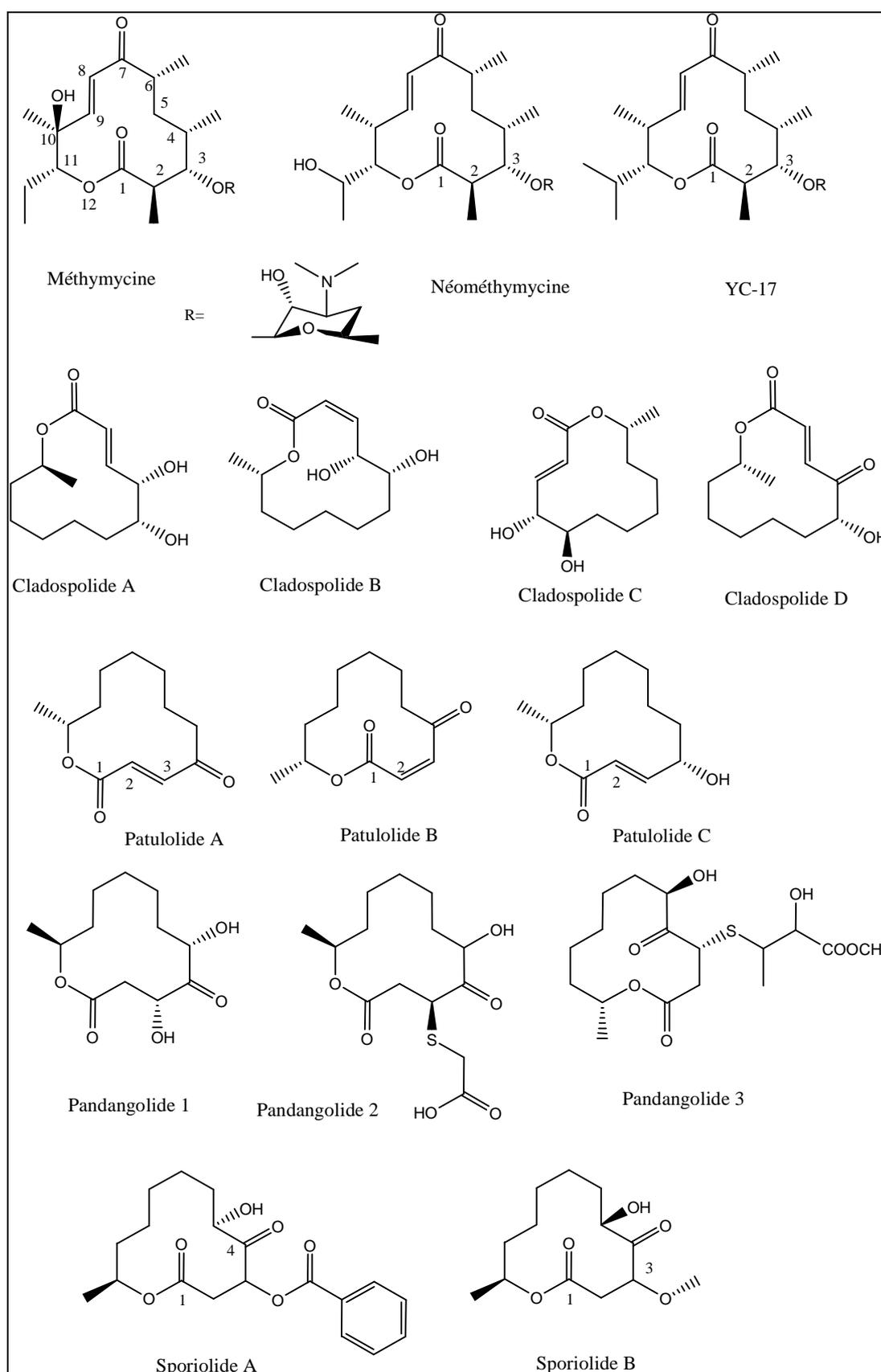
### **IV.3. 3. Structure - activity relationships and rules of five for 12-membered macrolides**

In the present work, we have applied Structure-Activity Relationship (SAR) and rules of thumb (Lipinski's rules) on thirteen conformers of 12-membered macrolides with respect to their antibiotics activity. The chemical structures of the studied macrolide are shown in Figure 4. 6 and for example, Figure 4. 5 shows the favored conformation in 3D of cladospolide B.



**Figure 4.5:** 3D Conformation of cladospolide B (Gauss View 3.09).

**Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**



**Figure 4.6:** Chemical structure of the studied macrolides.

**Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

The properties involved are: Partition coefficient octanol/water (log P), molecular weight (MW), hydrogen bond donors (HBD), hydrogen bond acceptors (HBA), Surface area grid (SAG), molar volume (V), hydration energy (HE), molar refractivity (MR) and polarizability (Pol). The results using HyperChem 8.0.8 software are shown in Table 4.6.

**Table 4.6:** QSAR proprieties for 12-membered macrolides

Macrolides	HBA	HBD	Volume (Å <sup>3</sup> )	Surface (Å <sup>2</sup> )	Mass (uma)	Log P	Energy of hydratation (Kcal/mol)	Polarizability (Å <sup>3</sup> )
1-Cladospolide A	4	2	696.538	414.994	228.288	1.591	-7.656	23.825
2-Cladospolide B	4	2	699.040	418.608	228.288	1.591	-10.029	23.825
3-Cladospolide C	4	2	689.945	414.379	228.288	1.591	-8.770	23.825
4-Cladospolide D	4	1	683.482	408.871	226.273	1.189	-5.995	23.274
5-Methymycine	5	1	1324.521	709.552	469.619	3.376	-5.694	49.617
6-Neomethy- mycine	5	1	1332.030	711.049	469.619	3.483	-5.304	49.617
7-YC-17	4	0	1351.548	721.440	467.646	4.727	-1.478	50.815
8-Patulolide A	3	0	656.943	397.729	210.273	2.763	-0.330	22.637
9-Patulolide B	3	0	646.126	389.778	210.273	2.763	-0.981	22.637
10-Patulolide C	3	1	675.196	411.752	212.289	2.360	-5.437	23.188
11- Pandangolide 1	5	2	713.430	423.606	244.288	1.341	-9.242	24.103
12- Pandangolide 2	6	2	837.203	474.734	318.385	1.603	-3.760	30.859
13- Pandangolide 3	7	2	1015.292	572.231	376.465	1.800	-8.267	37.001
14-Sporiolide A	6	1	970.017	557.306	348.396	3.383	-5.375	35.684
15-Sporiolide B	5	1	748.108	438.928	258.315	1.619	-4.113	25.938

## **Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

In the first part, we have studied Lipinski rules to identify "drug-like" compounds. Rich absorption or permeability is more likely when: [35]

- 1- There are less than 5 H-bond donors (expressed as the sum of OHs and NHs);
- 2- The molecular weight is under 500 DA;
- 3- The log P is under 5;
- 4- There are less than 10 H-bond acceptors (expressed as the sum of Ns and Os);

We used the Lipinski rules to identify compounds posing problems of absorption and permeability if these compounds don't validate at least two of its rules.

The rules are based on a strong physicochemical rationale. Hydrogen bonds increase solubility in water and must be broken in order for the compound to permeate into and through the lipid bilayer membrane.

Thus, an increasing number of hydrogen bonds reduce partitioning from the aqueous phase into the lipid bilayer membrane for permeation by passive diffusion. The studied 12-membered macrolides are in accordance with rule 1 and rule 4 as shown in Table 4. 6 so we can say that they probably less polar and easily absorbed.

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

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

#### **Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

Increasing Log  $P$  also decreases aqueous solubility, which reduces absorption. Thus, membrane transporters can either enhance or reduce compound absorption by either active uptake transport or efflux, respectively. 12-membered macrolides satisfy also the rule number 3 so it has a consequence of a better solubility in aqueous and lipidic solutions too [36].

The majority ( $\log P$ ) of studied molecules have optimal values. For good oral bioavailability, the  $\log P$  must be greater than zero and less than 3 ( $0 < \log P < 3$ ). For  $\log P$  too high, the drug has low solubility and a  $\log P$  too low, the drug has difficulty penetrating the lipid membranes.

Compound 4 presents the low coefficient of division (1.189). When the coefficient of division is rather low, it has as a consequence a better gastric tolerance. Compound 3 which has higher value (4.727), has capacity to be dependent on plasmatic proteins.

In the second part we perform the structure-activity relationship where we found that the polarizability values are generally proportional to the values of surfaces and volumes, the decreasing order of polarizability for these studied Macrolides (Table 4.6). The order of polarizability is approximately the same one for volume and surface. This also is explained by the relation between polarizability and volume, for the relativity non polar molecules. They are directly linked, for the centers of gravity of negative and positive charges in the absence of external fields to coincide, and the dipole moment of the molecule is zero.

The polarizability of a molecule depends only on its volume, which means that the thermal agitation of non-polar molecules does not have any influence on the appearance

#### **Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

of dipole moments in these molecules. On the other hand, for the polar molecules, the polarizability of the molecule does not depend solely on volume but also depends on other factors such as the temperature because of the presence of the permanent dipole [37].

Surface and distribution volume of these molecules are definitely higher than those of more polar molecules like the lipopeptides. For example, Deleu et al. used Tammo software on the surfactins C13, C14 and C15 having cores similar to the macrolides [38]. They found that their surfaces vary from 129 to 157 Å<sup>2</sup>, contrarily for these macrolides derivatives, surfaces vary from 389.778 to 721.440 Å<sup>2</sup>.

The most important hydration energy in the absolute value is that of the compound 2: Cladospolide B (10.029 kcal/mol) and the weakest is that of compound 8: Patulolide A (0.330 kcal/mole). Indeed, in the biological environments the polar molecules are surrounded by water molecules.

Compound 2 has two donor site of proton (2OH), and four acceptor sites of proton (4O). Moreover, compound 8 has only three acceptor sites of proton (3O). The first having higher value, it has one more acceptor site of protons.

This property supports the first compound, not only by fixing on the receiver, but also by activating. It is thus about an agonist. It has as a consequence i.e. a better distribution in fabrics.

They are established hydrogen bonds between a water molecule and these molecules. The donor sites of the proton interact with the oxygen atom of water and the acceptor sites of the proton interact with the hydrogen atom. The first corresponds to the complex

## **Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

with the strongest hydrogen bond. These hydrated molecules are dehydrated at least partially before and at the time of their interaction. These interactions of weak energy, which we observe in particular between messengers and receivers, are generally reversible [39].

### **IV.4. CONCLUSION**

The present work studied the geometric and the electronic proprieties of macrolides using different methods MM, PM3 and ab-initio methods that shown similarity between the results which explain that these methods can be used quite satisfactorily in predicting the chemical reactivity of the macromolecules. The present study, offers the ability to guide design and selection using Lipinski's rules to quickly identify compounds from the 12 membered macrolides series, which are likely to achieve outcome in the clinic and occupy a strong market position. Also it provides a discussion of several qualitative approximations of the structure activity relationship to search the preferred conformations to establish correlations between structural parameters and the various properties of the investigated macromolecules and improving the conception of new therapeutic drugs.

### **References**

- [1] Salah Belaidi, Zineb Almi and Djemoui Bouzidi, *J. Comput. Theor. Nanosci.* 11 (2014) 2481-2488.
- [2] Dalal Harkati, Salah Belaidi, Aicha Kerassa, and Noureddine Gherraf, *Quantum Matter* 4 (2015) in press.

**Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

- [3] A. Srivastava, N. Saraf, and A. K. Nagawat, *Quantum Matter* 2 (2013) 401-407.
- [4] Malika Mellaoui, Salah Belaidi, Djemoui Bouzidi, and Noureddine Gherraf, *Quantum Matter* 3 (2014) 435-441.
- [5] Nadjib Melkemi and Salah Belaidi, *J. Comput. Theor. Nanosci.* 11 (2014) 801-806.
- [6] Salah Belaidi, Houmam Belaidi and Djemoui Bouzidi, *J. Comput. Theor. Nanosci.* 12 (2015) in press.
- [7] M. Narayanan and A. John Peter, *Quantum Matter* 1 (2012) 53.
- [8] Gregorio H. Coccoletzi and Noboru Takeuchi *Quantum Matter* 2 (2013) 382-387.
- [9] Medhat Ibrahim and Hanan Elhaes, *Rev. Theor. Sci.* 1 (2013) 368-376.
- [10] E. Chigo Anota, H. Hernández Coccoletzi, and M. Castro, *J. Comput. Theor. Nanosci.* 10 (2013) 2542-2546.
- [11] Faranak Bazooyar, Mohammad Taherzadeh, Claes Niklasson, and Kim Bolton, *J. Comput. Theor. Nanosci.* 10 (2013) 2639-2646.
- [12] X. C. Still and I. Galynker, *Tetrahedron Lett.* 37 (1981) 3981.
- [13] R. E. Taylor, Y. Chen, A. Beatty, *J. Am. Chem. Soc.* 125 (2003) 26.
- [14] R. Mazri, S. Belaidi, A. Kerassa, T. Lanez, *Journal of International Letters of Chemistry, Physics and Astronomy* 14(2) (2014) 146-167.
- [15] S. Belaidi, M. Omari, T. Lanez et A. Dibi, *J. Soc. Alger. Chim.* 14 (2004) 27.
- [16] M. Joydeep, C. Raja, S. Saikat, V. Anjay, D. Biplab, T. K. Ravi, *Der. Pharma. Chemica* 1(2) (2009) 188-198.

**Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

- [17] M. Narayanan and A. J. Peter, *Quantum Matter* 1 (2012) 53.
- [18] H. Languer, K. Kassali and N. Lebga, *J. Comput. Theor. Nanosci.* 10 (2013) 86-94.
- [19] Arockia Jayalatha and John Bosco B. Rayappan, *Rev. Theor. Sci.* 3, (2015) 245-256.
- [20] Zineb Almi, Salah Belaidi, and Lamri Segueni, *Rev. Theor. Sci.* 3 (2015) 264-272
- [21] Vladimir A. Belyakov and Vladimir A. Burdov, *Quantum Matter* 4, 85-93 (2015),  
*Quantum Matter* 4 (2015) 85-93 .
- [22] Hyung-Joong Yun, Yi-Rang Lim, Jouhahn Lee, and Chel-Jong Choi, *J. Comput. Theor. Nanosci.* 12 (2015) 712-718
- [23] N. Paitya and K. P. Ghatak, *Rev. Theor. Sci.* 1 (2013) 165-305.
- [24] Davide Fiscaletti, *Rev. Theor. Sci.* 1 (2013) 103-144.
- [25] Medhat Ibrahim and Hanan Elhaes, *Rev. Theor. Sci.* 1 (2013) 368-376.
- [26] HyperChem (Molecular Modeling System) Hypercube, Inc., 1115 NW, 4th Street, Gainesville, FL 32601; USA, (2008).
- [27] Gaussian 09, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V.N.

**Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

- Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. 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, *Gaussian Inc.* Wallingford CT(2010).
- [28] K. Mehmet, S. Leena, O. Prasad, A. M. Asiri, M. Cinar, *J. Spectrochimica. Acta. Part A: Molecular and Biomolecular Spectroscopy*, 115 (2013) 753.
- [29] N. Bodor, Z. Gabanyi, C. K. Wong, *J. Am. Chem. Soc.* 111 (1989) 3783.
- [30] A. K. Ghose, A. Pritchett, G. M. Crippen, *J. Comp. Chem.* 9 (1988) 80.
- [31] K. J. Miller, *J. Am. Chem. Soc.* 112 (1990) 8533.
- [32] S. Belaidi, T. Lanez, M. Omari, and A. Botrel, *Asian Journal of Chemistry* 17 (2) (2005) 859–870.
- [33] A. K. Sachan, S. K. Pathak, S. Chand, R. Srivastava, O. Prasad, S. Belaidi, L. Sinha, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 132 (2014) 568–581
- [34] Z. Almi, S. Belaidi, T. LANEZ, N. Tchouar, *International Letters of Chemistry, Physics and Astronomy* 18 (2014) 113-122.
- [35] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *J. Adv. Drug. Delivery. Rev.* 64, (2012) 4.

**Chapter IV: In silico Approach for Conformational Analysis, drug-likeness properties and Structure activity Relationships of 12-Membered Macrolides.**

---

- [36] Toufik Salah, Salah Belaidi, Nadjib Melkemi and Nouredine Tchouar, Rev. *Theor. Sci.* 3, (2014) 355-364
- [37] B. Yavorski, A. Detlaf, *Checklist of Physics*, Mir, Moscow, Russia (1980), 376.
- [38] M. Deleu Synthesis of Surfactin Derivatives and Study Their Properties, *Thesis Ph.D, FUSAGX*, Belgium (2000).
- [39] L. B. Kier, *Molecular Orbital Theory in Drug Research*, Academic Press, New York USA (1981).

## **General Conclusion**

In this work, we applied the methods of computational chemistry on two types of molecules: macrolides and three molecules containing in DNA.

Our objective is to predict chemical reactivity and biological activity of macrolides realized photochemical qualitative study on (1;3) sigmatropic hydrogen migration for the conformations of molecules: thymine, thymidine and thymidylic acid in DNA.

This study comported:

- Conformational analysis of 12 membered macrolides.
- Qualitative study of the structure-activity relationship 12 membered for macrolides.
- Study sigmatropic hydrogen migration in thymine, thymidine and thymidylic acid. This approach is potentially performed to predict the interactions weak, which play vital role chemistry in DNA.

Different methods of molecular modeling were used in our work. Molecular mechanics has been used in the study of analysis conformations for 12 membered macrolides. Eight stables conformers families were determined in a field energy of 13

## General Conclusion.

---

kcal / mol above global minimum. The dissymmetric conformer type 3 is the most favored in all conformers obtained a Boltzmann distribution of 35.0%.

The obtained results, for substitution of the symmetrical macrolactone, which represents basic nucleus of most of the 12 membered antibiotic macrolides, it has identified structural units involved in the antibiotic and tensioactives properties of new macrolides. Indeed in these substituted macrocycles, the compound (2), 3-methoxy macrolide has the lowest gap HOMO-LUMO (09.7948 ev), so it is the most chemically active.

Quantum mechanics methods were used in the study of the chemical reactivity of thymine, thymidine and thymidylic acid and such conformers. Methods: PM3, ab initio / DFT and HF were used to determine the structural, electronic and energy parameters associated with the molecules studied.

Six stable thymine tautomers were determined; the conformer (e) is the most favored with energy of -53.5455Kcal / mol. The efficacy of these methods was confirmed by comparison of the structural parameters calculated by the two theoretical ab initio and DFT with experimental data.

The nature of type of substituent (donor, acceptor) affects the electronic and energetic parameters of basic nucleus 12d (1-3). Indeed, this study allows us to predict the chemical reactivity of these molecules.

The qualitative study of structure-property/activity relationship (QSPR / QSAR) was performed on a series of 12-membered antibiotics macrolides, most recently synthesized, and some are already on the drugs market.

**ملخص:** العمل المقدم يتضمن دراسة اساسية و اصلية على قسمين من المركبات: حلقات ماكروليدية و جزيئات من الحمض النووي. لهدف توقع الفعالية الكيميائية و النشاط البيولوجي للمكروليدات. دراسة فوطوكيميائية نوعية على هجرة الرابطة البسيطة الهيدروجينية (1؛ 3) للجزيئات المتماكبة: التيمين, التميمدين و الحمض التيميديلي. هذا التقريب محقق بقوة للتنبؤ بالتفاعلات الضعيفة, التي تلعب دور حيوي في كيمياء الحمض النووي.

التحليل الأمتثالي للحلقات الكبيرة ذات 12 ذرة تم تحقيقه باستخدام الميكانيكي الجزيئي و طرق ab initio/HF تم تطبيقها في دراسة كيفية تكميلية لعلاقة بينية – فعالية/ نشاط على سلسلة من الحلقات الكبيرة من المضادات الحيوية ذات 12 ذرة.

طرق النمذجة الجزيئية المتبعة في هذا العمل هي: PM3, ab initio /HF. هذه الطرق استعملت لتعيين العوامل البنوية, و الالكترونية و الطاقة المرتبطة بالجزيئات المدروسة. طبيعة و نوعية المستبدل (مانح, مستقبل) تؤثر على العوامل للالكترونية و الطاقة للنواة الأساسية. فعليا, هذه الدراسة تسمح لنا بتنبؤ الفاعلية الكيميائية لمشتقات هذه الحلقات الكبيرة.

تمت دراسة "SAR" على سلسلة من المضادات الحيوية و هي الحلقات الكبيرة ذات 12 ذرة لتعيين نشاطاتها البيولوجية.

**Résumé :** Le présent travail comporte une recherche fondamentale et originale sur deux classes de composés, les macrolides et des molécules contenant dans l'ADN. Le but est de prédire la réactivité chimiques et l'activité biologique des macrolides. Une étude qualitative photochimique sur la migration sigmatropique (1; 3) d'hydrogène pour les conformations de molécules ; la thymine, la thymidine et l'acide thymidylique. Cette approche est potentiellement effectuée pour prédire les interactions faibles, qui jouent un rôle vital dans la chimie de l'ADN.

L'analyse conformationnelle des macrolides à 12 chaînons a été réalisée en utilisant la mécanique moléculaire et la méthode ab initio/HF. Une étude qualitative complémentaire de la relation structure-propriété/activité a été réalisée sur une série de macrolides antibiotiques à 12 chaînons. Les méthodes de modélisation moléculaire utilisées dans notre travail sont : PM3, ab initio/HF. Ces méthodes ont été utilisées pour déterminer les paramètres structuraux, électroniques et énergétiques associés aux molécules étudiées. La nature de type de substituant (donneur, accepteur) influe sur les paramètres électroniques et énergétiques de noyau de base. En effet, cette étude nous permet de prédire la réactivité chimique des dérivés de ces macrolides.

Une étude "SAR" a été effectuée également pour une série de macrolides antibiotiques à 12 chaînons pour déterminer leurs activités biologiques.

**Abstract:** This work involves a fundamental and original research on two classes of compounds, macrolides and containing molecules in DNA. The aim is to predict the chemical reactivity and biological activity of macrolides. A qualitative study on the photochemical sigmatropic migration (1; 3) Hydrogen for the conformations of molecules; thymine, thymidine and thymidylic acid. This approach is potentially performed to predict weak interactions, which play a vital role in the chemistry of ADN. The conformational analysis of 12-membered macrolides has been performed using molecular mechanics and ab initio method. A complementary qualitative study of the structure-property / activity was performed on a series of 12-membered macrolide antibiotics. The molecular modeling methods used in our work are: PM3, ab initio/HF. These methods were used to determine the structural parameters, electronics and energy associated with molecules studied. The nature of such substituent (donor, acceptor) affects the electronic and energy parameters of basic core. Indeed, this qualitative study allows us to predict the chemical reactivity of derivatives of this macrolides. SAR studies have been carried out for a series of 12-membered macrolide antibiotics.