

**THEORETICAL STUDIES OF
BIOFERROELECTRICITY IN DNA/RNA
NUCLEOBASES AND THEIR BASE PAIRS**

YAM SEE CHUAN

**FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2019

**THEORETICAL STUDIES OF
BIOFERROELECTRICITY IN DNA/RNA
NUCLEOBASES AND THEIR BASE PAIRS**

YAM SEE CHUAN

**DISSERTATION SUBMITTED IN FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF MASTER
OF SCIENCE**

**DEPARTMENT OF CHEMISTRY
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2019

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **YAM SEE CHUAN**

Matric No: **SGR 160043**

Name of Degree: **MASTER OF SCIENCE**

Title of Dissertation ("this Work"):

**THEORETICAL STUDIES OF BIOFERROELECTRICITY IN DNA/
RNA NUCLEOBASES AND THEIR BASE PAIRS**

Field of Study: **THEORETICAL AND COMPUTATIONAL CH**

EMISTRY

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature

Date:

Subscribed and solemnly declared before,

Witness's Signature

Date:

Name:

Designation:

THEORETICAL STUDIES OF BIOFERROELECTRICITY IN DNA/RNA NUCLEOBASES AND THEIR BASE PAIRS

ABSTRACT

We have performed computational molecular modelling to study the ferroelectric polarization switching and hysteresis loop behaviours of DNA/RNA nucleobases and their base pairs. All the nucleobases and base pairs - adenine (A), thymine (T), guanine (G), cytosine (C), uracil (U), guanine-cytosine (G-C), adenine-thymine (A-T) and adenine-uracil (A-U) are modelled. Our simulation indicated that all the nucleobases and base pairs exhibit switchable polarization through the application of the electric field. The shape of polarization (P) versus electric field (E) hysteresis loop is square, implying a typical ferroelectrics behaviour. We also found an interesting relationship between the minimum electric field required for polarization switching E_c and the ratio of the topological polar surface area (TPSA) to the total surface area (TSA) of a nucleobase and base pair. In particular, the E_c of a nucleobase and base pair are inversely proportional to the TPSA/TSA ratio. Furthermore, the total length of the base pair against electric field shows a nonlinear relationship. We believe that our findings may contribute to the fundamental understanding of the possible existence of bioferroelectricity in the biological system.

Keywords: bioferroelectricity, nucleobases, polarization, hysteresis loop, molecular modelling.

KAJIAN TEORI BIOFERROELEKTRIK TERHADAP DNA/RNA NUKLEOTIDA DAN PASANGAN BES

ABSTRAK

Dalam kajian ini, pengiraan molekul DNA/RNA nukleotida dan pasangan bes secara komputasi telah dilaksanakan untuk mengkaji peralihan pengutuban dan ciri-ciri gelung histeresis. Kesemua nukleotida dan pasangan bes - adenina (A), timina (T), guanina (G), sitosina (C), urasil (U), guanina-sitosina (G-C), adenina-timina (A-T), dan adenina-urasil (A-U) telah dimodelkan. Hasil simulasi ini menunjukkan bahawa kesemua nukleotida dan pasangan bes mempamerkan pengutuban boleh dialih melalui aplikasi elektrik. Gelung histeresis, pengutuban (P) berbanding elektrik (E) yang berbentuk segi empat sama menunjukkan ciri-ciri ferroelektrik. Kami juga mendapati hubungan yang menarik di antara elektrik minima yang diperlukan untuk beralih pengutuban (E_c) dan nisbah topologi permukaan berkutub (TPSA) kepada keseluruhan luas permukaan (TPSA) nukleotida. Secara khususnya, E_c nukleotida dengan nisbah TPSA/TSA adalah berkadar songsang. Tambahan pula, hasil kajian menunjukkan bahawa jumlah panjang pasangan bes terhadap elektrik menunjukan hubungan tidak selari. Secara keseluruhannya, kajian ini dapat memberi penerangan tentang kemungkinan adanya bioferroelektrik dalam sistem biologi.

Kata Kunci: bioferroelektrik, nukleotida, pengutuban, gelung histeresis, pemodelan molekul.

ACKNOWLEDGEMENTS

When I started my master's degree in research, I knew that I would experience and enjoy the work which would keep me busy for two years. And indeed, in the research journey, it is full of discovering and exploring, attending conferences and finally, ending by writing a scientific manuscript was a great pleasure. This was only possible due to the opportunity that was given to me during the past two years. Thereby, I would like to thank all who contributed and given this fulfilling opportunity.

My first thoughts go to my supervisors, Assoc. Prof. Dr Chew Khian Hooi, and Assoc. Prof. Dr Vannajan Sanghiran Lee whose splendid guidance, authentic supervision, assiduous cooperation, moral support and constant encouragement enabled us to make out our research problem in the present form.

Great thanks to Prof. Dr Sharifuddin Md Zain for his encouragement and insight through these years. It was his guidance and support that helped me through all the difficult time.

A special word of thanks all my lab mates from computational chemistry, Dr Chin Sek Peng, Chong Wei Lim, Vertika Gautam, Ameerul Hazeeq, Hadieh Monajemi for their thoughtful discussions as well as sharing their knowledge and experience with me.

My master project was largely computational based and would not have been possible without access to the excelled high-performance computing resources available at the Data Intensive Computing Centre maintained by Dr Liew Chee Sun and Mr Amjad Kotobi.

Above all, I would deeply thankful to my family members and friends for their love, support and sacrifices. This thesis would have been impossible without their perpetual moral support.

TABLE OF CONTENTS

ORIGINAL LITERARY WORK DECLARATION	ii
ABSTRACT	iii
ABSTRAK	iv
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF FIGURES	ix
LIST OF TABLES	xii
LIST OF SYMBOLS AND ABBREVIATIONS	xiii
LIST OF APPENDICES.....	xv
CHAPTER 1: INTRODUCTION	1
1.1 Research background.....	1
1.2 Research objectives	6
1.3 Research scope.....	6
1.4 Organization of the dissertation.....	6
CHAPTER 2: LITERATURE REVIEW AND BASIC CONCEPTS	8
2.1 Introduction	8
2.2 Ferroelectricity.....	8
2.3 Basic features of ferroelectricity.....	11
2.3.1 Spontaneous polarization.....	11
2.3.2 Ferroelectric hysteresis loop and polarization switching	13
2.3.3 A short history of ferroelectricity	14
2.4 Dielectric	15
2.5 General properties of dielectric	17

2.5.1	Polarization.....	17
2.5.1.1	Electronic Polarization (P_E)	20
2.5.1.2	Atomic Polarization (P_A).....	21
2.5.1.3	Dipolar Polarization (P_D)	22
2.5.1.4	Interfacial Polarization (P_I)	22
2.5.2	Dipole moment	23
2.5.3	Polarizability.....	24
2.6	Introduction to deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)	25
2.7	Emergence of bioferroelectricity	29
2.7.1	Bioferroelectricity in glycine.....	31
2.7.2	Bioferroelectricity in porcine aortic walls.....	34
2.7.3	Bioferroelectricity in microtubules.....	35
2.7.4	Bioferroelectricity in voltage-gated ion channels.....	37
CHAPTER 3: METHODOLOGY.....		39
3.1	Introduction	39
3.2	Molecular modelling theory	39
3.2.1	Geometry optimization.....	39
3.2.2	Molecular mechanics (MM)	41
3.2.3	Quantum mechanics (QM)	42
3.2.4	Semi-empirical	44
3.2.4.1	Complete neglect of differential overlap (CNDO).....	45
3.2.4.2	Intermediate neglect of differential overlap (INDO)	46
3.2.4.3	Zero differential overlap (ZDO).....	46
3.2.4.4	Neglect of diatomic differential overlap (NDDO)	47
3.2.4.5	Parameterized Model 3 (PM3)	48
3.2.5	Hartree-Fock Equation	49

3.2.6	Kohn-Sham Equation	50
3.3	Implementation method	51
3.3.1	Molecular structures preparation	52
3.3.2	Physical and structural properties calculation	55
3.3.3	Applied an external electric field in molecular simulation	55
CHAPTER 4: DNA AND RNA NUCLEOBASES		58
4.1	Introduction	58
4.2	Results and discussion	58
4.2.1	Physical and structural properties	58
4.2.2	Electric field effects on electrical, structural and physical properties	62
4.3	Summary	68
CHAPTER 5: DNA AND RNA BASE PAIRS		69
5.1	Introduction	69
5.2	Results and Discussion	69
5.2.1	Physical and structural properties	69
5.2.2	Electric field effects on electrical, structural and physical properties	73
5.3	Summary	77
CHAPTER 6: CONCLUSION AND FUTURE WORKS		78
6.1	Conclusion	78
6.2	Future works	78
REFERENCES		80
LIST OF PUBLICATIONS AND PAPERS PRESENTED		89
APPENDIX		90

LIST OF FIGURES

Figure 2.1: Relations among piezoelectric, pyroelectric and ferroelectric.....	10
Figure 2.2: Classification of piezoelectric, pyroelectric, and ferroelectric crystals.	11
Figure 2.3: Polarization (P) versus temperature (T) curve.....	12
Figure 2.4: Ferroelectric (P - E) hysteresis loop (Reprinted with permission from Sellmann, 2016).....	14
Figure 2.5: A dielectric material showing the orientation of positive and negative charges creating a polarization effect.....	17
Figure 2.6: Unpolarized atoms become polarized when an electric field (\vec{E}) is applied.....	18
Figure 2.7: The frequency dependence of dielectric constant in the electronic, atomic, dipolar, and interfacial polarization (Reprinted with permission from American Geophysical Union) (Chen & Or, 2006)...	20
Figure 2.8: Schematic diagram of electronic polarization.....	21
Figure 2.9: Schematic diagram of atomic polarization.....	21
Figure 2.10: Schematic diagram of dipolar polarization.....	22
Figure 2.11: Schematic diagram of interfacial polarization.....	23
Figure 2.12: Dipole moment of a water molecule.....	24
Figure 2.13: Double-stranded DNA structure (Reprinted with permission from Scitable by Nature Education) (Pray, 2008).....	29
Figure 3.1: A 3D potential energy surface (PES) map.....	40
Figure 3.2: Chemical structure of DNA and RNA nucleobases. The nitrogen atom is in blue colour, oxygen in red, hydrogen in white, and carbon in green.....	53

Figure 3.3: Chemical structure of DNA and RNA base pairs. The nitrogen atom is in blue colour, oxygen in red, hydrogen in white, carbon in green and the yellow dotted line represent the hydrogen bonds	54
Figure 3.4: Implementation process flow chart.....	57
Figure 4.1: Structural formula and atom numbering for purines and pyrimidines of DNA and RNA nucleobases. Blue arrows indicate the magnitude of dipole moments. An electric field is applied along the <i>Y</i> -direction (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).....	59
Figure 4.2: Electric field (<i>E</i>) dependence of (a) polarization (<i>P</i>) and (b) total energy (<i>U</i>) of cytosine. Insets in (a) show the molecular structure of cytosine at certain electric fields. The initial state of the molecule and the direction of an applied field is along the <i>Y</i> -axis, as shown in Fig 4.1 (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).....	64
Figure 4.3: Dependence of (a) electron charges, (b) bond lengths and (c) bond angles of cytosine on an applied electric field (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).....	65
Figure 4.4: Electric field (<i>E</i>) dependence of polarization (<i>P</i>) for DNA and RNA nucleobases (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).....	66
Figure 4.5: Electric field (<i>E</i>) dependence of total energy (<i>U</i>) for (a) adenine, (b) thymine, (c) guanine, and (d) uracil (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).....	66
Figure 5.1: Optimized structures of Guanine-Cytosine, Adenine-Thymine and Adenine-Uracil base pairs. Blue arrows indicate the magnitude of dipole moments and the dotted line represent the hydrogen bonds. The carbon atom is in grey colour, nitrogen in blue, oxygen in red and hydrogen in white.....	71

Figure 5.2: Electric field (E) dependence of polarization (P) for DNA and RNA base pairs.....	74
Figure 5.3: Electric field (E) dependence of total energy (U) for (a) Guanine-Cytosine (b) Adenine-Thymine (c) Adenine-Uracil.....	75
Figure 5.4: Changes of hydrogen bonds distance between the base pairs under an applied electric field: (a) Guanine-Cytosine (b) Adenine-Thymine (c) Adenine-Uracil.....	76
Figure 5.5: Changes of the total length of DNA and RNA base pairs under an applied electric field.....	77

LIST OF TABLES

Table 2.1: pKa values for ionise atoms in nucleobases (McLennan & Turner, 2012).....	27
Table 4.1: Physicochemical properties of DNA and RNA nucleobases (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).....	60
Table 4.2: Calculated dipole moments of DNA and RNA nucleobases (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).....	61
Table 4.3: Calculated topological polar surface area (TPSA), total surface area (TSA), and TPSA/TSA ratio for DNA and RNA nucleobases. Red and black colours represent the polar and non-polar surface areas, respectively (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).....	61
Table 4.4: Coercive field and zero-field polarization of DNA and RNA nucleobases. The values are obtained from the hysteresis loops Figs 4.2a and 4.4 (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).....	67
Table 5.1: Physicochemical properties of DNA and RNA base pairs.....	72
Table 5.2: Calculated dipole moments of the DNA and RNA base pairs.....	72
Table 5.3: Calculated topological polar surface area (TPSA), total surface area (TSA), and TPSA/TSA ratio for DNA and RNA base pairs.....	73
Table 5.4: Calculated interaction energy of the DNA and RNA base pairs.....	73
Table 5.5: Coercive field and zero-field polarization of DNA and RNA base pairs as obtained from the hysteresis loops in Fig. 5.2.....	75

LIST OF SYMBOLS AND ABBREVIATIONS

3D	:	Three-dimensional
A	:	Adenine
A-T	:	Adenine-Thymine
A-U	:	Adenine-Uracil
a.u.	:	Atomic units
AFM	:	Atomic force microscope
C	:	Cytosine
E	:	Electric field
E_c	:	Coercive field
G	:	Guanine
G-C	:	Guanine-Cytosine
MM	:	Molecular mechanics
P	:	Polarization
P_s	:	Spontaneous polarization
P_r	:	Zero-field polarization
P_{sat}	:	Saturation polarization
\vec{p}	:	Dipole moment
PM3	:	Parameterized model number 3
PES	:	Potential energy surface
PFM	:	Piezoresponse force microscopy
T	:	Thymine
TPSA	:	Topological polar surface area
TSA	:	Total surface area
T_c	:	Curie temperature

QM : Quantum mechanics
QSAR : Quantitative structure activity relationship
RNA : Ribonucleic acid
U : Uracil
U : Total energy

University of Malaya

LIST OF APPENDICES

Appendix A: Nucleobases Geometries and Charge Distributions.....	89
--	----

University of Malaya

CHAPTER 1: INTRODUCTION

1.1 Research background

Ferroelectrics are materials that exhibit spontaneous electric polarization, which was first discovered in 1921 in sodium potassium tartrate tetrahydrate, a Rochelle salt, by Joseph Valasek (Valasek, 1921). A significant breakthrough in the ferroelectric research occurred in the period of 1940 during the Second World War, with the discovery of ferroelectric behaviour in barium titanate (BaTiO_3) by Von Hippel et al. (von Hippel et al., 1946). Ever since then, thousands of materials in various forms including inorganic ceramics, oxides, crystals, liquid crystals, polymer, as well as organic materials have been known to exist ferroelectric properties rendering the application of capacitor and transducer.

The most prominent and notable features of ferroelectric materials are the ability to reverse their polarization or dipole under the influence of an external applied electric field, reminiscent of ferromagnetism. The reversal or switching process typically occurs below a Curie temperature, leading to a characteristic hysteresis loop of polarization versus the applied electric field and a butterfly-like loop in the deformation voltage graph resembles the basis of commercial ferroelectric memories. Ferroelectric materials usually belong to pyroelectrics and piezoelectrics as well, where their polarization behaviour can be regulated in response to applied mechanical stress or temperature, respectively. Such versatile properties and the ability to control them through several parameters such as pressure, temperature and electric field make them useful in the development of various smart device applications with multifunctionalities in the area of medical (foetal heart monitors), engineering (pressure sensors), military (depth sounders) and metrology (ultrasonic measurement technology).

Due to its attractive properties and applications, it is tempting to investigate if ferroelectricity, pyroelectricity and piezoelectricity, has also evolved in biological materials through natural selection. Indeed, piezoelectricity was first discovered and described on bones by Fukada and Yasuda in 1957 (Fukada & Yasuda, 1957) and shortly after that reported in muscle (Fukada & Ueda, 1970). From then on, piezoelectricity has been demonstrated in various biological materials including skin (Rossi et al., 1986), tendon (Williams & Breger, 1975), protein (Lemanov et al., 2002), myosin and actin (Hiroyoshi & Eiichi, 1971), polypeptides (Date et al., 1970) as well as nucleic acid (Fukada & Ando, 1972). It was hypothesized that the existence of piezoelectricity in biological systems might theoretically control cell division, growth, matrix production, intra, and extracellular signalling, energy transfer, enzyme reaction, and intracellular pH (Bassett, 1968). For instance, in vivo studies conducted at the University of Pennsylvania by Brighton et al. shown that sustained applications of an external electrical potential could stimulate bone growth and development (Brighton et al., 1979). Most recently, it has also been revealed that the inherent polarization switching in seashells associated with piezoelectricity may trigger energy dissipation when external mechanical pressure is applied (T. Li & Zeng, 2011). Analogously, the quantitative pyroelectric effect measurements made on the palmlike gymnosperm leaves suggested that the effect can enhance and improve the efficiency of thermal energy conversion into electricity (Lang & Athenstaedt, 1977).

The piezoelectricity and pyroelectricity have been confirmed in a wide variety of biological materials during the past century. However, there is very limited evidence for biological ferroelectricity. Scientific validations of the ferroelectricity phenomenon existence in organic will provide valuable insight to understand the possibility of the existence of ferroelectricity in biological materials, specifically bioferroelectricity. In

fact, bioferroelectricity is a branch of ferroelectricity concerning studies of ferroelectric behaviours in living biological species.

Over the past years, numerous theoretical and experimental investigations of the electrical properties of certain DNA and RNA seemed to suggest the existence of ferroelectricity in these biomaterials. In the early of 1996, Polonsky et al. observed the first ferroelectric effect in DNA. The sodium salt DNA sample extracted from cock erythrocytes were prepared in the form of oriented films and rolled between two cylinders up to 1.5 times of their original length and observed in the temperature range from 21 to 70 °C. The ferroelectric hysteresis loops were measured using a super-regenerative oscillator (SRO) at a frequency of 50 Hz and with a maximum electric field of 1000 Vcm⁻¹. Based on the hysteresis result, the Curie temperature reported being in the vicinity of 60 °C (Polonsky et al., 1960). Thereafter in 1968, the anomalous dielectric constant behaviours were reported to be presented in RNA. The experiment involved the sodium salt calf thymus RNA sample containing 7 percent of moisture was first dissolved in water within 1 inch square of the brass substrates for crystallization. During the drying process, moisture was allowed to remain in the crystals, a condition that manifests itself in the conductivity of the capacitor formed. A thermostat was used to vary the temperature from 30 to 120 °C. Hysteresis loop was displayed and observed at 26 °C when the capacitors in the formed with the crystallized sodium RNA were placed in the Sawyer-Tower circuit. Unfortunately, the sample decomposed and hysteresis loop disappeared at a recorded temperature of 50 °C when it placed in an oven. They suggested that RNA was seen to have the attributes of a ferroelectric substance and possesses the capability of storing information (Stanford & Lorey, 1968). However, in 2009, Yarmarkin et al. repeated a similar experiment as they disagreed of the hypothesis that the ferroelectric effect exists in DNA and RNA. The studied was carried out using sodium salt DNA from calf thymus with 20 µm thick in fiber-shaped.

Perhaps, it is well known that relative humidity plays a significant role and might possible influence the ferroelectric properties of biological materials, thus in order to understand the effect of humidity toward the sample of DNA, they conducted the experiment under air, humid and vacuum environment with heating and cooling process at the temperature range between 20 to 60 °C. Surprisingly, they recorded the narrow shape of dielectric hysteresis loop at the air, humidity and vacuum chamber within the room temperature, similar to those observed in Polonsky and Stanford. But, at the same time, they also argued that the dielectric hysteresis loop observed in several temperatures and environment conditions could be accounted from the moistening and drying of the DNA rather than by a manifestation of intrinsic ferroelectric as none pre-polled and polarized in a field of ~ 1 kV/cm was measured in their results. Therefore, they assumed that the DNA sample used in theirs, and Polonsky et al. do not possess ferroelectric properties (Yarmarkin et al., 2009). On the other hand, the recent comprehensive experimental and theoretical studies carried out by Bdikin et al. indicate the presence of the ferroelectric spontaneous polarization and local piezoelectric properties in thymine nucleobases. They examined both of the thymine anhydrate and monohydrate phase microcrystals using the atomic force microscopy (AFM) and piezoresponse force microscopy (PFM) method. In order to examine the ferroelectric polarization, they performed PFM switching analysis to obtain local piezoresponse hysteresis loop in which the polarization state is changed by the application of a sequence of poling pulse. In the poling process, the tip was fixed at a predefined position on the surface and external bias pulse was applied. After that, the area was scanned again and the polarization was probed using PFM imaging. The longitudinal piezocoefficient value obtained from the ac voltage dependences of deformations was determined to be 4.8 pm/V. On the theoretical side, the structural model which made up of four thymine molecule in one single layer was constructed for the computational

molecular modelling and physical properties calculation. Single point energy and geometry optimization of the total energy were performed using first-principles density functional theory (DFT) and semi-empirical parametrization PM3. It was found that the piezoelectric coefficient was 10 pC/N for the applied electric field ~ 5.14 Mv/cm. Furthermore, they calculated the dipole moment as being ~ 4.5 Debye for an individual thymine molecule, directed along the *Y*-axis, corresponding to spontaneous polarization of $P = 0.114$ C/m² (Bdikin et al., 2015; Bystrov et al., 2015).

Enormous numbers of studies have been made in the last decades on the subject, devoted to the exploration of fundamental electrical properties of DNA and RNA, and their potential applications as a biosensor for DNA damage and mutation detector. However, due to the complexity of DNA and RNA, the origin and mechanism behind their ferroelectricity are still not fully understood and whether the DNA and RNA possess ferroelectric properties are remain controversial. Thus, we are motivated to study the possible existence of ferroelectric properties in DNA and RNA nucleobases: Adenine (A), Thymine (T), Guanine (G), Cytosine (C), Uracil U) and their base pairs: Guanine-Cytosine (G-C), Adenine-Thymine (A-T), Adenine-Uracil (A-U) using computational molecular modelling. Since nucleobases and their base pairs are the essential building blocks of DNA and RNA, an understanding of the principles within the nucleobases at the structural level would lead to a better apprehension of their generic features and function. In this project, we perform a molecular modelling on the reversal polarization and hysteresis loop behaviours of all the five nucleobases and their base pairs under the variation of the external electric field. From this point forward we will always list the nucleobases and their base pairs in alphabetical order (e.g. Adenine as A or Guanine-Cytosine as G-C)

1.2 Research objectives

This research embarks the following objectives:

- I. To model the structure of DNA and RNA nucleobases (A, T, G, C, U) and their base pairs (G-C, A-T, A-U).
- II. To investigate the electronic and structural properties of the DNA and RNA nucleobases and their base pairs using quantum mechanics (QM) calculations.
- III. To identify the spontaneous polarization of the DNA and RNA nucleobases and their base pairs.
- IV. To examine the polarization switching and hysteresis loop behaviours of the modelled DNA and RNA nucleobases and their base pairs.

1.3 Research scope

In general, the research scope can be divided into four categories.

- I. Physical and structural properties of DNA and RNA nucleobases, and their complementary base pairs.
- II. Effect of electric field on physical properties.
- III. Polarization and hysteresis loop behaviours.

1.4 Organization of the dissertation

Overall, this dissertation consisting of six chapters. Chapter 1 describes the background and recent progress of the relevant research while highlighting the objectives of this research, followed by the scope of the present study. Chapter 2 mainly cover exhaustive literature review contributions from the theoretical and experimental works of the bioferroelectricity. The basic concepts regarding the ferroelectric properties, dipole moment, polarization, and polarizability are also included. Chapter 3

outlines the methodology employed in this research. It begins by describing geometry optimization, molecular mechanics (MM), quantum mechanics (QM), and underlying theory that give rise to the Parameterized Model 3 (PM3) method, followed by the description of the computational and implementation method used in this study. In chapter 4, the results of the calculations determining the physical and structural properties of the DNA and RNA nucleobase in the vacuum are discussed. The study of the absence and applied electric field on the electronic structure and physical properties are also provided. The polarization switching and hysteresis loop behaviours obtained from the simulations using the quantum mechanics (QM) are analysed. A brief summary of the DNA and RNA nucleobases works are presented at the end of this chapter. Chapter 5 reported the studies on the DNA and RNA base pairs, a conceptual extension from the nucleobases calculation. Similar to the nucleobases, the results for the base pairs are divided into two parts. The first part of the physical and structural properties calculations was investigated in the absence of an electric field. The effect of the electric field on the hydrogen bonds distance, and length of the molecule are also explored and discussed. Chapter 6 concludes the present research works by highlighting the key findings and suggests possible avenues for future work. Additional contents are provided through the appendices.

CHAPTER 2: LITERATURE REVIEW AND BASIC CONCEPTS

2.1 Introduction

In this chapter, we will begin by presenting the fundamental concepts regarding the ferroelectric, dielectric, deoxyribonucleic acid (DNA), and ribonucleic acid (RNA). We have also summarized the emergence and current state of development in the field of ferroelectricity in biological materials. We further review four most relevant biological materials such as glycine, porcine aortic walls, microtubules, and voltage-gated ion channel that reported to exhibit bioferroelectricity.

2.2 Ferroelectricity

The prefix ferro- part in the word of ferroelectrics is derived by the Greek and Latin word ferrum indicating the presence of iron. In fact, it seems to be an inappropriate misnomer as it does not refer to the presence of iron in the ferroelectric phenomena. Instead, the term ferroelectric arose from the analogous of ferromagnetic, which has shared many similar characteristics with ferroelectricity. Analogously, a ferromagnetic also presents a spontaneous magnetization in the absence of electric field. However, the fundamental mechanisms of ferroelectricity are totally different from ferromagnetism. For example, ferroelectrics undergo a phase transition at Curie temperature that involves the rearrangement of ions and change of crystal symmetry, and the materials change between the paraelectric and ferroelectric state. In the case of ferromagnetic materials, the transition between the spontaneously magnetized and magnetically disordered states that occurs at the Curie point does not involve a change in the crystal structure. Spontaneous electric polarization in ferroelectric is due to the noncentrosymmetric crystal structure, while spontaneous magnetization is due to unpaired and unsaturated electron spin. Above the Curie point, spontaneous electric polarization disappears since the materials structure changes to a centrosymmetric system, while individual magnetic

moments or unpaired electron spins do exist, but are randomized and cancel one another completely due to thermal motion (Dragan, 1998).

Ferroelectricity is a composition phenomenon where all ferroelectrics materials are both piezoelectric and pyroelectrics, but not all piezoelectric and pyroelectrics are ferroelectric. Additionally, ferroelectric materials are distinguished from ordinary dielectrics by their extremely high permittivities, the possibility of retaining some residual polarization after an applied electric field has been switched off, and the reversibility of the polarization or dipole moment, under the action of an applied electric field (Heinrich, 2012). To be more precise, they are a special group of polar materials that have noncentrosymmetric structure and possess at least two equilibrium orientations of the spontaneous electric polarization or dipoles vector in the absence of an electric field. These spontaneous electric polarization or dipoles vector can switch directions by the influence of an external applied electric field. As mentioned before all ferroelectrics are pyroelectric, they both have spontaneous polarization, also display hysteresis loop. The different between them is that the polarization of ferroelectric can be reversed by applying an electric field, but pyroelectric cannot be reserved. They would break down before the reversal could occur. While piezoelectric is a class of material which can be polarized in response to the external mechanical stress by converting mechanical into electrical energy. A schematic diagram illustrates the relations between piezoelectric, pyroelectric, and ferroelectric are shown in Fig. 2.1.

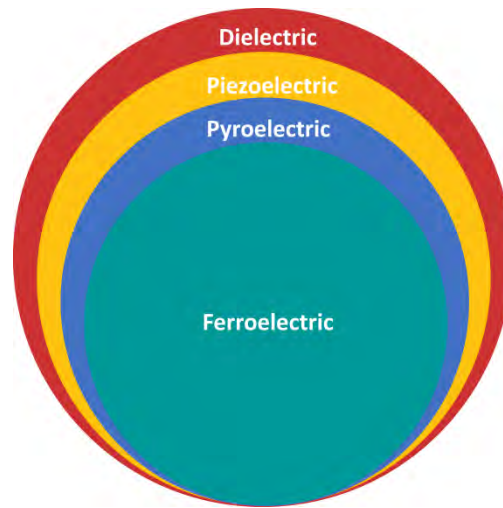


Figure 2.1: Relations among piezoelectric, pyroelectric and ferroelectric.

Based on the symmetry elements of translational position and orientation, crystalline lattices can be classified into 32 point group according to their crystallographic symmetry. These 32 crystallographic point groups can be further divided into two subgroups, one with a centre of symmetry and the other without. There are 11 of them belongs to centrosymmetric and the remaining 21 point group which belong to non-centrosymmetric. Out of a total of 21 non-centrosymmetric groups, 20 of them presented special direction and called piezoelectric. 10 point group among the 20 piezoelectric point groups belong to polar crystals and exhibit a spontaneous polarization in the absence of electric field. These 10 crystals are known as pyroelectric. Among the pyroelectric, those presences of a permanent polarization and can be reversed by the application of an external electric field are known as ferroelectric. This subgroup can exhibit all three piezoelectric, pyroelectric, and ferroelectric effect (Cowley & Coombs, 1973). The classification of piezoelectric, pyroelectric, and ferroelectric crystals can be illustrated in Fig. 2.2.

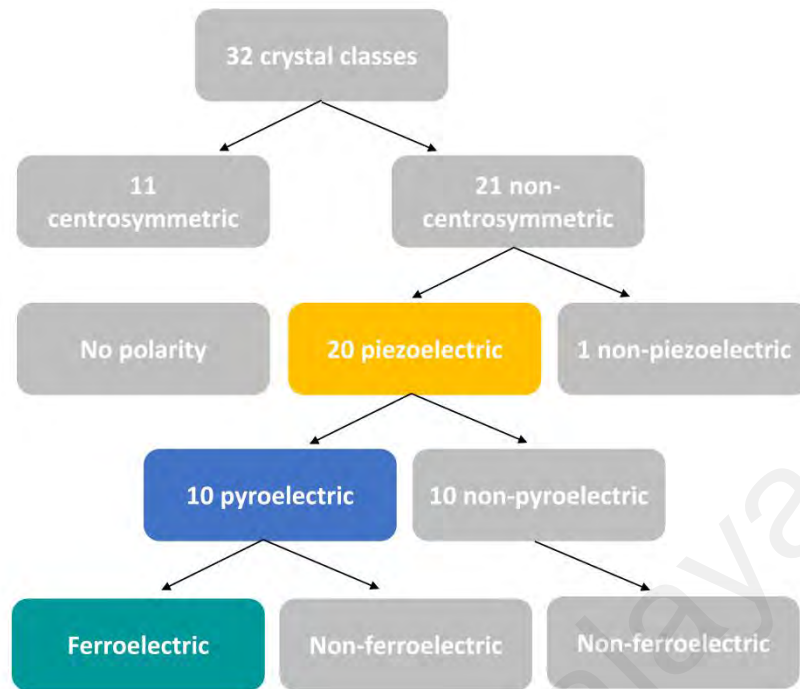


Figure 2.2: Classification of piezoelectric, pyroelectric, and ferroelectric crystals.

2.3 Basic features of ferroelectricity

2.3.1 Spontaneous polarization

Spontaneous polarization (P_s) is defined as the magnitude of polarization has a nonzero value in the absence of an applied electric field. A material is said to exhibit spontaneous polarization if the centres of the positive and negative charges do not coincide naturally even without the application of an electric field. The simplest method for measuring spontaneous polarization is using the experimental Sawyer and Tower circuit (Sawyer & Tower, 1930). Spontaneous polarization is a fundamental property for all pyroelectric and ferroelectric materials. But ferroelectric differ with pyroelectric in the sense that the direction of the spontaneous polarization vector in ferroelectrics can be switched between by the application of the electric field. Additionally, both pyroelectricity and ferroelectricity is an ordering phenomenon which the value of spontaneous polarization disappears at a critical temperature known as the Curie

temperature (T_c). Below the T_c , the material structure is slightly deformed with cations shifting in one direction and anions in the opposite direction, leading to the formation of an electric dipole moment hence results in spontaneous polarization. Along with the transition temperature or T_c , the material undergoes a phase transition, which is called a displacement phase transition. Whereas, above the T_c or also known as the paraelectric phase, the spontaneous polarization is disappeared for both pyroelectric and ferroelectric materials and this could be due to the dipole moment are randomly oriented and resulting in a zero net dipole moment (Gupta & Gupta, 2015). In general, therefore the change in spontaneous polarization vector ΔP_s lead by the temperature is given by

$$\Delta P_s = \lambda \Delta T \quad (2.1)$$

where ΔT is defined as a change in the temperature, and λ is a coefficient

Figure 2.3 shows the variation of the polarization (P) as the temperature decreases below the Curie point. In order to gain a better understanding of the ferroelectric phenomenon, an understanding of ferroelectric hysteresis loop and polarization switching are essential, which will be discussed in the next section 2.3.2.

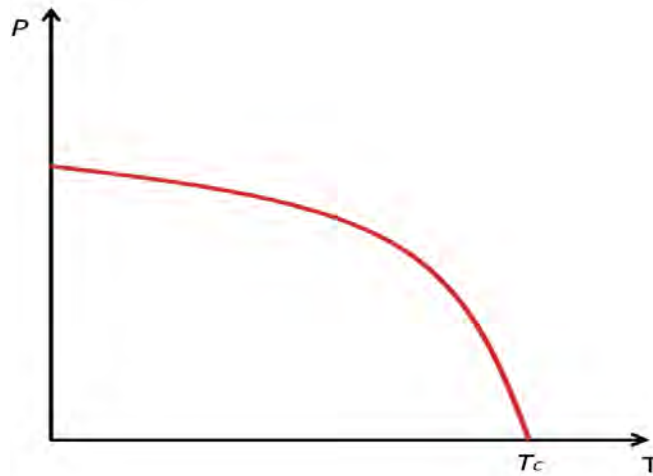


Figure 2.3: Polarization (P) versus temperature (T) curve.

2.3.2 Ferroelectric hysteresis loop and polarization switching

The most important feature that distinguishes the ferroelectrics materials from other pyroelectric materials is the reversal polarization or switching by an applied electric field. The reversal and switching mechanism are understood to take place when a sufficiently strong electric field is applied. By plotting, the change of polarization against an applied electric field, a ferroelectric hysteresis loop which resembles the magnetic hysteresis loop is displayed. Figure 2.4 illustrates a typical idea of hysteresis loop driven by a continuous applied electric field. Before the discussion, we first define the notations. The coercive field (E_c) is the required applied electric field to bring the polarization to zero. The value of polarization at zero fields is called the spontaneous polarization (P_s) or zero-field polarization (P_r). Starting at point O, by applying a small electric field, a weak polarization is induced in segment AB, being linearly proportional to the applied electric field. At this stage, the ferroelectric behaves like an ordinary dielectric because the applied field is not large enough to switch domains with the unfavourable direction of polarization. As the strength of the electric field increases further, the polarization starts to increase abruptly, due to nucleation and domain switching within the materials in response to the applied electric field. After most of the domains switched their direction to achieve the closet alignment with the externally applied electric field, polarization will slowly increase and eventually reach the saturation state as the electric field rises further (point B). The saturation polarization (P_{sat}) is determined by extrapolating the DC segment until it intersects the vertical axis at $E=0$ and the polarization at this region is strongly nonlinear. Reducing the electric field to zero leaves the material in a polarized condition with a zero-field polarization (P_r) which is usually slightly less than P_s , since not all domains have enough energy to switch back. Increasing the negative electric field reduces the polarization by reversing the direction of the domain (point E to G). The polarization first reaches zero at E_c , the

coercive field of the material, and then further to the saturation state in the reverse direction. The polarization will reduce with a reducing electric and reach a zero-field polarization as the field reaches zero (point F). If the negative electric field is applied again, the domains reverse their polarization direction in response to the electric field and a complete hysteresis loop is then obtained when a saturation state reach (point G). Although it is possible to back switch the material to the state of zero polarization, the material is no longer same as the starting condition, which can only be reinstalled by thermally depoling such as heating the material to temperatures above the Curie point.

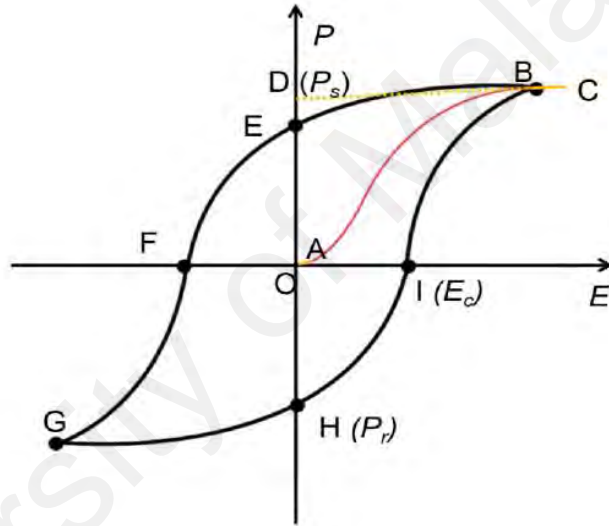


Figure 2.4: Ferroelectric (P - E) hysteresis loop (Reprinted with permission from (Sellmann, 2016).

2.3.3 A short history of ferroelectricity

The historical roots of ferroelectrics can be traced back to 1880 when two French physicists Jacques and Pierre Curie discovered quartz crystal or even some certain materials produce electrical charges in response to pressure, a phenomenon they referred to as piezoelectric. The discovery of piezoelectric properties has provided the driving force for Erwin Schrodinger and Walter Cady to further study the ferroelectric phenomenon. Using the Langevin's and Weiss's theories extending from the Debye's

theory of liquid dielectrics, they speculated that in solids an electric polarization similar with the magnetization, ought to be presented in the absence of electric field at a sufficiently low temperature or nowadays refers as the Curie temperature. Shortly, thereafter in 1912, Erwin Schrodinger coined the term ferroelectricity. Soon afterwards in 1921, Joseph Valasek demonstrated that the direction of the spontaneous polarization of Rochelle salt could be reversed by the application of an external electric field (Valasek, 1921). He observed a normal dielectric behaviour when the sample of Rochelle salt is placed at the two capacitor plates with the applied an electric field. After an applied electric field has been switched off, a residual polarization, as well as the hysteresis behaviour was detected between 255 to 297 K, indicates the existence of two stable polarization states in the Rochelle salt. One of the pioneer studies in the 1940s with the discovery ferroelectric effect in barium titanate (BaTiO_3) by Von Hippel et al. (von Hippel et al., 1946). Since then, many others oxides have been demonstrated ferroelectric properties including lead titanate (PbTiO_3) (Bhide et al., 1968) and strontium titanate (SrTiO_3) (Kim et al., 2007). Another exciting advancement related to polymer is the observation of ferroelectric behaviour in polyvinylidene fluoride (PVDF) by Bystrov et al. in 2013. PVDF has a spontaneous polarization around $8 \mu\text{C cm}^{-2}$, which is almost on the same order with BaTiO_3 , has opened the door for many novel applications with the characteristics of soft, flexible, lightweight, lead-free and biocompatible (Bystrov et al., 2013). Today, ferroelectricity remains one of the most active areas in solid materials.

2.4 Dielectrics

A dielectric is a poor conductor of electric current but efficient to sustain an electric field. Dielectrics are also known as insulators because they are having opposite properties as conductors. In practice, most dielectric materials are made to support an electrostatic field while dissipating with minimum energy in the form of heat (Frohlich,

1949). Dielectric materials are common to present in the form of solid, fluids as well as gases and these materials are widely used in many areas such as in capacitors, and radio-frequency transmission lines (Baker-Jarvis & Kim, 2012). An example of solid dielectric material including rubber, glass, ceramics while example for gases such as nitrogen and sulfur hexafluoride. Lastly, distilled water represents the dielectric liquid. In general, the dielectric can be classified into non-polar and polar dielectric categories. The major different between non-polar and polar dielectric is in nature all the non-polar dielectrics belong to symmetric shape and the centre of mass of positive charges as well as negative charge will coincide with each other (e.g. O_2 , N_2 , H_2) while polar dielectric consists of asymmetric shape therefore the positive does not coincide with negative charges (e.g. H_2O , CO_2 , NO_2). In the absence of an external electric field, the non-polar dielectric molecules are neutral, whereas molecules for polar dielectric are randomly oriented, resulting in zero net dipole moment. When an external electric field is applied to a non-polar dielectric, the molecules tend to separate into positive and negative charge, consequently forming electric dipole moment called induced dipoles. If the polar dielectric is placed with an applied electric field, all the permanent dipole will align themselves in the direction of the external electric field (Wallace, 2017).

An applied field to a dielectric material, cause the nuclei to push by the field, resulting in an increase the positive charge toward the electric field while the electron clouds with negative charges are pulled against to the opposite. This process creates two centres of charge are known as polarization and the dielectric material in such a state is said to be polarized as shown in Fig. 2.5. Basically, the electric field can distort the charge distribution of a molecule by stretching and rotating. When a uniform electric field is applied to a dielectric material, the dielectric molecules are bent and stretched by the field resulting in an induced dipole moment. Rotation mechanism only occurs in polar dielectric materials, as the permanent dipole moment rotates in the direction of the

applied electric field. The rotation process is more dominant than the stretch mechanism, as it is easier for the electric field to rotate a dielectric molecule rather than stretch (Kroh & Felderhof, 2000). Typically, polar dielectric materials polarize more strongly than non-polar materials, for instance, the polar water molecule has dielectric strength 70 times as compared with a non-polar nitrogen molecule. Because an external electric field stretches both polar and non-polar molecules, but polar molecules stretch and rotate in response to the field (Bottcher, 1973).

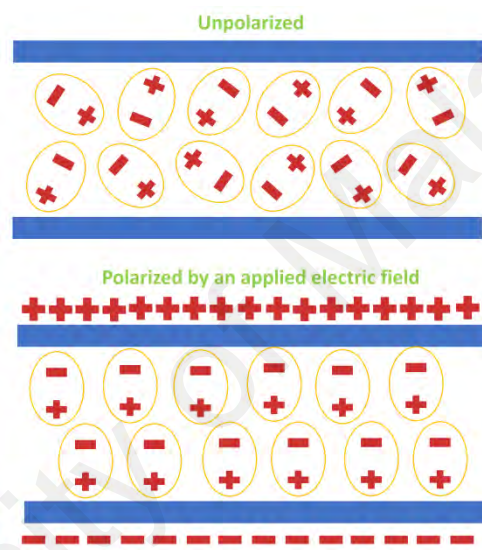


Figure 2.5: A dielectric material showing the orientation of positive and negative charges creating a polarization effect.

2.5 General properties of dielectric

2.5.1 Polarization

The formation of polarization can be described when polar or non-polar dielectric materials are kept in an electric field, therefore occur a slight displacement of the negative and positive charges of the dielectric's atom or molecules (Raju, 2003). The dielectric materials are said to be polarized or in a state of polarization when the dipole moments (\vec{p}) are present. As a simple example, consider that a polarized atom of

dielectric material is represented by an electric dipole where the positive charges ($+Q$) are representing the nucleus and negative charges ($-Q$) representing the electronic charge, the two charges (Q) being separated by a small distance (\vec{d}). The electrons orbit the nucleus and act like a negatively charged cloud surrounding the nucleus. When the atom is in the unpolarized state, the cloud surrounds the nucleus symmetrically as shown in Fig. 2.6. Under the applied electric field, the electron cloud becomes slightly displaced and asymmetrical.

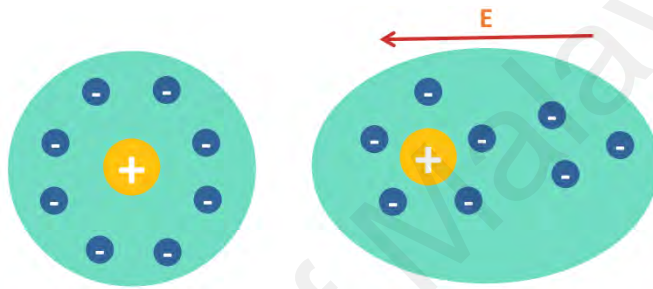


Figure 2.6: Unpolarized atoms become polarized when an electric field(\vec{E}) is applied.

The effect of all this can be described by the polarization(\vec{P}) or dipole moment(\vec{p}) per unit volume with the unit of C/m^2 . Thus,

$$\vec{P} = \frac{N\vec{p}}{V} = \frac{NQ\vec{d}}{V} \quad (2.2)$$

where N is the number of dipoles in volume V and the net dipole moment in volume is equal to $\vec{p} = Q\vec{d}$

In most cases, the polarization(\vec{P}) is directly proportional to the applied electric field but independent of temperature. This linear dielectric can be expressed as

$$\vec{P} = \varepsilon_0 X_e \vec{E} \quad (2.3)$$

where X_e called for electric susceptibility, ε_0 for permittivity of free space and \vec{E} is equal to the electric field.

Furthermore, the net induced dipole moment (\vec{p}) of an atom is proportional to the applied electric field (\vec{E}) with its direction parallel to the field.

$$\vec{P} = \alpha \vec{E} \quad (2.4)$$

where α called atomic polarizability $\alpha = \frac{\vec{p}}{\vec{E}}$

Generally, there are four different types of mechanism which polarization can occur in dielectric materials and it can be distinguished by their reactions when it is subjected to an external electric field. Figure 2.7 summarized the frequency dependence of various polarization mechanisms, represented by dielectric constant. The electronic polarization (P_E) is the fastest and typically occurs at frequencies between to $\sim 10^{13}$ to 10^{15} Hz. In contrast, atomic polarization (P_A) typically persists in ionic compounds and responds to field frequencies between $\sim 10^{10}$ to 10^{13} Hz while dipolar polarization (P_D) arises at the intermediate frequency range between $\sim 10^8$ to 10^9 Hz involving the reorientation of the permanent molecular dipole moments. Interfacial or space charge polarization (P_I) can only contribute when there is an accumulation of charge at interfaces at the frequencies below 10Hz (Bain & Chand, 2017). Further details regarding these four polarizations are discussed below. The total polarization (P_{Total}) may be obtained by summing from all the four different types of polarizations and can be conveyed as the following equation:

$$\vec{P}_{total} = \vec{P}_E + \vec{P}_A + \vec{P}_D + \vec{P}_I \quad (2.5)$$

where P_E is the electronic polarization, P_A is the atomic polarization, P_D is the dipolar polarization, P_I is the interfacial polarization

In fact, polarization can also arise through other external factors such as mechanical pressure as shown in the case of piezoelectric materials or spontaneously exist in particular ferroelectric materials.

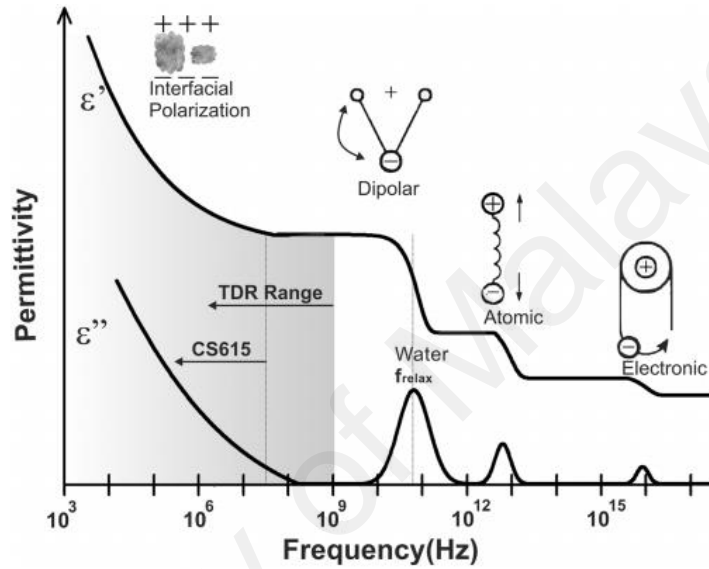


Figure 2.7: The frequency dependence of dielectric constant in the electronic, atomic, dipolar, and interfacial polarization (Reprinted with permission from American Geophysical Union) ((Chen & Or, 2006).

2.5.1.1 Electronic Polarization (P_E)

Electronic polarization occurs in an atom when the electric field distorts the negative cloud of the electron relative to its positive atomic nuclei in a direction opposite the field. The shifting of electron clouds created a dipole moment where one side of the atom slightly positive and the opposite side slightly negative. This process happens instantly when the voltage is applied due to the electrons is very light, and thus able to follow the high-frequency field. The polarization displacement is proportional to the volume of the atoms but independent to temperature (Wang et al., 2018).

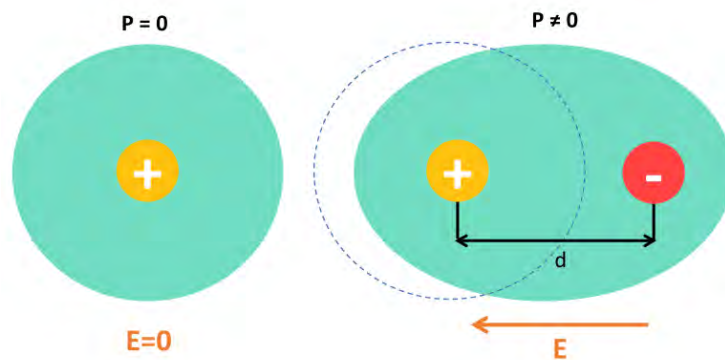


Figure 2.8: Schematic diagram of electronic polarization.

2.5.1.2 Atomic Polarization (P_A)

Atomic polarization arises due to relative displacement between cations and anions in an ionic crystal such as sodium chloride (NaCl), potassium chloride (KCl) and lithium bromide (LiBr). The dipole moment is generated when an external field is applied, which leads to ions displaced thereby giving rise to a net polarization. While in the absence of an external electric field, the positive and negative ions are cancelled each other resulting net zero polarization. As the atomic polarization mechanism involving more heavy ions, therefore, the process is slightly slower than electronic polarization. Like electronic polarization, atomic polarization is resonant in nature and is independent of temperature. Atomic polarization occurs in compounds in which chemical bonds are not 100% covalent (Bonin & Kresin, 1997).

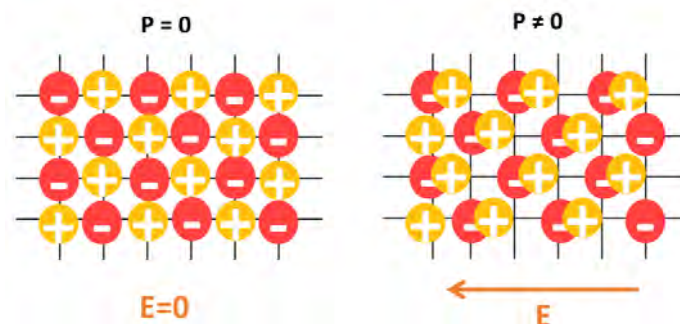


Figure 2.9: Schematic diagram of atomic polarization.

2.5.1.3 Dipolar Polarization (P_D)

This type of polarization typically occurs in polar materials with a permanent dipole moment and the process is slower compared with electronic and atomic polarization (Boswarva & Franklin, 1965). For instance, a water molecule (H_2O) is bent structured where oxygen atom will have negative charges and the two hydrogens will have positive charges. This resultant permanent dipole moment in each water molecule even in the absence of applied electric field. However, in the presence of an external electric field, the molecules with permanent dipole moment orient themselves along to the direction of the applied electric field. This process called dipolar polarization, which is dependent over temperature.

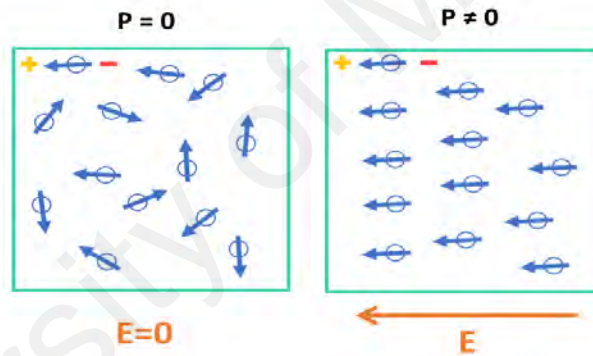


Figure 2.10: Schematic diagram of dipolar polarization.

2.5.1.4 Interfacial Polarization (P_I)

Interfacial or also known as space charge polarization consider the slowest process among the four polarizations as it occurs when there is an accumulation of charge at the interface or grain boundary in the amorphous or polycrystalline solids material. It often occurs when there are two electrodes connected to materials and the free charge response to the applied electric field, giving rise to redistribution of charges with respect

to the direction of the applied field. Without the applied electric field, the charges are orderly oriented (Wang et al., 2018).

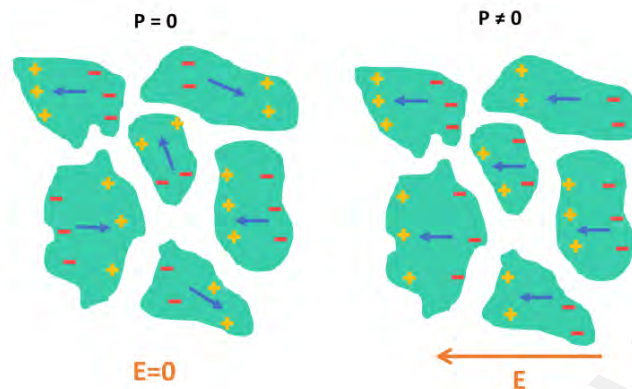


Figure 2.11: Schematic diagram of interfacial polarization.

2.5.2 Dipole moment

In physics, the electric dipole moment (\vec{p}) is defined as a measurement of two equal point charges (Q) of opposite sign separated by a distance (\vec{d}) and the product of $Q\vec{d}$ is its electric dipole moment. If the two charges are superposed, the resultant field is zero, but if they are separated by even a small distance, there is a finite resultant field. Dipole moment arises when atoms in a molecule share electrons unequally and this commonly occurs in ionic bond or atoms in a covalent bond (Sidgwick, 1936). A popular example is the water molecule which is made up of one negative charge oxygen atom and two hydrogen atoms with positive charges, as shown in Fig. 2.12. The differences in electronegativity and lone electrons give bent in the middle, the O-H axes making an angle of about 105° with a 1.8 Debye net dipole moment (Talman, 2017).

Although the SI unit of dipole moment is coulomb metre ($\text{C}\cdot\text{m}$), it is still commonly reported in the non-SI unit Debye (D), named after Peter Debye, a pioneer in the study of dipole moments of molecules where $1\text{D} = 3.335 \times 10^{-30} \text{ Cm}$ (Ramsey, 1982). The

electric dipole moment form by two separated charges ($-Q$ and $+Q$) can be calculated by equation

$$\vec{p} = Q\vec{d} \quad (2.6)$$

where \vec{p} is the dipole moment, Q is a total charge and \vec{d} is the vector that defines the position of the two charges.

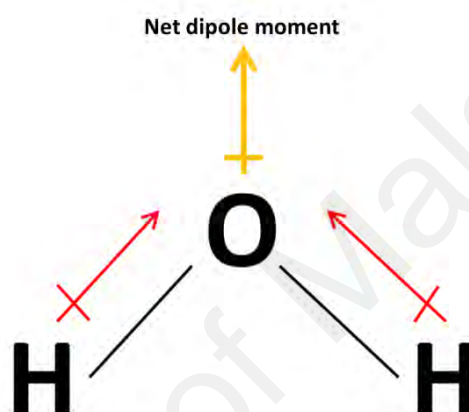


Figure 2.12: Dipole moment of a water molecule.

2.5.3 Polarizability

The electric polarizability (α) can be defined in terms of the relative tendency of an atom to be distorted from its original shape and become polarized in the presence of an external electric field. This distortion causes the electron cloud of a non-polar molecule or atom to acquire a dipole moment. The greater the polarizability, the larger is the induced dipole moment for a given applied field. In ordinary, the interaction of the atom's electrons and the nucleus is correlated with the polarizability. For an atom with fewer electrons and a small number of electron clouds, the interaction between the positive charge nucleus and negative charges electrons in the atom's orbitals is stronger. In contrast, large atoms with excess electrons negative ions are more easily to become

polarized because the excess electrons will cancel a portion of the nucleus positive charge and thereby increase the repulsive interaction between itself and electron farther away. In another word, the higher the number of the electron, the less control of the nuclear charge and therefore the greater polarizability of the atom. This phenomenon is described as the shielding effect (Miller, 1990). The following equation describes as the ration of induced dipole moment(\vec{p})to the applied external electric field(\vec{E}) and the unit for polarizability can be written as $C^2m^2V^{-1}$.

$$\alpha = \frac{\vec{p}}{\vec{E}} \quad (2.7)$$

where \vec{E} denotes the strength of the applied electric field and \vec{p} is the induced dipole moment.

2.6 Introduction to deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)

The two nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) comprise the genetic makeup of all living organism on earth. This repository stores all information needed to specify cellular functions and can be considered a blueprint for life. Besides storing or transmitting information, RNA forms structural parts of units such as the ribosome, and in some systems has a catalytic function. Both DNA and RNA are long, chainlike polymers assembled from repeating subunits, the nucleotides. A nucleotide consists of three major parts, a five-carbon pentose sugar, a flat heterocyclic nucleobases, and a negatively charged phosphate group, all linked together by covalent bonds. In RNA, the pentose sugar is ribose, and in DNA, it is deoxyribose, in which hydrogen replaces the hydroxyl group at the 2' carbon. A chain of one, two or three phosphates links to the ribose or deoxyribose sugar at its 5' carbon to complete the mono-, di- or triphosphate form of a nucleotide (Sinden, 1994). The nucleobases in naturally occurring nucleotides, pyrimidines, and purines, are ring-shaped molecules

containing both carbon and nitrogen atom. Pyrimidines contain one carbon-nitrogen ring, whereas purines contain two rings. The purines adenine (A) and guanine (G) and the pyrimidine cytosine (C) are present in both DNA and RNA. While the pyrimidine thymine (T) present in DNA is replaced by the pyrimidine uracil (U) in RNA. Thus, DNA and RNA differ in the sugar present, either deoxyribose or ribose and the presence of either T in DNA or U in RNA. The nucleobases are often abbreviated as A, T, G, C, and U, respectively (Chargaff et al., 1950). The nucleobases are planar, conjugated rings that are typically uncharged under physiological conditions, but can gain or lose a proton at near physiological pH. The pK_a values for each nucleobase, which are listed in Table 1.1, correspond to the pH at which this transition generally occurs at a given atom is protonated at pH values below the pK_a and unprotonated at pH above the pK_a (Krishnamurthy, 2012). This property of nucleobases is particularly important for the ability of RNA molecules to participate in a chemical reaction.

DNA serves as the storehouse, or cellular library, that contains all the information required to build the cells and tissues of an organism. The exact duplication of this information in any species from generation to generation assures the genetic continuity of that species. In the process of transcription, the information stored in DNA is copied into RNA, which has three distinct roles in protein synthesis. Messenger RNA (mRNA) carries the instructions from DNA that specify the correct order of the amino acids during protein synthesis. The remarkably accurate, stepwise assembly of amino acids into protein occurs by translation of mRNA. In this process, the information in mRNA is interpreted by the second type of RNA called transfer RNA (tRNA). tRNA is also necessary synthesizing proteins and helps translate the sequence of nucleic acids in a gene into the correct sequence of amino acids. Ribosomal RNA (rRNA) works as an enzyme to form the peptide bonds between amino acids in a polypeptide. A wide range of other RNA molecules also performs important functions within the cell.

Table 2.1: pK_a values for ionise atoms in nucleobases (McLennan & Turner, 2012).

Nucleobases	Atom	pK_a
A	N1	3.88
T	N3	10.5
G	N7	3.60
C	N3	4.56
U	N3	10.1

DNA exists in cells as double-stranded containing two intertwined chain of nucleotides, as illustrated in Fig. 2.13. In DNA the two strands are held together by hydrogen bonds between the nucleobases. When coiled, DNA resembles a spiral staircase. When unwound, it resembles a stepladder. The side of the ladder is made entirely of phosphate and sugar molecules and the rungs of the ladder exhibit complementary base pairing. A always pairs with T and G always pairs with C, often called Watson-Crick base pairs (Watson & Crick, 1953). Base pairs have shapes that fit together. Base pairing allows DNA to replicate in a way that ensures that the sequence of nucleobases will remain the same. This is important because it is the sequence of nucleobases that determine the sequence of amino acids in a protein. Unlike DNA, which exists primarily as a very long double-stranded helix, most RNA normally occurred as a single-stranded, and they exhibit a variety of conformations. The fundamental structural unit of folded RNA molecules is short. Differences in the sizes and conformations of the various types of RNA permit them to carry out specific functions in a cell. RNA molecules can be described in terms of three levels of structural organization. The primary structure of RNA refers simply to the sequence of nucleobases along an RNA strand, reading from 5' to 3' direction. The secondary structures in a single-stranded RNAs are formed by pairing of nucleobases. Hairpins are formed by pairing of nucleobases within about five to ten nucleotides of each other, and stem-loops by the pairing of bases that are separated by eleven to several hundred

nucleotides. These simple folds can cooperate to form more complicated tertiary structures, one of which is termed a pseudoknot (Alberts et al., 1989).

Single strands of RNA are relatively easily digested by enzymes called ribonucleases unless they possess much intrastrand base pairing (making them double-stranded over localized regions) or extensively complexes with proteins such as those in ribosomes. By contrast, DNA molecules are much more stable because they are double-stranded throughout their length and enzymes are normally present in the nucleus to repair damage to DNA. There are no comparable enzymes either in the nucleus or in the cytoplasm to repair RNA damage. RNA molecules must be continuously replaced as they are being broken down at various rates.

RNA molecules range in length from less than one hundred to many thousands of nucleotides, the nucleotide (nt) or bases is used as a measure of length. The smallest RNA molecules (75-80 bases) are tRNAs. These serve as adaptor molecules in position each amino acid in proper alignment on the ribose for polymerization into polypeptide chains. tRNAs are unique in that they contain a few unusual bases in addition to the standard ones. Since there can be 20 different kinds of amino acids in biological proteins, there must be at least 20 different species of tRNA molecules, one specific for each amino acid. DNA molecules can be as long as several hundred million nucleotides. The length of a double-stranded DNA is often measured in the number of base pairs (bp). The average unit of length for 1 kilobase pair (kb) is corresponding to 1000 base pairs and 1 megabase pair (Mb) is equivalent to 1,000,000 base pairs (Turner & McLennan, 2005).

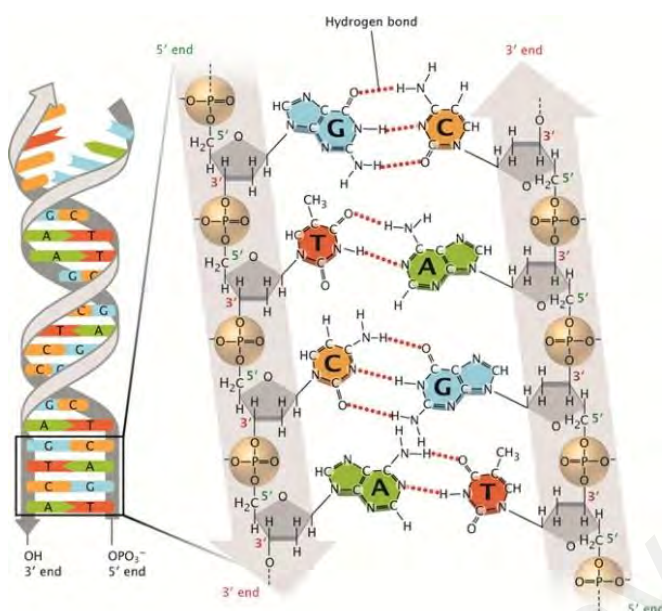


Figure 2.13: Double-stranded DNA structure (Reprinted with permission from Scitable by Nature Education) (Pray, 2008).

2.7 Emergence of bioferroelectricity

Piezoelectricity, where mechanical deformation produced polarization, was first discovered in bone samples in 1957 (Fukada & Yasuda, 1957) and subsequently widespread observed in a number of biological materials or systems including teeth (Athenstaedt, 1971), tendon (Williams & Breger, 1975), skin (Rossi et al., 1986), muscle (Fukada & Ueda, 1970), protein (Lemanov et al., 2002), myosin and actin (Hiroyoshi & Eiichi, 1971). It is well known now that the piezoelectricity is a fundamental essential property of biological materials (Shamos & Lavine, 1967). Shortly thereafter, in 1966, the subset of piezoelectricity known as pyroelectricity was also discovered in bones (Lang, 1966) and later in many other biological materials as well, leading the possible existence of biological ferroelectricity (Athenstaedt, 1974; Athenstaedt et al., 1982). As mentioned previously, all ferroelectrics are pyroelectrics, and all pyroelectrics are piezoelectrics, so it is only logical to determine biological

ferroelectricity after both piezoelectricity and pyroelectricity are reported in biological materials.

Since the 1970s, the significant levels of interest for ferroelectricity in biological materials have been remarkable growth and the peaking period happened in the duration of 1980s with the discovery of voltage-dependent ion channels as the first ferroelectric biological membranes (H. Richard Leuchtag, 1988). However, the interest of the ferroelectricity, particularly in biological materials has reduced at the beginning of the 1990s. Later on, this interest has experienced a revival with the new era particularly in the field of nanoscale electromechanical studies owing to the high-resolution imaging of piezoresponse force microscopy (PFM) (Tuszynski et al., 2008). Although, ferroelectricity has attracted intensive research interests in the past, due to their unique versatile properties and applications but the interest of ferroelectricity in biological materials are still less (Blázquez-Castro et al., 2018). One of the primary factors for this rarity is mainly attributable to the poor understanding of the physiological and pathological significance of biological ferroelectricity. Moreover, the biological sample which containing water provided an additional fundamental challenge for the sample preparation. In reality, there are a large number of ions exist in the water that may lead to difficulties in the direct measurement of the true ferroelectricity, particularly when the sample required additional heating process. Heating is the most essential and important aspect in the ferroelectric hysteresis measurements, but also may cause physical alteration including denature, soften, evaporate or deform of the biological sample. Furthermore, setting up electrodes on biological materials especially biological tissues can be a challenging task because they are formed from a robust porous hierarchical structure with biological fluid in between them, which can easily trigger a leakage current that complicated the measurement process (J. Li et al., 2013; Pal et al., 2015; Zelisko et al., 2015).

Even though the study of bioferroelectricity can be challenging, but there are still a number of studies has been published in the past decades, dedicated to discover the principle of bioferroelectricity and to understand the origin of spontaneous polarization and ferroelectric switching of biological materials. To the best of our knowledge, the idea to apply ferroelectricity with biological materials was first proposed by von Hippel in 1970, during the 2nd International Meeting on Ferroelectricity, where he introduced “A variety of polar coupling phenomena can be expected to produce ferro- and antiferroelectric transitions in non-living and in biological systems” (Hippel, 1970). Afterwards, several studies from Bystrov, Leuchtag, Polonsky, Stanford, Liu and Heredia have determined the existences and significates of bioferroelectricity in biological materials (Bystrov et al., 2012; Heredia et al., 2012; H. Richard Leuchtag & Bystrov, 1999; Liu et al., 2013; Polonsky et al., 1960; Stanford & Lorey, 1968). In the following section, we will briefly discuss and review four reported biological materials relevant to the present study.

2.7.1 Bioferroelectricity in glycine

Traditionally, it was believed that amino acids are non-centrosymmetric, though their piezoelectric coefficients were shown to be extremely small (Lemanov, 2000). However, recent studies found that these amino acids especially glycine exhibits a strong piezoelectric response, suggesting the possible existence of ferroelectricity in amino acids. Glycine, also known as glycocoll, is a colourless solid crystalline that is commonly found within the loop and coil of protein. Although glycine is considered as nonessential amino acids, it can be still conditionally essential participates in both stimulatory and depressant roles in brain (Weinberg et al., 2016). It is interesting to note that glycine has been one of the most studied crystalline amino acids since the discovery in 1820 by Henri Braconnot (Vickery & Schmidt, 1931). Structurally, glycine can exhibit at least four kinds of polymorphic form in a different condition, namely α -

glycine, β -glycine, γ -glycine, and under high pressures it transforms to be δ phases. Among the four, γ -glycine is the most stable at room temperature (Bull et al., 2017). But, when γ -glycine is heated up to 165 °C, the conformation transform spontaneously into α form (Mailoud et al., 2018). An ordinary α -glycine was identified by Albrecht and Corey in 1939 (Albrecht & Corey, 1939). Though the unstable β polymorphic form was originally determined by Fischer and Fournau at the beginning of 1901 (Fischer & Fournau, 1901). The third form, γ -glycine was described in 1954 by Iitaka (Iitaka, 1954). A new polymorphic, δ -glycine, occur at the pressure of 2.75 GPa was discovered by Dawson et al. (Dawson et al., 2005). Lately, glycine has drawn the major attention and subsequently become a hot research topic, due to its essential role as an inhibitory transmitter in the brain and the ability to exhibits several distinguished polymorphic forms at different temperatures and pressures. Moreover, its strong piezoelectric response and ability to reverse polarization with the action of electric field adds important functionality in the development of bioelectronics and biomechanical applications.

Recent works by Heredia et al. have determined γ -glycine for the first time as the nanoscale ferroelectric at a single molecular level (Heredia et al., 2012). The glycine microcrystals were grown by a slow evaporation method with the presence of supersaturated solution at different temperatures. Fourier transform IR (FTIR) was used to confirm the presence of the γ -glycine. The measurements of the piezoelectric and ferroelectric at a nanoscale level were then carried out by PFM. The high-resolution phase image of the PFM clearly reveals distinct regions of polarization, where the dark and bright correspond to the polarization direction either toward the substrate or surface. A piezoelectric coefficient of 10 pmV⁻¹ was observed, indicating its excelled ferroelectricity. In addition, the dielectric constant showed a distinct maximum at temperature 210 °C under a frequency of about 1 MHz with an applied ac electric field,

which is the characteristic of phase transition from ferroelectric to paraelectric state. Later on, Heredia et al. fitted the data to a Curie-Weiss relationship, finding a first-order transition at 300 ± 1 °C and a Curie constant of 2000 ± 10 °C. More interestingly, another similar set of experiments performed by Vasileva et al. have demonstrated a significant local polarization switching in the β -glycine domain via an electrically biased PFM tip (Vasileva et al., 2019). The reversal of the domain was visualized by applying an electrical voltage with an amplitude in the range of 40V to 200V and a frequency range from 20 kHz to 2 MHz, through the conductive tip while scanning the surface. An obvious phase change occurs when the voltage is reached at 100 V, but shortly after the removal of the voltage, a 180° phase flip is observed. It has been further verified that such switching behaviour is repeatable throughout the sample.

In the theoretical work, made by Hu and his colleagues, performed using the density functional theory (DFT) calculations and molecular dynamics (MD) simulations have determined that the spontaneous polarization for γ -glycine is approximately $70.86 \mu\text{C}/\text{cm}^2$ at 300 K (Hu et al., 2019). The piezoelectric coefficient was estimated to be around 10.4 pmV^{-1} which is in good agreements with those reported Vasileva. In addition, Hu et al. also discovered that the γ -glycine exhibits the order-disorder phase transition with a high transition temperature T_c of 630 ± 5 K, which is much higher than the experimentally measured temperature. An interesting molecular model constructed by Bystrov et al. in 2016, consisting of 64 elementary of β -glycine molecules or 640 atoms in a single cluster (Bystrov et al., 2016). The optimization through DFT in local density approximation (LDA) method revealed the molecular cluster dipole moment of $D=150.6$ Debye, directed along the Y -axis, corresponding to a polarization $P=0.127 \text{ C}/\text{m}^2$.

2.7.2 Bioferroelectricity in porcine aortic walls

The piezoelectric effect has been observed in porcine aortic walls since 1969 by Fukada and Hara (Fukada & Hara, 1969). However, the existence of ferroelectric properties in porcine aortic walls was unknown until 2012 when Liu et al. first reported the measurement of permanent polarization in aortic walls (Liu et al., 2012). In this work, the aortic walls from pig, which composed of three different layers of the intima, media, and adventitia, were cut into the size of $2\text{ cm} \times 1\text{ cm}$ with the thickness of 0.82 mm, and then soaked in graded distilled ethanol for dehydration, followed by drying out. Afterward, the tissue samples with a size of $8\text{ mm} \times 8\text{ mm}$ were cut and glued at the silicon sputtered with 100 nm thick gold using silver paint for PFM studies. In order to enhance the sensitivity of the measurement, they applied an ac voltage near the resonant frequency of the cantilever-sample through the conductive atomic force microscope (AFM) tip with the employment of PFM to measure the piezoelectric effect of the inner wall. In doing so, the aortic wall image overlaid on the 3D topology of vertical and lateral amplitude were mapped as 250 pm and 90 pm respectively, both taken in the same region with a 3 V ac voltage. At the same time, Liu et al. attempted to verify the ferroelectricity by switching aortic wall polarization through applying a series of dc voltage to the sample using a triangular sawtooth form, with the piezoresponse measured simultaneously. By comparing PFM amplitude images, a clear 180° switching was found in the phase signal, indicating opposite distributions of polarization. The corresponding hysteresis loop of the phase-voltage and the butterfly loops of the amplitude-voltage were obtained throughout the sample, with a coercive voltage of 8.4 V on the positive side and -10.8 V on the negative side.

Observed from the switching spectroscopy and relaxation studies show that the aortic wall polarization was more preferred to stay at outward. This is because the induced polarization remains stable and permanent after the positive voltage is removed from the

sample. However, with a negative voltage, the inward polarization becomes unstable and would reverse the switching to the more stable outward orientation upon the removal of the external voltage. Further PFM analysis revealed that the aortic wall has a piezoelectric coefficient of $\sim 1 \text{ pmV}^{-1}$, which is two orders of magnitude are larger compared to the reported by Fukada and Hara for blood vessel walls. Surprisingly, a year after, Liu et al. reported that the ferroelectricity in aortic elastin can be suppressed by glucose (Liu et al., 2013). With this finding, researchers could open up the clue to understanding the important physiological and pathological implications from aging to arteriosclerosis that is closely related to aortic elastin.

2.7.3 Bioferroelectricity in microtubules

Another example of interesting biomaterials that demonstrating bioferroelectricity can be found inside the structure of the cell, called microtubules. Both eukaryotic and prokaryotic cells including archaea and bacteria are structurally organized by the cytoskeleton, serve as a scaffold that helps cells maintain their shape and provide resistance to mechanical stresses. The cytoskeletons of eukaryotes cell consisting of three major types of cytosolic fibers. The microfilaments responsible for cell shape and cytokinesis, whereas the functions of intermediate filaments are to control cell cohesion and to provide mechanical support. Microtubules are assembled from α and β tubulin as hollow cylindrical ferroelectric polymers, mainly control the intracellular signaling and chromosome segregation. Each microtubule is a polar structure with a diameter of approximately 25 nm and a length ranging from 50-60 μm (Satarić et al., 1993). Furthermore, microtubules also contribute by providing structures support and transportation for the interior of the cell. It transports membrane-bound vesicles and other cellular materials across the cell, playing a critical role in cell migration as well as cell division (Sataric et al., 2008). There have been several speculations that these functions can be understood based on the hypothesis that microtubules exhibit

ferroelectric-like effect and its potential physiological implications also have been suggested (Brown & Tuszynski, 1999).

The idea that microtubules possess ferroelectric behaviour was first proposed in 1985 by Hol when he investigated the α -helix dipole in protein function (Hol, 1985). Lately, the ferroelectric properties measurement performed by Vassilev et al. (Vassilev et al., 1982), Kaimanovich et al. (V. Kaimanovich et al., 1990; V. A. Kaimanovich et al., 1989) and White et al. (White et al., 1990) have demonstrated that the tubulin consist of a permanent electric dipole that contributes to the overall polarity of the microtubules. Therefore, we believe that the measurement of the dipole moment in tubulin could stimulate further study of microtubules, aiming at deeper understanding of the actual polymerization mechanism as well as enhance knowledge for the experiment parameter setting. For similar reason, it will also benefit our future understanding on how to employ the tubulin self-assembly mechanism into novel biomaterials nanofabrication application. In the computational simulation studies conducted by Mershin *et al.*, they reported that the α and β tubulin dipole moment as $\alpha=552$ Debye and $\beta=1193$ Debye, respectively (A. Mershin et al., 2004). While the whole tubulin molecule are in the order of $\alpha\beta=1740$ Debye, agreed well to the experimental result obtained by Schuessler et al. (Schuessler et al., 2003). It also has been observed that both α and β tubulin are strongly negatively charged at physiological pH 7.2 (Andreas Mershin et al., 2006). This could be due to the presence of the carboxyl terminus at the tail of the tubulin, and in fact, there are many suggestions that this C-terminal is important for polymerization, protein interactions and charge conduction (D. L. Sackett, 1995). Moreover, Sanabria et al. studied the dielectric constant and refractive index using the surface plasmon resonance (SPR) technique. The measured dielectric constant and refractive index were $\epsilon(=n^2) = 8.4$ and $n=2.9$, respectively (Sanabria et al., 2006).

Apart from the above observations, there is also some experimental evidence in support of microtubules exhibit ferroelectricity, came from the study with an applied electric field. In the works of Stracke *et al.*, tubulin has been purified from porcine brain homogenates through two cycles of temperature-dependent polymerization/depolymerisation, followed by a phosphocellulose column chromatography to separate the tubulin from microtubules associated proteins (Stracke et al., 2002). Later on, the tubulin with the weight of 6.3 mg/ml was supplemented into 10 μ M paclitaxel and 1mM GTP, followed by incubation at 37 °C for 20 min. A special observation chamber was then constructed and two microscopic slide materials covered the sample as the support with the equipped of Axiophot microscope. The direct magnitude of dc voltage range between 0 and 100V was constantly applied throughout the platinum wires. At \sim pH 6.8, the microtubule moved from the negative to the positive electrode, indicating that microtubule orientation is accompanied by substantial changes of charge distributions inside the tubulin. Similarly, Sackett et al. showed that the microtubules net charge strongly depends on pH (D. Sackett, 1997). Timmons et al. theoretical studied the effect of an applied electric on the structural and electrical properties of tubulin-based on molecular simulation. In the study, a tubulin heterodimer was subjected to an external electric field in the range of 50 to 750 kV/cm for 10ns simulation (Timmons et al., 2018). They found an obvious mean squared displacement within the flexible loop and C-terminal of tubulin.

2.7.4 Bioferroelectricity in voltage-gated ion channels

Voltage-gated ion channels are remarkable transmembrane proteins, which are commonly found in cell plasma membranes and involved in the transmission impulse of muscle and nerve cells. These glycoproteins consist of four subunits that arrange in circular symmetry around a central pore and they are embedded at the lipid membrane, controlling the flow of ions according to the voltage across the secretory and epithelial

cells. Both sodium and potassium ion channels are composed of two different conformations states: open and close. They can switch from a resting potential to an action potential as the inner cell membrane voltage change from a negative to the positive charge. The transmissions of nerve impulses along the neuron totally depending on these two types channels, where sodium ions move into the cell and potassium ions move out through these ion channels. In 1988, Leuchtag reported that voltage-gated ion channels, including sodium and potassium ion channels exhibit ferroelectric effects (H. Richard Leuchtag, 1988). The detailed structural-analysis suggested that these ion channels undergo a first-order phase transition, from a closed conformation ferroelectric state to an open conformation paraelectric state. He further added that the ion channels are liquid crystalline and the excitable membrane possess a heat-block temperature which is associated with the Curie temperature (H. R. Leuchtag, 1987). The typical current-voltage (I-V) hysteresis was obtained from both of the membrane and channel. The temperature-dependent currents process further confirms the existence of pyroelectric behaviours as well. In addition, the sodium channels showed distinct surface charge density in the range of $\sim 2.2 \mu\text{C}/\text{cm}^2$, which is the ferroelectric spontaneous polarization (Bystrov et al., 1994; H. Richard Leuchtag & Bystrov, 1999).

CHAPTER 3: METHODOLOGY

3.1 Introduction

This chapter will begin with an introduction to the molecular modelling theory and briefly describe different types of computational chemistry methods. We also present the concept of geometry optimization. At the end of this chapter, we will describe in details the computational techniques and approach that we used in the present study.

3.2 Molecular modelling theory

3.2.1 Geometry optimization

It is important to have an optimal structure (lowest energy) or stable conformer of a molecule before precede calculating the properties and studying the behaviours based upon the molecular structure. Because when a molecule is a model or design using a modelling program, the starting structure or initial geometry does not always correspond to stable conformation. Thus, geometry optimization or more commonly referred to as energy minimization, energy optimization, and geometry minimization procedures are required to determine the equilibrium geometry structure. An optimization or minimization involves several changes in the atom position through rotation, chemical bond stretching and balancing of the internal atomic force of each atom. This procedure will be iterated and repeated many times until the molecule energy reaches a minimum or stationary point of the potential energy surface (PES) (Lipkowitz & Boyd, 2009). The minimum energy of a molecule in a stable conformation will then be calculated and accomplished by either molecular mechanics (MM) or quantum mechanics (QM) methods with the applying of force field.

A potential energy surface (PES) usually used to describe the energy of a molecule with the changes in its geometry. In principle, different minima or maxima energy implies to different conformation structure of the molecule. A minimum or maximum is

defined as local if the energy located at the lowest or highest in the particular region of the PES, whereas, global minimum or maximum determine as the lowest or highest point in the entire PES, as shown in Fig. 3.1. In PES, every single atom is given by a position called vector, (r) which in turn can be represented as the Cartesian Coordinates and the energy as a function of the atom position is determined as $E(r)$. Due to the gradient is the negative of the net force and the net force is also zero at the stationary point, the first derivative of the energy or gradient, $E'(r_0)$ with respect to the r is zero. Similarly, for the second derivative, $E''(r_0)$ or also known as the Hessian matrix, which specifies the curvature of the PES at r has positive definite with all positive eigenvalues.

$$E(r) = E(r_0) + E'(r_0)(r - r_0) + \frac{1}{2}E''(r_0)(r - r_0)^2 \quad (3.1)$$

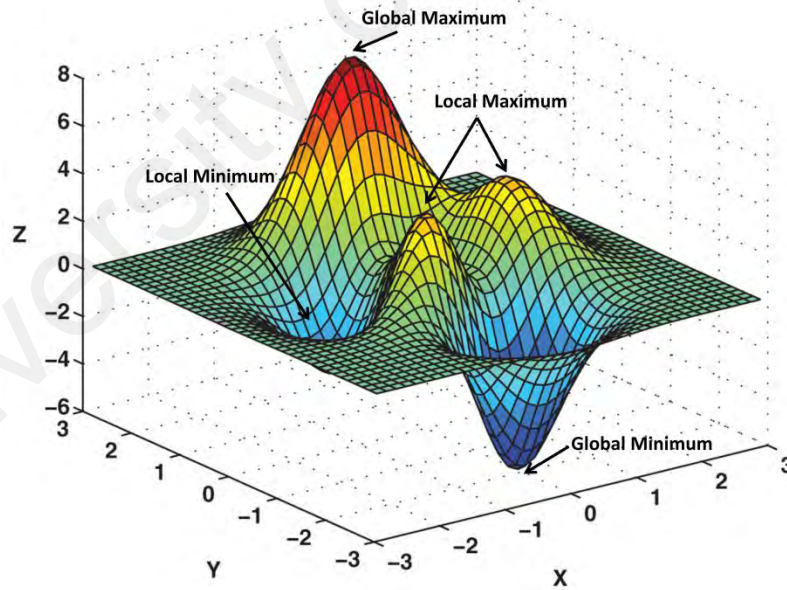


Figure 3.1: A 3D potential energy surface (PES) map.

3.2.2 Molecular mechanics (MM)

Molecular mechanics (MM) use the classical mechanics along with numerous classical equations of motion to predict the structure and properties of a molecule. It works by considering the force field acting on each atom and minimizing the strain energy of the structure. The MM calculation is based on several assumptions. The first of these is all the electrons are not explicitly included and the energy of a molecule is calculated based upon the nuclear position. Secondly, every single atom is treated as a particle and each particle is assigned with a van der Waals radius, net charge, and polarizability which are determined from experimental and theoretical. In addition, the bond interaction between the atoms is based on the spring model and other classical potentials (Atkins & Friedman, 2011). The total potential energy (E_{total}) as a sum of bond stretching energy (E_{str}), angle bending energy (E_{bend}), torsional rotation energy (E_{tor}), lastly the energy from the non-bonded atom (E_{nb}).

$$E_{total} = E_{str} + E_{bend} + E_{tor} + E_{nb} \quad (3.2)$$

$$E_{str} = \sum \frac{1}{2} k_b (r - r_0)^2 \quad (3.3)$$

$$E_{bend} = \sum \frac{1}{2} k_\theta (\theta - \theta_0)^2 \quad (3.4)$$

$$E_{tor} = \sum A \{1 + \cos(n\tau - \phi)\} \quad (3.5)$$

$$E_{nb} = \sum_{i>j} \left\{ -\frac{c_{ij}}{r_{ij}^6} + \frac{D_{ij}}{r_{ij}^{12}} + \frac{q_i q_j J_0}{\epsilon_r r_{ij}} \right\} \quad (3.6)$$

where k_b is the empirical force constant, k_θ is stiffness of the angular spring, r_0 for equilibrium bond length, θ_0 is the equilibrium angle, A is amplitude, n is the periodicity factor, ϕ the displacement in the rotation angle τ . The first and second terms of the non-

bonded atom energy E_{nb} equation 3.6, represent van der Waals and repulsion interaction derive by empirical coefficient C_{ij} and D_{ij} , where each non-bonded pair such as C and H, C and C, C and O are assigned together. The last term serves as the Coulombic interaction of particles of charges $q_i e$ and $q_j e$ in an environment of local relative permittivity ϵ_r .

MM can be very useful to perform calculations on molecular systems ranging in size and complexity from small to large biological systems such as protein and nucleic acids or materials assemblies with containing thousands to millions of atoms. Certainly, the MM method requires less computational time than quantum mechanics (QM) and their computational requirement don't affect by the molecular size. Nonetheless, limitations of MM undoubtedly still exist as it can't readily study properties dependent on electron distribution and not applicable for bond breaking transition states.

3.2.3 Quantum mechanics (QM)

In 1926, Erwin Schrödinger and Werner Heisenberg formulated the wave-particle theory (Bohr model) proposed by Louis de Broglie (Schrödinger, 1926). In doing so, the modern understanding of the electron behaviours within atom based on wavelike behaviours which are developed by Schrödinger is commonly known as quantum mechanics (QM) or wave mechanical model. QM model used the Schrödinger equation to explain the motion of electron on the scale of atoms and molecules. Indeed, the Schrödinger equation can be useful to describe a moving electron as a matter wave. In the Bohr model, the electron moves in well-defined circular orbits around the central nucleus of the atom. Unlike the Bohr model, the QM model electron acts like a fuzzy cloud around a positively charged nucleus, called atomic orbitals (Kumar, 2018). The Schrödinger wave equation for a single particle is shown below.

$$i\hbar \frac{\partial}{\partial t} \psi(x,t) = \frac{-\hbar^2}{2m} \nabla^2 \psi(x,t) + V(x,t) \psi(x,t) \quad (3.7)$$

where i is the imaginary number $\sqrt{-1}$, \hbar is the Planck's constant divided by 2π , $\psi(x, t)$ is the wavefunction, m is the mass of the particle, $V(x, t)$ is the potential energy influencing the particle.

In equation 3.8, the electron describes as a wave that separated both displacement and time. But, in most wave behaviours studies, we are mostly interested in the probability density of finding an electron in a particular region within an atom. Therefore, the probability of one-dimensional system can be defined as,

$$\int \psi(x,t)^* \psi(x,t) dx dt \quad (3.8)$$

where $\psi(x, t)$ is the wavefunction and $dxdt$ represent the density at time t for finding the electron at position x .

In the QM model, the electron of an atom is represented in three-dimensional space (3D). Thus, three quantum numbers (principle, angular momentum, magnetic quantum number) which corresponding to three coordinates are required to specify the position of the electron in an atom. The principle quantum number (n) may describe the energy of the particular electron and the average distance of an electron from the nucleus. It is always assigned with positive integer values ($n=1, 2, 3, 4, \dots$). As n increases, more likely the electron energy and distance of an electron from the nucleus is also increased. The angular momentum quantum number (ℓ) could be used to determine the shape of the orbital and the shape depends on the value of n ($\ell=0, 1, 2, 3, \dots, n-1$). For magnetic quantum number (m_ℓ), it contains information about the position of electron and their behaviour under the influence of a magnetic field. The m_ℓ depend upon the value of ℓ ($m_\ell = -\ell, -\ell+1, \dots, 0, \dots, \ell-1$) (Griffiths, 2016).

3.2.4 Semi-empirical

Semi-empirical is a very cost-effective tool when it comes to the properties studies such as structural, stability, spectroscopy as well as the chemical reaction of a molecule. Semi-empirical is based on the Hartree-Fock and Kohn-Sham density functional theories (DFT), but in many parameters, approximations, or generalizations design including dipole moments of molecules, ionization energies of atoms, heats of formation and geometries are obtained from empirical parameters. It simplified the Hartree-Fock procedure by introducing distinct approximations to the Hamiltonian, parameterizing or omitting core electron to speed up the calculations. In order to correct for these loss caused by the approximations, the empirical function with parameters is incorporated into the formalism and calibrated with a large number of accurate experimental or high-level theoretical reference data to produce results that best possible agreement with experimental data (Thiel, 2005). Although, semi-empirical are less accurate than higher level calculations, but they are typically much faster than ab initio calculations, especially for large molecular systems and may give accurate results when applied to molecules that are similar to the molecules used for parameterization.

Traditional semi-empirical models are based on the simpler integral schemes of complete neglect of differential overlap (CNDO) and intermediate neglect of differential overlap (INDO). In contrast, modern models such as modified neglect of diatomic overlap (MNDO), austin model 1 (AM1) and parameterized model 3 (PM3) are all based on the neglect of differential diatomic overlap (NDDO) integral approximation where the unit matrix been replaced with the overlap matrix S . This enables the replacement of the Hartree-Fock secular equation $|\mathbf{H}-\mathbf{E}\mathbf{S}| = 0$ with a simpler equation $|\mathbf{H}-\mathbf{E}| = 0$. Among all the methods, NDDO currently retains the most popular and frequently used approach due to it can sustain the higher multipoles of charge distributions in the two-centre interaction. By contrast, CNDO and INDO shorten charge distribution after

the monopole. Another similar method which also reduces and neglects the two-electron integrals is the zero differential overlap (ZDO) approximations. In the ZDO approximations, all the two-electron integrals involving two-centre charge distributions are ignored. In particular, all the product of atomic orbitals (AOs) functions $\chi_u \chi_v$ is set to zero and the overlap integral $S_{uv} = \delta_{uv}$.

3.2.4.1 Complete neglect of differential overlap (CNDO)

In the complete neglect of differential overlap (CNDO) approach, all the integrals are neglected except for the Coulomb one-centre and two-centre two-electron integrals. However, when two different AOs χ_u and χ_v belong to the same atom, the two-electron integral will be set to zero. The one-central integrals approximation for the CNDO approach remains the same as the INDO. The approximations and parameterization for the CNDO method were based on the Pariser-Pople-Parr (PPP) method, where only π -orbitals are treated and the sigma effects (σ) are neglected.

One-electron integrals are determined as

$$\begin{aligned} \langle \mu_A | h | \mu_A \rangle &= \left\langle \mu_A \left| -\frac{1}{2} \nabla^2 - V_A \right| \mu_A \right\rangle - \sum_{a \neq A}^{N_{\text{nuclei}}} \langle \mu_A | V_a | \mu_A \rangle \\ \langle \mu_A | h | \nu_A \rangle &= -\delta_{\mu\nu} \sum_{a \neq A}^{N_{\text{nuclei}}} \langle \mu_A | V_a | \mu_A \rangle \end{aligned} \quad (3.9)$$

where μ and ν refer to two atomic orbitals (AOs) ϕ_u and ϕ_v . While h is the one-electron part of the Fock matrix.

and the two-electron integrals can be written as

$$\langle \mu_A \nu_B | \lambda_C \sigma_D \rangle = \delta_{AC} \delta_{BD} \delta_{\mu\lambda} \delta_{\nu\sigma} \langle \mu_A \nu_B | \mu_A \nu_B \rangle \quad (3.10)$$

where μ and ν are two AOs on one atom, and λ and σ are AOs on the different atom.

3.2.4.2 Intermediate neglect of differential overlap (INDO)

The intermediate neglect of differential overlap (INDO) method contains similar characteristics as the CNDO method, where treats the valence electrons explicitly. Unlike the CNDO method, the differential overlap between AOs on the same atom is retained in the INDO method, if χ_u and χ_v belong to the same atom. But the entire two-centre two-electron integrals that are not from the Coulomb type are still neglected. Additionally one-centre exchange integrals are included in the INDO method, these calculations are very useful for helping to explain the splitting between the electronic states that come from the same configuration. Although both of the INDO and CNDO methods can provide a good measurement in bond lengths, bond angles, and electronic spectra. However, the prediction for dipole moment and energy are still poor.

3.2.4.3 Zero differential overlap (ZDO)

Most of the semi-empirical theories are based upon the zero differential overlap (ZDO) approximation. This approximation greatly simplifies the many electrons by neglecting certain integrals, usually two-electron repulsion integrals. The overlap between pairs of different orbitals is set to zero for all volume elements $d\nu$.

$$\phi_\mu \phi_\nu d\nu = 0 \quad (3.11)$$

and the overlap integrals can be written as

$$S_{\mu\nu} = \delta_{\mu\nu} \quad (3.12)$$

When the two different atomic orbitals (AOs) ϕ_u and ϕ_v are located on different atoms then the differential overlap is referred to as diatomic differential overlap. In contrast, if ϕ_u and ϕ_v are on the same atom then it becomes monoatomic differential overlap. When the ZDO approximation is applied to the two-electron integrals

$(\mu\nu | \lambda\sigma)$ then the integral will be equal to zero if $\mu \neq \nu$ or $\lambda \neq \sigma$. This can be written concisely using the Kronecker delta.

$$(\mu\nu | \lambda\sigma) = (\mu\mu | \lambda\lambda) \delta_{\mu\nu} \delta_{\lambda\sigma} \quad (3.13)$$

When the molecular orbitals ϕ_v are extended in terms of N basis function, χ_μ^A , it can be defined as.

$$\phi_v = \sum_{\mu=1}^N c_{\mu i} \chi_\mu^A \quad (3.14)$$

where $c_{\mu i}$ is defined as the coefficient and A is the atom of basis function.

3.2.4.4 Neglect of diatomic differential overlap (NDDO)

Many modern semi-empirical molecular orbital models such as MNDO, AM1, and PM3 are based on the neglect of diatomic differential overlap (NDDO) approximation, where the diatomic differential overlap would not be retained. In NDDO formalism, the differential overlap will be neglected if the basis functions belong to the different atoms. Therefore, the NDDO approximation, proposed by Pople in 1970, retains all the one-centre exchange integrals, one-centre two-electron integrals as well as the two-centre two-electron integrals (Pople & Beveridge, 1970). The integrals approximations for the NDDO approach are more accurate and have more adjustable parameters compares to the ZDO approximation. The NDDO approximation can be described as equation 3.15 and 3.19, by using μ and ν to determine either as S or P orbitals (S_x , S_y , S_z).

The overlap integrals can be defined as

$$S_{\mu\nu} = \langle \mu | \nu \rangle = \delta_{\mu\nu} \delta_{AB} \quad (3.15)$$

One-electron integrals are written by

$$\langle \mu_A | h | \nu_A \rangle = \left\langle \mu_A \left| -\frac{1}{2} \nabla^2 - V_A \right| \nu_A \right\rangle - \sum_{a \neq A}^{N_{nuclei}} \langle \mu_A | V_a | \nu_A \rangle \quad (3.16)$$

$$\langle \mu_A | h | \nu_B \rangle = \left\langle \mu_A \left| -\frac{1}{2} \nabla^2 - V_A - V_B \right| \nu_B \right\rangle \quad (3.17)$$

$$\langle \mu_A | V_C | \nu_B \rangle = 0 \quad (3.18)$$

$$\left\langle \mu_A \left| -\frac{1}{2} \nabla^2 - V_A \right| \nu_A \right\rangle = \delta_{\mu\nu} \left\langle \mu_A \left| -\frac{1}{2} \nabla^2 - V_A \right| \mu_A \right\rangle \quad (3.19)$$

Two-electron integrals are defined as

$$\langle \mu_A \nu_B | \lambda_C \sigma_D \rangle = \delta_{AC} \delta_{BD} \langle \mu_A \nu_B | \lambda_A \sigma_B \rangle \quad (3.20)$$

3.2.4.5 Parameterized Model 3 (PM3)

The Parameterized Model 3 (PM3) which developed in 1989 by James Stewart is the third parameterization of MNDO and the AM1 being considered as a second (Stewart, 1989). Both AM1 and PM3 implement the same NDDO approximation. In fact, PM3 contains the essential Hamiltonian similar to the AM1, but the parameterization strategy is different. For AM1, the parameters are largely based on a small number of atomic data and obtained by applying chemical knowledge. Whereas PM3 is develop based on a large number of molecular properties and all the quantities that enter the Fock matrix have been treated as pure parameters. A significant difference in the parameterization and nuclear repulsion treatment allow PM3 to assess hydrogen bond slightly better than AM1. In addition, PM3 is no doubt has more accuracy of thermochemical predictions than AM1 and widely used for the rapid estimation of molecular properties and yields the recently upgraded version allows the inclusion of transition metals.

3.2.5 Hartree-Fock Equation

In order to solve the many-body problems, Hartree introduced an effective method to approximate the electron-electron interactions through the one-electron model in which each electron move independently in the average density. He created a self-consistent field method to solve the Schrödinger equation that results from the time-independent Schrödinger equation after invoking the Born-Oppenheimer approximation. It is based on a simple approximation where the N -body wavefunction of the system can be approximated using a single Slater determinant or by a single permanent with one spin-orbital per electron (Hartree, 2008). The many-body wavefunction can be expanded as a product of one-electron wavefunctions.

$$\psi(r_1, r_2, \dots, r_n) = \psi(r_1)\psi(r_2)\dots\psi(r_n) \quad (3.21)$$

where the non-interacting electrons involving the interaction with the mean-field Coulomb potential coming from the other nuclei, electrons and external fields, yielding one-electron Schrödinger equation

$$E_i \psi_i(r) = \left(-\frac{1}{2} \nabla_i^2 + U_{en}(r) + U_H(r) \right) \psi_i(r) \quad (3.22)$$

$$U_{en}(r) = -\sum_{I=1}^N \frac{Z}{|r - R_I|} \quad (3.23)$$

where U_{en} is the attractive potential between electrons and nuclei.

$$U_H(r) = \int \frac{\rho(r')}{|r - r'|} dr' \quad (3.24)$$

and the U_H is the Hartree potential obtain from the classical Coulomb repulsive interaction between each electron and mean field arising from the $N-1$ other electrons.

$$\rho(r') = \sum_i^N |\psi_i(r)|^2 \quad (3.25)$$

where $\rho(r')$ is the charge density at r' , determined from the summation of the probability density associated with electron i

3.2.6 Kohn-Sham Equation

In QM calculation, specifically density function theory (DFT), attempts in calculates the total energy of an interacting electron as a function of the electron density. However, by adopting the electron density without including the wavefunctions were not very correct due to lack of accurate approximations of the kinetic energy functional. In order to overcome this issue, Kohn and Sham introduced an alternative approach for finding the exact ground state energy with a standard non-interacting system of electrons (Kohn & Sham, 1965). The Kohn-Sham equation basically builds on directly solving the one electron Schrödinger equation by involving a fictitious non-interacting particles system. This fictitious non-interacting system, which is intended in some way to mimics the interacting N -electron system, is known as the Kohn-Sham (KS) system. The Kohn-Sham ground state energy in the term of electron density can be written as.

$$E[\rho(r)] = T[\rho(r)] + \int \rho(r)v(r)dr + E_{ee} \quad (3.26)$$

The equation 3.26 can be separated into three parts, where the first term involved the kinetic energy, second is the electron-nucleic interaction energy and the final term which included the electron-electron interaction energy. For single particles wavefunctions (ψ) or also known as Kohn-Sham orbitals are related to the electron density of N electrons.

$$\begin{aligned}\rho(r) &= \sum_i^N |\psi_i(r)|^2 \\ &= N \int |\psi|^2 dr\end{aligned}\quad (3.27)$$

While in multiple particles, the Kohn-Sham orbitals electron density is revealed as the product of one particle densities.

$$\rho(r, r') = \rho(r)\rho(r') + \Delta\rho(r, r') \quad (3.28)$$

The kinetic energy is given by

$$T = -\frac{1}{2} \sum_i^N \int \phi_i^*(r) \nabla^2 \phi_i(r) dr + \Delta T \quad (3.29)$$

And the electron-electron Coulomb repulsion which can be expressed as

$$E_{ee} = \frac{1}{2} \iint \frac{\rho(r)\rho(r')}{|r-r'|} dr dr' + \Delta E_{ij} \quad (3.30)$$

By combining all these equations, the ground state energy (E) of an electron (N) can be defined as

$$E = -\frac{1}{2} \sum_i^N \int \phi_i^*(r) \nabla^2 \phi_i(r) dr + \int U_{ext} \rho(r) dr + \frac{1}{2} \iint \frac{\rho(r)\rho(r')}{|r-r'|} dr dr' + \Delta T + \Delta E_{ij} \quad (3.31)$$

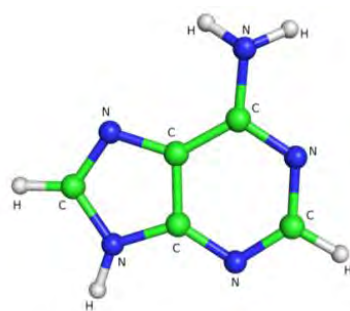
3.3 Implementation method

Our implementation methodology can be classified into three main categories, the first describing the process of molecular structure preparation, followed by the physical and structural properties calculation. The remainder of the section is devoted to the molecular simulation, where the process involves the application of an external electric field.

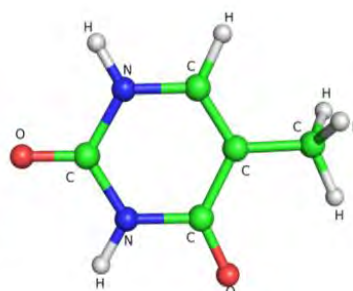
We have used various level of computational method, ranging from molecular mechanics (MM) to quantum mechanics (QM). Both DNA and RNA nucleobases and their base pairs were initially optimized by using first MM, with BIOCHARM force field and then the optimized structures were again optimized with QM using the PM3 method. All the optimization calculations were performed using Hyperchem 7.5 software (Hyperchem, 2002).

3.3.1 Molecular structures preparation

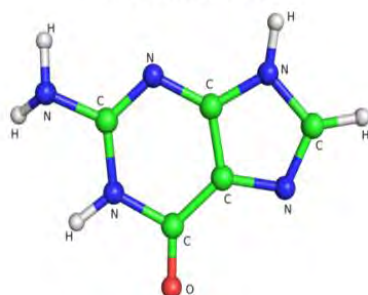
The initial three-dimensional (3D) molecular structures of DNA and RNA nucleobases (A, T, G, C, U) and their base pairs (G-C, A-T, A-U) were modelled using commercial software package Hyperchem7.5 (Hyperchem, 2002). After that, the modelled molecular structures were then further optimized using molecular mechanics (MM) specifically BIOCHARM with 2000 cycles to obtain a stable and minimized structure. The optimized molecular structures of nucleobases (A, T, G, C, U) and their base pairs (G-C, A-T, and A-U) are depicted in Fig. 3.2 and 3.3.



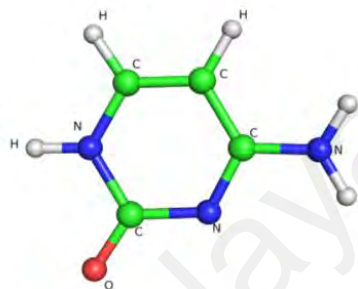
Adenine (A)



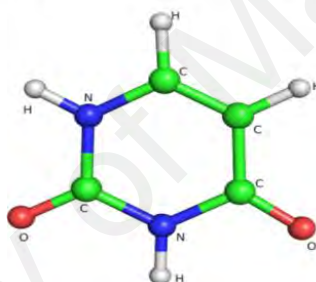
Thymine (T)



Guanine (G)

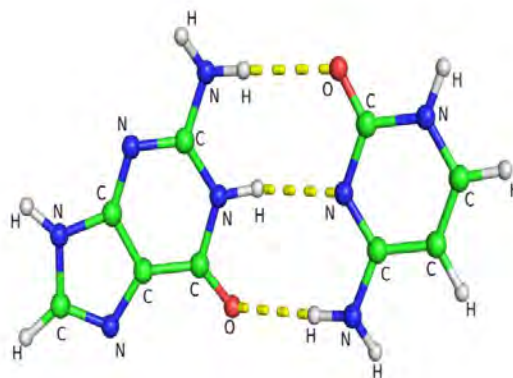


Cytosine (C)

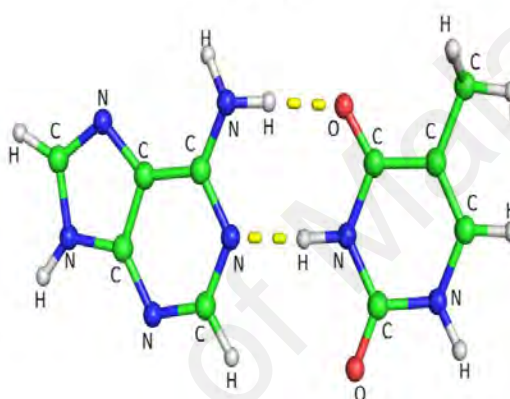


Uracil (U)

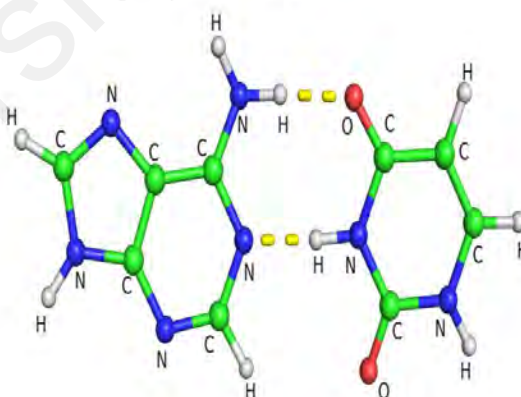
Figure 3.2: Chemical structure of DNA and RNA nucleobases. The nitrogen atom is in blue colour, oxygen in red, hydrogen in white, and carbon in green.



Guanine-Cytosine
G-C



Adenine-Thymine
A-T



Adenine-Uracil
A-U

Figure 3.3: Chemical structure of DNA and RNA base pairs. The nitrogen atom is in blue colour, oxygen in red, hydrogen in white, carbon in green and the yellow dotted line represent the hydrogen bonds.

3.3.2 Physical and structural properties calculation

The physicochemical properties such as a heavy atom, molecular weight, van der Waals surface volume, and polarizability were calculated by QSAR properties module from Hyperchem7.5. Calculation of the van der Waals surface volume is carried out based on the atomic radii gavezotti method derived by Bodor et al. (Bodor et al., 1989). The scheme given by Miller with 3% in precision was used to predict the polarizability with different increments is associated with different atom types (Miller, 1990). Other physicochemical parameters including topological polar surface area (TPSA) and total surface area (TSA) were determined using the fragment-based method by Ertl et al. (Ertl et al., 2000) through the online tool Molinspiration Cheminformatics server (<http://www.molinspiration.com>). The visualization map of TPSA was viewed and constructed in the Molinspiration Galaxy 3D Generator.

3.3.3 Applied an external electric field in molecular simulation

By applying computational molecular modelling and simulation methods, various principal electrical and physical properties including dipole moment, polarization, total energy, atomic charges, and bond lengths, as well as bond angles, were calculated in both absences and under an applied external electric field. An external electric field was introduced in the *Y*-axis direction with the strengths between 0.00 to 0.011 atomic units (a.u.) (1 a.u. \sim 514V/nm). Geometry optimizations have been computed with a first principle quantum mechanics (QM), semi-empirical method with parametrization number 3 (PM3) approach to obtain the minimum of the total energy or potential energy surface (PES). These molecular geometry calculations were executed using the Polak-Ribere conjugate gradient algorithm with energy relative convergence limit up to 0.001 accuracy. Generally, PM3 which is developed by Stewart, (Stewart, 1989) is the most accurate and quickest method based on the neglecting diatomic differential overlap (NDDO) approximation with utilizes several experimental parameters. It employs

empirical parameter sets for calculating ground-state electronic properties of open and closed-shell molecule, total charge distribution and polar properties for biological materials such as DNA and RNA nucleobases which consist of carbon, oxygen, nitrogen and hydrogen atoms, which is readily available for the purpose of present calculation. The interaction energy or repulsion core between the nuclei A and the charges B including nuclear charges and electronic charges in the PM3 method were indicated in the equation (Stewart, 2007). The overview of the strategic planning process and implementation method are depicted in Fig. 3.4.

$$\Delta E_N(A, B) = \sum_{(A, B)} \left\{ Z_A Z_B (S_A S_A | S_B S_B) \left[1 + e^{-\alpha_A R_{AB}} + \frac{e^{-\alpha_A R_{AB}}}{R_{AB}} \right] - Z_A Q_B (S_A S_A | S_B S_B) \right\} \quad (3.32)$$

$$(S_A S_A | S_B S_B) = \frac{1}{\left[R_{AB}^2 + \frac{1}{2} \left(\frac{1}{AM_A} + \frac{1}{AM_B} \right)^2 \right]^{1/2}} \quad (3.33)$$

where $Z_A Z_B$ are the nuclear charges, R_{AB} is the distance between them, Q_B is the net charge on the atom B and AM are the monopole interaction parameters.

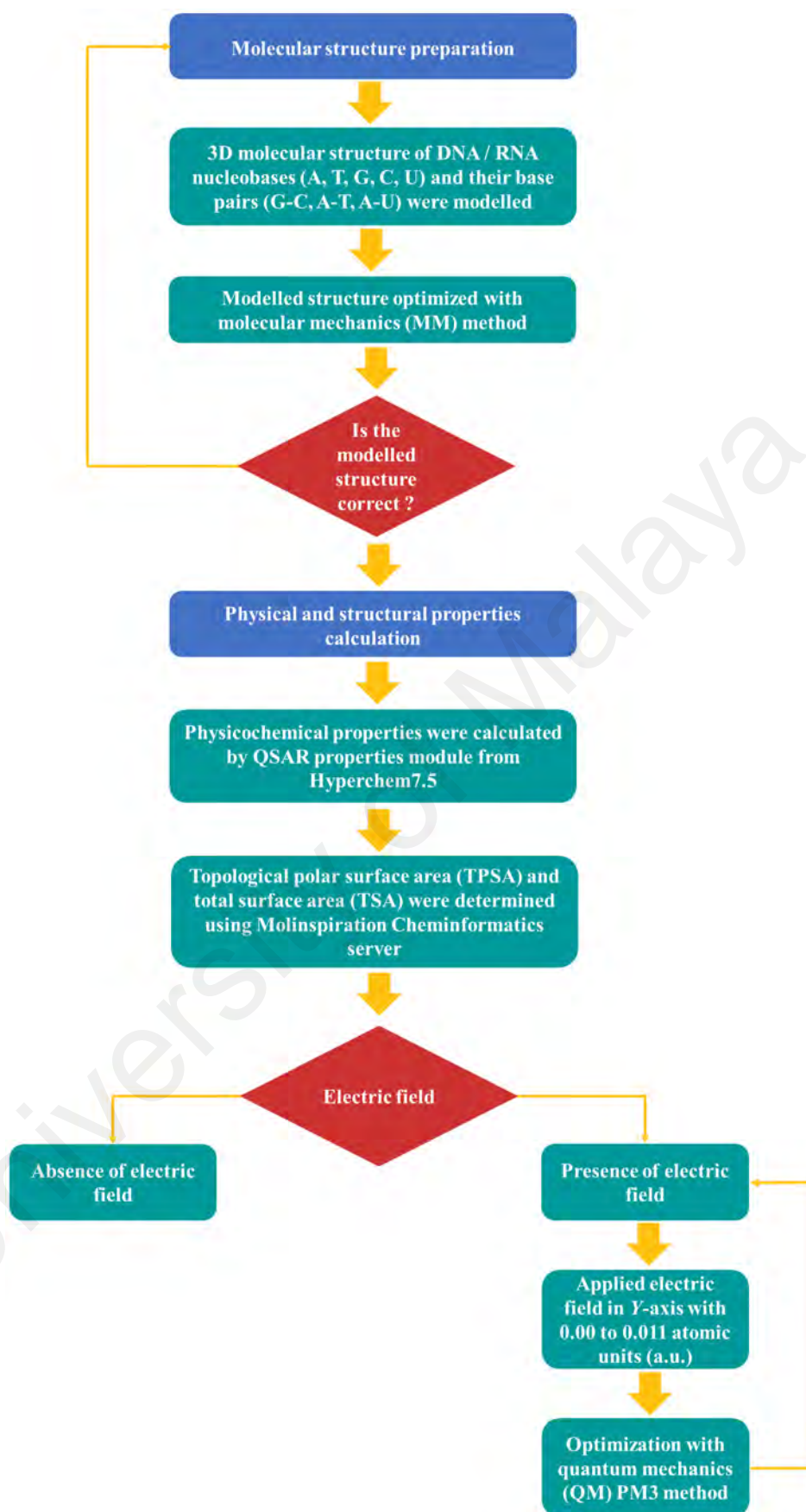


Figure 3.4: Implementation process flow chart.

CHAPTER 4: DNA AND RNA NUCLEOBASES

4.1 Introduction

In this chapter, we discuss the systematic computational molecular modelling performed on the study of polarization switching and hysteresis loop behaviours of all the five nucleobases. The calculation results will be divided into two sections. We will first discuss the physical and structural properties of the nucleobases in the absence of an electric field, followed by the effect of an applied electric field.

4.2 Results and discussion

4.2.1 Physical and structural properties

We first examined the molecular structures of the nucleobases under the absence of an applied electric field. In the calculations, the initial stage of the nucleobases is set based on the coordinate system, as shown in Fig. 4.1. In Fig. 4.1, the chemical structures of the nucleobases (A, T, G, C, U) as well as the magnitude of dipole moment are shown. The two purine nucleobases, A and G consist of the double-ringed structure formed by a five-membered ring attached to a six-membered ring. The T, C, and U tend to be smaller in size, as pyrimidines contain only a single heterocyclic six-membered ring structure, which made up of two nitrogen and four carbon atoms. The nitrogen atom at position 1 in both purine and pyrimidine is covalently bonded to the carbon at position 1 of the deoxyribose or ribose.

Four physicochemical properties and the chemical formula, number of atoms, and heavy atom associated with the nucleobases are listed in Table 4.1. It is observed that the polarizability is proportional to the molecular weight and volume (see Table 4.1). The increasing order of polarizability for the nucleobases is $U > C > T > A > G$ correlated with the increasing molecular volume but it is slightly different from the molecular weight.

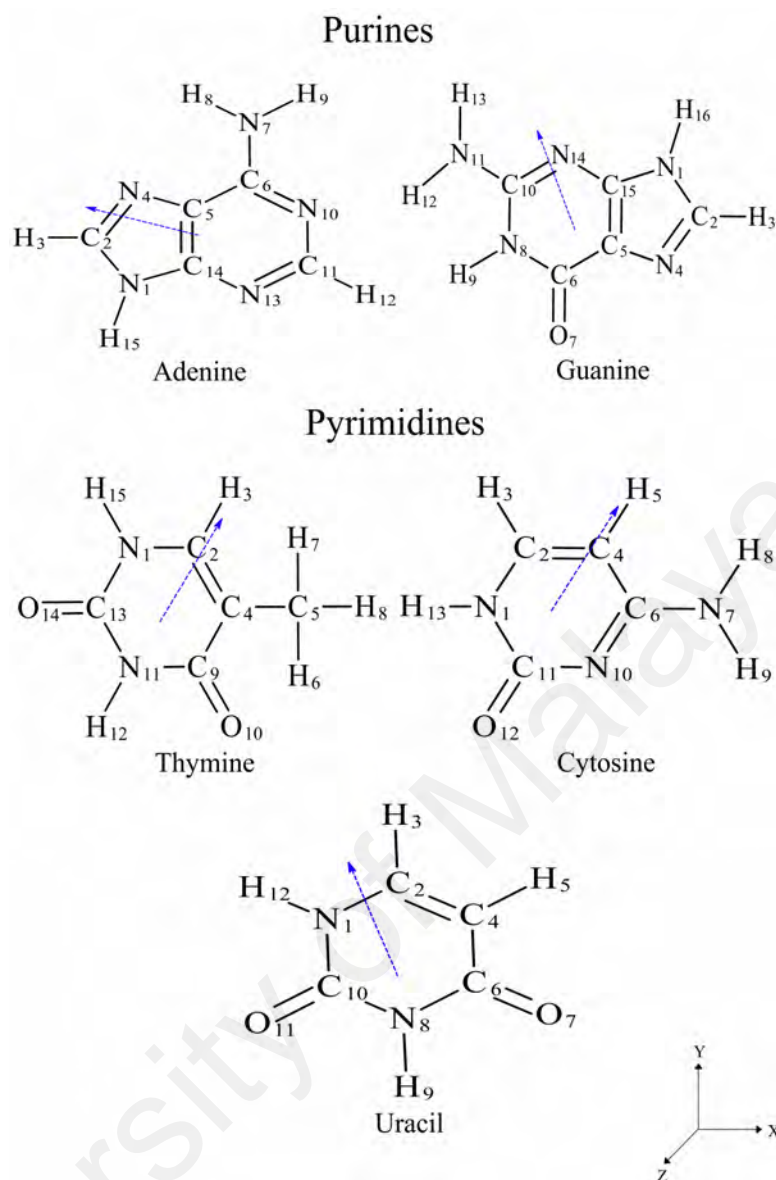


Figure 4.1: Structural formula and atom numbering for purines and pyrimidines of DNA and RNA nucleobases. Blue arrows indicate the magnitude of dipole moments. An electric field is applied along the Y-direction (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).

For example, guanine has the largest values of polarizability (14.22 \AA^3), volume (118.77 \AA^3) and molecular weight (159.19 g/mol). In contrast, uracil is the smallest molecule in the series, which has a small value of polarizability (10.02 \AA^3) and volume (88.53 \AA^3). In the present study, we consider carbon (C), nitrogen (N) and oxygen (O) as heavy atoms. The purine nucleobases A and G possess of more nitrogen atoms than

their complement pyrimidines nucleobases T, C, and U. On the other hand, both pyrimidines nucleobases T and U possess the highest number of the oxygen atom. The main difference between A and G is that G contains an extra oxygen atom. Additionally, G has the largest whereas both C and U have the smallest number of heavy atoms count.

Table 4.1: Physicochemical properties of DNA and RNA nucleobases (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).

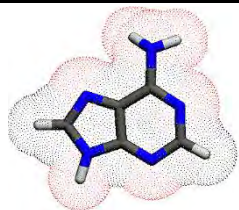
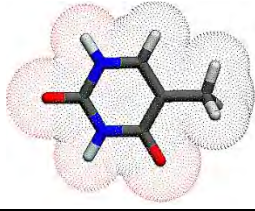
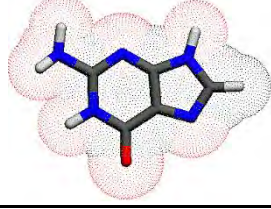
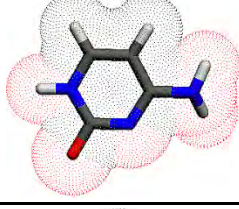
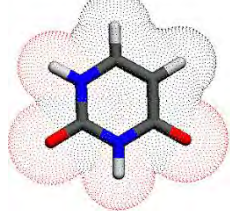
Nucleobases	Chemical Formula	Number of Atoms	Number of heavy atoms	Molecular Weight [g/mol]	Van der Waals surface volume [\AA^3]	Polarizability [\AA^3]
Adenine	$\text{C}_5\text{H}_5\text{N}_5$	15	10	143.19	110.42	13.71
Thymine	$\text{C}_5\text{H}_6\text{N}_2\text{O}_2$	15	9	132.16	105.10	11.86
Guanine	$\text{C}_5\text{H}_5\text{N}_5\text{O}$	16	11	159.19	118.77	14.22
Cytosine	$\text{C}_4\text{H}_5\text{N}_3\text{O}$	13	8	117.15	93.32	10.87
Uracil	$\text{C}_4\text{H}_4\text{N}_2\text{O}_2$	12	8	118.14	88.53	10.02

In Table 4.2, we summarized the dipole moments (without an applied electric field) of the nucleobases (A, T, G, C, and U). In general, the calculated dipole moments of nucleobases based on the PM3 method agrees well to those obtained by Tasi et al. as well as those reported by Bdikin et al. (Bdikin et al., 2015; Tasi et al., 1993). The values of dipole moments from the highest to lowest is $C > G > U > T > A$. Generally, this is primarily due to the difference in the electronegativity of the bonded atoms and other factors such as the distance between the charge separations. In other words, the greater the electronegativity difference, the greater the dipole moments. The TPSA and the TPSA/TSA ratio of the nucleobases are given in Table 4.3. Among the 5 nucleobases, G has the largest TPSA and TSA, as well as the TPSA/TSA ratio. Both T and U possess small TPSAs, and the U has the smallest TSA of 131.02\AA^2 . The value of the TPSA/TSA ratio in the order from high to low is $G > C > A > U > T$.

Table 4.2: Calculated dipole moments of DNA and RNA nucleobases (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).

Nucleobases	Dipole Moment [Debye]		
	Present work	Tasi et al.	Bdikin et al.
Adenine	2.492	2.49	2.495
Thymine	3.989	3.88	4.51
Guanine	5.443	5.45	5.45
Cytosine	6.069	5.68	6.074
Uracil	3.990	3.90	Not available

Table 4.3: Calculated topological polar surface area (TPSA), total surface area (TSA), and TPSA/TSA ratio for DNA and RNA nucleobases. Red and black colours represent the polar and non-polar surface areas, respectively (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).

Nucleobases	TPSA [\AA^2]	TSA [\AA^2]	TPSA/TSA $\times 100\%$	Visualization of TPSA
Adenine	80.50	165.64	48.60	
Thymine	58.20	161.75	35.98	
Guanine	96.20	174.74	55.05	
Cytosine	67.50	131.02	51.52	
Uracil	58.20	130.23	44.69	

4.2.2 Electric field effects on electrical, structural and physical properties

We first investigated and discussed the effects of an applied electric field (E) on the polarization behaviours of the cytosine. The dependence of polarization (P) and total energy (U) on an applied electric field (E) of cytosine nucleobase are presented in Fig. 4.2 (a) and (b), respectively. The electric field is applied along the direction in the Y -axis of the cytosine molecule, which is illustrated as in Fig. 4.2 (a). The applied field is varied between -6 and $+6$ V/nm. In general, the shape of the polarization (P) versus electric field (E) hysteresis loop is square and symmetry (see Fig. 4.2 (a)). If the applied electric field $E = 0$, the polarization (P) is $\sim \pm 0.200$ C/m². Switching occurs at the coercive field $E_c = \pm 0.006$ a.u. which is equivalent to 3.085 V/nm (1 a.u. ~ 514 V/nm). Going from a low to a high negative electric field, the point I ($E = 0$) corresponds to a positive polarization state with $P = 0.202$ C/m² and the corresponding total energy $U = -30701.113$ kcal/mol. As the applied electric field increases (at point II), $E < E_{c-}$, the magnitude of the polarization (P) decreases while the total energy (U) increases. Upon further increasing $E > E_{c-}$ (at point III), the molecule switches from a positive to a negative polarization state. The reversal of polarization is accompanied by an abrupt decrease of U at $E \sim -3.085 = E_{c-}$. Similarly, the point IV ($E = 0$) corresponds to a negative polarization state with $P = -0.201$ C/m² and $U = -30701.117$ kcal/mol. The cytosine molecule reverses from a negative to a positive state (from V to VI) at $E \sim 3.085 = E_{c+}$.

Insets in Fig. 4.2 (a) show the molecular structure of cytosine under a particular electric field. It is seen that the cytosine molecule rotates from a positive (point II) to a negative (point III) state and from a negative (point V) to a positive (point VI) state at $E = E_c$. Hysteresis is not only found in the electric field dependence of polarization but also in the dependence of the total energy (U) on an applied electric field with a

butterfly-like loop behaviour as illustrated in Fig. 4.2 (b). From the butterfly-like $U-E$ loop, two small jumps can be found at $E = E_c$.

In Fig. 4.3, we show the dependence of electron charges, bond lengths and bond angles of cytosine on an applied electric field. Note here that only atoms that are sensitive to an applied electric field are shown for the clarity of the discussion. In the following discussion, the label of the atom for the nucleobase is based on Fig. 4.1 unless otherwise specified. The net atomic charges of N₁, H₃, H₅, C₆, N₇, H₈, H₉, C₁₁, and H₁₃ are positive. However, the atomic charge for C₂, C₄, N₁₀, and O₁₂ are found to be negative. Among the atoms, the electron-rich oxygen O₁₂ atom exhibits the highest negative charge whereas the carbon C₁₁ atom carries the highest positive charge. These values are in good agreements with those reported by Santamaria et al. (Santamaria & Vázquez, 1994).

As observed from the TPSA visualization, atoms N₁, N₁₀, O₁₂, H₁₃ and -NH₂ amino group (N₇, H₈, and H₉) can be attributed as the TPSA polar atom. Therefore, these atoms are particularly sensitive to the applied electric field. As shown in Fig. 4.3 (a), O₁₂ is very sensitive to the electric field where an obvious change in the electron charge can be found at $E \sim E_c$. Similar trends can be seen in the electric-field dependence of bond lengths and angles in cytosine. Figure 4.3 (b) shows the bond length of the cytosine as a function of the electric field. It is clearly seen that the bond length of the partial-polar TPSA C₆-N₇, N₁₀-C₆, C₁₁-N₁, and C₁₁-O₁₂ bond length are sensitive to the applied electric field. The variation of bond length and angles are particularly obvious near $E \sim E_c$.

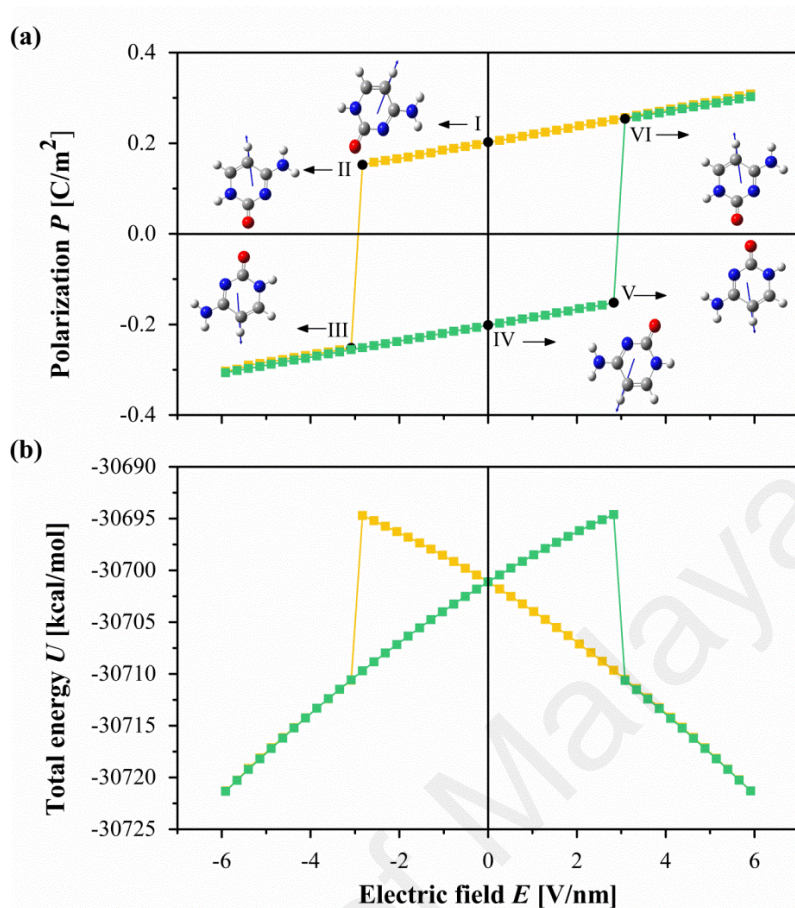


Figure 4.2: Electric field (E) dependence of (a) polarization (P) and (b) total energy (U) of cytosine. Insets in (a) show the molecular structure of cytosine at certain electric fields. The initial state of the molecule and the direction of an applied field is along the Y -axis, as shown in Fig 4.1 (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).

On the other hand, the bond length or bond angle of the non-polar TPSA atoms is almost independent of an applied electric field (see Fig. A1(b) in the appendix). Figure 4.3 (c) depicts the predicted bond angles variation of cytosine under an applied electric field. The bond angles for non-polar atoms $\text{H}_5\text{-C}_4\text{-C}_6$, $\text{H}_3\text{-C}_2\text{-C}_4$, and $\text{C}_2\text{-C}_4\text{-C}_6$ are fairly insensitive to the applied electric field, as compared with -NH_2 amino group $\text{H}_8\text{-N}_7\text{-H}_9$. A complete picture for the dependence of atomic charge distributions and geometrical parameters on an applied electric field for the cytosine can be found in Fig. A1 in the appendix. Figures A2–A5 in the appendix showed the results for adenine, thymine, guanine, and uracil, respectively.

Figure 4.4 shows the P - E hysteresis loops for five nucleobases (A, T, G, C, and U). The corresponding total energy U - E loops for the adenine, thymine, guanine, and uracil nucleobases are shown in Fig. 4.5. All the nucleobases are found to have butterfly-like loops, as expected. The coercive field and zero-field polarization of the nucleobases obtained from the hysteresis loop are summarized in Table 4.4. Our calculated molecular polarization (at $E = 0$) and coercive field of the nucleobases are all in the same order of magnitude as reported by Bystrov et al. (Bystrov et al., 2015). In general, thymine exhibits a broader hysteresis loop, and therefore, a higher coercive field. Among them, guanine and cytosine have large dipole moments. The magnitude of an average coercive field from high to low is $T > U > A > C > G$, whereas the average value of zero-field polarization from high to low is $C > G > U > T > A$.

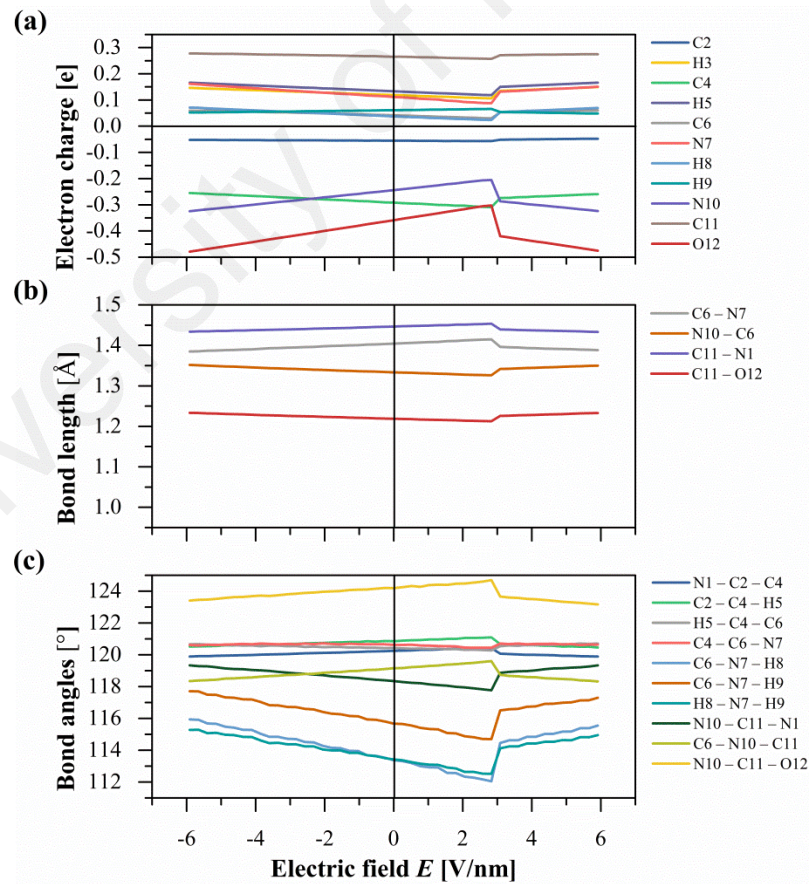


Figure 4.3: Dependence of (a) electron charges, (b) bond lengths and (c) bond angles of cytosine on an applied electric field (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).

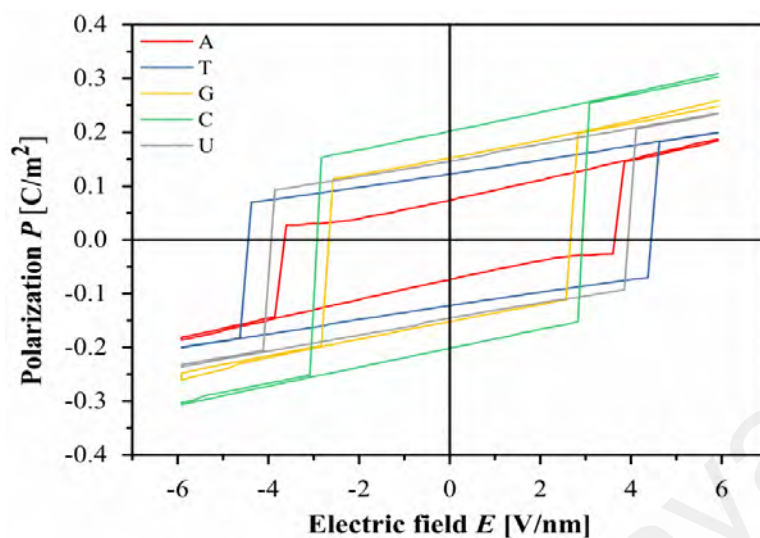


Figure 4.4: Electric field (E) dependence of polarization (P) for DNA and RNA nucleobases (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).

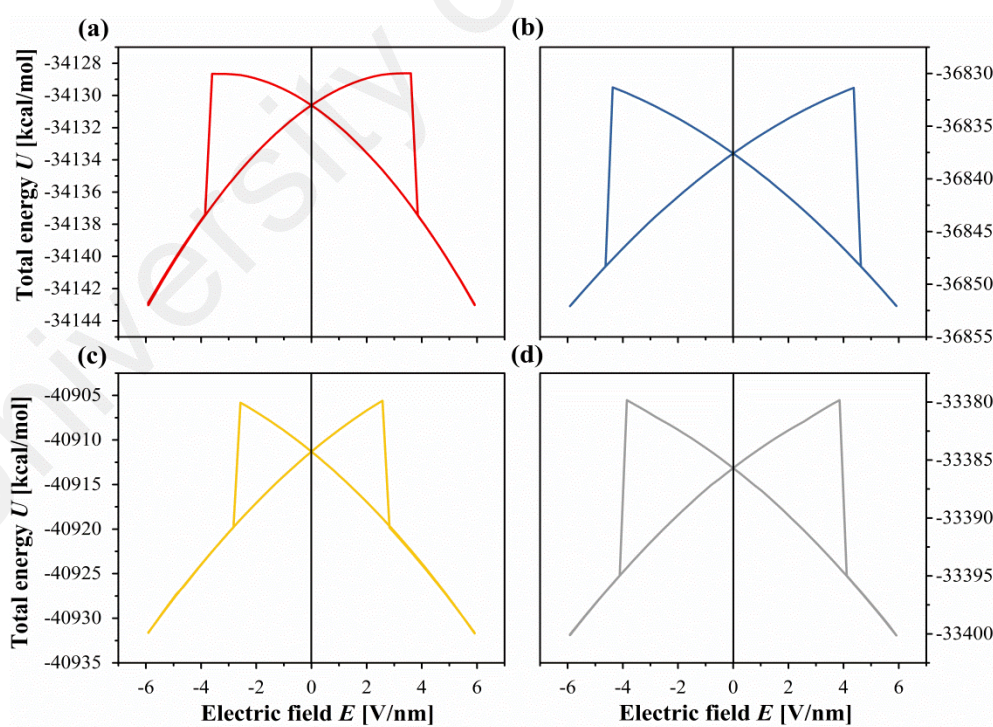


Figure 4.5: Electric field (E) dependence of total energy (U) for (a) adenine, (b) thymine, (c) guanine, and (d) uracil (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).

Table 4.4: Coercive field and zero-field polarization of DNA and RNA nucleobases. The values are obtained from the hysteresis loops Figs 4.2a and 4.4 (Reprinted with permission from European Physical Journal E) (Yam et al., 2018).

Nucleobases	Coercive Field [V/nm]	Zero-Field Polarization [C/m ²]
Adenine	3.857	0.073
Thymine	4.628	0.122
Guanine	2.828	0.152
Cytosine	3.085	0.202
Uracil	4.114	0.146

In order to further understand the hysteresis loop behaviour of the nucleobases, we analyse the relationship between the TPSA, polarization, and coercive field. TPSA is a very useful physiochemical descriptor that gives information about the polarity of the molecule. The calculation of this parameters is based on the summation of the surfaces of the polar atoms in a molecule, primarily oxygen, nitrogen and their attached hydrogen atoms bound to these electronegative atoms (Pranitha & Lakshmi, 2014). An interesting result implies the inverse correlation between the TPSA/TSA ratio and the coercive field of the nucleobases. In particular, a nucleobase with a higher number of TPSA/TSA ratios has a smaller coercive field (refer to Table 4.3). For example, guanine has a higher value of TPSA/TSA ratio compared to other nucleobases and resulting in a narrow coercive field. On the contrary, thymine and uracil with a lower number of TPSA/TSA ratios require a higher electric field to switch the molecules. The inverse relationship between the TPSA/TSA ratio and the coercive field can be interpreted as follows. Amino -NH₂ group and oxygen on the carbonyl are polar atoms in the nucleobases of DNA and RNA. A higher value of TPSA/TSA ratio indicates that the total surface area of a nucleobase that is sensitive to an applied electric field is larger. Therefore, the minimum electric field required to switch the nucleobases is smaller. The

analysis of other parameters such as molecular weight, charge, and size does not have an obvious relationship with the polarization and coercive field.

4.3 Summary

In this chapter, we have studied the polarization switching and hysteresis loop behaviour of the DNA and RNA nucleobases based on the PM3 method. Our studies revealed that all nucleobases exhibit zero-field polarization and square-shape hysteresis loops. The total energy (U) as a function of the electric field (E) also showed a butterfly-like hysteresis loop feature. The minimum field required for the polarization switching of a nucleobase is inversely proportional to the ratio of TPSA to TSA, though the zero-field polarization does not have an obvious correlation with the TPSA/TSA ratio.

CHAPTER 5: DNA AND RNA BASE PAIRS

5.1 Introduction

In the previous chapter, we focused on the physical and structural properties of the DNA and RNA nucleobases in the absence and present of an applied electric field. We then extended the study on DNA/RNA nucleobases to Watson and Crick base pairs. In this chapter, we report the investigation of the physical and structural properties of these base pairs as well as the polarization switching and hysteresis behaviours.

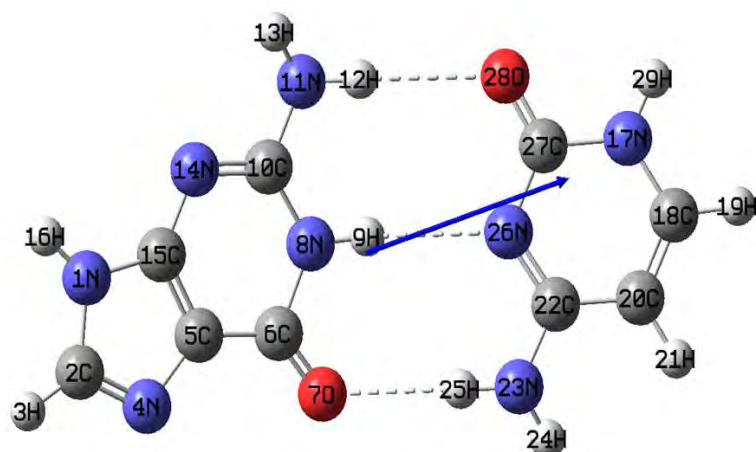
5.2 Results and Discussion

5.2.1 Physical and structural properties

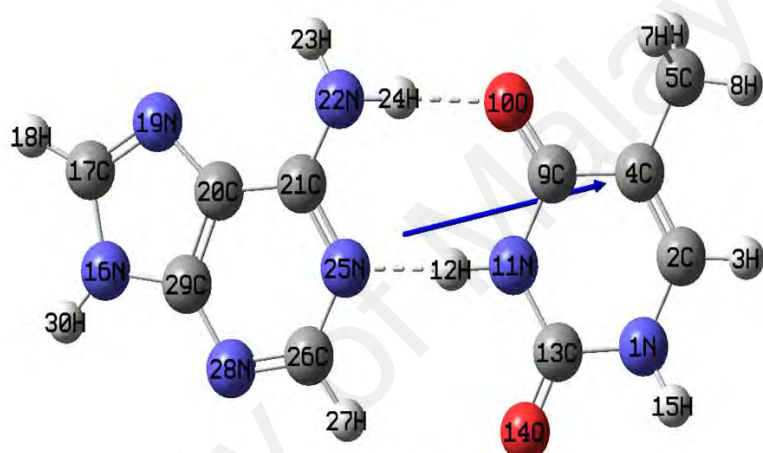
First of all, we look at the DNA and RNA base pairs in the absence of an applied electric field. Fig. 5.1 illustrates the optimized structure of G-C, A-T, and A-U base pairs. The atom number, hydrogen bond, and magnitude of dipole moment for each base pair are also shown in the Fig. 5.1. In the G-C base pair, G joins with C through three hydrogen bonds. One of the hydrogen bonds is formed between the amino hydrogen (H_{12}) of G (as an electron donor) and the carbonyl oxygen (O_{28}) of C (as an electron acceptor). The second hydrogen atom forms between the N_8H_9 site of G with the N_{26} of C ($N-H\cdots N$) and lastly, the third hydrogen bond ($O\cdots H-N$) is formed between the carbonyl oxygen (O_7) of G and the amino hydrogen (H_{25}) of C. For the A-T base pair, two hydrogens are formed between the A and T nucleobases. In this case, the first hydrogen bond is formed between the amino hydrogen (H_{24}) of A and the carbonyl oxygen (O_{10}) of T ($N-H\cdots O$). Whereas the second bond forms between the N_{25} of A and the $N_{11}H_{12}$ of T. The A-U base-pair also possesses two hydrogen bonds like the A-T base pair. The hydrogen bonding pattern linking between A and U is similar to the A-T base pair.

In the present work, several fundamental physicochemical properties, including the chemical formula, number of atoms, and heavy atom of the base pairs have been calculated and listed in Table 5.1. It can be seen that the G-C base pair has the largest molecular weight and van der Waals surface volume. On the other hand, the A-U base pair, has the lowest molecular weight, van der Waals surface volume, and polarizability.

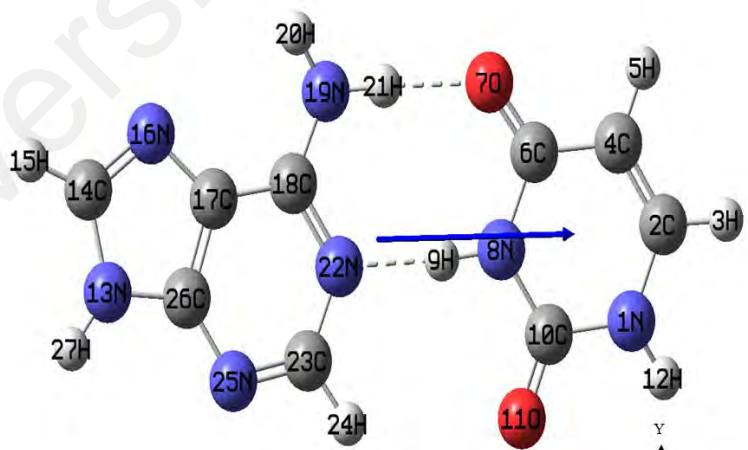
In Table 5.2, we present the dipole moments of the base pairs (G-C, A-T, A-U), which obtained without an applied electric field. For comparison purposes, the results obtained from Santamaria et al. and Sharma *et al.*, are also included. Altogether, the dipole moment is in the increasing order of A-U > A-T > G-C. The dipole moment of G-C is about four times larger than that of the A-T and A-U base pairs. It is interesting to note that in our previous calculations on single nucleobases, the G and C nucleobases also demonstrated the greatest dipole moments among the five nucleobases. On the contrary, A-U base pair exhibits the smallest dipole moment (1.211 Debye), which is almost identical to the A-T base pair (1.283 Debye). It is clear that our calculated dipole moment in the present work shows a good agreement with those reported by Santamaria et al. and Sharma et al. (Santamaria & Vázquez, 1994; Sharma et al., 2008). Nevertheless, we have observed that all the magnitude of dipole moment are pointing from purines nucleobases (A and G) toward the pyrimidines nucleobases (T, C, and U), indicating the purines nucleobases are more likely to hold the excess charge.



Guanine-Cytosine
G-C



Adenine-Thymine
A-T



Adenine-Uracil
A-U

Figure 5.1: Optimized structures of Guanine-Cytosine, Adenine-Thymine and Adenine-Uracil base pairs. Blue arrows indicate the magnitude of dipole moments and the dotted line represent the hydrogen bonds. The carbon atom is in grey colour, nitrogen in blue, oxygen in red and hydrogen in white.

Table 5.1: Physicochemical properties of DNA and RNA base pairs.

Base pairs	Chemical Formula	Number of Atoms	Number of heavy atoms	Molecular Weight [g/mol]	Van der Waals surface volume [\AA^3]	Polarizability [\AA^3]
Guanine-Cytosine	$\text{C}_9\text{H}_{10}\text{N}_8\text{O}_2$	29	19	262.23	266.97	25.09
Adenine-Thymine	$\text{C}_{10}\text{H}_{11}\text{N}_7\text{O}_2$	30	19	261.25	262.39	25.57
Adenine-Uracil	$\text{C}_9\text{H}_{10}\text{N}_4\text{O}_4$	27	18	247.22	242.11	23.74

Table 5.2: Calculated dipole moments of the DNA and RNA base pairs.

Base pairs	Dipole Moment [Debye]		
	Present work	Santamaria et al.	Sharma et al.
Guanine-Cytosine	5.242	6.34	5.48
Adenine-Thymine	1.283	1.39	Not available
Adenine-Uracil	1.211	Not available	2.89

Table 5.3 summarizes the TPSA and the TPSA/TSA ratio of all the three base pairs. The TPSA value for A-T and A-U are identical and considerably smaller than G-C. The resulting TPSA/TSA ratio for the G-C becomes the largest among all the base pairs. The value of the TPSA/TSA ratio sorted in the increasing order is $\text{A-T} > \text{A-U} > \text{G-C}$. In order to gain a deeper insight into the role of the hydrogen bond in governing the pairing between the base pair, we further evaluate the interaction energy of the base pair. The interaction energy of two nucleobases A and B was calculated as ΔE_{int} , where the energy difference between the total electronic energy of the base pair E_{AB} and the isolated nucleobases electronic energies, E_A and E_B .

$$\Delta E_{int} = E_{AB} - E_A - E_B \quad (5.1)$$

Based on the calculated interaction energy as presented in Table 5.4, we observed that the G-C base pair demonstrates the greatest interaction energy value of -7.35 kcal/mol. This result, which in principle is not surprising and can be attributed to the fact that the G pairing C has an additional hydrogen bond, compared to the A-T and A-U base pairs. In the A-T base pair, the total interaction energy is -5.70 kcal/mol, which

is about 1.65 kcal/mol weaker than the value of G-C. However, for the A-U base pair, the calculated interaction energy is -5.83 kcal/mol, which is only 0.13 kcal/mol stronger than the A-T base pair. The interaction energy hierarchy in the order from low to high is A-T > A-U > G-C.

Table 5.3: Calculated topological polar surface area (TPSA), total surface area (TSA), and TPSA/TSA ratio for DNA and RNA base pairs.

Base pairs	TPSA [\AA^2]	TSA [\AA^2]	TPSA/TSA \times 100%
Guanine-Cytosine	163.64	310.43	52.71
Adenine-Thymine	138.68	333.49	41.58
Adenine-Uracil	138.68	300.67	46.12

Table 5.4: Calculated interaction energy of the DNA and RNA base pairs.

Base pairs	Interaction Energy [kcal/mol]
Guanine-Cytosine	-7.35
Adenine-Thymine	-5.70
Adenine-Uracil	-5.83

5.2.2 Electric field effects on electrical, structural and physical properties

We now examined the molecular structures of the base pairs under the influence of an applied electric field. Fig. 5.2, shows the polarization (P) of the G-C, A-T, and A-U base pairs as the function of an applied electric field (E). An electric field range between -6 and $+6$ V/nm was applied along the Y -axis. The polarization (P) against the electric field (E) exhibits a typical feature of square-shaped hysteresis loops, in which the coercive field and zero-field polarization can be obtained directly from the curve. The corresponding total energy (U) versus an electric field (E) of the base pairs is depicted in Fig. 5.3. From Fig. 5.3, it clearly showed that all the base pairs displaying a symmetric butterfly loop. Table 5.5 shows the coercive field and zero-polarization of the base pairs obtained from the hysteresis loop in Fig. 5.2. Surprisingly, we found that the coercive field of the base pairs are also inversely proportional to the TPSA/TSA

ratio, which is consistent with our earlier finding in the nucleobases study. However, the minimum switching field required by base pairs are much smaller as compare with the nucleobases. Principally, this could be possible due to the fact that the fusion nucleobases containing extra polar atoms than a nucleobase, which are more susceptible to the applied electric field. For example, the A-T and A-U have the smallest number in polar atom and TPSA/TSA ratio, therefore require the highest electric field to switch the molecules. On contrary, the G-C base pair with the largest ratio of TPSA/TSA required least electric field to flip or switch the molecules into the opposite. This is also resulting in a narrow coercive field and hysteresis loop as shown in Fig. 5.2. The magnitude of an average coercive field sorted in the increasing order is $G-C > A-U > A-T$, while the average zero-field polarization in the decreasing order is $G-C > A-T > A-U$.

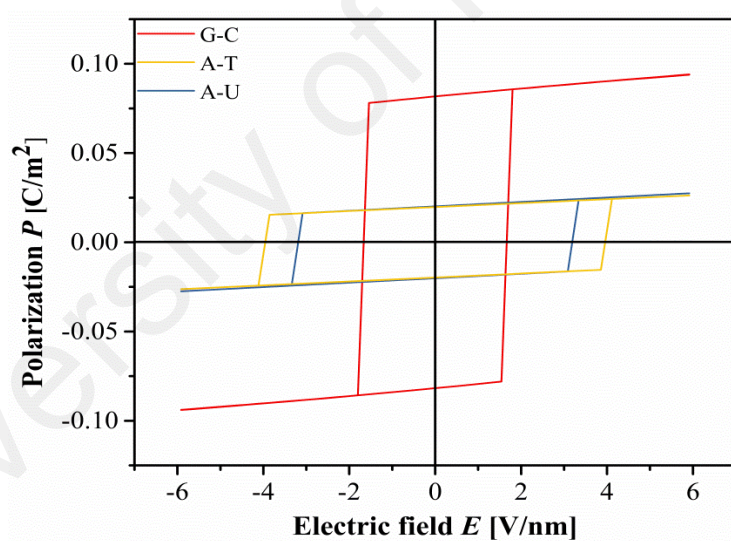


Figure 5.2: Electric field (E) dependence of polarization (P) for DNA and RNA base pairs.

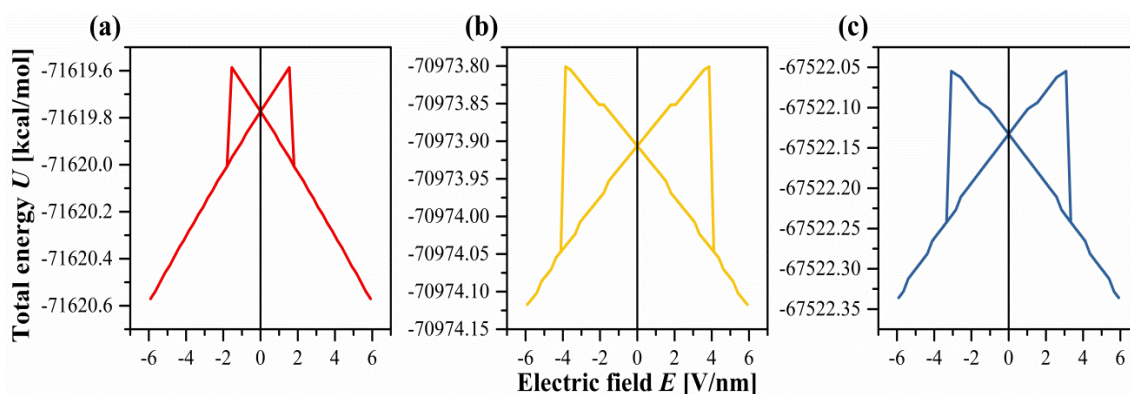


Figure 5.3: Electric field (E) dependence of total energy (U) for (a) Guanine-Cytosine (b) Adenine-Thymine (c) Adenine-Uracil.

Table 5.5: Coercive field and zero-field polarization of DNA and RNA base pairs as obtained from the hysteresis loops in Fig. 5.2.

Base pairs	Coercive Field [V/nm]	Zero-Field Polarization [C/m ²]
Guanine-Cytosine	1.80	0.081
Adenine-Thymine	4.11	0.020
Adenine-Uracil	3.34	0.019

Fig. 5.4 shows the hydrogen bond distances dependence on the electric field (E). Since the base pairs interact primarily via the hydrogen bond, the calculations were made in order to analyse the applied electric field effect on the hydrogen bond distance of the base pair. Here, we observed that the G-C hydrogen bond N-H---N exhibit the longest bond distance among all the three hydrogen bonds. In contrast, the O---H-N are bonded with the shortest bond distance. The bond distance for N-H---N and O---H-N at $E = 0$ are calculated to be 2.48 and 2.44 Å, respectively. As the applied electric field increases, both N-H---N and O---H-N hydrogen bond distance show a gradually increase, which indicates an opening of the hydrogen bond at the middle and down site of the base pairs. For the N-H---O hydrogen bond, the bond distance shorten by 0.06 Å, when the electric field strength is increasing from 0 to 6 V/nm. A similar trend is observed for the N-H---O hydrogen bond in A-U base pair, where a nonlinear relation is

observed between the distance of hydrogen bond and electric field. On the contrary, the N-H...N hydrogen bond from A-T base pair show dramatically increase in response to the applied electric field, which indicates a closing of the hydrogen bond at the upper site of the base pair and opening at the down site. The effect of the applied electric field in the A-T base pair gives a similar trend to the hydrogen bond in the A-U base pair. Our result shows good agreement with other DFT calculations (Ceron-Carrasco & Jacquemin, 2013). We further investigate the total length of the base pairs under different strengths of the applied electric field. Our result revealed interesting geometric changes of the base pair depicted in Fig. 5.5, where the total length of base pairs is plotted as a function of the electric field. As illustrated in Fig. 5.5, the applied electric field produces compression of the total lengths of the base pairs, indicating the base pair slowly gets shorter and smaller. We have found that the changes in hydrogen bond do directly affect the length of the base pairs.

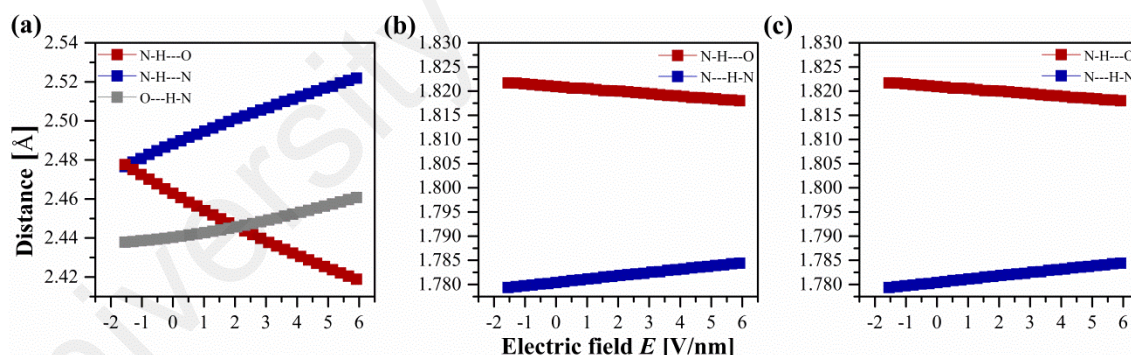


Figure 5.4: Changes of hydrogen bonds distance between the base pairs under an applied electric field: (a) Guanine-Cytosine (b) Adenine-Thymine (c) Adenine-Uracil.

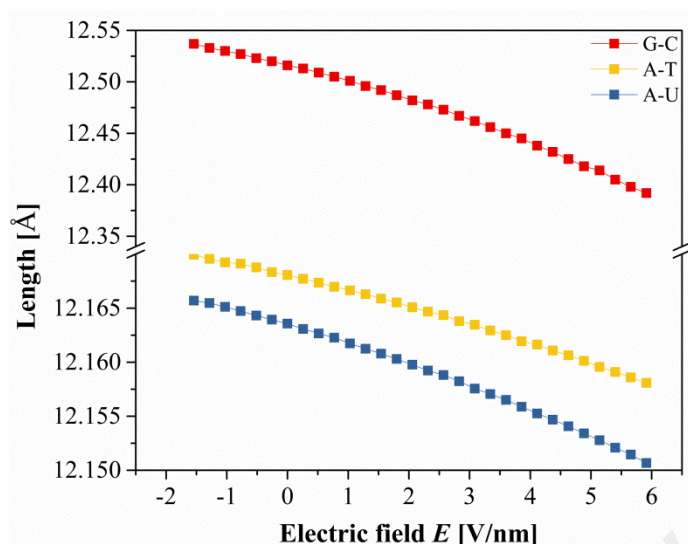


Figure 5.5: Changes of the total length of DNA and RNA base pairs under an applied electric field.

5.3 Summary

We have studied the physical and structural properties of G-C, A-T, and A-U base pairs with and without the presence of an applied electric field. Our results demonstrated that all the DNA and RNA base pairs have zero-field polarization and exhibit a square-shaped P - E hysteresis loop under the electric field. We found that the minimum field required for the polarization switching of a base pair is inversely proportional to the ratio of TPSA to TSA. Interestingly, the minimum switching field for a base pair was lower than the value needed for a nucleobase to switch.

CHAPTER 6: CONCLUSION AND FUTURE WORKS

6.1 Conclusion

In summary, we have systematically studied the physical and structural properties of the DNA/RNA nucleobases and their base pairs based on quantum mechanics (QM) PM3 method. We have calculated and studied the dipole moment, polarizability, topological polar surface area (TPSA) of the nucleobases and their base pairs in the absence of applied electric field. The results show that all five nucleobases and their three base pairs exhibit spontaneous dipole moment or polarization in the absence of an applied electric field.

Under the presence of an applied electric field, all nucleobases and their base pairs exhibit square-shaped hysteresis loops and butterfly-like U - E loop, suggesting a typical ferroelectric behaviour. The results imply the possible existence of bioferroelectricity in DNA and RNA. Moreover, we found an interesting relationship where the minimum field required for the polarization switching of a nucleobase and their base pair, which is inversely proportional to the ratio of TPSA to TSA. The present study may provide insight into the development of new nanoelectronics, nanomedicine, biomedical applications, and open up an intriguing set of applications based on the DNA/RNA nucleobases and their base pairs (e.g. DNA based conductors).

6.2 Future works

Almost all living biological molecules, including DNA and RNA are surrounded by a cluster of a water molecule. The water acts as a solvent to dissolve a variety of substances and also helps transport oxygen into the cell. Because of its polarity, water is able to form electrostatic interactions with other biological molecules. Hence, we can extend our works to study the important role and behaviours of water in DNA/RNA nucleobases and their base pairs under an applied electric field.

Furthermore, the study using complete double-stranded DNA or single-stranded RNA structure will provide more comprehensive information on the ferroelectricity in biological system. Ultimately, our theoretical calculation must progressively combine with the experimental measurements such as piezoresponse force microscopy (PFM), atomic force microscopy (AFM) or ultrasonic force microscopy for further validation.

University of Malaya

REFERENCES

- Alberts, B., Wilson, J. H., Bruce Alberts, D. B. J. L., Bray, D., Lewis, J., Raff, M., Roberts, K., & Watson, J. D. (1989). *Molecular Biology of the Cell*: Garland Pub.
- Albrecht, G., & Corey, R. B. (1939). The crystal structure of glycine. *Journal of the American Chemical Society*, 61(5), 1087-1103.
- Athenstaedt, H. (1971). Pyroelectric and piezoelectric behaviour of human dental hard tissues. *Archives of Oral Biology*, 16(5), 495-501.
- Athenstaedt, H. (1974). Pyroelectric and piezoelectric properties of vertebrates. *Annals of the New York Academy of Sciences*, 238, 68-94.
- Athenstaedt, H., Claussen, H., & Schaper, D. (1982). Epidermis of human skin: pyroelectric and piezoelectric sensor layer. *Science*, 216(4549), 1018-1020.
- Atkins, P. W., & Friedman, R. S. (2011). *Molecular Quantum Mechanics*: OUP Oxford.
- Bain, A., & Chand, P. (2017). *Ferroelectrics: Principles and Applications*.
- Baker-Jarvis, J., & Kim, S. (2012). The interaction of radio-frequency fields with dielectric materials at macroscopic to mesoscopic scales. *Journal of research of the National Institute of Standards and Technology*, 117, 1-60.
- Bassett, C. A. (1968). Biologic significance of piezoelectricity. *Calcified Tissue International*, 1(4), 252-272.
- Bdikin, I., Heredia, A., Neumayer, S. M., Bystrov, V. S., Gracio, J., Rodriguez, B. J., & Kholkin, A. L. (2015). Local piezoresponse and polarization switching in nucleobase thymine microcrystals. *Journal of Applied Physics*, 118(7), Article#072007.
- Bhide, V. G., Hegde, M. S., & Deshmukh, K. G. (1968). Ferroelectric properties of lead titanate. *Journal of the American Ceramic Society*, 51(10), 565-568.
- Blázquez-Castro, A., García-Cabañes, A., & Carrascosa, M. (2018). Biological applications of ferroelectric materials. *Applied Physics Reviews*, 5(4), Article#041101.
- Bodor, N., Gabanyi, Z., & Wong, C. K. (1989). A new method for the estimation of partition coefficient. *Journal of the American Chemical Society*, 111(11), 3783-3786.
- Bonin, K. D., & Kresin, V. V. (1997). *Electric-dipole Polarizabilities of Atoms, Molecules, and Clusters*: World Scientific.
- Boswarva, I. M., & Franklin, A. D. (1965). Orientation polarization of defect pairs in crystals. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, 11(110), 335-345.

- Bottcher, C. J. F. (1973). *Theory of electric polarization: Dielectrics in static fields*: Elsevier.
- Brighton, C. T., Black, J., & Pollack, S. R. (1979). *Electrical properties of bone and cartilage: experimental effects and clinical applications*: Grune & Stratton.
- Brown, J. A., & Tuszynski, J. A. (1999). A review of the ferroelectric model of microtubules. *Ferroelectrics*, 220(1), 141-155.
- Bull, C. L., Flowitt-Hill, G., de Gironcoli, S., Küçükbenli, E., Parsons, S., Pham, C. H., Playford, H. Y., & Tucker, M. G. (2017). ζ -Glycine: insight into the mechanism of a polymorphic phase transition. *IUCrJ*, 4(Pt 5), 569-574.
- Bystrov, V. S., Lakhno, V. D., & Molchanov, M. (1994). Ferroelectric active models of ion channels in biomembranes. *Journal of Theoretical Biology*, 168(4), 383-393.
- Bystrov, V. S., Paramonova, E., Bdikin, I., Kopyl, S., Heredia, A., Pullar, R. C., & Kholkin, A. L. (2012). BioFerroelectricity: diphenylalanine peptide nanotubes computational modeling and ferroelectric properties at the nanoscale. *Ferroelectrics*, 440(1), 3-24.
- Bystrov, V. S., Paramonova, E. V., Bdikin, I. K., Bystrova, A. V., Pullar, R. C., & Kholkin, A. L. (2013). Molecular modeling of the piezoelectric effect in the ferroelectric polymer poly(vinylidene fluoride) (PVDF). *Journal of Molecular Modeling*, 19(9), 3591-3602.
- Bystrov, V. S., Seyedhosseini, E., Bdikin, I., Kopyl, S., Neumayer, S. M., Coutinho, J., & Kholkin, A. L. (2015). Bioferroelectricity in nanostructured glycine and thymine: molecular modeling and ferroelectric properties at the nanoscale. *Ferroelectrics*, 475(1), 107-126.
- Bystrov, V. S., Seyedhosseini, E., Bdikin, I. K., Kopyl, S., Kholkin, A. L., Vasilev, S. G., Zelenovskiy, P. S., Vasileva, D. S., & Shur, V. Y. (2016). Glycine nanostructures and domains in beta-glycine: computational modeling and PFM observations. *Ferroelectrics*, 496(1), 28-45.
- Ceron-Carrasco, J. P., & Jacquemin, D. (2013). Electric-field induced mutation of DNA: a theoretical investigation of the GC base pair. *Physical Chemistry Chemical Physics*, 15(13), 4548-4553.
- Chargaff, E., Zamenhof, S., & Green, C. (1950). Composition of human desoxypentose nucleic acid. *Nature*, 165(4202), 756-757.
- Chen, Y., & Or, D. (2006). Effects of Maxwell-Wagner polarization on soil complex dielectric permittivity under variable temperature and electrical conductivity. *Water Resources Research*, 42(6), Article#W06424.
- Cowley, R. A., & Coombs, G. J. (1973). Paraelectric, piezoelectric and pyroelectric crystals: II Phase transitions. *Journal of Physics C: Solid State Physics*, 6(1), Article#143.

- Date, M., Takashita, S., & Fukada, E. (1970). Temperature variation of piezoelectric moduli in oriented poly(γ -methyl L-glutamate). *Journal of Polymer Science Part A-2: Polymer Physics*, 8(1), 61-70.
- Dawson, A., Allan, D. R., Belmonte, S. A., Clark, S. J., David, W. I. F., McGregor, P. A., Parsons, S., Pulham, C. R., & Sawyer, L. (2005). Effect of high pressure on the crystal structures of polymorphs of glycine. *Crystal Growth & Design*, 5(4), 1415-1427.
- Dragan, D. (1998). Ferroelectric, dielectric and piezoelectric properties of ferroelectric thin films and ceramics. *Reports on Progress in Physics*, 61(9), Article#1267.
- Ertl, P., Rohde, B., & Selzer, P. (2000). Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *Journal of Medicinal Chemistry*, 43(20), 3714-3717.
- Fischer, E., & Fourneau, E. (1901). Ueber einige Derivate des Glykocolls. *Berichte der deutschen chemischen Gesellschaft*, 34(2), 2868-2877.
- Frohlich, H. (1949). *Theory of dielectrics; dielectric constant and dielectric loss*. Oxford: Clarendon Press.
- Fukada, E., & Ando, Y. (1972). Piezoelectricity in oriented DNA films. *Journal of Polymer Science Part A-2: Polymer Physics*, 10(3), 565-567.
- Fukada, E., & Hara, K. (1969). Piezoelectric Effect in Blood Vessel Walls. *Journal of the Physical Society of Japan*, 26(3), 777-780.
- Fukada, E., & Ueda, H. (1970). Piezoelectric effect in muscle. *Japanese Journal of Applied Physics*, 9(844), Article#844.
- Fukada, E., & Yasuda, I. (1957). On the Piezoelectric Effect of Bone. *Journal of the Physical Society of Japan*, 12(10), 1158-1162.
- Griffiths, D. J. (2016). *Introduction to Quantum Mechanics*: Cambridge University Press.
- Gupta, K. M., & Gupta, N. (2015). *Advanced Electrical and Electronics Materials: Processes and Applications*.
- Hartree, D. R. (2008). The wave mechanics of an atom with a non-coulomb central field. part I. theory and methods. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(1), 89-110.
- Heinrich, J. (2012). *Measurement of Ferroelectric Hysteresis Loops in Multiferroic Materials*: Verlag nicht ermittelbar.
- Heredia, A., Meunier, V., Bdikin, I. K., Gracio, J., Balke, N., Jesse, S., Tselev, A., Agarwal, P. K., Sumpter, B. G., Kalinin, S. V., & Kholkin, A. L. (2012). Nanoscale Ferroelectricity in Crystalline γ -Glycine. *Advanced Functional Materials*, 22(14), 2996-3003.

- Hippel, A. v. (1970). *Do we really understand ferroelectricity? Proceedings Second Int Meeting on Ferroelectricity 1969* (Vol. 28).
- Hiroyoshi, U., & Eiichi, F. (1971). Piezoelectricity in myosin and actin. *Japanese Journal of Applied Physics*, 10(11), Article#1650.
- Hol, W. G. (1985). The role of the alpha-helix dipole in protein function and structure. *Progress in Biophysics & Molecular Biology*, 45(3), 149-195.
- Hu, P., Hu, S., Huang, Y., Reimers, J. R., Rappe, A. M., Li, Y., Stroppa, A., & Ren, W. (2019). Bioferroelectric properties of glycine crystals. *The Journal of Physical Chemistry Letters*, 10(6), 1319-1324.
- Hyperchem. (2002). *Tools for Molecular Modeling* Hypercube, Inc.
- Iitaka, Y. (1954). A New Form of Glycine. *Proceedings of the Japan Academy*, 30(2), 109-112.
- Kaimanovich, V., Krupitskii, Y. M., & Spirov, A. V. (1990). *Possible role of intercellular electric fields in the orientation of an assembly of microtubules* (Vol. 35).
- Kaimanovich, V. A., Krupitski, E. M., & Spirov, A. V. (1989). The possible contribution of intracellular electric fields to oriented assemblage of microtubules. *Journal of Bioelectricity*, 8(2), 243-245.
- Kim, Y. S., Kim, D. J., Kim, T. H., Noh, T. W., Choi, J. S., Park, B. H., & Yoon, J.-G. (2007). Observation of room-temperature ferroelectricity in tetragonal strontium titanate thin films on SrTiO₃ (001) substrates. *Applied Physics Letters*, 91(4), Article#042908.
- Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A), A1133-A1138.
- Krishnamurthy, R. (2012). Role of pK(a) of nucleobases in the origins of chemical evolution. *Accounts of Chemical Research*, 45(12), 2035-2044.
- Kroh, H. J., & Felderhof, B. U. (2000). Force and torque on a sphere with electric dipole moment moving in a dielectric fluid in the presence of a uniform magnetic field. *Physica A: Statistical Mechanics and its Applications*, 280(3), 256-265.
- Kumar, A. (2018). *Fundamentals of Quantum Mechanics*. Cambridge: Cambridge University Press.
- Lang, S. B. (1966). Pyroelectric Effect in Bone and Tendon. *Nature*, 212(5063), 704-705.
- Lang, S. B., & Athenstaedt, H. (1977). Pyroelectricity and Induced Pyroelectric Polarization in Leaves of the Palmlike Plant *Encephalartos villosus*. *Science*, 196(4293), 985-986.

- Lemanov, V. V. (2000). Piezoelectric and pyroelectric properties of protein amino acids as basic materials of soft state physics. *Ferroelectrics*, 238(1), 211-218.
- Lemanov, V. V., Popov, S. N., & Pankova, G. A. (2002). Piezoelectric properties of crystals of some protein aminoacids and their related compounds. *Physics of the Solid State*, 44(10), 1929-1935.
- Leuchtag, H. R. (1987). Indications of the existence of ferroelectric units in excitable-membrane channels. *Journal of Theoretical Biology*, 127(3), 321-340.
- Leuchtag, H. R. (1988). A proposed physical explanation of the activation of sodium channels. *Ferroelectrics*, 86(1), 105-113.
- Leuchtag, H. R., & Bystrov, V. S. (1999). Theoretical models of conformational transitions and ion conduction in voltage-dependent ion channels: bioferroelectricity and superionic conduction. *Ferroelectrics*, 220(1), 157-204.
- Li, J., Liu, Y., Zhang, Y., Cai, H.-L., & Xiong, R.-G. (2013). Molecular ferroelectrics: where electronics meet biology. *Physical Chemistry Chemical Physics*, 15(48), 20786-20796.
- Li, T., & Zeng, K. (2011). Piezoelectric properties and surface potential of green abalone shell studied by scanning probe microscopy techniques. *Acta Materialia*, 59(9), 3667-3679.
- Lipkowitz, K. B., & Boyd, D. B. (2009). *Reviews in Computational Chemistry*: Wiley.
- Liu, Y., Wang, Y., Chow, M.-J., Chen, N. Q., Ma, F., Zhang, Y., & Li, J. (2013). Glucose suppresses biological ferroelectricity in aortic elastin. *Physical Review Letters*, 110(16), 168101-168101.
- Liu, Y., Zhang, Y., Chow, M.-J., Chen, Q. N., & Li, J. (2012). Biological ferroelectricity uncovered in aortic walls by piezoresponse force microscopy. *Physical Review Letters*, 108(7), Article#078103.
- Mailoud, O. M., Elsayed, A. H., Abo-Elazm, A. H., & Fetouh, H. A. (2018). Synthesis and study the structure, optical, thermal and dielectric properties of promising Glycine Copper Nitrate (GCN) single crystals. *Results in Physics*, 10, 512-520.
- McLennan, A., & Turner, P. (2012). *Molecular Biology*: Garland Science.
- Mershin, A., Kolomenski, A. A., Schuessler, H. A., & Nanopoulos, D. V. (2004). Tubulin dipole moment, dielectric constant and quantum behavior: computer simulations, experimental results and suggestions. *Biosystems*, 77(1-3), 73-85.
- Mershin, A., Sanabria, H., Miller, J. H., Nawarathna, D., Skoulakis, E. M. C., Mavromatos, N. E., Kolomenskii, A. A., Schuessler, H. A., Luduena, R. F., & Nanopoulos, D. V. (2006). Towards experimental tests of quantum effects in cytoskeletal proteins. In J. A. Tuszynski (Ed.), *The Emerging Physics of Consciousness* (pp. 95-170). Berlin, Heidelberg: Springer Berlin Heidelberg.

- Miller, K. J. (1990). Additivity methods in molecular polarizability. *Journal of the American Chemical Society*, 112(23), 8533-8542.
- Pal, M., Guo, R., & Bhalla, A. (2015). *Ferroelectricity and Ferroic Like Signature in Biological Species: 'Bio-Multiferroics'—An Overview* (Vol. 166).
- Polonsky, J., Douzou, P., & Sadron, C. (1960). Demonstration of the ferroelectric properties in desoxyribonucleic acid (DNA). *Comptes Rendus Hebdomadaires Des Seances Academie Sciences*, 250, 3414-3416.
- Pople, J. A., & Beveridge, D. L. (1970). *Approximate molecular orbital theory*: McGraw-Hill.
- Pranitha, A., & Lakshmi, P. (2014). Towards a Correlation between Polar Surface Area of Drugs with Ex-vivo Transdermal Flux Variability. *Iranian Journal of Pharmaceutical Sciences*, 10(2), 47-60.
- Pray, L. A. (2008). Discovery of DNA structure and function: Watson and Crick. *Nature Education*, 1(1), Article#100.
- Raju, G. G. (2003). *Dielectrics in Electric Fields*: Taylor & Francis.
- Ramsey, N. F. (1982). Electric-dipole moments of elementary particles. *Reports on Progress in Physics*, 45(1), Article#95.
- Rossi, D. D., Domenici, C., & Pastacaldi, P. (1986). Piezoelectric properties of dry human skin. *IEEE Transactions on Electrical Insulation*, EI-21(3), 511-517.
- Sackett, D. (1997). *pH-induced conformational changes in the carboxy terminal tails of tubulin*. Paper presented at the Banff workshop molecular biophysics of the cytoskeleton, Banff, Alberta, Canada.
- Sackett, D. L. (1995). Structure and function in the tubulin dimer and the role of the acidic carboxyl terminus. *Subcell Biochem*, 24, 255-302.
- Sanabria, H., Miller, J. H., Jr., Mershin, A., Luduena, R. F., Kolomenski, A. A., Schuessler, H. A., & Nanopoulos, D. V. (2006). Impedance spectroscopy of alpha-beta tubulin heterodimer suspensions. *Biophysical Journal*, 90(12), 4644-4650.
- Santamaria, R., & Vázquez, A. (1994). Structural and electronic property changes of the nucleic acid bases upon base pair formation. *Journal of Computational Chemistry*, 15(9), 981-996.
- Sataric, M. V., Budinski-Petkovic, L., Loncarevic, I., & Tuszynski, J. A. (2008). Modelling the role of intrinsic electric fields in microtubules as an additional control mechanism of bi-directional intracellular transport. *Cell Biochemistry and Biophysics*, 52(2), 113-124.
- Satarić, M. V., Tuszynski, J. A., & Žakula, R. B. (1993). Kinklike excitations as an energy-transfer mechanism in microtubules. *Physical Review E*, 48(1), 589-597.

- Sawyer, C. B., & Tower, C. H. (1930). Rochelle Salt as a Dielectric. *Physical Review*, 35(3), 269-273.
- Schrödinger, E. (1926). An Undulatory Theory of the mechanics of atoms and molecules. *Physical Review*, 28(6), 1049-1070.
- Schuessler, H. A., Mershin, A., Kolomenskii, A. A., & Nanopoulos, D. V. (2003). Surface plasmon resonance study of the actin-myosin sarcomeric complex and tubulin dimers. *Journal of Modern Optics*, 50(15-17), 2381-2391.
- Sellmann, J. (2016). *Impact of strain and composition on structural and piezo-/ferroelectric properties of epitaxial NaNbO₃ and KxNa1-xNbO₃ thin films and superlattices grown by PLD* (Doctoral dissertation, University of Berlin). Retrieved from <https://depositonce.tu-berlin.de/handle/11303/5312>
- Shamos, M. H., & Lavine, L. S. (1967). Piezoelectricity as a fundamental property of biological tissues. *Nature*, 213(5073), 267-269.
- Sharma, P., Mitra, A., Sharma, S., & Singh, H. (2008). *Base pairing in RNA structures: A computational analysis of structural aspects and interaction energies* (Vol. 119).
- Sidgwick, N. V. (1936). Dipole moment and molecular structure. *Chemical Reviews*, 19(3), 183-194.
- Sinden, R. R. (1994). *DNA Structure and Function*: Elsevier Science.
- Stanford, A. L., & Lorey, R. A. (1968). Evidence of Ferroelectricity in RNA. *Nature*, 219(5160), 1250-1251.
- Stewart, J. J. P. (1989). Optimization of parameters for semiempirical methods I. Method. *Journal of Computational Chemistry*, 10(2), 209-220.
- Stewart, J. J. P. (2007). Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular Modeling*, 13(12), 1173-1213.
- Stracke, R., Bohm, K. J., Wollweber, L., Tuszyński, J. A., & Unger, E. (2002). Analysis of the migration behaviour of single microtubules in electric fields. *Biochemical and Biophysical Research Communications*, 293(1), 602-609.
- Talman, R. (2017). *The Electric Dipole Moment Challenge*. San Rafael California: Morgan & Claypool Publishers.
- Tasi, G., Palinko, I., Nyerges, L., Fejes, P., & Foerster, H. (1993). Calculation of electrostatic potential maps and atomic charges for large molecules. *Journal of Chemical Information and Computer Sciences*, 33(3), 296-299.
- Thiel, W. (2005). Chapter 21 - Semiempirical quantum-chemical methods in computational chemistry. In C. E. Dykstra, G. Frenking, K. S. Kim, & G. E. Scuseria (Eds.), *Theory and Applications of Computational Chemistry* (pp. 559-580). Amsterdam: Elsevier.

- Timmons, J. J., Preto, J., Tuszyński, J. A., & Wong, E. T. (2018). Tubulin's response to external electric fields by molecular dynamics simulations. *PLoS ONE*, 13(9), Article#e0202141.
- Turner, P. C., & McLennan, A. G. (2005). *Molecular Biology*: Taylor & Francis.
- Tuszyński, J., Craddock, T., & J. Carpenter, E. (2008). *Bio-Ferroelectricity at the Nanoscale* (Vol. 5).
- Valasek, J. (1921). Piezo-Electric and Allied Phenomena in Rochelle Salt. *Physical Review*, 17(4), 475-481.
- Vasileva, D., Vasilev, S., Kholkin, A. L., & Shur, V. Y. (2019). Domain Diversity and Polarization Switching in Amino Acid beta-Glycine. *Materials (Basel)*, 12(8), Article#1223.
- Vassilev, P. M., Dronzine, R. T., Vassileva, M. P., & Georgiev, G. A. (1982). Parallel arrays of microtubules formed in electric and magnetic fields. *Bioscience Reports*, 2(12), 1025-1029.
- Vickery, H. B., & Schmidt, C. L. A. (1931). The History of the Discovery of the Amino Acids. *Chemical Reviews*, 9(2), 169-318.
- Von Hippel, A., Breckenridge, R. G., Chesley, F. G., & Tisza, L. (1946). High dielectric constant ceramics. *Industrial & Engineering Chemistry*, 38(11), 1097-1109.
- Wallace, R. M. (2017). Dielectric materials for microelectronics. In S. Kasap & P. Capper (Eds.), *Springer Handbook of Electronic and Photonic Materials* (pp. 1-265). Cham: Springer International Publishing.
- Wang, B., Huang, W., Chi, L., Al-Hashimi, M., Marks, T. J., & Facchetti, A. (2018). High- k Gate dielectrics for emerging flexible and stretchable electronics. *Chemical Reviews*, 118(11), 5690-5754.
- Watson, J. D., & Crick, F. H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356), 737-738.
- Weinberg, J. M., Bienholz, A., & Venkatachalam, M. A. (2016). The role of glycine in regulated cell death. *Cellular and Molecular Life Sciences* 73(11-12), 2285-2308.
- White, R., Hyde, G., & Overall, R. (1990). *Microtubule arrays in regenerating Mougeotia protoplasts may be oriented by electric fields* (Vol. 158).
- Williams, W. S., & Breger, L. (1975). Piezoelectricity in tendon and bone. *Journal of Biomechanics*, 8(6), 407-413.
- Yam, S.-C., Zain, S. M., Sanghiran Lee, V., & Chew, K.-H. (2018). Correlation between polar surface area and bioferroelectricity in DNA and RNA nucleobases. *The European Physical Journal E*, 41(7), Article#86.

Yarmarkin, V. K., Shul'man, S. G., & Lemanov, V. V. (2009). Is there ferroelectricity in DNA? *Physics of the Solid State*, 51(9), 1881-1885.

Zelisko, M., Li, J., & Sharma, P. (2015). What is the mechanism behind biological ferroelectricity? *Extreme Mechanics Letters*, 4, 162-174.

University of Malaya

LIST OF PUBLICATIONS AND PAPERS PRESENTED

LIST OF PUBLICATION

Yam, S.-C., Zain, S. M., Sanghiran Lee, V., & Chew, K.-H. (2018). Correlation between polar surface area and bioferroelectricity in DNA and RNA nucleobases. *The European Physical Journal E*, 41(7), Article#86.

CONFERENCE

Yam, S.-C., Zain, S. M., Sanghiran Lee, V., & Chew, K.-H. (2017). Molecular modelling of bioferroelectricity in DNA and RNA nucleobases. Poster presented at the 9th Conference of the Asian Consortium on Computational Materials Science (ACCMS-9), 8-11 August 2017, Kuala Lumpur, Malaysia.