MOLECULAR DYNAMICS SIMULATION STUDY ON THE INTERACTIONS BETWEEN DNA AND A CONJUGATED POLYELECTROLYTE (CATIONIC OLIGOTHIOPHENE)

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Chemistry

by
Nehir NALINCI BARBAK

July 2019
İZMİR

We approve the thesis of Nehir NALINCI BARBAK

Examining Committee Members:

_____
Prof. Dr. Nuran ELMACI IRMAK
Department of Chemistry, İzmir Institute of Technology

_____
Assoc. Prof. Dr. Ümit Hakan YILDIZ
Department of Chemistry, İzmir Institute of Technology

_____
Assoc. Prof. Dr. Ezgi KARACA EREK
Department of Biomedicine and Health Technologies,
İzmir International Biomedicine and Genome Institute

17 July 2019

_____
Prof. Dr. Nuran ELMACI IRMAK
Supervisor, Department of Chemistry
İzmir Institute of Technology

_____
Prof. Dr. Ahmet Emin EROĞLU
Head of Department of Chemistry

_____
Prof. Dr. Aysun SOFUOĞLU
Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGMENTS

ABSTRACT

MOLECULAR DYNAMICS SIMULATION STUDY ON THE INTERACTIONS BETWEEN DNA AND A CONJUGATED POLYELECTROLYTE (CATIONIC OLIGOTHIOPHENE)

The absorption spectra of the cationic polythiophenes shift to the red, or the color changes in the solution are visible to the naked eye, when single-strand DNA (ssDNA) is added, so that they can be used as a tool for DNA detection, theranostic applications, and biological sensors. The red shift or color change is explained by the fact that the ssDNA leads to conformational changes in the polythiophene, but the form of structural change remains to be elusive (i.e. flattening, twisting, stacking, etc.).

In this study, molecular dynamics (MD) simulations of complexes consisted by ssDNA sequences with different nucleotides and polythiophene containing cationic side group were performed to enlighten the experimental studies. For this purpose, force field parameters of polythiophene which are not present in the current databases, were generated. The interactions between them were analyzed to determine the nature of conformational changes in the polythiophene when ssDNA was added.

MD simulations has been carried out with the CHARMM-compatible force field parameters obtained in the content of this work. Radius of gyration of oligomer increases with addition of ssDNA but is more affected by homopurine strand. Planarity index gets larger upon complexation with homopurine and T_rich strand, does not change with others. H-O and electrostatic interactions which are almost doubled in non-planar complexes can be interpreted as the major sources of conformational changes in oligomer. Considering all types of interactions between atoms in duplexes, it was observed that planarity was high in structures with less interaction of oligomer side groups.

ÖZET

DNA VE BİR İLETKEN POLİELEKTROLİT (KATYONİK OLİGOTİYOFEN) ARASINDAKİ ETKİLEŞİMLER ÜZERİNE MOLEKÜLER DİNAMİK SİMÜLASYON ÇALIŞMASI

Katyonik politiyofenlerin absorpsiyon spektrumları, tek sarmallı DNA zinciri (ssDNA) eklendiğinde, kırmızı bölgeye kayar veya çözeltideki renk değişiklikleri çıplak gözle görülebilir, böylece DNA tespiti, teranostik uygulamaları ve biyolojik sensörler için bir araç olarak kullanılabilirler. Kırmızı bölgeye kayma ya da renk değişimi, ssDNA'nın politiyofende konformasyonel değişikliklere yol açtığı, ancak yapısal değişiklik biçiminin açıkça bilinmediği (örneğin düzleştirme, büküm, istifleme, vs.) şeklinde açıklanmaktadır.

Bu çalışmada, deneysel çalışmaları aydınlatmak amacıyla, farklı nükleotitlere sahip ssDNA dizileri ve katyonik yan grup içeren politiyofen komplekslerin moleküler dinamik (MD) simülasyonları yapılmıştır. Bu amaçla, mevcut veritabanlarında bulunmayan politiyofenin kuvvet alanı parametreleri üretilmiştir. Aralarındaki etkileşimler, ssDNA eklendiğinde politiyofendeki konformasyonel değişikliklerin doğasını belirlemek için analiz edilmiştir.

MD simülasyonlar bu çalışmanın içeriğinde elde edilen CHARMM uyumlu kuvvet alanı parametreleri ile gerçekleştirilmiştir. Oligomerin dönme yarıçapı (radius of gyration), ssDNA ilavesiyle artar, ancak homopürin sarmalından daha çok etkilenir. Düzlemsellik indeksi, homopürin ve T_rich sarmalları ile kompleksleşince büyür, diğerleri ile değişmez. Düzlemsel olmayan komplekslerde neredeyse iki katına çıkan elektrostatik ve H-O etkileşimler, oligomerlerdeki başlıca konformasyonel değişiklik kaynakları olarak yorumlanabilir. Dublekslerdeki atomlar arasındaki tüm etkileşimler göz önüne alındığında, tiyofenin yan gruplarıyla daha az etkileşmeye sahip yapılarda düzlemselliğin yüksek olduğu gözlenmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

PE	Polyelectrolyte
PT	Polythiophene
DNA	Deoxyribonucleic Acid
ss	Single Strand
ds	Double Strand
ATP	Adenosine triphosphate
ADP	Adenosine diphosphate
UTP	Uridine-5'-triphosphate
EDOT	3,4-ethylenedioxythiophene
UV	Ultraviolet
Vis	Visible
MD	Molecular Dynamics
RMSD	Root Mean Square Deviation
CHARMM	Chemistry at Harvard Macromolecular Mechanics
MC	Monte Carlo
NAMD	Nanoscale Molecular Dynamics
VMD	Visual Molecular Dynamics
QM	Quantum Mechanical
MM	Molecular Mechanics
AMBER	Assisted Model Building with Energy Refinement
GROMOS	Groenigen Molecular Simulation
OPLS	Optimized Parameters for Large-scale Simulations
MMFF	Merck Molecular Force Field
ffTK	Force Field Toolkit
Rg	Radius of Gyration
CIS	Configuration Interaction Singles
HF	Hartree Fock
DFT	Density Functional Theory
TD	Time-Dependent
RMSE	Root Mean Square Error

CHAPTER 1

INTRODUCTION

1.1. Definition of Polyelectrolytes

Polyelectrolytes (PEs) are semi-conductive polymers, consisting of hydrophilic ionic or ionizable side groups and hydrophobic π-conjugated backbone. They are macromolecules which have positive or negative charges which are called polycation and polyanion, respectively. Solutions of these two oppositely charged polymers can react and then form polysalt (Figure 1.1) (Ankerfors 2012).
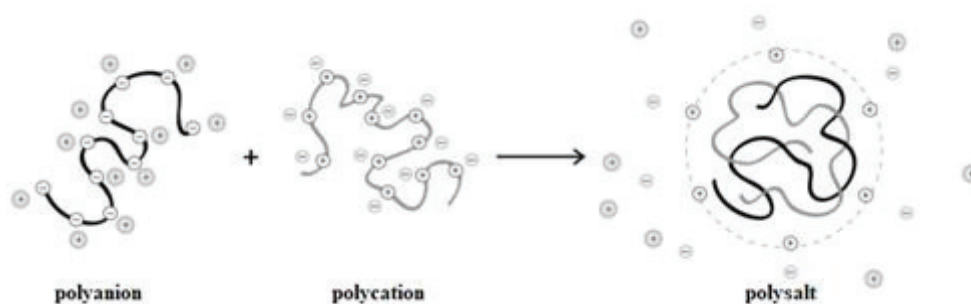
Figure 1.1. Polyelectrolyte complex formation. (Source: Ankerfors 2012)

The properties of PEs are similar to polymers (their molecular weight is high) and electrolytes (their solutions are conductive). The conjugated PEs are generally soluble in water hence they are used in biochemical and medical applications. They have an important place in the literature, not only good solubility in water but also sensitive optical responses to external factors (temperature, light, pH etc.) and increasing of fluorescence signals. Especially, water soluble polythiophenes (PTs), which are types of the conjugated polyelectrolytes, are widely synthesized because they show low toxicity and have a good photostability in living-cell experiments, and adjustable structural features for targeted molecule detection (Rubio-Magnieto et al. 2013; Carreon et al. 2014). That's why they provide a unique substructure for both biosensor and chemosensor developments.

1.2. Polythiophene Derivatives

One of the most important polyelectrolytes is polythiophene derivatives because they show low toxicity and have a good photostability in living cell experiments. That's why they provide a unique substructure for bio- and chemo- sensor developments.

1.2.1. Historical survey

In 1987, Wudl, Heeger and coworkers were firstly synthesized PTs sodium poly-(3-thiophene-5-ethanesulfonate) and sodium poly-(3-(thiophene-3-butanesulfonate) (Figure 1.2.). Their research and synthesizing methods have expanded the field of study in the water-soluble PTs (Patil et al. 1987; Li and Shi 2013).
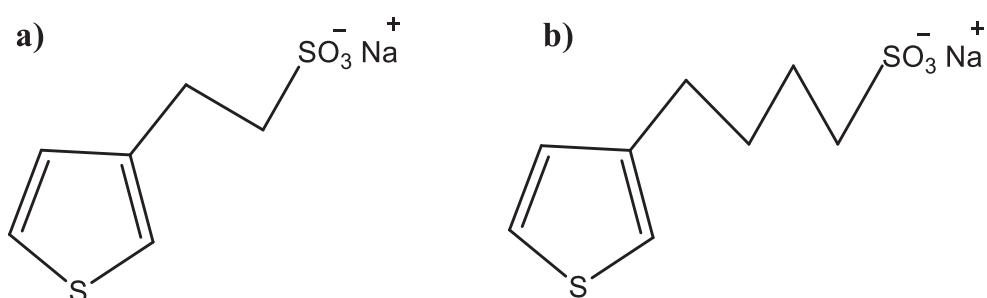
Figure 1.2. Chemical structures of the first synthesized polythiophene derivatives a) sodium poly-(3-thiophene-3-ethanesulfonate), b) sodium poly-(3-(thiophene-5-butanesulfonate)

In 1990's, the scientists thought that PTs would have a very widespread effect in the future due to their features (electrochemical, optical etc.) and development of their structural designs. However, to day, the relationship between structure and properties of the PTs could not be interpreted very well because there was not enough information about the properties of PTs (Roncali 1992).

Another study made by Leclerc's group in 1997 aimed to show how electrical and optical properties of different polythiophene derivatives could be controlled. They suggested that the side chains have the flexibility to cause some optical effects in the neutral polymer. Furthermore, their reversible structures (from non-planar to planar or less / highly conjugated, respectively) supported more advantages to use future research

and application areas such as development of antistatic and transparent coatings, light-weight batteries, conducting photoresists, light-emitting diodes, display devices, sensors, etc. (Relationships, Leclerc, and Faid 1997).

In a study conducted in 1998, the synthesis of polythiophenes has been thoroughly studied and the importance of polythiophenes in this field has been emphasized. The effects of the structure and function of the polymers on the physical properties have been shown, and accordingly the synthesis part has an important place in the literature. PTs have a significant advantage in good electronic and photonic properties, better conductivity, more stability and more efficient device designs (Mccullough 1998).

1.2.2. Complexation with DNA

Conjugated PEs have captured the attention of many researchers during the past few decades because of unique electronic and optical properties. They are important part of nanomaterials for biomedical applications, and also in many areas such as sensor, solar cells, light emitting diodes and photovoltaics (Rubio-Magnieto et al. 2013; Rajwar et al. 2016). In particular; cationic PTs have a great role in the biosensor development to be selective and sensitive probe due to their interaction with deoxyribonucleic acid (DNA).
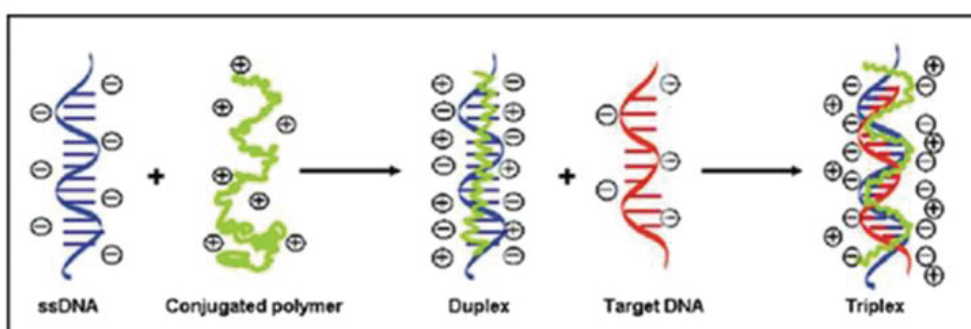
Figure 1.3. Demonstration of cationic polythiophene binding to ssDNA and then complementation to DNA during hybridization. (Source: Zheng and He 2014)

When single strand DNA (ssDNA) is added to conjugated PT, color of solution change from yellow to red which is visible to the naked eye and the red shift in the cationic PT absorption spectra is clearly observed in the experimental studies. ssDNA and cationic

PT are called as a duplex form. After that, the duplex mixture is mixed with complementary ssDNA, which is named as triplex, the color of solution turned to yellow again (Figure 1.3). This observation is explained by the ssDNA molecule causing conformational changes in the electrolyte polymer. However, the types of structural changes are not known exactly. Such response of polymer to ssDNA makes them a suitable tool for DNA detection, observation of DNA cleavage reaction, and theranostic polyplex applications.

1.3. Literature Work

There is a quite number of studies about cationic PTs in terms of their application areas and their capacity to be improved. Within the scope of this thesis, the studies in the literature will be considered as two parts: experimental / theoretical works about PTs and part of the force field parameters to create files needed to perform molecular dynamics simulations.

1.3.1. Studies on Polythiophenes
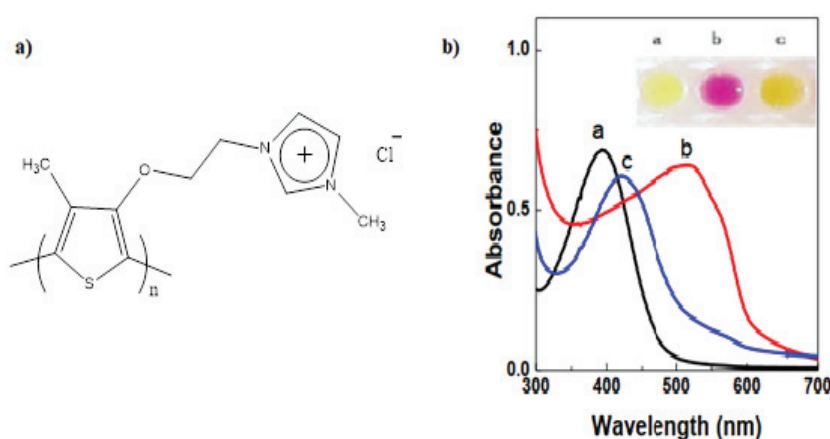
Figure 1.4. a) Chemical structure of poly (1H-imidazolium-1-methyl-3- {2- [(4-methyl-3-thienyl) -oxy] ethyl} chloride) and b) Absorption spectra of PT a, PT alone; b, addition of ssDNA; c, addition of complementary ssDNA (Source: Ho, Béra-Abérem, and Leclerc 2005).

In 2005, poly (1H-imidazolium-1-methyl-3- {2- [(4-methyl-3-thienyl) -oxy] ethyl} chloride) (Figure 1.4.a) was synthesized and improved by Leclerc et al. and coworkers to use as a sensor that target, or probe without any chemically labelling. Its maximum absorption was 397 nm in aqueous solution. When ssDNA and cationic PT was mixed, their maximum absorption was 527 nm, that means color change in solution which can be observed by naked eyes from yellow to red (Figure 1.4.b). Then, the complementary ssDNA was added this duplex to form triplex. Triplex maximum absorption was 421 nm (Figure 1.4.c) and solution color turned yellow again (Ho, Béra-Abérem, and Leclerc 2005).
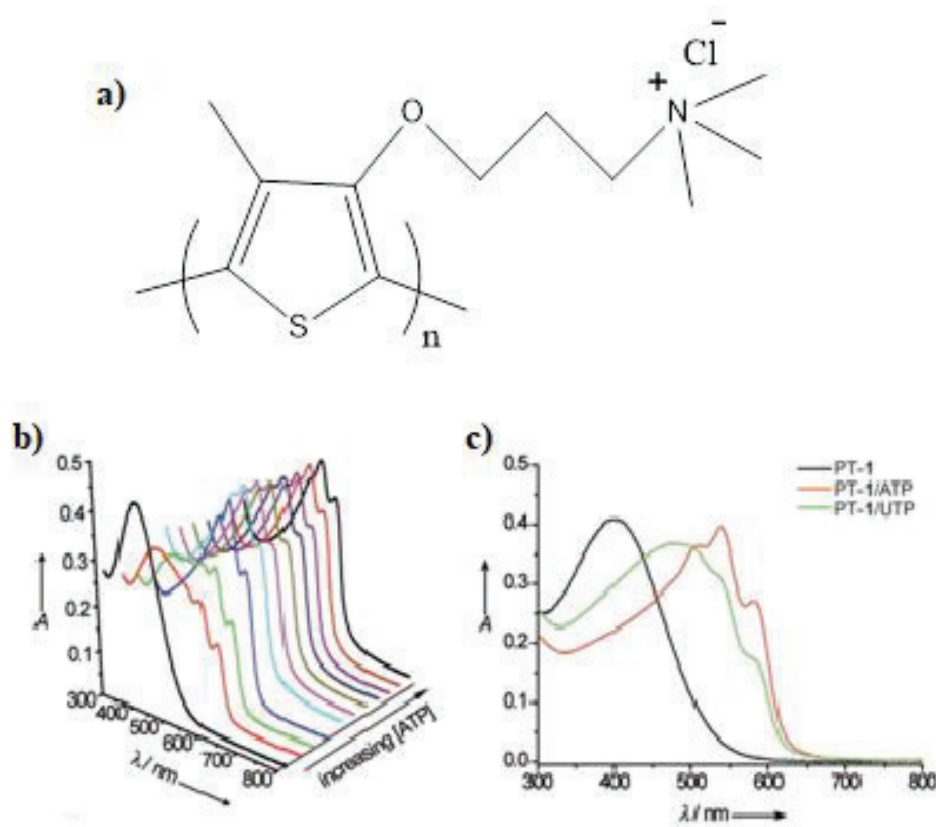
Figure 1.5. a) Chemical structure of PT-1, b) absorption spectra of PT-1 increasing ATP concentration, and c) absorption spectra of PT-1, PT-1/ ATP, PT-1/UTP (Source: Li et al. 2006)

Li et al. and coworkers paid attention to the complexation of optically active PTs derivative PT-1which is poly (N, N, N-trimethyl-3-((4-methylthiophen-3-yl) oxy) propan-1-aminium) (Figure 1.5. a) Chemical structure of PT-1, b) absorption spectra of PT-1 increasing ATP concentration, and c) absorption spectra of PT-1, PT-1/ ATP, PT-

1/UTP with small bio-anionic molecules such as adenosine triphosphate (ATP), adenosine diphosphate (ADP), uridine-5'-triphosphate (UTP). An only PT-1 (as a random coil) showed maximum absorption at 400 nm in water and color of solution was yellow (Figure 1.5c). Then, ATP (different concentration) was added, its maximum point was red-shifted to 538 nm, also the color of solution turned to purple (Figure 1.5b). In their opinion, this changing color and red shift in the absorption spectra was related to conformation (it may be more planar) and strong hydrophobic and electrostatic interactions (Li et al. 2006).
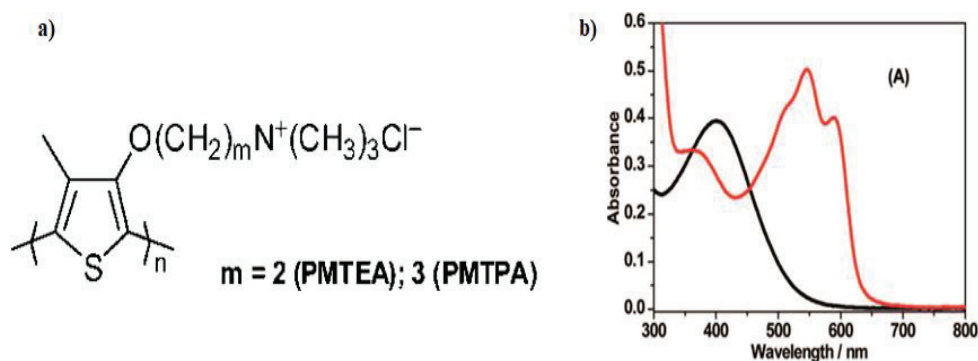
Figure 1.6. a) Chemical structures of PT derivatives, and b) absorption spectra of PT, PT/folic acid (Source: Yao, Li, and Shi 2008)

Water soluble and cationic PT derivative which is poly (N, N, N-trimethyl-3-((4-methylthiophen-3-yl) oxy) propan-1-aminium) (it is called as PMPTA in Figure 1.6a) and folic acid were formed complex, the result of this complexation the maximum absorbance shifted 400 nm to 546 nm and 588 nm (Figure 1.6b) and also at the same time the color of this solution turned to yellow to purple. They thought that the reason of this change backbone of the cationic PTs had a more planar conformation unlike the initial random coil structure (Yao, Li, and Shi 2008).

In 2011 article, Preat et al. and coworkers studied duplex formation between oxidized (3,4-ethylenedioxythiophene) (EDOT) and Dickerson's dodecamer sequence (5'-CGCGAATTCGCG-3') and examined interactions among them as a molecular level. They have carried out simulations in NPT conditions that is a fixed pressure P, temperature T, and number of atoms N, by taking different initial structures and production part of simulation with 10 ns. They have considered some specific interactions

(e.g. O-H and S-H hydrogen bond, π-π stacking, N-H-π interactions) for analysis of these simulation results (Figure 1.7).
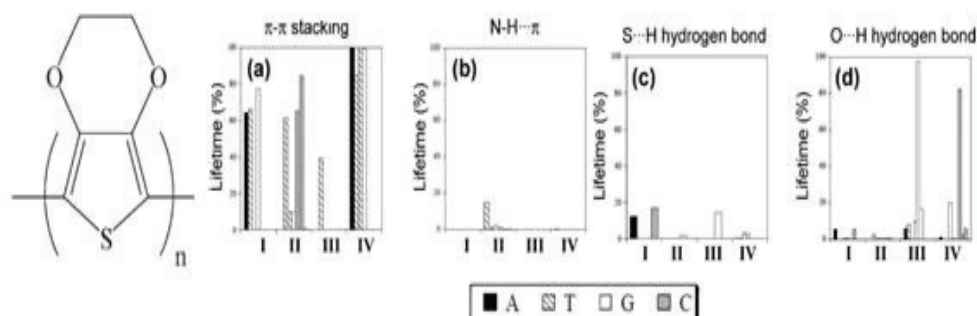
Figure 1.7. Chemical structure of EDOT and specific interactions in duplex (Source: Preat et al. 2011)

Moreover, they put a threshold when transferring these interactions to numerical values:

- Electrostatic interactions, distance between the EDOT units and the phosphate groups is shorter than 6.5 Å,
- Hydrogen bonds; distance is shorter than 3.0 Å,
- π-π stacking; distance is 4.0 Å,
- N-H-π interactions; distance is shorter than 3.0 Å.

Atomistic MD simulations results indicated that thymine and guanine bases were more effective than adenine and cytosine bases in all of these selected interactions; especially, four guanine bases made up hydrogen bonds with poly-EDOT chain in the initial structure where the DNA chain had a helical conformation in which it folded up the PEDOT chain (Preat et al. 2011).

Oligonucleotide	Polymer[a]	Near-UV band		Vis band	
		$\lambda_{abs.max}$ (nm)	$\Delta\lambda_{abs.max}$[b]	$\lambda_{abs.max}$ (nm)	$\Delta\lambda_{abs.max}$[c]
ssDNAd(R)$_{20}$	P3HT-Py	260	-6	490	+26
ssDNAd(T)$_{20}$	P3HT-Py	260,266	-6,0	476	+12
ssDNAd(R)$_{20}$	P3HT-PMe$_3$	261	-5	486	+36
ssDNAd(T)$_{20}$	P3HT-PMe$_3$	268	+2	466	+16
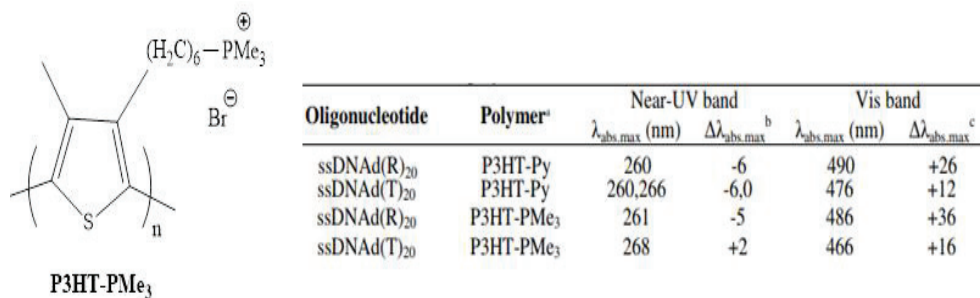
Figure 1.8. Chemical structure of PT and maximum absorption data for duplex (Source: Rubio-Magnieto et al. 2013)

The PTs are unique applicant for biomedical sensors or applications, through labeling cancerous cells in vitro, discrimination of proteins, monitoring anti-cancer drug, the intermolecular interactions with the DNA, optical properties and adjustable structure to identify addressed molecule. Rubio-Magnieto et al. and coworkers synthesized various containing different cationic side groups PTs and then mixed with DNA at a 1: 1 molar ratio. The aim of this study was to investigate the effects of PTs' backbone structures and the behavior of charged groups while binding to DNA. When PTs and DNA formed duplexes, they examined UV-Vis spectra of the duplexes and observed a red shift on the maximum absorption peak from 12 to 36 nm (Figure 1.8). They suggested that the polymer backbone was flattened as mentioned in previous studies (Ho, Béra-Abérem, and Leclerc 2005), and that the observed shift in Near-UV band belonged to the DNA sequence. In addition, containing the purine-rich ssDNA shifted more to the red region than the ssDNA containing the thymine, so they thought that the ssDNA sequence plays important role upon self-assembly with PTs (Rubio-Magnieto et al. 2013).
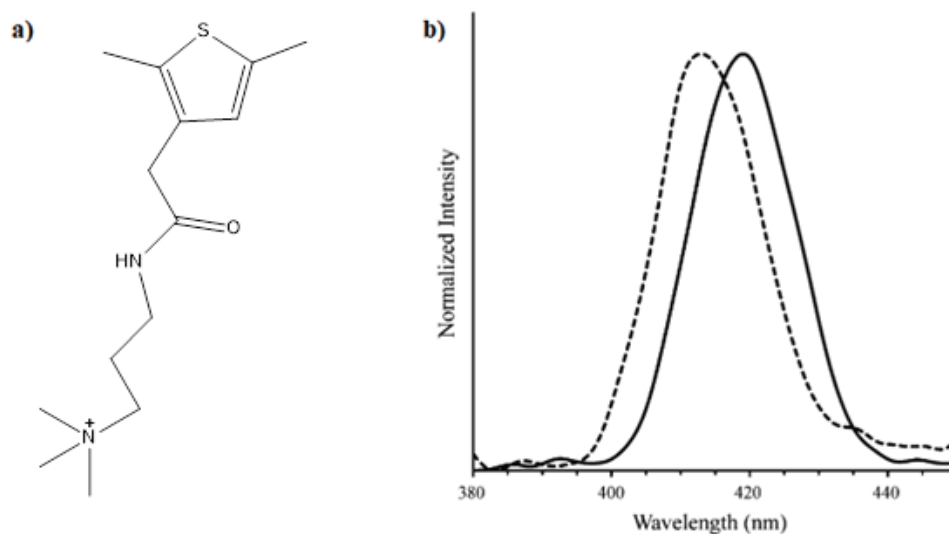
Figure 1.9. a) Chemical structure of polythiophene, b) PT and duplex fluorescence spectrum (Source: Carreon et al. 2014)

To use as a theranostic drug and gene delivery vehicle, a new cationic PT, which was poly(N,N,N-trimethyl-3-(2-(thiophen-3-yl)acetamido) propan-1-aminium iodide) (Figure 1.9a), was synthesized by Carreon et al. and coworkers because they have good quantum yield, photostability, water solubility and low toxicity compared to small molecule fluorophores. Furthermore, cationic PTs which contain tertiary and quaternary

amine are advisable for binding to DNA because their electrostatic interactions are favorable. Thus, they examined the fluorescence property and their fluorescence spectrum (Figure 1.9b) showed lightly red shift (from 413 to 419 nm). The reason for this red shift is thought to be changes in the structure of the imidazolium moiety of PT (Carreon et al. 2014).

Another experimental study performed to quantify the thermodynamic binding constants and the hybridization rates of DNA duplexes in the presence and absence of cationic PT (poly (1H-imidazolium, 1-methyl-3-[2-[(4-methyl-3-thienyl) oxy] ethyl]-bromide)). As they reported their results, the optical responses of the cationic PT are susceptible structural changes to bind ssDNA / dsDNA (Figure 1.3). In their UV absorbance measurements, the maximum value measured in the solution is 260 nm, while the maximum value obtained by adding the ssDNA is 500 nm, this fact was explained as the formation of complex and the shift to the red region due to the electrostatic interactions between them. According to their melting experiments, they thought that this duplex has a more stable structure by looking at the change in the melting curve of the polymer / ssDNA duplex (Zheng and He 2014).

Complex	UV band		Vis band	
	λ_{max} (nm)	$\Delta\lambda_{max}$	λ_{max} (nm)	$\Delta\lambda_{max}$
dT_{20}/CPT	268	+4	494	+43
dA_{20}/CPT	269	+12	442/520[d]	−9/+69
dR_{20}/CPT	261	+4	495	+44
dR_{rev20}/CPT	261	+4	490	+39

Figure 1.10. Maximum absorption data for duplex (Source: Rubio-Magnieto et al. 2015)

The other study of Rubio-Magnieto et al. and coworkers examined all parts of cationic poly[3-(60-(trimethyl phosphonium)hexyl)thiophene-2,5-diyl] that they named CPT or P3HT-PMe$_3$ (Figure 1.8) such as complexation with DNA, hybridization their complementary DNA, and melting experiments. Especially complexation part, for every duplex (e.g. dT_{20}: CPT, dA_{20}: CPT, …) actively exhibit remarkable red shift in their UV-Vis spectrums data (Figure 1.10). A larger number of red shift is dA/CPT (520 nm) compared to dT/CPT (494 nm) based on only-polymer maximum absorption peak (451 nm).

The theoretical part of same study, they were performed molecular dynamics (MD) simulations some duplexes: dT_{20}: CPT, dA_{20}: CPT, at 300 K and 15 ns. They looked at the radius of gyration plot, it showed that adenine-based ssDNA's line decline and it meant the structure of this complex has a more condensed, dT/CPT was vice versa. Another plot is root mean square deviation (RMSD), it demonstrated dA/CPT complex had more changes than dT/CPT. When the MD simulation's results (e.g. electrostatic interactions, π-type interactions, weak S-H hydrogen bonds) were analyzed as numerical (Figure 1.11), dA/CPT complex had large interaction number than dT/CPT complex (Rubio-Magnieto et al. 2015).
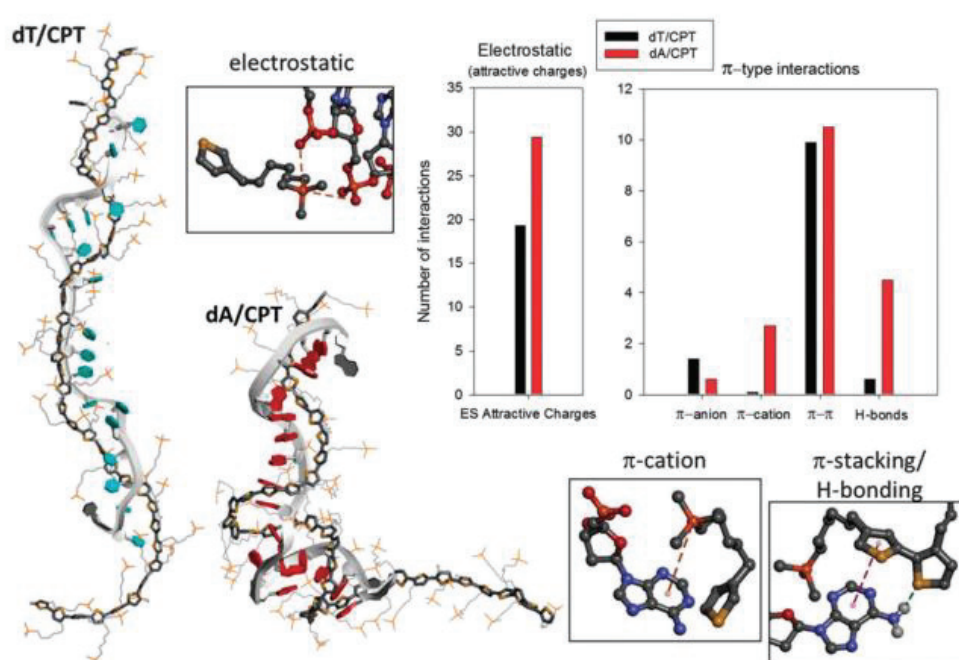
Figure 1.11. MD simulation results for DNA-PT duplex (Source: Rubio-Magnieto et al. 2015)

In another study, cationic PTs and their copolymers were synthesized by Rajwar et al. and coworkers to determine the importance of pendant groups, also microRNA detection as colorimetric in human plasma. This study results were given us PT copolymers have good optical responses for detection of microRNA, they were better than homopolymers. Their UV-Vis spectrum and maximum absorption changes, photochemical and chemical stabilities and excellent optic/electronic properties of PTs make a sign a good biosensing materials (Rajwar et al. 2016).

1.3.2. Studies on Force Field Parametrization

NAMD is a molecular dynamics program for UNIX platforms designed to perform high performance simulations, mostly in structural biology. Molecular dynamics simulations of atomic orbits are formed by solving the equations of motion using numerical force fields corresponding to the real atomic force of the biopolymer systems like the CHARMM (Chemistry at Harvard Macromolecular Mechanics) force field. Various versions of the CHARMM force field are available (Bernardi et al. 2018). For instance, CHARMM22 (all atom), CHARMM19 (united atom) for proteins, and also their atomic partial charges with quantum chemical calculations were calculated by using interactions between compounds and TIP3P explicit water model (Preat et al. 2011). Simulation of lipids, DNA and RNA are chosen one; CHARMM27 all-atom force field. It is calculated by empirical force field calculations for exploration of nucleic acids (Mackerell, Feig, and Brooks 2004). A general force field for drug-like molecules (CGenFF) was developed in 2009. It included various of chemical members such as biomolecules and drug like molecules. Any integration of chemical groups is in the general force field, this provides diminishing accuracy of their results. There are important warnings for the use of CGenFF parameters for molecules with special force fields (Vanommeslaeghe et al. 2009). These force fields created by the A. D. MACKERELL JR and his team can also be used in other molecular dynamics programs that support them like NAMD, GROMACS etc. Considering the literature examples, the parameters required to perform a molecular dynamics simulation are available for most biomolecules (proteins, nucleic acids, lipids, etc.) and many other biologically compatible small chemical compounds.

CHAPTER 2

COMPUTATIONAL METHODS

2.1. Molecular Dynamics Simulations

The use of computers to mimic the dynamic responses of one system by the behavior of another system which is described by a model or a mathematical description, is called as computer simulation. As a first, in the Manhattan Project in World War II the scientists used it to mimic the process of nuclear detonation using a Monte Carlo (MC) algorithm. Fermi, Pasta, Ulam and Tsingou developed another technique which is molecular dynamics (MD) in 1955, although the Monte Carlo simulation achieved high success.

MD is popular computer simulation method that model the behavior of molecules in a system what they do in real life over the time using Newton's laws. MD contains a pathway and includes dynamical properties of the system such as time-dependent responses to perturbations, transport coefficients, rheological properties and spectra. This feature gave advantage of MD comparative to MC (Allen and Tildesley 1989; Frenkel and Smit 2002).

2.1.1. History of MD

In the mid of 1950s, MD methods were originally designed in the theoretical physics groups. Alder and Wainwright performed the oldest MD simulation using the hard-sphere model, where atoms interact only with perfect collisions in 1957 (Alder and Wainwright 1957). To simulate real atomic interactions, Rahman applied a smooth, uninterrupted potential (Rahman 1964). In 1970 and later, while the development of computers, simulations were applied for more complex and large systems like protein, DNA, and macromolecules. The use of computer simulations in large systems has increased in more areas such as material science, biophysics, biochemistry etc. (Adcock

and McCammon 2006; Adcock and Mccammon 2008). Rahman and Stillinger were done first molecular dynamics simulations for a realistic system of liquid water. (Stillinger and Rahman 1974). In 1977, the simulation of the bovine pancreatic trypsin inhibitor was done as a first protein simulation (McCammon, et al. 1977). Up to now, predicting and understanding biological processes and complex chemical reactions was a challenging process, and molecular dynamics gained importance in order to overcome this challenge. The 2013 Nobel Prize in Chemistry was awarded jointly to Martin Karplus, Michael Levitt and Arieh Warshel for the development of multiscale models for complex chemical systems. This event highlighted the importance of MD simulations in providing unprecedented perspectives to complex chemical and biological systems (Chmiela et al. 2018).

2.1.2. The basic idea of MD

In some situations, doing the experiment may be impossible, high cost, and too dangerous (weather forecast, explosive reactions, under the high-pressure reactions), so simulations are a useful computational method. When simulations are performed, they may stimulate, explain and complement experiments.

What atoms are doing under certain conditions, MD stimulates these conditions with force field created by Newton's laws using potential energy definition. Generally, by solving the equations of motion of Newton (Equation 2.1) numerically for a system of interacting particles, the trajectories of molecules and atoms are obtained. Using interatomic potentials or molecular mechanics force fields, forces between these particles and their potential energies are evaluated.

$$\boldsymbol{F_i} = m_i \times \boldsymbol{a_i} = m_i \times \frac{d\boldsymbol{v_i}}{dt} = m_i \times \frac{d^2\boldsymbol{r_i}}{dt^2} = -\boldsymbol{\Delta_i}V \qquad (2.1)$$

Equation 2.1 shows the Newton's equations of motion. The force exerted on particle i, mass of particle i, and acceleration of particle i, represent as F_i, m_i, a_i, respectively. ΔV is the gradient of the potential energy of the system. Newton's equations

of motion formula prove that the derivative of the potential energy is related to the change in positions as a function of time (t).

MD requires initial positions and velocities of atoms. Then, it computes the momentum their velocities and masses. New positions of the atoms will have computed a short time later, it is called time step. Using the information obtained in these stages, Newton's equations of motion are solved and new velocities and accelerations for atoms are also calculated. These processes are repeated until the system reaches equilibrium (has a proper configuration for the given amount of energy). When the system reaches equilibrium, atomic coordinates saved every defined time. This saving of coordinate is called a trajectory (Table 2.1.). The trajectories are analyzed to obtain information about the system, for that reason of this to continue repeating and recording data until the data is collected (Kimber et al. 2011).

Table 2.1. Molecular Dynamics Basic Algorithm

Energy Calculation (Force Field)	Forces	Numerical Integration		
$V_{pot}(\mathbf{r}_i)$	$\mathbf{F}_i = -\dfrac{dV_{pot}}{d\mathbf{r}_i}$	$\mathbf{F}_i = \dfrac{d\mathbf{p}_i}{dt}$ $\mathbf{F}_i = m_i \times \dfrac{d\mathbf{v_i}}{dt}$ $\mathbf{F}_i = m_i \times \boldsymbol{a}_i$ $\boxed{\boldsymbol{a}_i = \dfrac{F_i}{m_i}}$	$\dfrac{F_i}{m_i} = \dfrac{d\mathbf{v_i}}{dt}$ $\dfrac{F_i}{m_i} \times dt = d\mathbf{v_i}$ $\dfrac{F_i}{m_i} \times dt = \mathbf{v_f} - \mathbf{v_i}$ $\mathbf{v_f} = \mathbf{v_i} + \dfrac{F_i}{m_i} \times dt$ $\boxed{\mathbf{v_f} = \mathbf{v_i} + \boldsymbol{a}_i \times dt}$	$\dfrac{d\mathbf{r}}{dt} = \dfrac{\boldsymbol{p}}{m}$ $\boldsymbol{p} = m \times \mathbf{v}$ $\mathbf{v} = \dfrac{d\mathbf{r}}{dt}$ $\mathbf{v} \times dt = d\mathbf{r}$ $\mathbf{v} \times dt = \boldsymbol{r}_f - \boldsymbol{r}_i$ $\boxed{\boldsymbol{r}_f = \boldsymbol{r}_i + \mathbf{v}_f \times dt}$
Trajectory				

Table 2.1. shows the basic algorithm of the molecular dynamics. E_{pot} potential energy; t, simulation time; dt, iteration time; For each spatial coordinate of the N

14

simulated atoms (i): x, atom coordinate; F, forces component; a, acceleration; m, atom mass; v, velocity; p, momentum.

The trajectory of a system of interest; the time evolution of atomic positions, velocities; at microscopic level are obtained from MD simulations. Then using the statistical mechanics this information is used to get macroscopic properties such as energy, pressure, etc., for instance, to calculate changes in the binding free energy of a drug candidate, or to examine the energetics and mechanisms of conformational change. Some definitions should be well understood in order to establish a relationship between microscopic and macroscopic systems. The thermodynamic state of a system can be defined by the temperature, T, the pressure, P, volume, V, and the number of particles, N. A collection of all possible systems is called as an ensemble which have different microscopic states but have an identical macroscopic or thermodynamic state. Some different ensembles are described in below table.

Table 2.2. Some ensembles and their properties

Microcanonical ensemble (NVE)	a fixed number of atoms, N, a fixed volume, V, and a fixed energy, E.
Canonical Ensemble (NVT)	a fixed number of atoms, N, a fixed volume, V, and a fixed temperature, T
Isobaric-Isothermal Ensemble (NPT)	a fixed number of atoms, N, a fixed pressure, P, and a fixed temperature, T.
Grand canonical Ensemble (μVT):	a fixed chemical potential, μ, a fixed volume, V, and a fixed temperature, T.

2.2. The Force Field

The mathematical function that gives the energy of the system as a function of the conformation of that system is termed as a force field. It is assumed that a molecule

consists of atoms held together by elastic forces, and these forces are written in terms of potential energy functions (bond lengths, bond angle, non-bonded interactions etc.). So, the force field becomes a combination of the potential energy terms as given in the Equation 2.2. Note that sometimes the force field called as potential (Guha 2006).

$$E_{TOTAL} = \underbrace{V_{BOND} + V_{ANGLE} + V_{DIHEDRAL}}_{E_{BONDED}} + \underbrace{V_{ELECTROSTATIC} + V_{VANDERWAALS}}_{E_{NON-BONDED}} \qquad (2.2)$$

Potential energy functions for force fields can be divided into two parts: bonded (bond, angle, dihedral) and non-bonded (electrostatic, van der Waals) energy terms (Equation 2.2.).
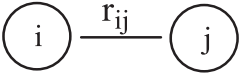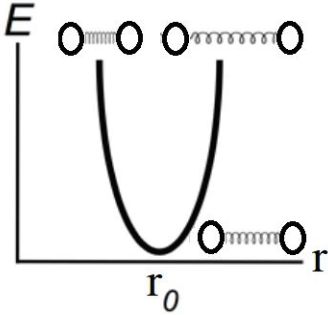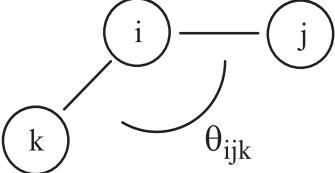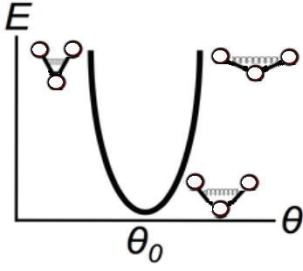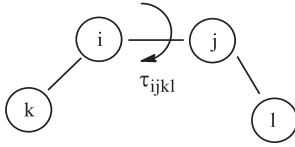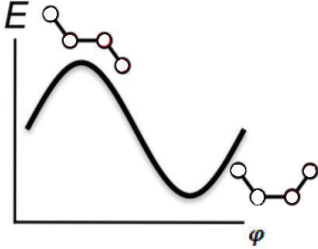
Two-body spring bond, three-body angular bond, four-body torsion (dihedral) angle are potential terms which symbolized interactions of covalently bonded atoms. All the atomic pairs are the non-bonded potential (Lennard-Jones and electrostatic potential etc.) terms including the interaction between (i, j) except for the atomic pairs that are generally incorporated in an already linked term. The equations of all the potential terms are given in the below Table 2.3.

There are many types of force fields, some of them come from the experimental sources (X-Ray diffraction) and the others come from theoretical studies (quantum mechanical (QM) calculations). They are generally used similar equations, but they contain some differences in the equations for generating parameters. AMBER (Assisted Model Building with Energy Refinement), CHARMM (Chemistry at HARvard using Molecular Mechanics), GROMOS (GROenigen Molecular Simulation), OPLS (Optimized Parameters for Large-scale Simulations), MMFF (the Merck Molecular Force Field) are some well-known force fields.

2.2.1. CHARMM Force Field
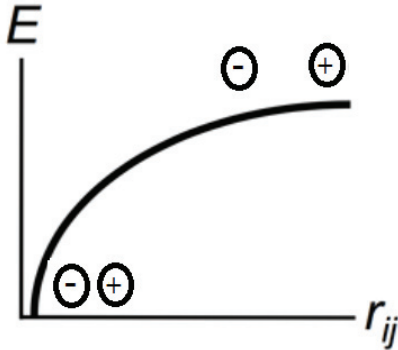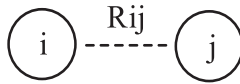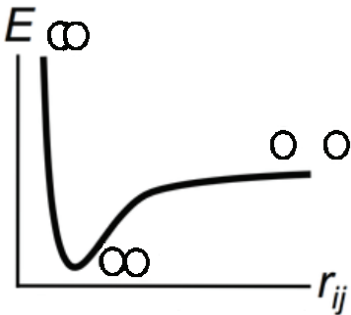
Chemistry at HARvard using Molecular Mechanics (CHARMM) is a computer software, which is quite common for molecular dynamics and which includes and generates force fields for the realization of the simulation. It has been developed by Martin Karplus and his group since 1983.

Table 2.3. Potential Energy Functions

Potential energy functions
Bonded potential energy terms

$$V_{bond} = k(r_{ij} - r_0)^2$$

shows the harmonic vibrational motion between an (i, j)-pair atoms,

$r_{ij} = \|r_j - r_i\|$ gives the distance between the atoms,

r_0 is the equilibrium distance,

k is the spring constant.

$$V_{angle} = k_\theta(\theta - \theta_0)^2$$

shows the angular vibrational motion between an (i, j, k)-triple atoms,

θ is the angle in radians between the atoms,

θ_0 is the equilibrium angle,

k_θ is the angle constant.

$$V_{dihedral} = k(1 + \cos(n\varphi + \emptyset))$$

shows the angular spring between the planes formed (i, j, k, l)-quadruple atoms by the first three and last three atoms,

φ is the angle in radians between the (i,j,k)-plane and the (j,k,l)-plane,

The integer constant n is nonnegative and indicates the periodicity,

$\emptyset$ is an equilibrium angle,

k is the multiplicative constant.

Table 2.3 (cont.)

Non-bonded potential energy terms	
$$V_{elec} = \epsilon_{14} \frac{Cq_iq_j}{\epsilon_0 \, r_{ij}}$$ The electrostatic potential $r_{ij} = \|r_j - r_i\|$ gives the distance between the pair of atoms, q_i and q_j are the charges on the respective atoms, C is the Coulomb's constant, ϵ_0 is the dielectric constant, ϵ_{14} is a unitless scaling factor	$$V_{LJ} = -E_{min}\left[\left(\frac{R_{min}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{min}}{r_{ij}}\right)^{6}\right]$$ The Lennard–Jones potential* accounts for the weak dipole attraction between distant atoms and the hard-core repulsion as atoms become close, $r_{ij} = \|r_j - r_i\|$ gives the distance between the pair of atoms.
*The simplest non-bonded interaction is the van der Waals interaction. In NAMD, van der Waals interactions are always truncated at the cutoff distance, specified by cutoff.	

The MacKerell Laboratory is actively updating and developing the CHARMM force fields by using and coordinating it. The CHARMM force fields include a wide database for biological molecules like peptides, proteins, prosthetic groups, small molecule ligands, nucleic acids, lipids, and carbohydrates etc. Examples of various versions of the CHARMM force field are CHARMM36 for proteins, CHARMM22, CHARMM27 for lipids and nucleic acids, respectively. The form of the potential energy function used is given by the following equation (Bernardi et al. 2018).

$$V = \sum_{bonds} k_b(b - b_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} k_\varphi[1 + \cos(n\varphi - \delta)]$$

$$+ \sum_{impropers} k_\omega(\omega - \omega_0)^2 + \sum_{Urey-Bradley} k_u(u - u_0)^2 \qquad (2.3)$$

$$+ \sum_{nonbonded} \epsilon\left[\left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{12} - \left(\frac{R_{min_{ij}}}{r_{ij}}\right)^6\right] + \frac{q_i q_j}{\epsilon_o r_{ij}}$$

In addition, the parameters required to perform a molecular dynamics simulation are available for biomolecules and many other biologically compatible small chemical compounds in CHARMM database. In the case of undefined molecules, parameters can be calculated using some ready-to-use plugins that also use QM calculations. If there is a molecule that is not defined in the system and it is desired to create the parameter (charges, bonds, angles, and dihedrals) of this molecule in a format compatible with CHARMM, then one of the plugins you can use ffTK (Force Field Toolkit) (Ecker 2016; Pavlova and Gumbart 2015; Mayne et al. 2013).

2.2.2. Generating Force Fields Using ffTK

When we look at the recent studies, the ffTK plugin has become quite popular for getting the force fields compatible with the CHARMM. In 2015, Pavlova and Gumbart generated macrolide antibiotics parameters using the force field toolkit. They divided molecule into 40 or less atom, this provided to ease QM calculations. Thus, it can be parametrized large molecules (Pavlova and Gumbart 2015). Another study, CHARMM-compatible force field parameters were generated by ffTK for atrazine molecule. In this way, atrazine molecule can be studied with other biomolecules in terms of their interactions using MD simulations (Ecker 2016). For cobalamin compatible- CHARMM force field parameters were generated by Pavlova, Parks and Gumbart by using ffTK in 2018. All bonded terms were optimized with the ffTK and they were validated by comparison with experimental data (x-ray crystal structures). This study showed that parametrization gave better agreement with the refences data and parametrization of

corrinoids or like large molecules with ffTK can be a suitable method (Pavlova, Parks, and Gumbart 2018).

With the ffTK plugin, the parameter creation flow includes the following steps and calculations. The initial structure was created with Avogadro, the formatted structure file and coordinate file are generated Molefacture and ffTK plugin, the targeted QM data is currently produced by Gaussian09 program.

➢ Preparation molecule

The partial charges of non-polar hydrogens are fixed to +0.09 for aliphatic and +0.15 for aromatic in CHARMM-compatible force field. While aromatic carbons not adjacent to a heteroatom are assigned the standard charge of -0.15 (Pavlova and Gumbart 2015). The remaining charges of the molecule are set to zero. Using Molefacture plugin in VMD, molecule's pdb and psf file are generated as properly formatted. All steps after this step will be performed in ffTK, except for QM calculations (are done with Gaussian09 program).
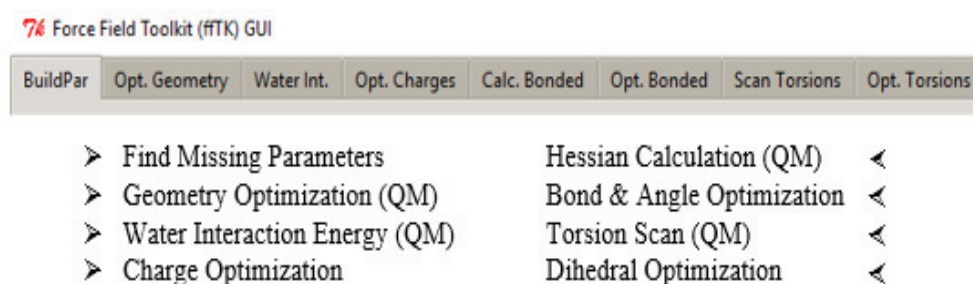
Figure 2.1. The illustration of Force Field Toolkit steps

➢ Determination of the missing vdW / LJ parameters

The topology and parameter files which are related to CHRAMM database, are selected and cross-checked after that, missing parameters are found. For example, if the structure is protein, CHARMM36 will be picked. Then, those with unknown parameters are identified and compared with existing parameters. Thus, the first parameter file is created with the existing information in the database.

➢ Optimization of the geometry

Gaussian09 software package is used for first QM calculation in this step. A geometry optimization run is given at MP2/6-31G* level of theory via Gaussian09 package. A low energy conformation is obtained, also it will be used for the following

calculations. As result of this calculation, new optimized coordinates will be written a pdf file.

> Interaction energy with water

When the non-polar hydrogens' charges are determined in the preparation molecule step, the other atoms' charges are fitted to interactions with explicit TIP3P water molecules. The structure has one or more interaction sites are divided into three categories: a) hydrogen bond donors (interact with oxygen of water molecule favorably), b) hydrogen bond acceptors (interact with hydrogen of water molecule favorably), and c) both. According these interaction sites, water molecules are placed to a possible hydrogen bonding geometry. In order to determine the partial charges of the remaining atoms of the molecule, two-dimensional optimization, which optimized both the distance between the water and its interaction partner as well as the rotation angle of the water about that hydrogen bond, is done with HF/6-31G* level of QM calculations.

> Charge optimization

After determination of water interaction sites manually, the first part of charge optimization step is done by the Gaussian09 program. The second part of charge optimization is manually set charge constraints or push the Guess button for optimized atoms except fixed charges atoms. The users can be assigned the total and net charge of the molecule. Then, optimization run is started until between the first and second iterations indicate negligible differences. The resulting charges are printed on a new psf file.

> Hessian calculation

QM calculations of the Hessian are the second derivative of potential energy function. These calculations reproduce the potential energy surface by distortions along the angles and bonds. MP2/6-31G* level of theory is applied to optimized structure for Hessian calculations fitting the targeted bonds and angles data.

> Bond & Angle optimization

Bonds and angles are expressed using a simple harmonic potential. NAMD is performed a short geometry optimization in the background during the ffTK optimization. To populate the parameters for optimizing, push the Guess button. The first change ensures good fit of MM-optimized geometry and QM-optimized geometry. The optimization iterations perform until increasing the current final objective value. When

bond and angle optimization are done, parameter file (in-progress) is updated and proceed to the next step.

> ➢ Scanning of the dihedrals

Dihedrals are defined as the multiplicity that each containing three parameters (periodicity (n), force constant (k), and phase shift angle (δ)). One must decide to the multiplicity and periodicity, but the force constant and phase shift angles will be optimized by ffTK. There are two methods that specify each dihedral scan. One method which is simple, is to push Read from PAR button. The parameter file from the previous step uses in this step. ffTK describes all dihedral of the input structure, eliminates the dihedrals of terminated hydrogen, and unnecessary dihedrals which are shared bond torsion are filtered by this method. The second method is to push Add button that dihedrals can be added manually. QM calculations to be made after deciding the dihedrals, the relaxed PES (potential energy surface) scan performed at the default MP2/6- 31G* level of theory. In this calculation, the dihedral is fixed and for the other part of the molecule, a geometry optimization is performed.

> ➢ Dihedral optimization

In this step, the last change ensures good fit of a multidimensional optimization MM calculations and high-level QM calculations. PES profiles (QME is the target and MME is the computed) compare and refine the dihedral parameters until showing a good match to QME. (Pavlova and Gumbart 2015; Mayne et al. 2013; Ecker 2016; Mayne, Muller, and Tajkhorshid 2015).

2.3. Nanoscale Molecular Dynamics (NAMD)

In structural biology, Nanoscale Molecular Dynamics is a parallel MD program for Unix platforms designed to perform high performance simulations. It is generally named as NAMD that has been developed by the Theoretical and Computational Biophysics Group (TCB) and the Parallel Programming Laboratory (PPL) at the University of Illinois at Urbana–Champaign. It is often used to mimic large systems which contain millions of atoms. To analyze trajectories and setup simulations, NAMD uses the VMD (is the more popular graphics program). It also uses files which are different, compatible formats like AMBER, CHARMM, and X-PLOR.

At least four things are required to initiate any molecular dynamic simulation using NAMD:

⟹ Protein data bank file (.pdb): It includes initial coordinates or velocities of the atoms that make up the system. These files can be obtained from the Protein Data Bank (http://www.pdb.org) or from other sources. It is important to be able to identify the atoms and their positions in three-dimensional (3D) space to calculate the movement of individual molecules under the influence of a force. Coordinate files contain only such information and can be used in many formats. In this study, coordinate files will be obtained with Avogadro program.

⟹ Protein Structure File (.psf): The structure of molecules of interest to all the information (name, charge, mass of the all atoms) is available. In this study, the required X-PLOR compatible-file will be obtained with the Automic PSF Generation (AutoPSF), which is an extension of VMD.

⟹ Parameters File (.par, .inp, etc.): The parameter of the CHARMM potential function is used in NAMD. For the missing potential parameters; the force constants (bond, angle and dihedral) required for the atoms will be calculated in accordance with the CHARMM database.

⟹ Configuration File (.namd, .conf): It is the file where all the conditions required to start and continue any molecular dynamics simulation are stored. The configuration files that will be created to make the simulations can be as follows: firstly, a configuration file is prepared to minimize the system. The aim is to reconcile the used force field with the observed structure, and the temperature is equal to zero. After the minimization is completed, it is aimed to increase the temperature of the system to the desired temperature in next configuration file created to perform the heating operation. Then another configuration file is created to guarantee that the system is stable, that is, the balance is reached, and this step is called the equilibrium step. Finally, the last configuration file is prepared to simulate the system under the desired conditions (NVT, NPT, etc.) and to collect the data which is obtained (Figure 2.). Each configuration file created consists of different options such as time steps (timestep), cut-off distance, temperatures, pressures, box sizes, periodic boundary conditions, thermostats, barostats etc.

2.4. Visual Molecular Dynamics (VMD)

A program, which is a molecular visualization and molecular modelling, is called as Visual Molecular Dynamics (VMD) for displaying, animating, and analyzing large biomolecular systems for MacOS X, Unix, or Windows platforms. It has been developed by Klaus Schulten in TCB group at the Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana–Champaign. VMD supports the visualization for 3-D graphics and built-in scripting. It also includes some plugins to view and analyze the results of MD simulations because it gets help from Tcl/Tk (is a dynamic programming language/is a toolkit for building a graphical user interface (GUI) in many programming languages).

2.5. Analysis Methods

MD has several steps as described and their results are analyzed by using some programs which hand writings that will be created by the project team, NAMD (Nanoscale Molecular Dynamics), VMD (Visual Molecular Dynamics) and its plugins, and Gaussian09. Firstly, drawing the initial geometry and define the initial coordinates of system. Second, preparation of needed files to run the simulation (parameter files, topology files, configuration files etc.). Then, when the simulation is done, trajectory files are obtained from the simulation program NAMD. The final step is the analysis of these trajectory files as conformational, interactions between selected or all atoms, and their thermodynamic and kinetic properties.

General description of RMSD (root mean square deviation) is measuring of the differences between two values which are predicted and observed. It is a comparative tool for two differ or similar atomic structures. When RMSD calculated for all atoms of molecule, interested or specific region of molecule (active site, surface loop, hydrogen atoms) is given attention and analyzed easily. Especially biomolecular simulation, RMSD calculation provides convenience for comparing pairs of structures, class with ensembles, examining clusters etc. (Gromiha 2011; Pitera 2014).

Formula of RMSD is given below:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\vec{r}_{1,i} - \vec{r}_{0,i})^2} \qquad (2.4)$$

At the end of the minimization step, RMSD will be calculated, and the RMSD graph (RMSD vs. time step) is almost fixed, minimization is enough.

In the heating process, the last frame from the minimization trajectories will be taken and the desired temperature will be reached, so this process will be followed by temperature - time step graph.

In the next step, the pressure/volume will be specified, and equilibration will be started, the system will be observed with the potential energy – time step graph.

At the last step of MD, the simulation step will be started and collected the data from obtaining trajectories. Conformal changes in the systems will be monitored by using the successive display of the course.

The interactions may cause structural changes of the oligomers will be examined in each snapshot obtained during the simulation. For electrostatic interaction, how many oligomer anions and cations interact in these snapshots. Maximum length of the anion and cation interacting will be 0.65 nm. This method will be used to plot H-bonds, π–π stacking, π-cation, π-anion interactions, and the lengths' threshold at which the interactions start will be taken as 0.3 nm, 0.4 nm and 0.65 nm, 0.65 nm respectively (Rubio-Magnieto et al. 2015; Preat et al. 2011).

The radius of gyration (Rg) refers to the center of mass of selected atoms or the mean square distance from a given axis. It is suitable for identifying branched chains and defining the dimensions of a polymer chain can also be used to measure the degree of folding in the chains.

$$R_g = \sqrt{\frac{\sum_{a=1}^{N} m_a(r_a - r_{com})^2}{\sum_{a=1}^{N} m_a}} \qquad (2.5)$$

Table 2.4. Definitions and thresholds of interactions between PT and ssDNA

	Image of Interactions:	Interacted Atoms:	Distance: (d)
Electrostatic interactions (N⁺- O⁻)		PT's nitrogen atom from ammonium group and ssDNA's oxygen atom from phosphate groups	Shorter than 6.5 Å
π-type interactions (π- cation)		PT's nitrogen atom from ammonium group and nucleobases 6 membered rings	Shorter than 6.5 Å
H - bonding	**H-O** The hydrogen atom from the PT and oxygen atom from the ssDNA	**H-bond**	Shorter than 3.0 Å
	O-H The oxygen atom from the PT and hydrogen atom from the ssDNA		
	S-H The sulfur atom from the PT and hydrogen atom from the ssDNA		

Under the simulation conditions, the condensed/condensation or expanded/expansion of oligomer chains will be examined by using radius of gyration. End-to-end distance for oligomer will be evaluated.
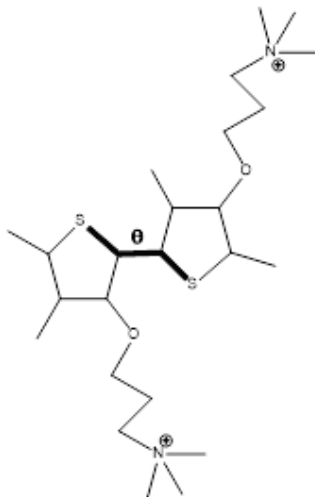
Figure 2.2. Presentation of dihedral angle
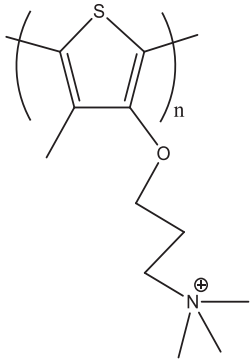
$$P_\theta = \frac{1}{N_\theta} \sum_{i}^{N_\theta} \frac{||\theta_i| - 90|}{90} \tag{2.6}$$

The dihedral angles (Figure 2.2) affecting the planarity of the structure will be analyzed based on the following formula (Sjöqvist et al. 2014). P_θ will be used as a planarity index, if the P_θ is a close to 1, it will be said that PT is a planar. If the P_θ is equal to zero, PT is random coil form.

2.6. Computational Details

In this work, a monomer (N, N, N-trimethyl-3-((4-methylthiophen-3-yl) oxy) propan-1-aminium) represents a repeating unit and each monomer has cationic part (ammonium group). 20-mer (the size that is the number of repeating units of oligomer is 20) is used in this thesis and symbolized as P0. There are 4 different ssDNA strands consisting 20 nucleobases which are shown in Table 2.5.

Table 2.5. Chemical Structure of cationic monomer and Sequences of the ssDNA

P0	
n=20	**Ade strand:** 5'- AAA AAA AAA AAA AAA AAA AA-3'
	Thy strand: 5'- TTT TTT TTT TTT TTT TTT TT-3'
	A_rich strand: 5'- CGT CAC GTA AAT CGG TTA AC -3'
	T_rich strand: 5'- GTT AAC CGA TTT ACG TGA CG -3'

In type of sequence, the containing only adenine base is homopurine strand and only thymine base is homopyrimidine strand, they are named as Ade strand and Thy strand, respectively. A_rich strand consists of a mixture of homopurine (6 adenine, 4 guanine) and homopyrimidine (5 thymine, 5 cytosine). T_rich strand consists of a mixture of homopurine (5 adenine, 5 guanine) and homopyrimidine (6 thymine, 4 cytosine). All initial structures were built using Avogadro. For MD simulations, P0 and ssDNA (Ade, Thy, A_rich, T_rich strands, respectively) form a complex, their names are P0-Ade, P0-Thy, P0-A_rich, P0-T_rich.

QM target data were calculated by Gaussian09 program. MM and MD simulations were studied with NAMD package. NAMD was also used for generating MD trajectories. All-atom force field provided to calculate the energy. The parameters for the P0 was created with contributions one of our research group member, Erman KIBRIS, and the parameters for the ssDNA were taken from CHARMM libraries (CHARMM36 for nucleic acids and CHARMM27 for TIP3 waters.). Each of complexes and only P0 were placed in the center of cubic simulation box with dimensions 90.0 x 90.0 x 90.0 Å, water added to fill the boxes. The water molecules were chosen TIP3 model. There is no addition of ions (counter ions) in these simulations because they have already reached the electrically neutral. The number of atoms in the systems are given the below Table 2.6.

For all the ssDNA/P0 complexes and P0, their energies were minimized by performing 20 ns simulation. After the minimization step, the systems were heated from 0 K to 310 K, gradually. Along 20 ns, MD was performed with (N, V, T) ensemble at a time step 2 fs, cut-off was set to 14 Å. Then, a 60 ns (N, P, T) ensemble production

simulations were carried. During the production step, the coordinates and energies were saved in every 20 ps (It provides that all steps have 1000 frame).

Table 2.6. Number of atoms in the systems

Systems \ Numbers	P0	P0-Ade	P0-Thy	P0-A_rich	P0-T_rich
Oligomer	642	642	642	642	642
Nucleobases	-	643	643	637	640
Water	67734	65366	65513	65462	65561

CHAPTER 3

RESULTS AND DISCUSSION

The aim of this thesis study is generating force field parameters for thiophene oligomers (P0) to perform MD simulations to;

i.   investigate interactions between ssDNA and cationic P0,

ii.   deduce the effects of different sequences of ssDNA on the response of P0,

iii. explore how P0 conformations are changing to response facing to ssDNA in order to provide a theoretical explanation for the change of experimental UV-VIS spectra of CPT when ssDNA is added to the solution.

3.1. Force Field Parameters

In the literature, the parameters of the cationic thiophene oligomer (in given Table 1) do not exist, so the unknown parameter sets which are necessary to perform MD simulations are calculated using QM calculations as mentioned in Generating Force Field Using ffTK part and they fit to the MM calculations.
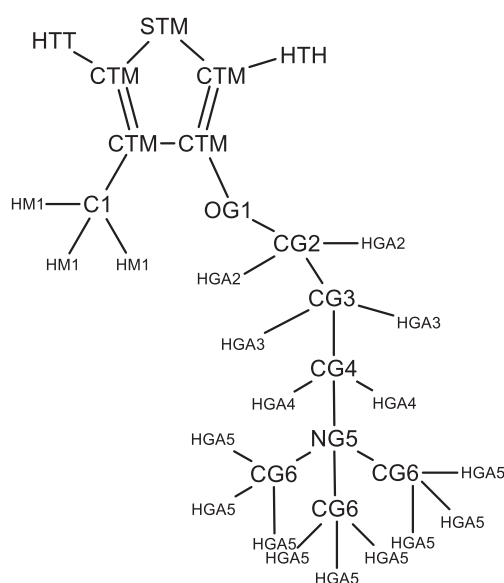

Figure 3.1. Naming atoms of monomer

With the four types of required files (.psf, .pdb, .par, .top), this oligomer can be part of CHARMM-based molecular dynamics simulations, so that interactions with biomolecules (like ssDNA, dsDNA) can be studied. Avogadro, VMD and its plugins (ffTK, Molefacture etc.), and Gaussian09 are programs that are used for generating CHARMM-compatible force field parameters within this work. The monomer of the oligomer was uniquely named as above figure (Figure 3.1).

The LJ/vdW parameters are identified in monomer as its constituent atoms are named in Fig. 3.1, then they compared with existing parameters in the literature, the matched ones are used. The optimum favorable structure of the monomer was obtained at MP2/6-31G* level of theory for obtaining QM data. QM calculations at HF /6-31G* level of theory was used to obtain charges of monomer interacting with a water molecule. Hessian calculations (MP2/6-31G* level of theory) were applied to optimized structure. This QM data is used to fit the MM bonds and angles data (Figure 3.2 and Figure 3.3).

Bond and angle parameters were fitted as described in the computational methods part. 0.02 Å (for bond values) and 3° (for angle values) are recommended maximum deviations for the compatible CHARMM force field (Vanommeslaeghe et al. 2009). The target QM bond and angle maximum deviations values were 0.012 Å and 4° respectively, which are close to the recommended values.
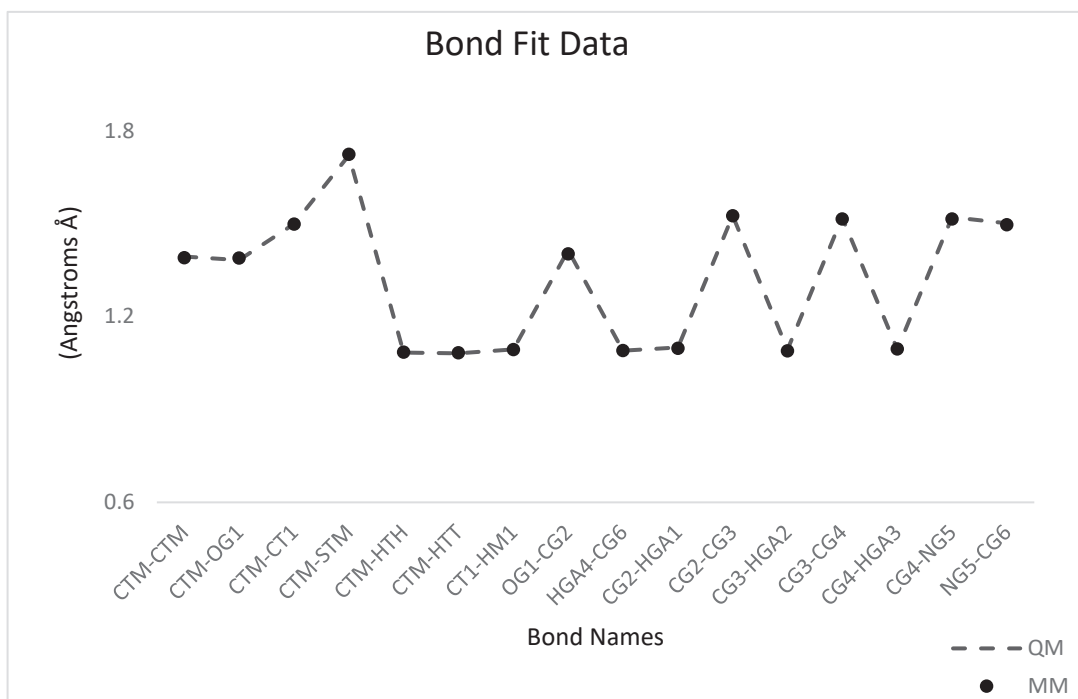

Figure 3.2. QM bond data fit the MM bond data
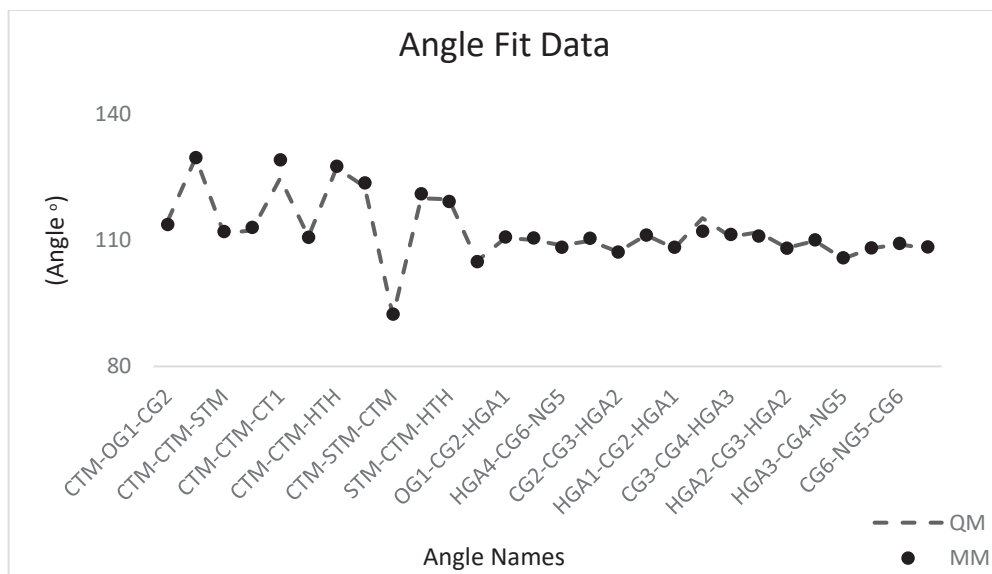
Figure 3.3. QM angle data fit the MM angle data

The dihedrals' QM calculations, the relaxed PES scan performed at the MP2/6-31G* level of theory and refinement was applied to catch good fitting QM-MM data (Figure 3.3). Dihedral scan parameters were fitted as described in the computational methods part. The root-mean square error (RMSE) value was obtained as 0.7870 kcal/mol. After all refinement, the result plot demonstrates that MM data nearly fit to the target QM data. 0.5 kcal/mol is recommended RMSE value for the compatible-CHARMM force field. Although the RMSE was not exact the same as recommended value, it was nearly well-matched (Vanommeslaeghe et al. 2009).

Figure 3.4. QM dihedral data fit the MM dihedral data

3.2. Molecular Dynamics Results

Figure 3.5. RMSDs for minization step

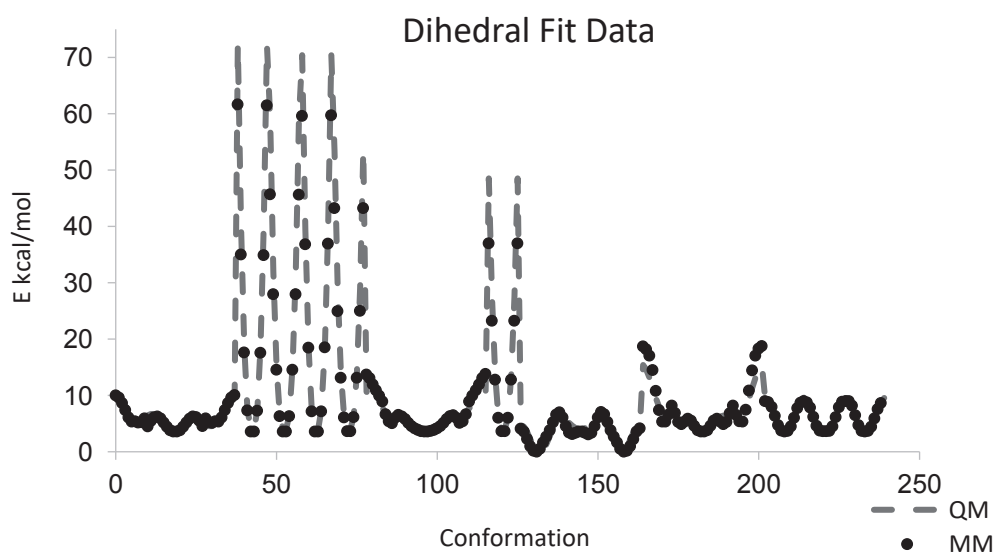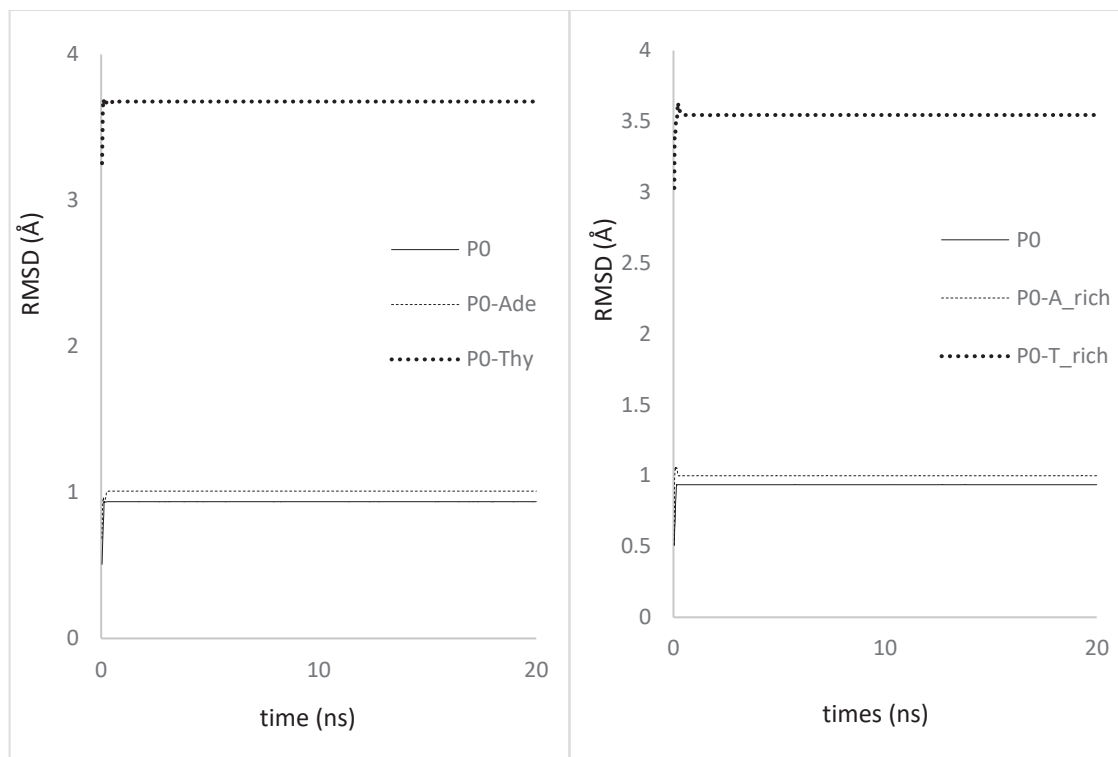Figure 3.5 shows the RMSD graphs for P0 and P0 in the formed complexes at the end of minimization step. It defines the deviations from a defined position in space to a last molecule position, also it gives information about the molecule is conformationally stable or not at the end of run. In another words, if the RMSD is increasing at the graph, the system is not yet stable which means that it is still searching for a lower energy state or a more stable conformation. When looked at the RMSD trends, P0 has a graph that does not display an increasing tendency, so they do not search any lower energy state.

The graphs of radius of gyration (Figure 3.6.) demonstrate the temporal evolution of the distance between the geometric centers of the P0 and ssDNA chains for all complex under study. These values are related to both the bending deformation of the cationic P0 and the corruption in which the DNA chain affects.

In the radius of gyration plots are given in Figure 3.6 using production part of MD simulations. When P0's radius of gyration trends compared; P0 and P0 in P0-Thy complex are similar each other, P0 in P0-Ade complex is different from the others.

Figure 3.6. Radius of gyration for P0; a) in water, in P0-Ade complex, in P0-Thy complex; b) in water, in P0-A_rich complex, in P0-T_rich complex.

The Rg values P0 in water continued to decline throughout the simulation, it has lowest Rg values after 10 ns. It is close to a random coil form and has a more condensed. On the other hand, Rg of P0 in P0-Ade complex is always largest. Whereas Rg of P0 in P0-Thy complex is in between P0 and P0-Ade complex. It can be said that P0 in P0-Ade complex is more stretched than in P0-Thy complex. At the beginning of simulation (about 5ns) all (P0, P0-Ade, P0-Thy) have almost the same Rg, the oligomer have similar structure, in the range of 5-30 ns P0 in P0-Ade complex has the largest one but P0 in P0-Thy complex and P0 have similar Rg's, in 30-60 ns oligomer without ssDNA have the lowest Rg, in the mid time range of simulation Rg values of P0 in P0-Thy complex and

34

P0-Ade complex are starting to approach each other, on the other hand in 5-30 ns P0 in P0-Thy complex and P0 have the same Rg's. The largest change of Rg is observed in P0, which shows the conformation of P0 fluctuating and mobile, however Rg's of the complexes have not changed significantly throughout simulation which indicates the polymer becomes immobilized (see Figure 3.6.a).

P0-A_rich and P0-T_rich contain adenine, thymine, cytosine, guanine nucleobases. P0, in P0-A_rich and P0-T_rich complex, is similar each other, P0 in water is different from the others. The Rg values P0 in water continued to decline throughout the simulation, it has lowest Rg values after 20 ns. It is close to a random coil form and has a more condensed. First 10 ns, all (P0, P0-A_rich, P0-T_rich) have almost the same Rg, the oligomer have similar structure, in the range of 10-20 ns P0 in P0-A_rich complex and P0 have nearly similar Rg's, in 20-50 ns oligomer without ssDNA have the lowest Rg, in the end of simulation Rg values of P0 in P0-T_rich and P0-A_rich complex are starting to approach each other. The largest change of Rg is observed in P0 as reflected by its standard deviation bar (see Figure 3.6.b).

Figure 3.7. Average radius of gyration for P0 in water and in P0-ssDNA complexes, bars represent standard deviation.

In production part of the MD simulations, it takes along 60 ns and trajectories, velocities and positions are recorded in every 0.06 ns. Every production part, 1000 frame are obtained for the analysis. Average radius of gyration is given in the above bar graph

with standard deviations (Figure 3.7). When looked at the average R_g value of P0 (more condensed) is lowest and P0-Ade (more extended) is highest among them.
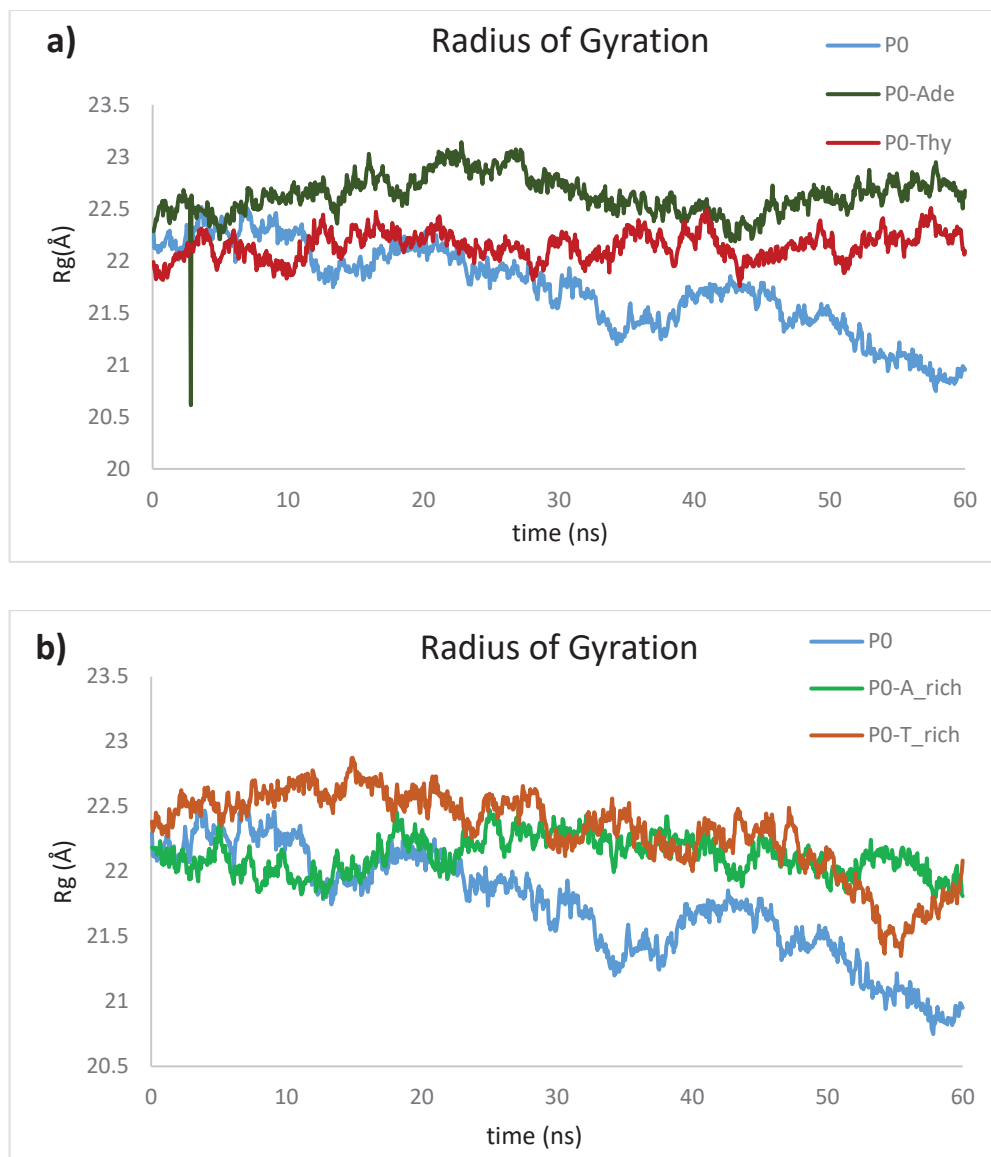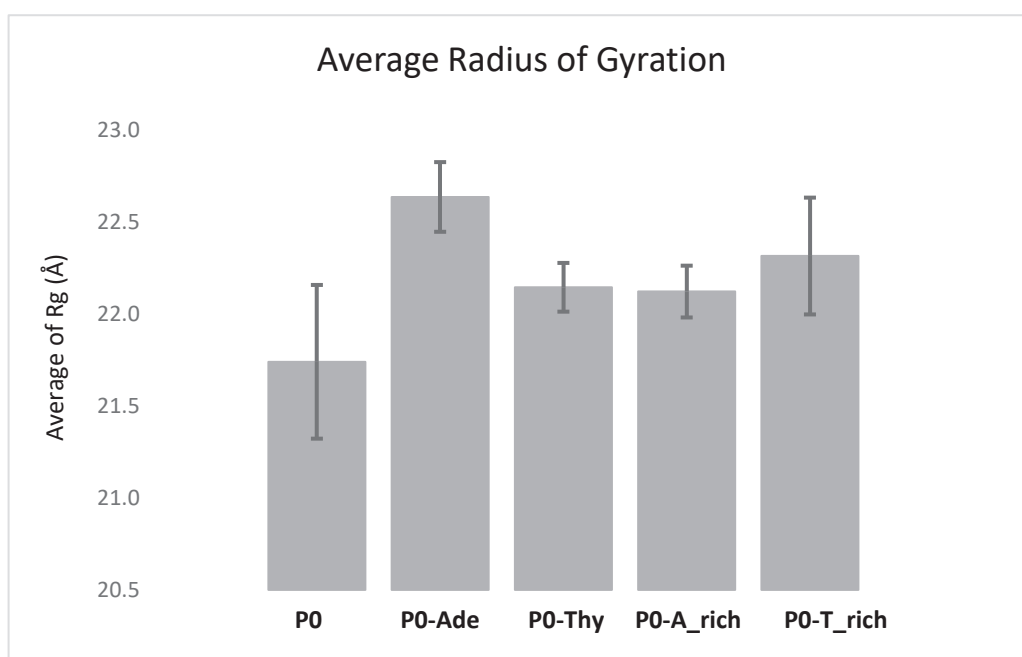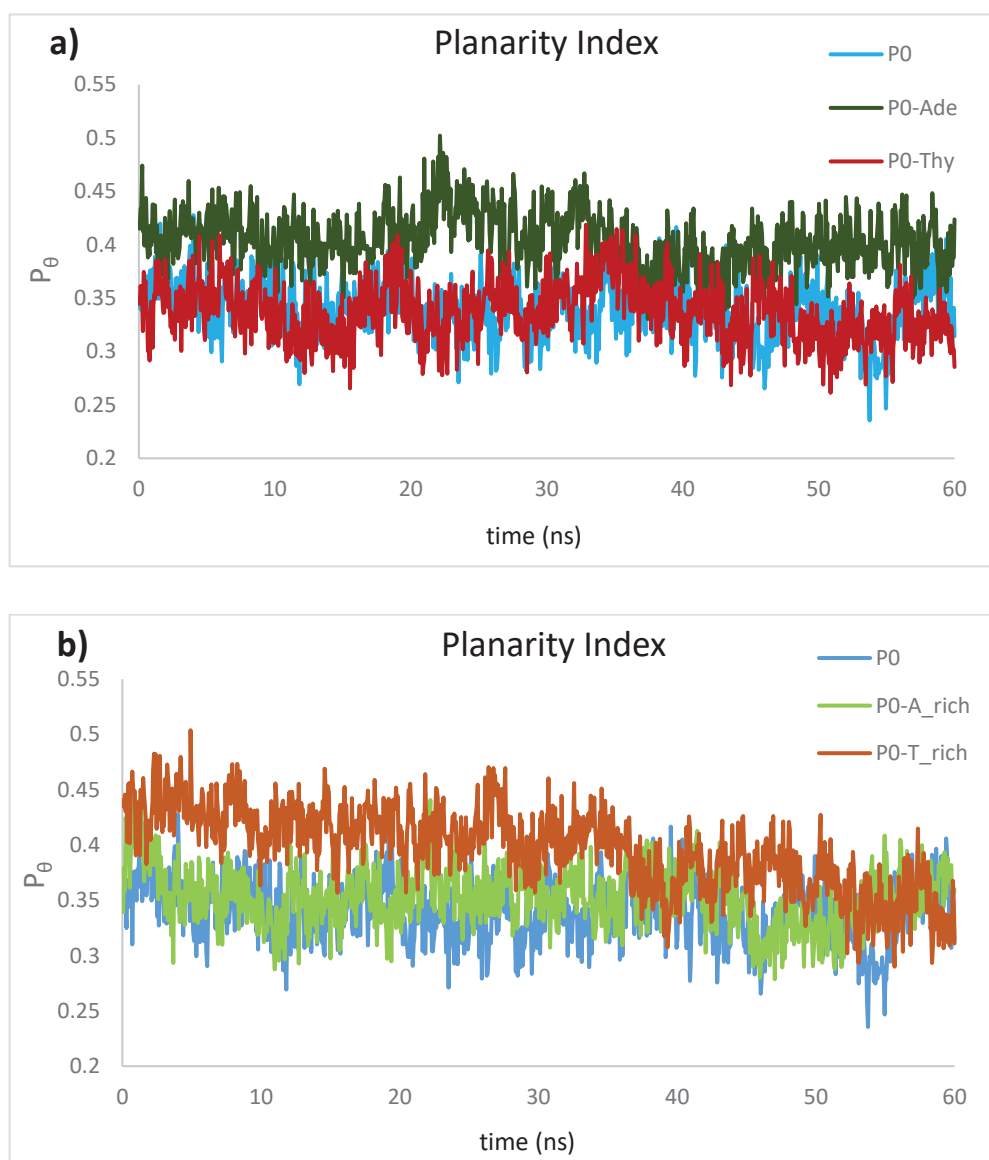

Figure 3.8. Planarity index for P0; a) in water, in P0-Ade complex, in P0-Thy complex; b) in water, in P0-A_rich complex, in P0-T_rich complex.

The planarity index, P_θ, is one of the strongest indices to show dihedral angles, indicating how planar it is. Figure 3.8.a, for P0 in water and P0-Thy have nearly same planarity trend, so they have same planarity proportion. The planarity index of P0 in P0-Ade complex is highest which means most planar. Figure 3.8.b shows that the planarity index for P0 in water and P0-A_rich displaying a similar trend. The P_θ value of P0 in P0-T_rich complex is highest than the others two (P0; in water and P0-A_rich), so it is flatter.

Figure 3.9. Average planarity index for P0; in water and in P0-ssDNA complexes, bars represent standard deviation.

0,06 ns is recording time in each frame. 1000 frame are used for calculation average value. Average planarity index is given in the above bar graph (Figure 3.9.). When looked at the average planarity index bar graph, P0 has 0.340 planarity 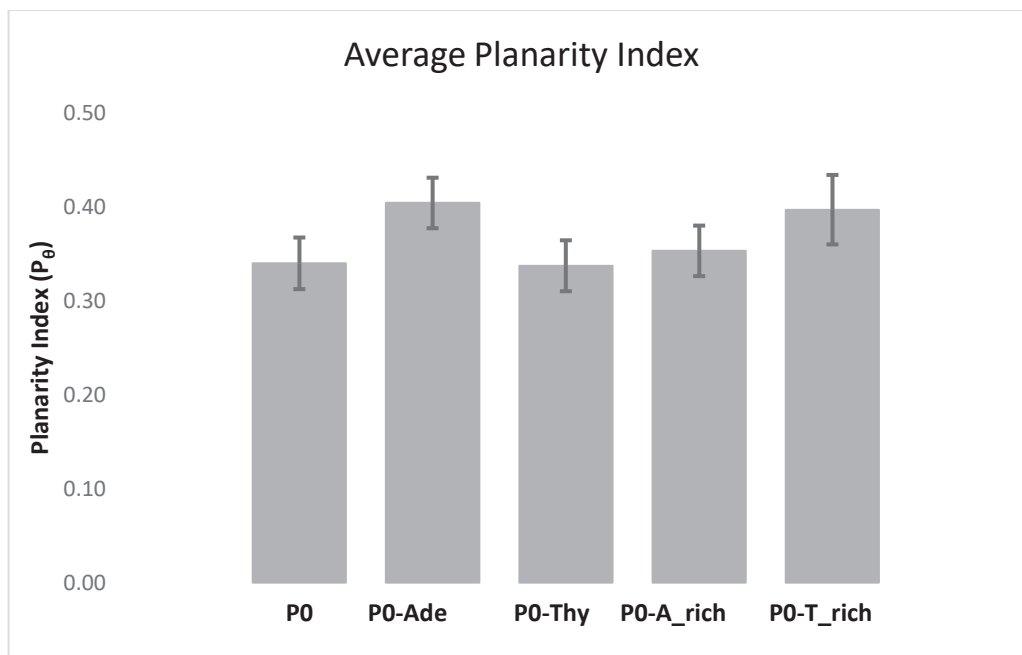index, it can say that it is a close to non-planar conformation or random coil form. When adding ssDNA, the planarity index of P0 becomes:

⇒ 0.404 for P0-Ade,

⇒ 0.337 for P0-Thy,

⇒ 0.353 for P0-A_rich,

⇒ 0.397 for P0-T_rich, in Figure 3.9.

Looking at the all indexes, they have been symbolized a non-planar conformation. No values were not represented a fully planar conformation. On the other hand, if they compared with P0's planarity indexes, they will be a meaningful. P0 in P0-Ade and P0-T_rich are close to be a planar.

It is reported that when ssDNA is added to PT, the color of solution (duplex) turns to red in experimental studies, they observed the changing color of solution from yellow to red is visible to the naked eye and the red shift in the cationic PT absorption spectra is clear (Figure 1.3.). They explained this situation as ssDNA molecule causing conformational changes (but the types of changes are not known), in the polymer. Some

researchers think that a having more planar conformations give the red shift in the spectra (Zheng and He 2014). In a study that included both experimental and computational studies, it was shown experimentally that adding ssDNA chain containing adenine bases, the shift to the red region in the absorption spectra was greater than ssDNA chain containing thymine bases (Rubio-Magnieto et al. 2015). Our simulation results support these experimental findings, P0 affected most upon the addition of pure adenine ssDNA (homopurine) since it has largest planarity index and Rg values.

3.3. Interactions between PT and ssDNA

In order to deduce the type of interactions which cause to structural changes of P0 under the influence of ssDNA are analyzed. The interactions that considered for analyses are π-cation (six membered ring of the nucleobases – nitrogen atom of the oligomer), O-H bond (oxygen atom of the oligomer – hydrogen atoms of the ssDNA), H-O bond (hydrogen atom of the oligomer – oxygen atoms of the ssDNA), S-H bond (sulfur atom of the oligomer – hydrogen atoms of the ssDNA) and electrostatic interaction N^+-O^- (nitrogen atom of the oligomer – oxygen atoms of the phosphate backbone). Some threshold distance values are used to find out the quantity of interactions as mentioned in the Analysis Methods part of this thesis (Table 2.3).
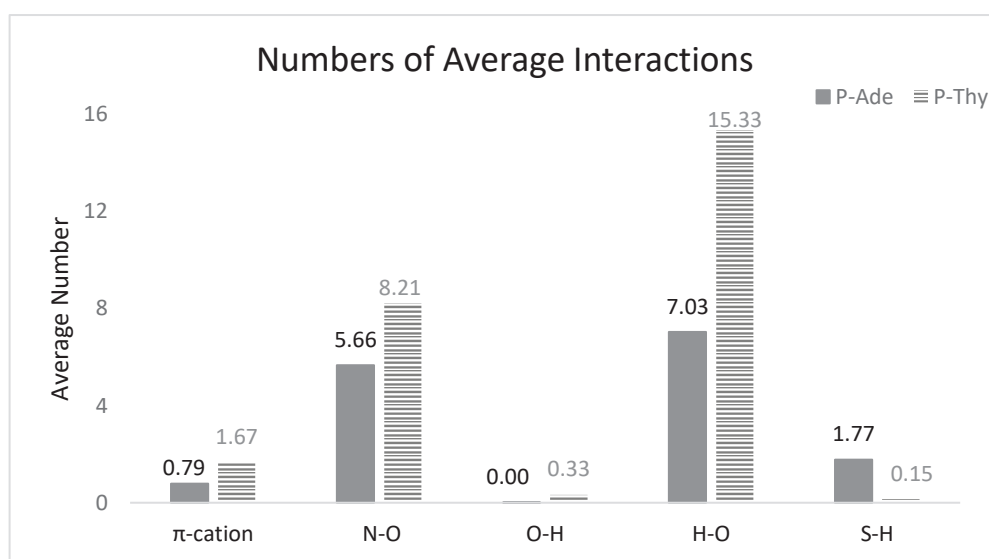

Figure 3.10. Average interaction numbers in P0-Ade and in P0-Thy

The amounts of interactions π-cation, electrostatic (N^+-O^-) interactions, hydrogen bonds (oxygen atom from P0 – hydrogen atom from ssDNA chain and oxygen atom from ssDNA chain – hydrogen atom from P0) are compared in P0-Ade and P0-Thy duplexes. All interactions are greater numerically in P0-Thy than in P0-Ade. If these interactions are associated with planarity, it is observed that the complex where the interactions are less, has more planar structure, and at the same time it has more in the S-H bonds (sulfur atom from the P0 in the backbone - hydrogen atom from the ssDNA) than in other interactions (Figure 3.10).

π-cation and O-H interactions have usually minor contributions for all conformations, S-H has slightly higher values but not most significant. Electrostatic interactions are moderately important, and the highest influence is coming from H-O interactions for homopurine and homopyrimidine complexes. However, for the complexation with the mixture of nucleobases, the electrostatic effect approaches to the H-O effect, but still H-O interactions role is higher (Figure 3.10).
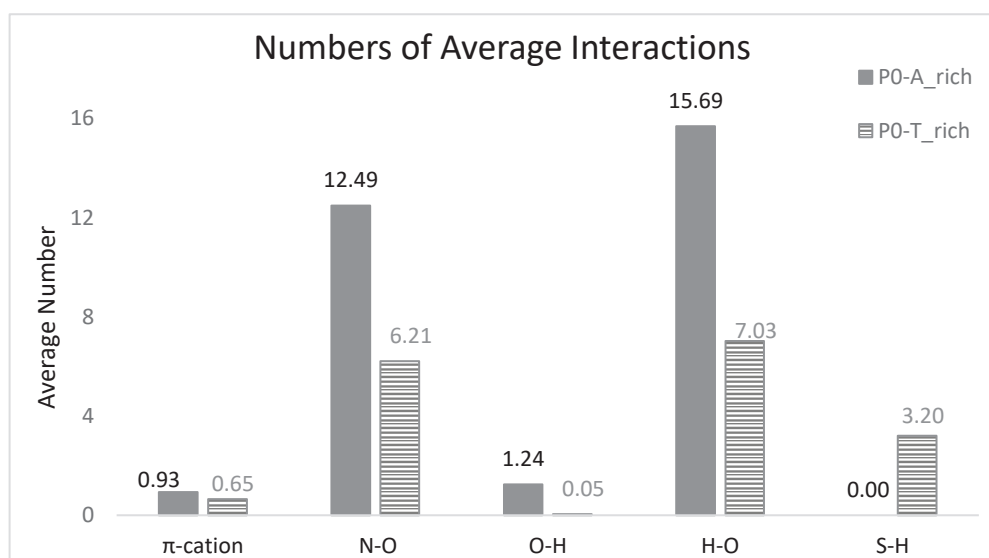

Figure 3.11. Average interaction numbers in P0-A_rich and in P0-T_rich

The second comparison is made between mixed nucleobases ssDNA which are P0-A_rich and P0-T_rich (have adenine, thymine, guanine and cytosine nucleobases). When examined numerically, the interactions of P0-T_rich, which is more planar, are higher than those of P0-A_rich except for S-H bond. (Figure 3.11).

These observations can be interpreted as less interactions (π-cation, O-H, H-O bond and electrostatic N^+-O^-) and more S-H interactions (which stretch the backbone of oligomer) leads to more planarization of the oligomer.
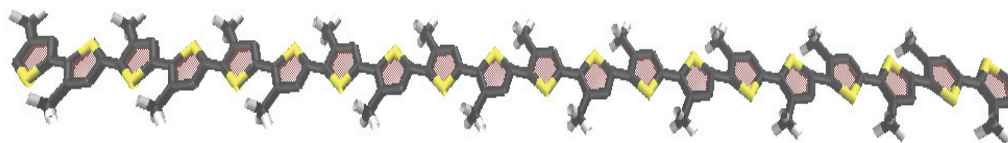
Figure 3.12. Backbone of the PT (carbon: grey, sulfur: yellow, hydrogen: white)
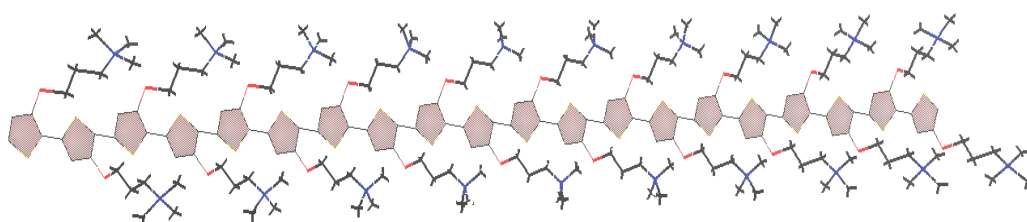
Figure 3.13. Side groups of the PT (carbon: grey, nitrogen: blue, hydrogen: white, oxygen: red)

To understand the role of side groups and backbone of the oligomer in the structural change upon complexation with ssDNA, the fractions of these interactions coming from side parts and oligomer skeleton are found by counting all the interactions without specifying by using 6.5 Å threshold. The backbone and side groups parts of P0 are shown in Figure 3.12 and Figure 3.13 respectively.
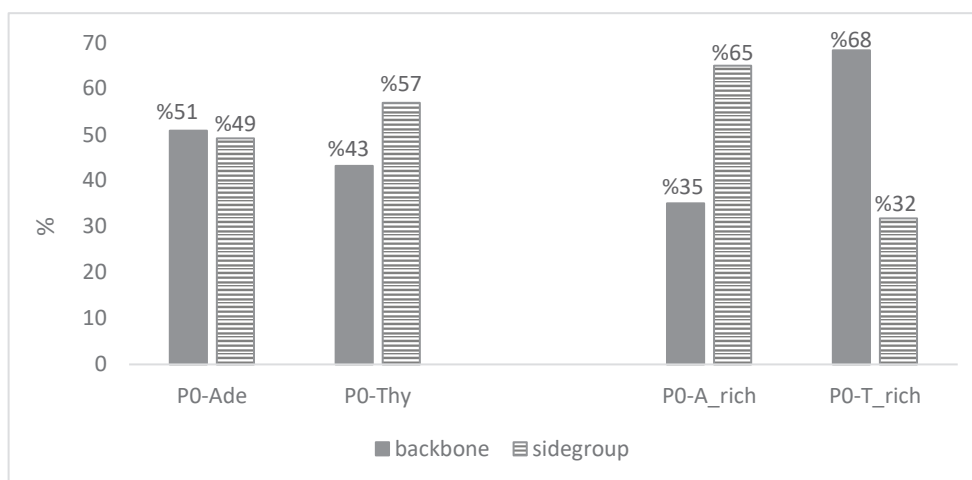
Figure 3.4. Percentage of backbone and side group effects on the all interactions

The percentages of interactions with the side groups and backbones are displayed in Figure 3.14. It is observed that, for the structures which are more planar (higher planarity index), backbone contributions of interactions are more effective than the side groups otherwise vice versa. This also agrees well with the literature (Rubio-Magnieto et al. 2015).

For every complex; 1000 frame are saved for the analysis and their statistics are given in Table 3.1. The numbers of frame are counted if complexes obey the order in the average values of properties listed (first two and 4-5 numerical columns) in the Table 3.1. We named the fraction of this situation as consistency.

Table 3.1. Average Properties over 1000 frames and consistency*

Complexes / Properties	P0-Ade	P0-Thy	consistency (%) in 1000 frame	P0-A_rich	P0-T_rich	consistency (%) in 1000 frame
Planarity	0,40	0,34	96,4	0,35	0,40	83,5
Rg	22,28	22,14	99,9	22,12	22,15	72,6
π-cation	0,79	1,67	57,5	0,93	0,65	39,8
electrostatic(N^+-O^-)	5,66	8,21	82,0	12,49	6,21	99,9
O-H	0	0,33	27,8	1,24	0,05	82,3
H-O	7,03	15,33	96,8	15,69	7,03	94,7
S-H	1,77	0,15	75,9	0	3,20	97,6

*The numbers of frame are counted if complexes obey the order in the average values of properties listed (first two and 4-5 numerical columns) in the table and the percentage of this situation is named as consistency.

For example, the average planarity index and Rg of the P0-Ade is higher than P0-Thy, the fractions of frames with higher planarity index and Rg of P0 in P0-Ade complex when compared to in P0-Thy complex are given in the table, for the interactions except S-H, P0-Ade complex has lower average value, the fractions of frames which has lower interaction value for P0-Ade complex are displayed in the Table 3.1 as consistency.

Within 1000 frame, the conformations with minimum and maximum planarity indexes of each complex are selected and these specific conformations are used for property analysis (mentioned in above Table), the results are shown in Table 3.2 and

Table 3.3. For both minimum and maximum planarity index conformations total number of interactions are greater in P0-Thy complex and P0-A_rich complex (π-cation, O-H, H-O and electrostatic N^+-O^- interactions), it seems that H-O interactions play important role for the structures of P0 in P0-Thy complexes since the H-O number changed much from P0-Ade to P0-Thy complex. By comparing higher planarity with lower planarity with the same complex, the instant conformation does not provide any correlation with the specific interactions.

Table 3.2. Minimum and maximum planarity index and their interaction numbers in one frame for P0, P0-Ade, P0-Thy
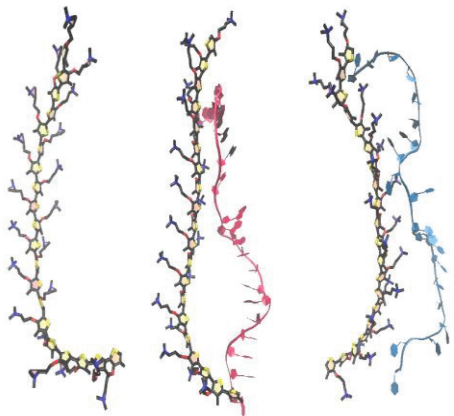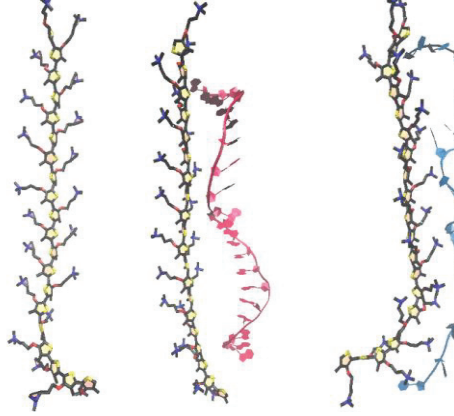
Name of Complex	Minimum			Maximum		
	P0	P0-Ade	P0-Thy	P0	P0-Ade	P0-Thy
Interaction types						
Planarity	0.24	0.33	0.26	0.43	0.50	0.42
Rg	21.14	22.26	22.02	22.44	22.93	22.09
π-cation	-	2	2	-	0	2
N^+-O^-	-	5	7	-	5	5
O-H	-	0	0	-	0	1
H-O	-	5	21	-	4	12
S-H	-	1	0	-	0	0

Table 3.3. Minimum and maximum planarity index and their interaction numbers in one frame for P0, P0-A_rich, P0-T_rich

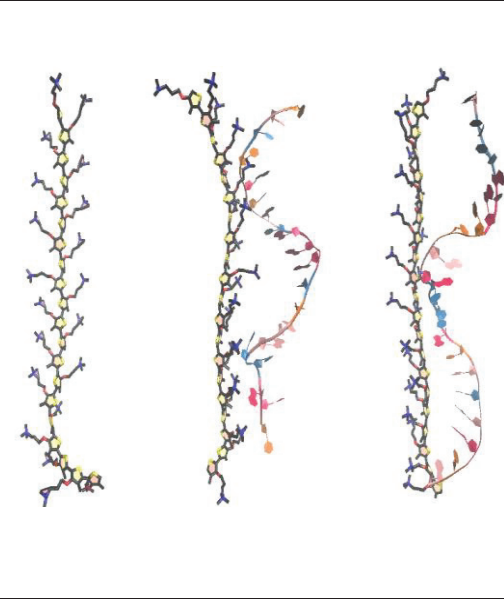| Name of Complex	Minimum			Maximum		
	P0 P0-A_rich P0-T_rich			P0 P0-A_rich P0-T_rich		
Interaction types						
Planarity	0.24	0.28	0.29	0.43	0.45	0.50
Rg	21.14	22.06	21.52	22.44	22.11	22.51
π-cation	-	1	0	-	1	0
N⁺-O⁻	-	12	6	-	17	8
O-H	-	1	0	-	1	0
H-O	-	12	8	-	21	10
S-H	-	0	5	-	0	2

The interactions on the complexes were also analyzed at the points where they had the least value of the corresponding distances, along the MD simulations (Table 3.4). By looking at these values, it will be possible to comment on which interaction has a stronger interaction. The strongest interaction (the smallest distance) belongs to the H-O bond (below 2 Å) that applies to all complexes, (1.75 Å - 1.91 Å). S-H and O-H interactions are in the range of 1.99 Å - 2.41 Å and 2.22 Å - 2.97 Å, respectively. The

minimum electrostatic interactions vary in between 2.77 Å and 3.11 Å. The π-cation has the weakest interaction with 4.44 Å - 4.71 Å minimum distance.

Table 3.4. The minimum distances of the interactions

Complexes / Interactions	P0-Ade	P0-Thy	P0-A_rich	P0-T_rich
π-cation	4.71	4.67	4.58	4.44
electrostatic(N^+-O^-)	3.04	3.11	2.77	3.11
O-H	2.97	2.28	2.22	2.42
H-O	1.91	1.82	1.75	1.81
S-H	2.41	2.41	-	1.99

CHAPTER 4

CONCLUSION

The essential thing for performing MD simulations is the force field definition of the interested system. In this thesis study, the CHARMM-compatible force field was generated by using ffTK and quantum mechanical (QM) calculations for the monomer of a cationic polythiophene (N, N, N-trimethyl-3-((2,4,5-trimethyltetrahydro thiophen-3-yl) oxy) propan-1-aminium). The dihedral parameters were found by using its dimer. The related parameters are calculated by QM techniques and then matched with the molecular mechanic (MM) results. The maximum deviations were obtained as 0.012 Å, 4.000° for the target QM bonds and angles, respectively, and RMSE is 0.787 kcal/mol for the dihedrals which are close to the recommended values given in the literature.

Molecular Dynamics (MD) simulations at the body temperature were performed using NAMD program to explain experimental observations on the spectra of PT and its complexes with ssDNA. MD simulation results were used for the analysis of specific interactions (π-cation, H-bonds (O-H, H-O, S-H), electrostatic interactions (N^+-O^-)) to compare the complexes of P0 with different single DNA strands (Ade, Thy, A_rich and T_rich).

The planarity index (P_θ), radius of gyration (R_g) were used to deduce how P0 response with the addition of ssDNAs. P_θ values were only calculated for the oligomer part of the complexes in order to compare the corresponding values of oligomer and complexes. The average of P_θ in P0-Ade complex was obtained as 0.404, which shows the oligomer has more planar conformation than P0 in water (0.340) and P0-Thy (0.337) complex as reported in the literature. P_θ in P0-A_rich complex is 0.353, which indicates the complex is more twisted than P0-T_rich (0.397) complex. In addition, different nucleobases (cytosine and guanine) influence the contributions of specific interactions, since adenine or thymine rich strands has different behavior from the homopurine and homopyrimidine strands. So, response of P0 depends on the sequence of ssDNA.

It has been observed that, the number interactions, except for S-H bond are decreasing with increasing planarity of the complexes. H-O and electrostatic interactions play an important role in all cases and their numbers increases as the complex more

coiled, in other word more strongly bound to ssDNA. Besides that, backbone contribution becomes greater on the averages as the planarity index raises. Based on all these remarks, it may be said that, the planarity index is a valuable parameter to explain the effects of the red shift in the absorption spectra.

The force field created within this thesis allows new simulations studies for the similar polymer to design new advanced materials which owns sensor ability (chemo or bio sensors). Our results support the experimental findings about the red shift in UV-VIS spectra upon complexation with ss-DNA with the change of structure of oligomer, we provide further information about the type of interactions that cause these geometric alterations like the electrostatic and H-O roles. In the literature, when ssDNA containing adenine was added to PT, it was found that a larger the red shift is observed in complexation with adenine than thymine strand, would be compatible with planarity and this work was also supporting these experimental data.

REFERENCES

Adcock, Stewart A., and J. Andrew McCammon. 2006. "Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins." *Chemical Reviews* 106 (5): 1589–1615. https://doi.org/10.1021/cr040426m.

Adcock, Stewart A, and J Andrew Mccammon. 2008. "Nihms60922" 106 (5): 1–53. papers://c33b182f-cf88-47e8-a9c5-ad67b5626483/Paper/p1859.

Alder, B. J., and T. E. Wainwright. 1957. "Phase Transition for a Hard Sphere System." *The Journal of Chemical Physics* 27 (5): 1208–9. https://doi.org/10.1063/1.1743957.

Ankerfors, Caroline. 2012. *Polyelectrolyte Complexes : Preparation , Characterization , and Use for Control of Wet and Dry Adhesion between Surfaces*.

Bernardi, R, M Bhandarkar, A Bhatele, E Bohm, R Brunner, F Buelens, C Chipot, et al. 2018. *NAMD User's Guide*. Urbana, Illinois, USA: Theoretical and Computational Biophysics Group, Beckman Institute, University of Illinois.

Carreon, Analyn C., Webster L. Santos, John B. Matson, and Regina C. So. 2014. "Cationic Polythiophenes as Responsive DNA-Binding Polymers." *Polym. Chem.* 5 (2): 314–17. https://doi.org/10.1039/C3PY01069D.

Chmiela, Stefan, Huziel E. Sauceda, Klaus Robert Müller, and Alexandre Tkatchenko. 2018. "Towards Exact Molecular Dynamics Simulations with Machine-Learned Force Fields." *Nature Communications* 9 (1). https://doi.org/10.1038/s41467-018-06169-2.

Ecker, János. 2016. "Generating CHARMM-Compatible Force Field Parameters for Atrazine." *Journal of Universal Science* 3 (1): 1–8. https://doi.org/10.17202/JUSO.2016.3.1.

Gromiha, M. Michael. 2011. "Protein Structure Analysis." *Protein Bioinformatics*, no. January: 63–105. https://doi.org/10.1016/b978-8-1312-2297-3.50003-5.

Guha, Rajarshi. 2006. "Force Fields / Molecular Mechanics." *Encyclopedic Reference of Molecular Pharmacology*, no. 9915607: 370–370. https://doi.org/10.1007/3-540-29719-0_627.

Ho, Hoang A., Maïté Béra-Abérem, and Mario Leclerc. 2005. "Optical Sensors Based on Hybrid DNA/Conjugated Polymer Complexes." *Chemistry - A European Journal* 11 (6): 1718–24. https://doi.org/10.1002/chem.200400537.

Kimber, Ian, Colin Humphris, Carl Westmoreland, Nathalie Alepee, Gianni Dal Negro, and Irene Manou. 2011. *Computational Chemistry, Systems Biology and Toxicology. Harnessing the Chemistry of Life: Revolutionizing Toxicology. A Commentary*. *Journal of Applied Toxicology*. Vol. 31. https://doi.org/10.1002/jat.1666.

Li, Chun, Munenori Numata, Masayuki Takeuchi, and Seiji Shinkai. 2006. "Unexpected Chiroptical Inversion Observed for Supramolecular Complexes Formed between an Achiral Polythiophene and ATP." *Chemistry, an Asian Journal* 1 (1–2): 95–101. https://doi.org/10.1002/asia.200600039.

Li, Chun, and Gaoquan Shi. 2013. "Polythiophene-Based Optical Sensors for Small Molecules." *ACS Applied Materials & Interfaces* 5 (11): 4503–10. https://doi.org/10.1021/am400009d.

Mackerell, Alexander D., Michael Feig, and Charles L. Brooks. 2004. "Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulation." *Journal of Computational Chemistry* 25 (11): 1400–1415. https://doi.org/10.1002/jcc.20065.

Mayne, Christopher G., Jan Saam, Klaus Schulten, Emad Tajkhorshid, and James C. Gumbart. 2013. "Rapid Parameterization of Small Molecules Using the Force Field Toolkit." *Journal of Computational Chemistry* 34 (32): 2757–70. https://doi.org/10.1002/jcc.23422.

Mayne, Christopher G, Melanie Muller, and Emad Tajkhorshid. 2015. "Parameterizing Small Molecules Using the Force Field Toolkit (FfTK)." *University of Illinois at Urbana-Champaign*, no. September.

Mccullough, By Richard D. 1998. "<McCullough R.D. - 1998 - Adv. Mater..Pdf>," no. 2: 93–116.

Patil, A. O., Y. Ikenoue, F. Wudl, and A. J. Heeger. 1987. "Water-Soluble Conducting Polymers." *Journal of the American Chemical Society* 109 (6): 1858–59. https://doi.org/10.1021/ja00240a044.

Pavlova, Anna, and James C. Gumbart. 2015. "Parametrization of Macrolide Antibiotics Using the Force Field Toolkit." *Journal of Computational Chemistry* 36 (27): 2052–63. https://doi.org/10.1002/jcc.24043.

Pavlova, Anna, Jerry M. Parks, and James C. Gumbart. 2018. "Development of CHARMM-Compatible Force-Field Parameters for Cobalamin and Related Cofactors from Quantum Mechanical Calculations." *Journal of Chemical Theory and Computation* 14 (2): 784–98. https://doi.org/10.1021/acs.jctc.7b01236.

Pitera, Jed W. 2014. "Expected Distributions of Root-Mean-Square Positional Deviations in Proteins." *Journal of Physical Chemistry B* 118 (24): 6526–30. https://doi.org/10.1021/jp412776d.

Preat, Julien, David Zanuy, Eric A. Perpete, and Carlos Aleman. 2011. "Binding of Cationic Conjugated Polymers to DNA: Atomistic Simulations of Adducts Involving the Dickerson's Dodecamer." *Biomacromolecules* 12 (4): 1298–1304. https://doi.org/10.1021/bm200022n.

Rahman, A. 1964. "Correlations in the Motion of Atoms in Liquid Argon." Physical Review 136 (2A): A405–11. https://doi.org/10.1103/PhysRev.136.A405.

Rajwar, Deepa, Gopal Ammanath, Jamal Ahmed Cheema, Alagappan Palaniappan, Umit Hakan Yildiz, and Bo Liedberg. 2016. "Tailoring Conformation-Induced Chromism of Polythiophene Copolymers for Nucleic Acid Assay at Resource Limited Settings." *ACS Applied Materials & Interfaces* 8 (13): 8349–57. https://doi.org/10.1021/acsami.5b12171.

Relationships, Structure-property, By Mario Leclerc, and Karim Faid. 1997. "Electrical and Optical Properties of P :% E" 357 (14): 1087–94.

Roncali, Jean. 1992. "Conjugated Poly(Thiophenes): Synthesis, Functionalization, and Applications." *Chemical Reviews* 92 (4): 711–38. https://doi.org/10.1021/cr00012a009.

Rubio-Magnieto, Jenifer, Elias Gebremedhn Azene, Jérémie Knoops, Stefan Knippenberg, Cécile Delcourt, Amandine Thomas, Sébastien Richeter, et al. 2015. "Self-Assembly and Hybridization Mechanisms of DNA with Cationic Polythiophene." *Soft Matter* 11 (32): 6460–71. https://doi.org/10.1039/c5sm01484k.

Rubio-Magnieto, Jenifer, Amandine Thomas, Sébastien Richeter, Ahmad Mehdi, Philippe Dubois, Roberto Lazzaroni, Sébastien Clément, and Mathieu Surin. 2013. "Chirality in DNA–π-Conjugated Polymer Supramolecular Structures: Insights into the Self-Assembly." *Chemical Communications* 49 (48): 5483. https://doi.org/10.1039/c3cc42108b.

Sjöqvist, Jonas, Jerôme Maria, Rozalyn A. Simon, Mathieu Linares, Patrick Norman, K. Peter R. Nilsson, and Mikael Lindgren. 2014. "Toward a Molecular Understanding of the Detection of Amyloid Proteins with Flexible Conjugated Oligothiophenes." *Journal of Physical Chemistry A* 118 (42): 9820–27. https://doi.org/10.1021/jp506797j.

Vanommeslaeghe, K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, et al. 2009. "CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields." *Journal of Computational Chemistry* 8 (8): NA-NA.

https://doi.org/10.1002/jcc.21367.

Yao, Zhiyi, Chun Li, and Gaoquan Shi. 2008. "Optically Active Supramolecular Complexes of Water-Soluble Achiral Polythiophenes and Folic Acid: Spectroscopic Studies and Sensing Applications." *Langmuir* 24 (22): 12829–35. https://doi.org/10.1021/la802086d.

Zheng, Weiming, and Lin He. 2014. "Quantitative Measurements of Thermodynamics and Kinetics of Polythiophene–DNA Complex Formation in DNA Detection." *Biomaterials Science* 2 (10): 1471. https://doi.org/10.1039/C4BM00210E.

APPENDIX A

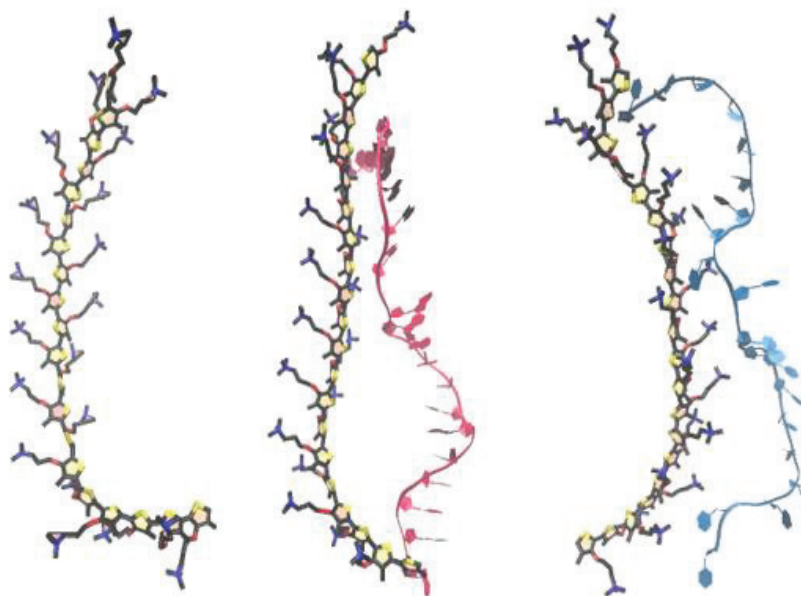## SNAPSHOTS OF COMPLEXES AT THE HAVING MINIMUM AND MAXIMUM PLANARITY INDEX

Figure A.1. The snapshots of P0, P0-Ade complex, and P0-Thy complex, respectively at having minimum planarity index
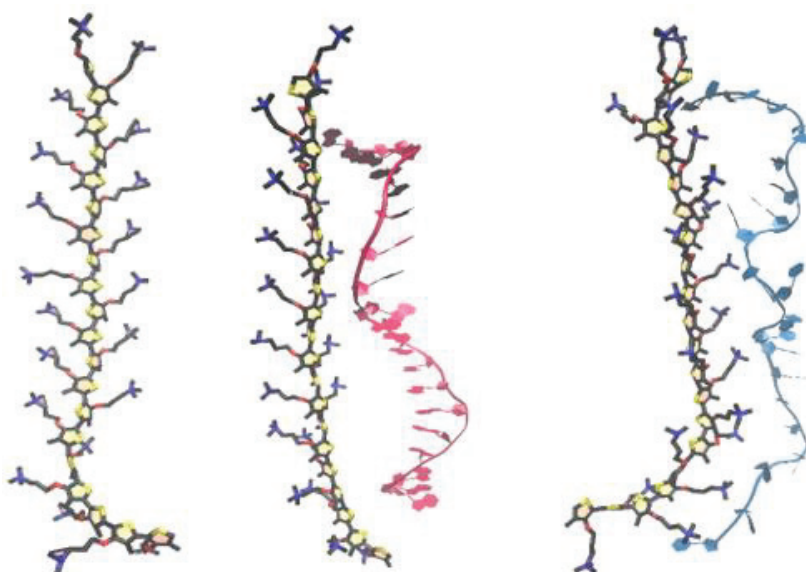

Figure A.2. The snapshots of P0, P0-Ade complex, and P0-Thy complex, respectively at having minimum planarity index
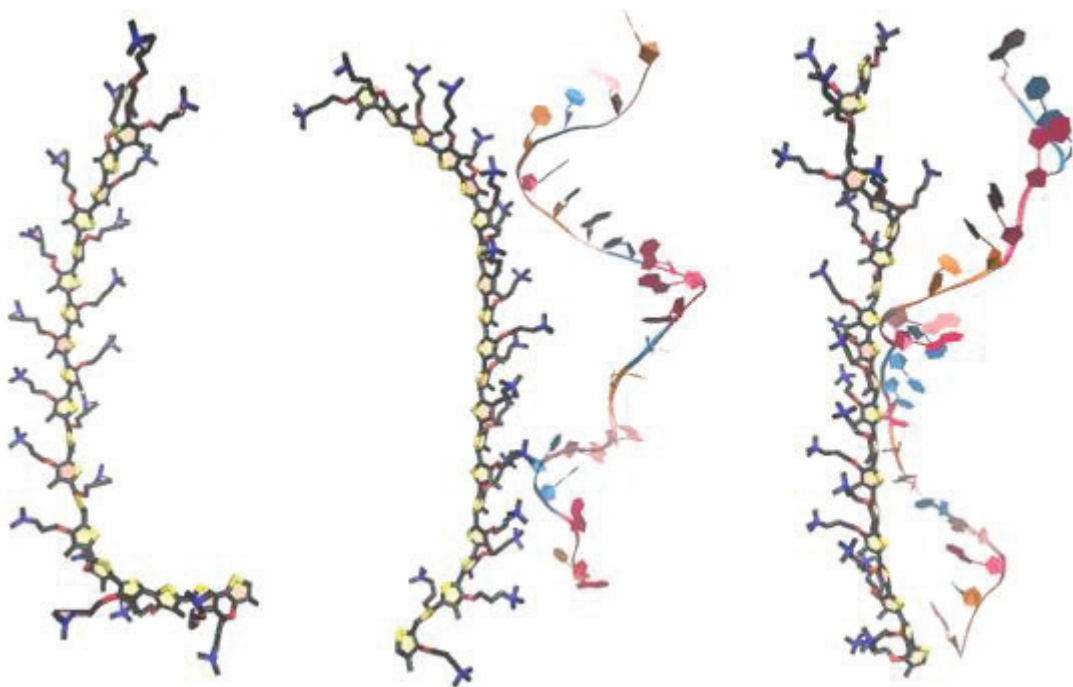
Figure A.3. The snapshots of P0, P0-A_rich complex, and P0-T_rich complex at having minimum planarity index, respectively
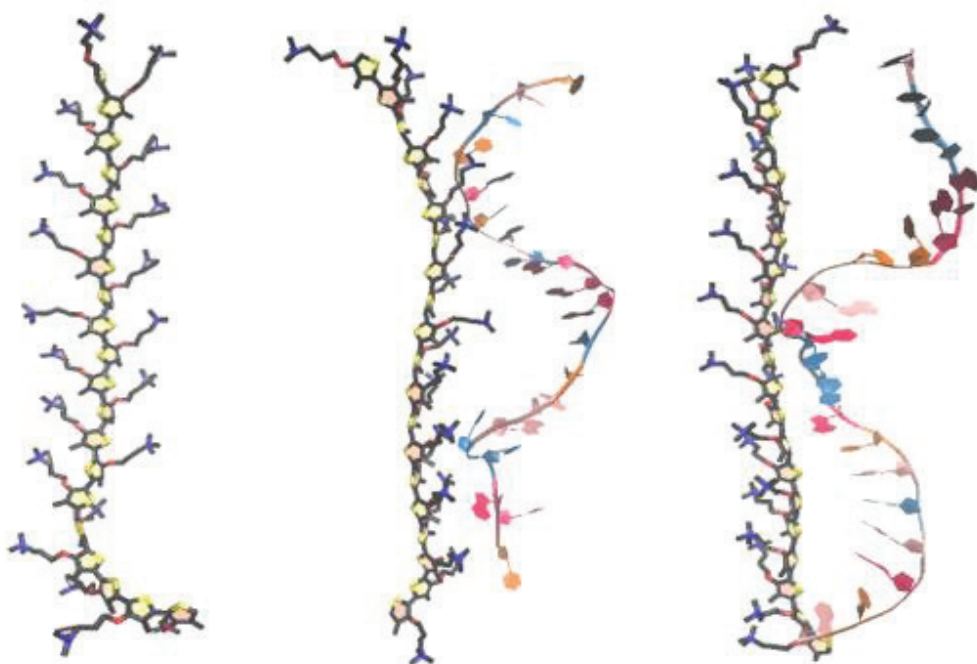

Figure A.4. The snapshots of P0, P0-A_rich complex, and P0-T_rich complex at having maximum planarity index, respectively