MOLECULAR DYNAMICS SIMULATIONS OF A CATIONIC THIOPHENE OLIGOMER AND A NUCLEOTIDE COMPLEX

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
Fethi Can DEMİRCİ

July 2022
İZMİR

ACKNOWLEDGEMENT

I would like to thank everyone who supported me during this work. Firstly, I would like to thank my supervisor Prof. Dr. Nuran ELMACI IRMAK, for her support, patience, guidance, and trust. It was a pleasure to work with her.

I also would like to thank Ph.D. candidates of the ELMACI Research group, Erman KIBRIS and Nehir NALINCI BARBAK for their help during this thesis and contribution to the force field parameters. Besides, I would like to thank everyone in the ELMACI research group.

Thirdly, I am grateful to my committee members, Assoc. Prof. Dr. Ümit Hakan YILDIZ and Prof. Dr. Elif ŞAHİN IŞGIN, for participating and reviewing my thesis. Special thanks to TÜBİTAK for this project (119Z100) and their financial support.

Lastly, I would like to thank my family and my friends for their endless love and support.

ABSTRACT

MOLECULAR DYNAMICS SIMULATIONS OF CATIONIC THIOPHENE OLIGOMER AND A NUCLEOTIDE COMPLEX

In this thesis, parametrization of cationic polythiophene (CPT) and molecular dynamics (MD) simulations of CPT with DNA complexes were performed to understand the behaviors of the CPT with DNA complex and CPT DNA complexes in different salt solutions (NaCl, KCl, $MgCl_2$, $CaCl_2$).

The results of MD simulations show that the end-to-end distance of CPT is affected by both the type sequences and length of the DNA, and the addition of 20T elongates the backbone of the oligomer while 20A and MIX ssDNAs almost have no significant effect. When the complementary DNA chain is added to the duplex solutions, the backbone structure of the oligomer becomes very similar to its structure without ssDNAs since R_{ee} values in both cases are almost the same. It was observed that the CPT-20A complex has a more random coil form than the CPT-20T complex.

According to the interaction analysis of MD simulations, all the CPT-DNA duplexes except CPT-20A prefer electrostatic interaction rather than π-cation interaction. DNAs like to interact with the oligomer's side chain rather than its backbone in all systems. Thus, electrostatic interactions and the side chain of oligomer play an important role in the structure of duplexes with thymine which gets the highest response from the oligomer. The addition of 20T makes backbone of F0 more elongated and less compact. 20T has higher electrostatic and π-cation interactions. Thus, F0 is more sensitive to 20T than 20A and MIX.

ÖZET

BİR KATYONİK TİYOFEN OLİGOMERİ VE NÜKLEOTİT KOMPLEKSİNİN MOLEKÜLER DİNAMİK BENZETİMLERİ

Bu tez çalışmasında, uzun yan zincirli bir katyonik politiyofen için simülasyon veri tabanlarında olmayan kuvvet alanı parametreleri elde edildi. Katyonik politiyofenin farklı ortamlarda (Tuzsuz, NaCl, KCl, $MgCl_2$, $CaCl_2$) DNA ile kompleksleşmesini anlamak amacıyla moleküler dinamik simülasyonlar gerçekleştirilmiştir.

Yapılan analizlerin sonucuna göre, DNA'nın sekansı ve uzunluğu CPT'nin baş-son uzunluğunu (R_{ee}) etkilemiştir. CPT yanına 20T ssDNA eklenmesi polimer omurgasını uzatırken 20A ve MIX ssDNAlarının hemen hemen hiç etkisi olmadığı gözlenmiştir. Dublekslerin üzerine tamamlayıcı DNA zinciri eklendiğinde ise oligomerin yapısının kompleksleşmeden önceki haline dönüştüğü izlendi (R_{ee}'leri yaklaşık olarak aynı). CPT-20A kompleksi COPT-20T kompleksine göre daha büzülmüş bir yapıya sahip olduğu görülmüştür (Rg değerine göre).

MD simülasyonları sonuçlarına yapılan etkileşim analizlerine göre, CPT-20A hariç tüm CPT-DNA kompleksleri π-katyona göre elektrostatik etkileşimleri tercih etmektedir. DNA'ların çoğunlukla oligomerin iskeleti yerine yan zinciriyle etkileşmeyi seçtikleri gözlenmiştir. 20T'nin eklenmesi F0'ın omurgasını uzatıp daha az kompakt bir hale getirmiştir. 20T'de polimer ile daha fazla π-katyon ve elektrostatik etkileşimlere sahiptir. Dolayısıyla, F0 20T'ye 20A ve MIX e göre daha duyarlıdır.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

Figure **Page**

Figure **Page**

LIST OF TABLES

ABBREVIATIONS

PE	Polyelectrolyte
PT	Polythiophene
CPT	Cationic Polythiophene
ATP	Adenosine triphosphate
AMP	Adenosine monophosphate
ADP	Adenosine diphosphate
DNA	Deoxyribonucleic Acid
UV	Ultra-Violet
VIS	Visible
ss	Single stranded
ds	Double-stranded
Rg	Radius of Gyration
QM	Quantum Mechanics
MM	Molecular Mechanics
MD	Molecular Dynamics
CHARMM	Chemistry at Harvard Macromolecular Mechanics
AMBER	Assisted Model Building with Energy Refinement
GROMOS	Groningen Molecular Simulation
NAMD	Nanoscale Molecular Dynamics
VMD	Visual Molecular Dynamics
ffTK	Force Field Tool Kit
CGenFF	CHARMM General Force Field
VDW	Van Der Waals

LJ	Lennard-Jones
HF	Hartree–Fock
R_{ee}	End-to-End Distance
RMSE	Root Mean Square Error

CHAPTER 1

INTRODUCTION

1.1 Polyelectrolytes

Polyelectrolytes (PEs) which are polymers with charged groups have semi-conductor behavior because of their ionic side moieties. PEs have been used in different areas such as detecting genetic disorders, solar cells, biomedical applications, photovoltaic studies, and light-emitting diodes (Feng 2008, Rubio-Magnieto 2013, Ozenler 2019, Ammanath 2022).

1.1.1 Cationic Polythiophenes

Cationic polythiophenes (CPT) are the most used polymer form of PEs due to their electronic and optical properties. In living cell experiments, CPTs have low toxicity and good photostability. Thus, CPTs are widely used in chemo- and biosensors such as sensing AMP (Kıbrıs 2021), ATP (Li 2005, Yildiz 2012), and DNA (Leclerc 2002, Thomas 2007, Carreon 2014, Rubio-Magnieto 2015).

1.2 Literature Work

This thesis will be evaluated in two parts. One part is the parametrization of cationic polythiophene using the force field tool kit (ffTK) in the Virtual Molecular Dynamics (VMD) program. The second is the complexation of CPT and DNA. Thus, in this section of the thesis, the literature work is separated into two parts: First, studies on cationic polythiophenes and their complexes with DNA. Second, studies on force field tool kit with CPT (ffTK).

1.2.1 Studies on Cationic Polythiophenes

The first charged polythiophene (PT) was synthesized by Patil and his friends in 1987 (Patil 1987).

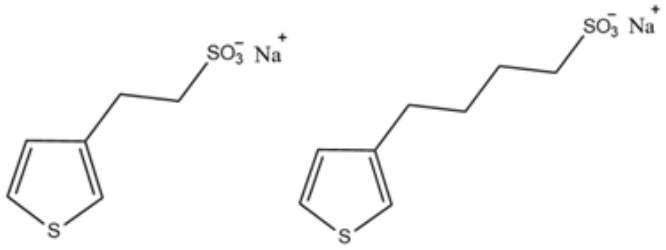


Figure 1.1 Structures of charged polythiophenes: Sodium poly(3-thiophene-fl-ethanesulfonate) (P3-ETSNa) (Left) and sodium poly(3-(thiophene-6-butanesulfonate) (P3-BTSNa) (Right) (Source: Patil, 1987)

In 1997, Leclerc and Faid investigated the optical and electrical properties of polythiophene derivatives. They stated that the side chain of the polymer affects not only solubility but also optical properties (M. F. Leclerc 1997).

There is a review article written by McCullough in 1998. Syntheses of different CPTs were reported in this article. It was stated very different CPTs could be created from PTs in the future. The physical properties and sensing ability of CPTs toward metals were investigated (McCullough 1998).

In 2005, Li et al. studied the detection of ATP with CPT.

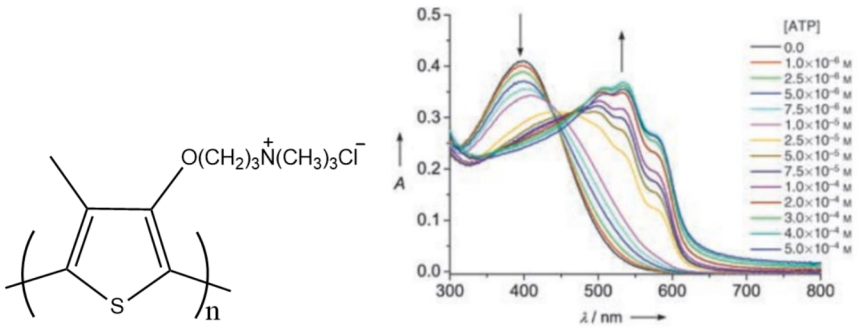


Figure 1.2 Structure of CPT (Left) and UV spectra of CPT and ATP complex at different concentrations (Right) (Source: Li, 2005)

In that article, increasing the concentration of ATP causes a red shift in the absorption spectra. They also stated this shift was caused by the PT backbone. They also studied the sensing ability of PT with different anionic molecules such as AMP, ADP, chloride, carboxylate, phosphate, and triphosphate ions. The color change can be seen with the naked eye (see Figure 1.3). It was stated that the development of this sensor for ATP would open new doors to new research about CPTs and sensors (Li 2005).

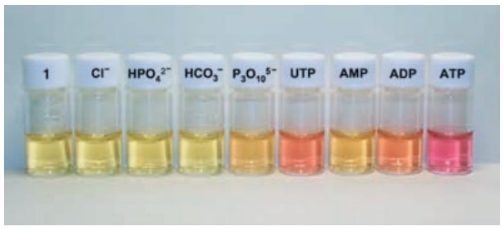


Figure 1.3 Color change upon the addition of different anionic groups to the PT solution (1.0 x 10^{-4}) (Source: Li, 2005)

In 2012, Wang et al. did a similar experiment with Poly[N, N, N-trimethyl-4-(thiophene-3-ylmethylene)cyclohexanaminium chloride] (PTCA-Cl) with different analytes such as K^+, Na^+, Ca^{2+}, Mg^{2+}, F^-, Cl^-, Br^-, I^-, SO_4^{2-}, CO_3^{2-}, NO_3^-, PO_4^{3-} and HPO_4^{2-} (Wang 2012).

Lots of research were done by U.H. Yıldız and his research team about the cationic polythiophenes. In 2019, Ozenler et al. synthesized a tetramer to investigate the C-H/C-H coupling in thiophene. After tetramer, reaction stopped due to low solubility. Conformation of the tetramer was found HT-TT-HT (Ozenler 2019a).

In 2019, Ozenler et al. synthesized poly[1,4-dimethyl-1-(3-((2,4,5-trimethylthiophen-3-yl)oxy)propyl)piperazin-1-ium bromide] to detect hepatocellular carcinoma cells. Single chain cationic polymer dots (Pdots) with their small radius size (<10 nm) have great advantages over conjugated polymer nanoparticles (CPNs) (>30 nm). In this article, Pdots were formed with ethylene glycol which is a better solvent than water. This could help surgeons to remove all the cancer cells (see Figure 1.4) (Ozenler 2019b).

In another study, familial Mediterranean fever (FMF) were investigated with CPT. Using the sensing ability of CPT, another testing technique was found with a 96% accuracy and 90% cheaper than the cheapest technique (Yucel 2021).

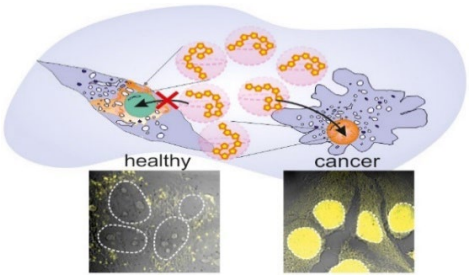


Figure 1.4 Pdots labeling cancer cells. (Source: Ozenler, 2019b)

1.2.1.1 Studies on complexation of CPT and DNA

Leclerc et al. investigated the detection of nucleic acids with CPT using colorimetric and fluorometric analysis in 2002. The CPTs and DNA sequences that they used are given in Figure 1.5 and Table 1.1, respectively.

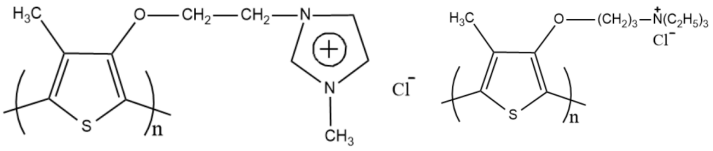


Figure 1.5 Structure of CPTs (Source: Leclerc, 2002)

Table 1.1 Sequences of the DNAs (Source: Leclerc, 2002)

Names	Sequences
X1	5′-CATGATTGAACCATCCACCA-3′
Y1	3′-GTACTAACTTGGTAGGTGGT-5′
Y2	3′-GTACTAACTTCGAAGGTGGT-5′
Y3	3′-GTACTAACTTCGTAGGTGGT-5′

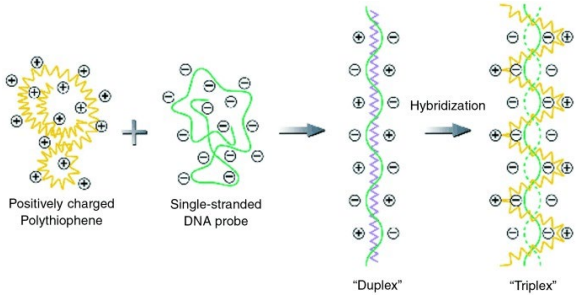


Figure 1.6 Schematic of the complexation of CPT and DNA. (Source: Leclerc, 2002)

As given in Figure 1.6, first, a complexation of CPT and X1 was created. Then three different triplex was made with Y1, Y2, and Y3. Two CPTs gave similar results for colorimetric and fluorometric analysis. They concluded that for P1 (Figure 1.7), nucleic acids could be detected with CPTs. CPTs can even detect a very small amount or with a change in the sequence fast and precisely (Leclerc 2002).

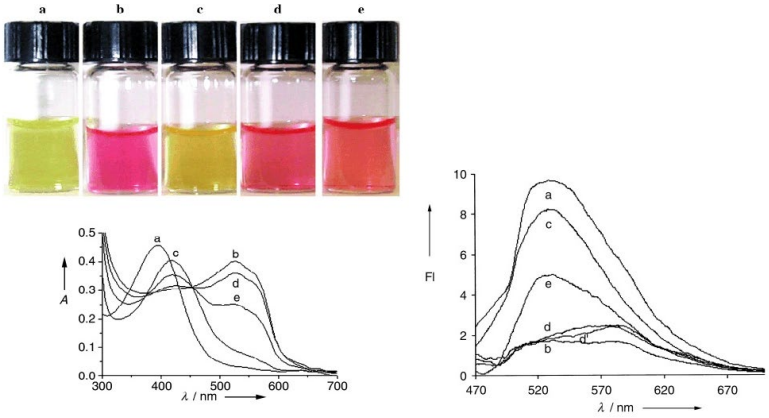


Figure 1.7 Photographs of solutions (Top Left), UV/VIS spectrum ($7.9x10^{-5}$M) (Bottom Left), and Fluorescence spectrum ($2.0x10^{-7}$M) (Right) analysis of CPTs with DNA complexes. a) P1, b) P1X1 duplex, c) P1X1Y1 triplex d) P1X1Y2 triplex e) P1X1Y3 triplex. (Source: Leclerc, 2002)

In 2013, Carreon et al. synthesized poly (N, N, N-trimethyl-3-(2-(thiophene-3-yl)acetamido)propan-1-aminium iodide) as a sensor for DNA detection. They observed a small but observable red shift when DNA was added to the system (Carreon 2014).

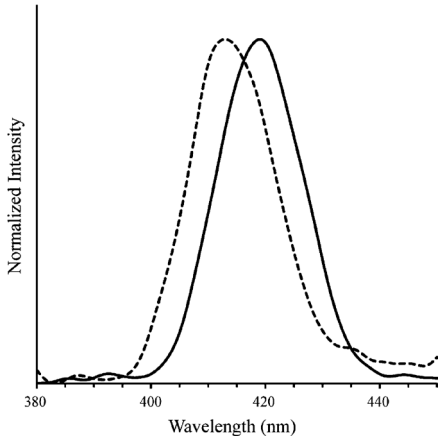


Figure 1.8 Fluorescence spectrum of polymer only (dashed line) and polymer with DNA (solid line) (Source: Carreon, 2014).

Rubio-Magnieto et al. investigated the complexation of CPT and DNA in 2015. In that study, cationic poly[3-(60-(trimethylphosphonium)hexyl)thiophene-2,5-diyl was synthesized, and molecular dynamics simulations were done.

Table 1.2 Sequence of the DNAs that were used in the article (Source: Rubio-Magnieto, 2015).

ssDNA	$d(T)_{20}$	5'-TTT TTT TTT TTT TTT TTT TT-3'
	$d(A)_{20}$	5'-AAA AAA AAA AAA AAA AAA AA-3'
	$d(T)_{40}$	5'- TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT T-3'
	$d(A)_{40}$	5'-AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA A-3'
	$d(R)_{20}$	5'-CGT CAC GTA AAT CGG TTA AC-3'
	$d(R)_{rev20}$	5'-GTT AAC CGA TTT ACG TGA CG-3'
	$d(R)_{43}$	5'-CGT CAC GTA AAT CGG TTA ACA AAT GGC TTT CGA AGC TAG CTT C-3'
dsDNA	$d(A)_{20}d(T)_{20}$	5'-AAA AAA AAA AAA AAA AAA AA-3' 5'-TTT TTT TTT TTT TTT TTT TT-3'
	$d(R)_{20}d(R)_{rev20}$	5'-CGT CAC GTA AAT CGG TTA AC-3' 3'-GCA GTG CAT TTA GCC AAT TG-5'

Sequences of DNAs that were used in this article are given in Table 1.2. In the experimental section, different CPT and DNA complexes were obtained at different concentrations and temperatures.

Complex	λ_{max}
CPT	451
dT20/CPT	494
dT40/CPT	497
dT20/CPT/dA20	477
dAdT/CPT	473
dA20/CPT	442/520d
dA40/CPT	497
dA20/CPT/dT20	469
dR20/CPT	495
dR43/CPT	497
dR20/CPT/dRrev20	494
dRdRrev/CPT	493
dRrev20/CPT	490
dRrev20/CPT/dR20	491

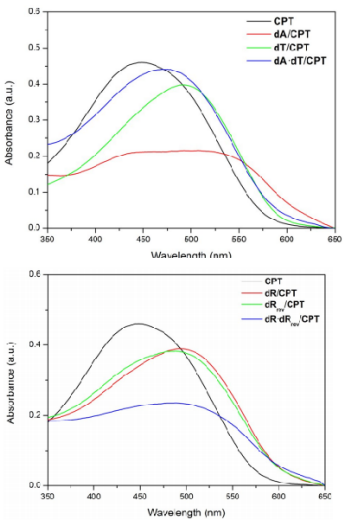


Figure 1.9 UV/VIS absorption wavelengths of the CPT and DNA complexes (Left). UV/VIS comparison graphs for some of the complexes (Right) (Source: Rubio-Magnieto, 2015)

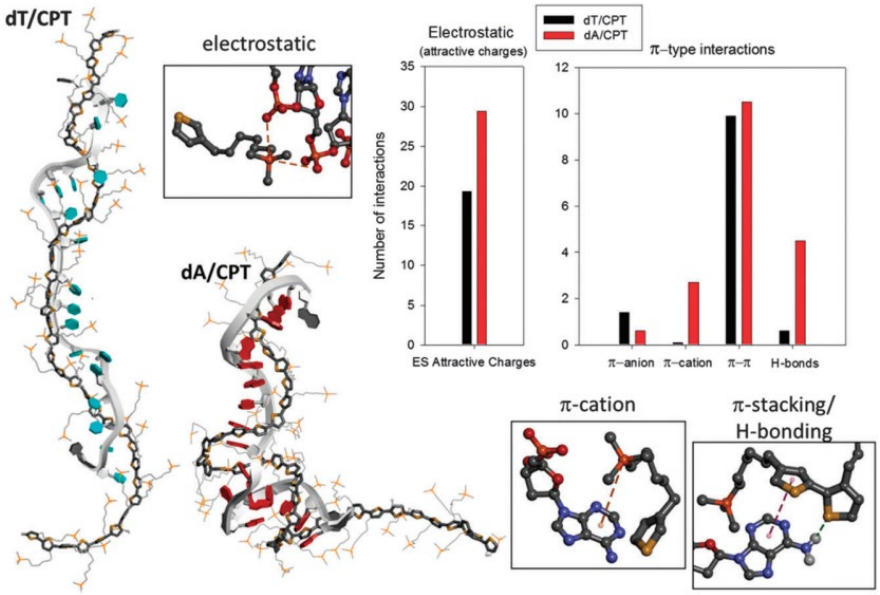


Figure 1.10 Snapshots of adenine and thymine complexes with CPT (Left). Average interaction number of last 10 frames and illustration of interactions (Electrostatic, π-cation, pi-stacking, and hydrogen bonding) between DNA and CPT (Right) (Source: Rubio-Magnieto, 2015)

Different from Leclerc's article, λ_{max} of the dsDNA and CPT complex were not close to λ_{max} of the CPT (see Figure 1.9). In that article, the red shift was observed when DNA was added to the system. In the theoretical part, MD simulations were performed with three different lengths of CPT (CPT6, CPT40, CPT60). 15ns simulations were performed except for CPT60, which was 10ns. After production simulations, the radius of gyration (Rg) and interaction analysis were done. Snapshot of the complexes at the last frame and illustrations of interactions are given in Figure 1.10.

According to the article, complexation with thymine have higher R_g than adenine complex. It was stated that the self-assembly between CPT and DNA is related to the sequence of DNA, length, and interactions (Rubio-Magnieto 2015).

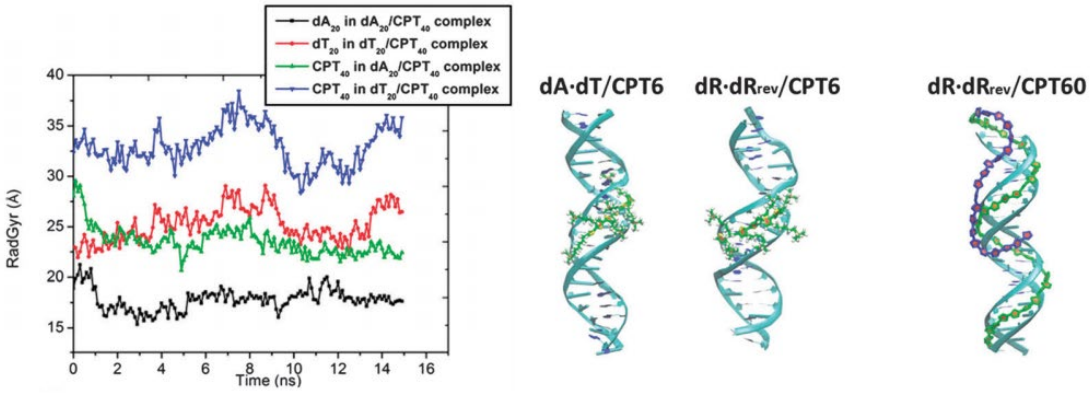
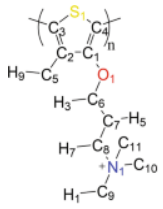


Figure 1.11 Radius of gyrations of complexes (Left) and snapshots of triplexes at the last frame (Right) (Source: Rubio-Magnieto, 2015)

1.2.2 Studies on Force Field Parametrization for CPT

In 2021, Kıbrıs et al. did theoretical research on complexation CPT with AMP and ATP. First, they obtained the absent force field parameters of CPT with ffTK. Parametrized atom types and charges are given in Figure 1.12.



Atom	Type	Charge	Atom	Type	Charge	Atom	Type	Charge
S_1	SG2R50	−0.16	C_5	CG331	−0.24	C_8	CG321	−0.11
C_1	CG2R51	0.03	H_9	HGA3	0.09	H_7	HGP5	0.25
C_2	CG2R51	0.33	C_6	CG321	−0.01	N_1	NG3P0	−0.60
C_3	CG2R57	−0.14	H_3	HGA2	0.09	C_9	CG334	0.25
C_4	CG2R57	0.07	C_7	CG321	−0.17	H_1	HGP5	0.09
O_1	OG301	−0.33	H_5	HGA2	0.09			

Figure 1.12 Atom types and charges (Source: Kıbrıs, 2021)

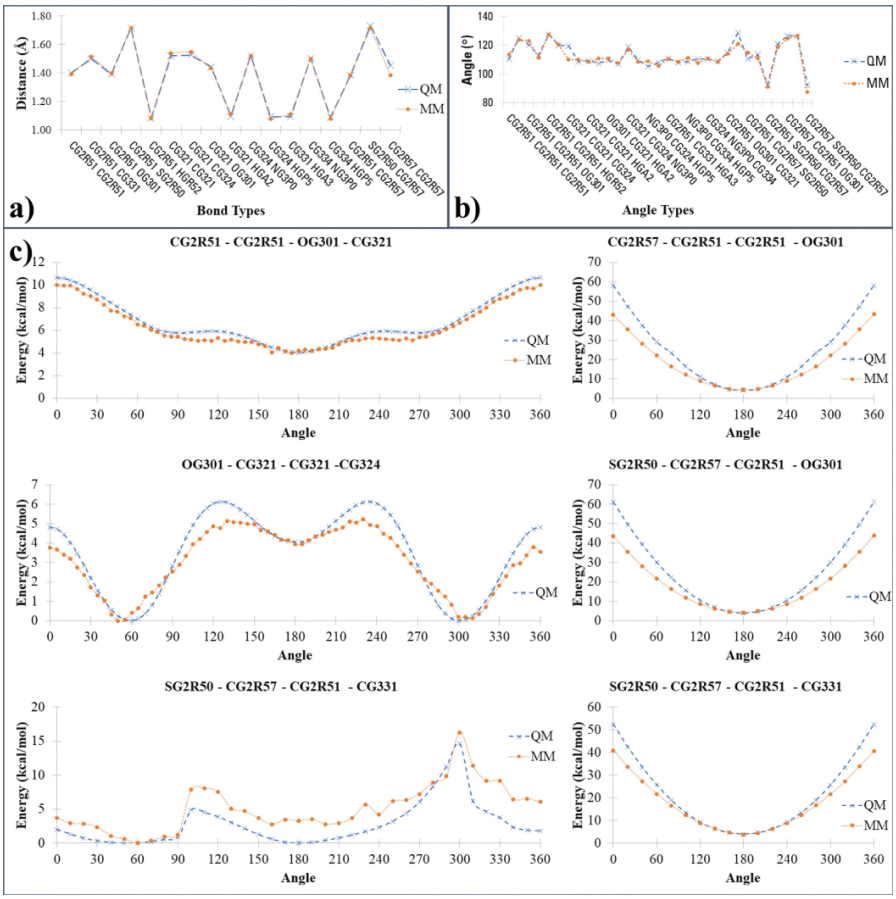


Figure 1.13 QM and MM fitting for a) Bond, b) Angle, and c) dihedral parameters (Source: Kıbrıs, 2021)

Then, bond, angle, and dihedral parameters were calculated (see Figure 1.13). After parametrization was done, CPT was simulated in water, with ATP in water, and

with AMP in water. According to the article, the complexation of CPT with AMP and ATP increases end to end distance of CPT. Also, in the article, it was stated that the red shift in UV/VIS spectra was caused by the π-cation interactions in the ATP-CPT duplex (Kıbrıs 2021).

1.3 Aim of this thesis

The aim of this thesis is to carry out a molecular dynamics (MD) simulation study for the complexation of a CPT (Poly- N, N, N-trimethyl-3-(4-methylthiophen-3-yl)oxy)hexan-1-aminium (F0)) in order to understand the experimental observations given in the literature part about the spectroscopic changes upon the addition of DNAs to the CPT at different conditions. For this purpose, the missing force field parameters for the CPT used in this work will be obtained and then used in MD simulations for various systems which contain different DNA sequences and conditions. The effect of the size of the side group in the thiophene monomer on the complexation and its structure will be investigated and compared with a CPT with a shorter side chain. The simulation results will be analyzed to find out the dominant type of interactions, such as electrostatic, hydrogen bonding, and π-cation, which play an essential role in the structure of the CPT backbone in the duplexes.

CHAPTER 2

COMPUTATIONAL METHODS

2.1 MD Simulations

Molecular dynamics simulations involve the time evolution of atoms or molecules in a system for a fixed period. It can be used for an experiment that is dangerous, need extreme conditions, or understand the molecule's behavior. MD calculates forces using the potential energy definition of the system and uses Newton's equation of motions (see Equation 2.1) with the force field to create trajectories for a period. In MD simulations, three different equations, acceleration (a), velocity (v), and coordinates (q), are calculated by the use of the forces acting on the system through Equations 2.3, 2.4, and 2.5, respectively. The force components are found by Equation 2.2 through the potential definition described in section 2.2.

$$\mathrm{F_i} = \mathrm{m_i a_i} = \mathrm{m_i}\frac{\mathrm{dv_i}}{\mathrm{dt}} = \mathrm{m_i}\frac{\mathrm{d^2 q_i}}{\mathrm{dt^2}} \tag{2.1}$$

$$F_i = -\frac{dV}{dq_i} \tag{2.2}$$

$$\mathrm{F_i} = \mathrm{m_i a_i} \Rightarrow \mathbf{a_i} = \frac{\mathrm{F_i}}{\mathrm{m_i}} \tag{2.3}$$

$$\frac{F_i}{m_i} = \frac{dv}{dt} \Rightarrow \frac{F_i}{m_i}dt = dv \Rightarrow \boldsymbol{v_f} = v_i + \frac{F_i}{m_i}dt \tag{2.4}$$

$$\mathrm{v} = \frac{dq}{dt} \Rightarrow v\mathrm{dt} = \mathrm{dq} \Rightarrow \boldsymbol{q_f} = \mathrm{q}_i + \mathrm{vdt} \tag{2.5}$$

The first molecular simulation was introduced in 1953 with Monte Carlo simulation by Metropolis et al. (Metropolis 1953) and in 1957, with the hard-sphere system by Alder and Wainwright (Alder 1957). Since then, interest in molecular dynamics simulation has been increasing. In 1964, Rahman achieved a realistic potential

for liquid argon (Rahman 1964). With the development in technology and computers, more complex and larger systems have been simulated. In 1977, McCammon performed the first protein simulation (McCammon 1977). In 2013, Martin Karplus, Michael Levitt, and Arieh Warshel were awarded the Nobel Prize for the development of multiscale models for complex chemical systems.

2.1.1 Thermodynamic Ensembles

Different ensembles can be used in MD simulations. Different ensembles have different properties.

Microcanonical ensemble (NVE): Fixed number of atoms, N, volume, V, and energy, E. It is the adiabatic process for MD simulations.

Canonical Ensemble (NVT): Fixed number of atoms, N, volume, V, and temperature, T.

Isobaric-Isothermal Ensemble (NPT): Fixed number of atoms, N, pressure, P, and temperature, T.

Grand Canonical Ensemble (μVT): Fixed chemical potential, μ, volume, V, and temperature, T.

Isobaric-Isoenthalpic Ensemble (NPH): Fixed number of atoms, N, pressure, P, and enthalpy, H.

2.2 Force Fields

In MD, atoms connect each other with flexible forces. A force field computes these forces in the form of potential energy equations.

$$E_{Total} = V_{Bond} + V_{Angle} + V_{Dihedral} + V_{LJ} + V_{Electrostatic} \quad (2.6)$$

Equation 2.6 shows the potential energy equations. This equation can be separated into two parts: Intramolecular (bond, angle, and dihedral) and intermolecular (LJ and electrostatic) forces. Every force field has a different force constant for each potential equation. These constants can be calculated from QM and MM calculations or experimental data.

Some of the popular force fields are Chemistry at Harvard Molecular Mechanics (CHARMM), Assisted Model Building and Energy Refinement (AMBER), Groningen Molecular Simulation (GROMOS)

2.2.1 CHARMM Force Field

CHARMM is the most common force field that is used for MD simulations. CHARMM has been started and developed by Martin Karplus for almost 30 years. It has a very wide database of force constants for proteins, lipids, nucleic acids, and other biological molecules. It has been actively updated and developed by the MacKerell Lab (Vanommeslaeghe 2010).

$$V = \sum_{\text{bond}} k_b(b - b_0)^2 + \sum_{\text{angle}} k_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedral}} k_\phi\left(1 + cos(n\phi - \delta)\right) + \sum_{\text{improper}} k_\omega(\omega - \omega_0)^2 + \sum_{\text{Urey-Bradley}} k_u(u - u_0)^2 + \sum_{\text{nonbonded}} \left(\epsilon\left[\left(\frac{R_{\min_{ij}}}{r_{ij}}\right)^1 2 - \left(\frac{R_{\min_{ij}}}{r_{ij}}\right)^6\right] + \frac{q_i q_j}{\epsilon r_{ij}}\right)$$

2.3 Nanoscale Molecular Dynamics (NAMD)

Nanoscale Molecular Dynamics (NAMD) is the software for high-performance simulation of large biomolecular systems. NAMD was introduced by the Theoretical and Computational Biophysics Group (TCB) and the Parallel Programming Laboratory (PPL) at the University of Illinois at Urbana–Champaign. It solves equations of the force field to approximate the trajectories of the atoms of the molecule. NAMD uses VMD to visualize the trajectories. 5 files are needed to run a molecular dynamics simulation via NAMD (Phillips 2020).

- **Protein Data Bank File (.pdb):** It contains the coordinates of the molecule and sometimes contains every bond on the molecule. It can be downloaded from the Protein Data Bank (http://www.pdb.org) or created from software such as Avogadro, GaussView, or Molefacture.

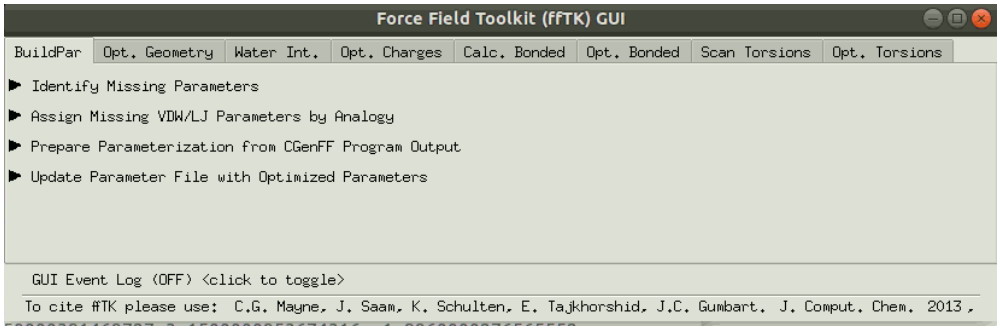
- **Protein Structure File (.psf):** It contains the charges and molecular weights of the atoms in the molecule. It also includes every bond, angle, and dihedral. It can be created from Molefacture with a pdb file or can be created via VMD with a topology file.
- **Topology File (.top):** It includes every charge, molecular weight, and bond in the molecule.
- **Parameter File (.par):** It contains potential parameters for every bond, angle, and dihedral.
- **Configuration File (.namd):** This contains every condition for the molecular dynamics simulation, such as temperature, pressure, box sizes, cut-off, velocities, timestep, and the number of steps.

2.4 Visual Molecular Dynamics (VMD)

VMD is software to visualize and analyze molecular dynamics simulations. It was developed by the Theoretical and Computational Biophysics group at the Beckman Institute for Advanced Science and Technology, the University of Illinois at Urbana–Champaign (Humphrey 1996).

2.5 Force Field Tool Kit (ffTK)

Force Field Tool Kit (ffTK) is a module in VMD for the parametrization of molecules. ffTK has a procedure to follow to find missing parameters (Mayne 2013).

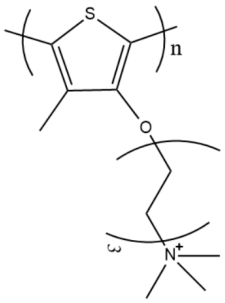


- **Preparing Molecule:** In this part, the molecule will be prepared. Molecules can be drawn via GaussView, Avogadro, or Molefacture. After the molecule is drawn, atoms should be named as in the database of CHARMM General Force Field (CGenFF), and charges should be set for the fixed charge atoms (non-polar hydrogens should be set to 0.09, hydrogens in the ammine group should be set to 0.25 according to CHARMM).
- **Assigning missing VDW/LJ Parameters:** In this part, VDW/LJ parameters will be assigned to all names in the molecule from the CGenFF database.
- **Optimizing Geometry:** In this part, the molecule will be optimized via the Gaussian09 software at MP2/6-31G* level. This part finds the lowest energy conformation of the molecule. After this part, a new pdb will be created.
- **Water Interaction:** In this part, the interaction of donors (Interaction with the oxygen of water) and acceptors (Interaction with the hydrogen of water) with water molecule will be optimized via the Gaussian09 software at HF/6-31G* level. To optimize water interaction, water molecules be placed at interaction sites, and optimization is done.
- **Optimization of Charges:** After optimization of water interaction sites, charges will be optimized except for fixed charged molecules (H atoms) by using NAMD software. After this part, a new psf will be created with optimized charges.
- **Bond and Angle Optimization:** The missing bond and angle parameters will be calculated in this part. After a short optimization process of charges, bonds and angles will be optimized via Gaussian09 software at MP2/6-31G* level, then via NAMD. Then, missing parameters are added to the parameter file.
- **Scan and Optimization Torsion:** In this part, missing dihedral parameters were first calculated via Gaussian09 software at MP2/6-31G* level. Then a small optimization was done with NAMD software to match QM and MM data. Missing parameters were added to the parameter file.

2.6 Computational Details

In this thesis, poly- N, N, N-trimethyl-3-(4-methylthiophen- 3-yl) oxy) hexan-1-aminium (F0) is modeled with 20 mer of this polymer. Three types of ssDNA sequences

that were used in the complexation and two dsDNAs that were used in the triplexes with 20mer are given in Figure 2.1.



20A	5'- AAA AAA AAA AAA AAA AAA AA - 3'
20T	5'- TTT TTT TTT TTT TTT TTT TT - 3'
MIX	5'- CGCGAATTCGCG - 3'
20AT	5'- AAA AAA AAA AAA AAA AAA AA - 3' 3'- TTT TTT TTT TTT TTT TTT TT - 5'
dMIX	5'- CGCGAATTCGCG - 3' 3'- GCGCTTAAGCGC - 5'

Figure 2.1 Structure of F0 and DNAs used in MD simulations

In the MD simulations, first, a 10 ns vacuum minimization was done. After the vacuum minimization, the molecule was solvated in a cubic box whose size depended on the size of the complex after the vacuum minimization process. If necessary, Na^+ and Cl^- ions (counter ions) were added to set the net charge of the system to zero. Then, a 50 ns simulation was done at the NPT ensemble as the production. Timestep and the cut-off were taken as 2 fs and 14 Å, respectively. The system temperature was taken as 310K. Verlet algorithm and Langevin thermostat were used as integrators. For every 0.05 ns, coordinates and energies were saved during the simulation. Box dimensions changed according to complex size after the vacuum process. Cubic box determined by the minimum box size possible. All simulations were doubled with the same input. In the analysis of these simulations, the averages of them were taken.

In the MD simulations, different complexations were used. F0 was first simulated with water. Then, complexations of F0 with ssDNAs (20A, 20T, and MIX) were simulated. For the dsDNA, 20AT and dMIX were used. To see the chain length effect of ssDNA, chain lengths were doubled (40A, 40T, and 2MIX). In the complexation, the center of the mass of the DNA was approximately placed 20 Å away from the center of mass of the F0. 4 salts (NaCl, KCl, $CaCl_2$, and $MgCl_2$) were added additionally to the cubic box of the complexation F0 with ssDNA to see the salt effects. To see the salt effect, 400 Cl^- were added, and their counterions were also added according to the salt.

2.7 MD Simulation Results Analysis

There are many ways to analyze the simulations. In this work, end-to-end distance (R_{ee}), the radius of gyration (Rg), and interactions between F0 and the DNA strands (π-cation, N-O interactions, and hydrogen bonds) have been analyzed.

End-to-end Distance (R_{ee}): The distance between the two hydrogens of the first and the last rings of the thiophene was taken. This shows how foiled the polymer is.

The radius of gyration (R_g): The radius of gyration was measured from the distance of each atom to the center of mass or an axis. In this case, the center of mass was taken. R_g measures the compactness of the polymer.

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

Interactions: Three types of interactions were investigated. π-cation interaction is the interaction of nitrogen of the oligomer and the 6-membered ring of nucleic acids. Electrostatic interaction (N-O) interaction is the interaction between the nitrogen of oligomer and oxygens of DNA's phosphate groups. The threshold value of π-cation and electrostatic interactions was taken as 6.5 Å. The last type of interaction was hydrogen bonding. Three types of hydrogen bonding (S-H, O-H, and H-O) were analyzed. S-H interaction is the interaction between the sulfur atom of the oligomer and the hydrogens of DNA. O-H interaction is between the oxygen atom of the oligomer and the hydrogens of DNA. H-O interaction is between the hydrogen of oligomer and oxygens of DNA (see Figure 2.2). The threshold value for hydrogen bonding was set to 3.0 Å.

Interactions were found by counting interactions below the threshold value. After that, to normalize the interaction value, the total value is divided by the possible number of atoms for the specified interaction and the total frame number to find the normalized interaction value for each frame. Another analysis was done with hydrogen interactions to understand the side chain's role against different DNA strands. Oligomer's interaction region (from backbone or side chain) with DNA strands calculated by summing of O-H and H-O (Excluding hydrogens of the methyl group in the oligomer backbone) for the side chain and S-H for the oligomer backbone.

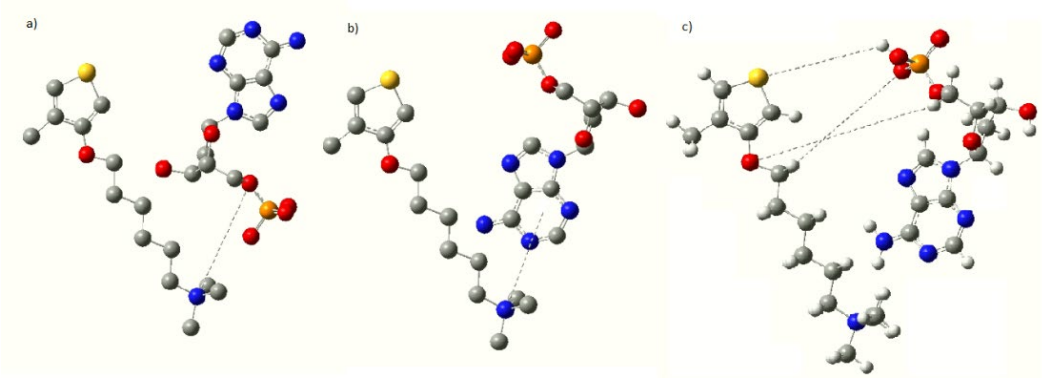


Figure 2.2 Representation of a) Electrostatic interaction (N-O), b) π-cation interaction, c) Hydrogen bonding (S-H, O-H, and H-O)

CHAPTER 3

RESULTS AND DISCUSSION

In this thesis, CHARMM force field parametrization of the CPT and molecular dynamics simulations of the CPT with different DNA strands were studied.

3.1 FORCE FIELD RESULTS

Parametrization of the monomer started with naming atoms according to CGenFF (see Figure 3.1). These atom names were given from the database of CGenFF. For example, the amine group's names already exist in the CGenFF database as tetramethylammonium. For F0, parameters that existed were taken from the CGenFF database, and some other parameters were obtained from Kıbrıs et al. work (Kıbrıs 2021). Missing parameters were calculated in the context of this thesis.

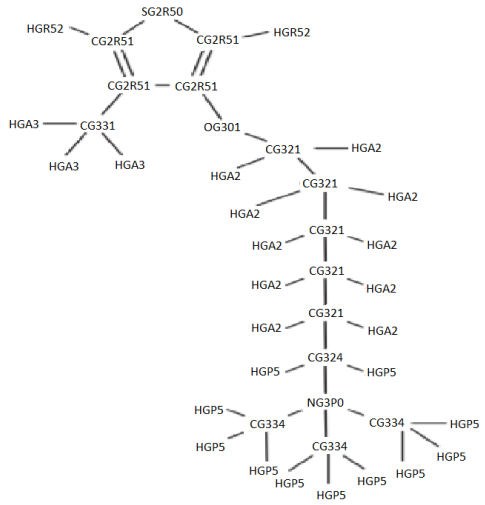


Figure 3.1 Naming atoms

For the parametrization of the monomer and dimer, missing parameters were found, and to find these parameters, ffTK's procedure was followed. After all the

parameters were found, a simulation was done, and validations between QM and MM were done.

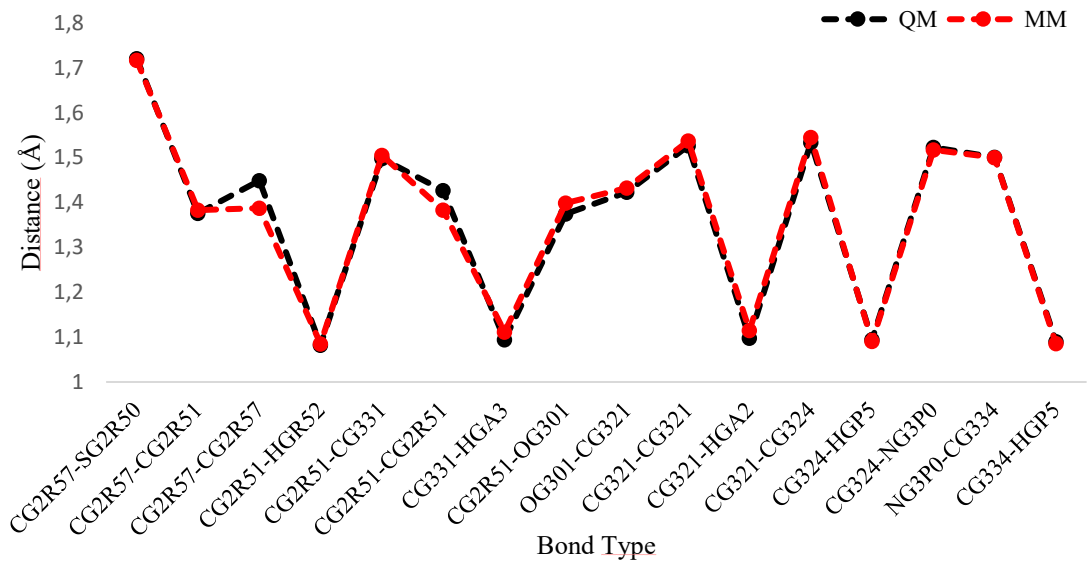


Figure 3.2 Comparison of the bond lengths from the validation of QM and MM.

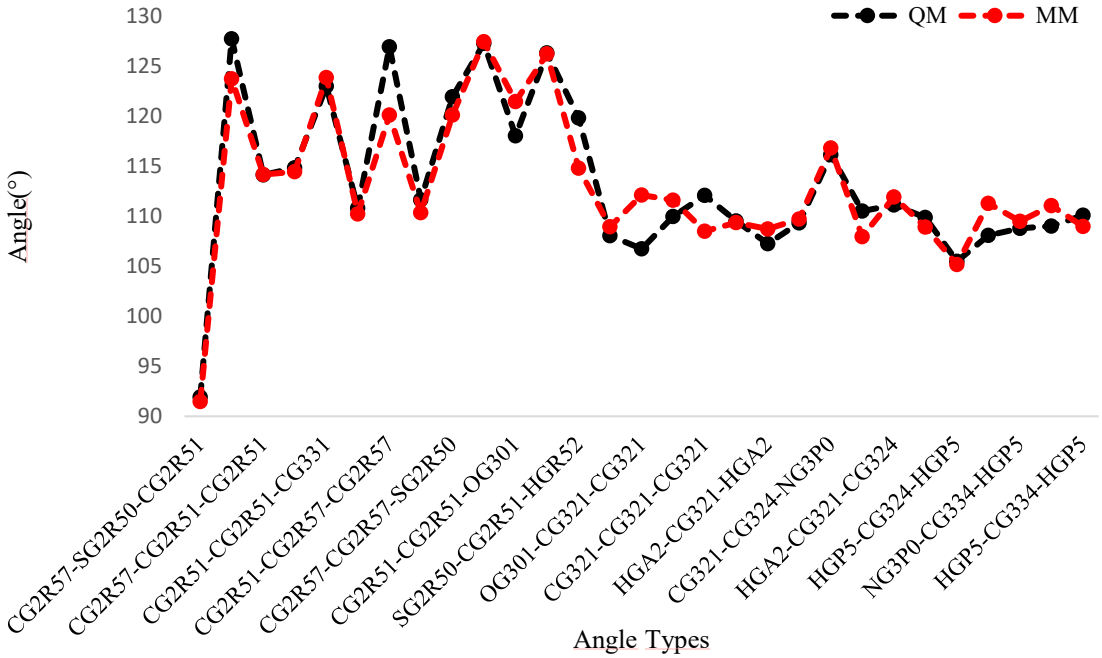


Figure 3.3 Comparison of the angles from the validation of QM and MM.

Validation graphs for QM and MM were given for bond lengths (Figure 3.2) and angles (Figure 3.3). RMSE values for bond lengths and angles were obtained as 0.02 Å

and 2.49˚, respectively. They fall within the acceptable range according to the CHARMM parameters (under 0,03 Å for bond lengths and under 3˚ for angles).

A comparison of the dihedral angles was given in Figure 3.4. RMSE value for the first four graphs was 0.4 kcal/mol, and for the last one is 0.8 kcal/mol. The acceptable value for the CHARMM is 0.5 kcal/mol. The calculated parameters are given in Appendix B.

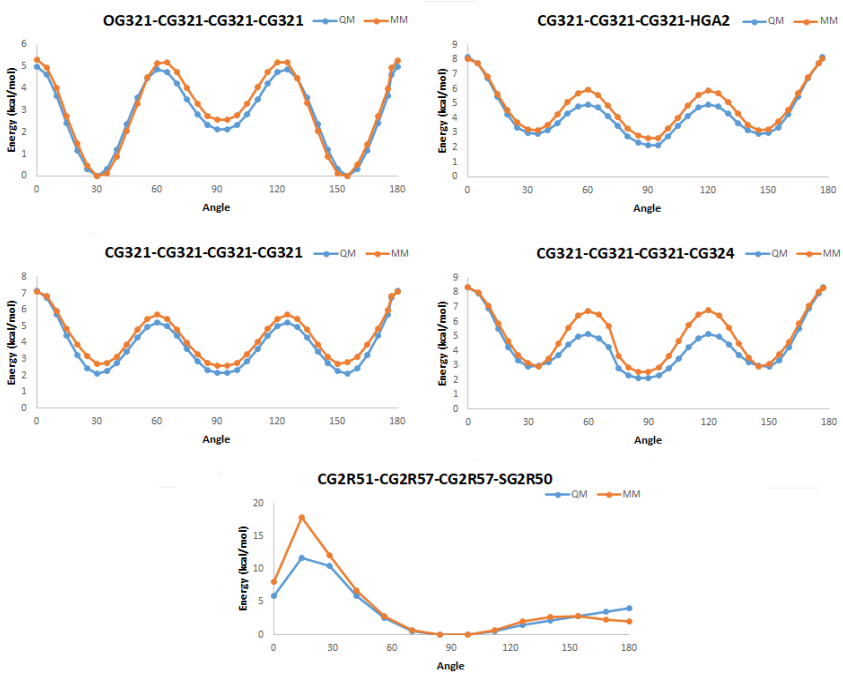


Figure 3.4 Comparison of the dihedral angles from the validation of QM and MM.

3.2 MOLECULAR DYNAMICS SIMULATION RESULTS

The Ree and Rg values are obtained from the analysis of the simulation data for 1000 frames. Ree values are found in the range of [29.9 Å;68.3 Å], [38.4;66.9], [47.7 Å;65.9 Å], [36.3 Å;65.6 Å] for F0, F0-20A, F0-20T, F0-MIX respectively. Their average Ree values are 53.3 Å, 54.3 Å, 59.0 Å and 54.18 Å for F0, F0-20A, F0-20T, F0-MIX respectively. The structures of complexes that have minimum and maximum Ree are given in Figure 3.5, and the structure of complexes that have average Ree are given in appendix A. F0 have the broadest Ree range compared to its complexes. It can be said that it has freely taken a coil or a strand form without DNA.

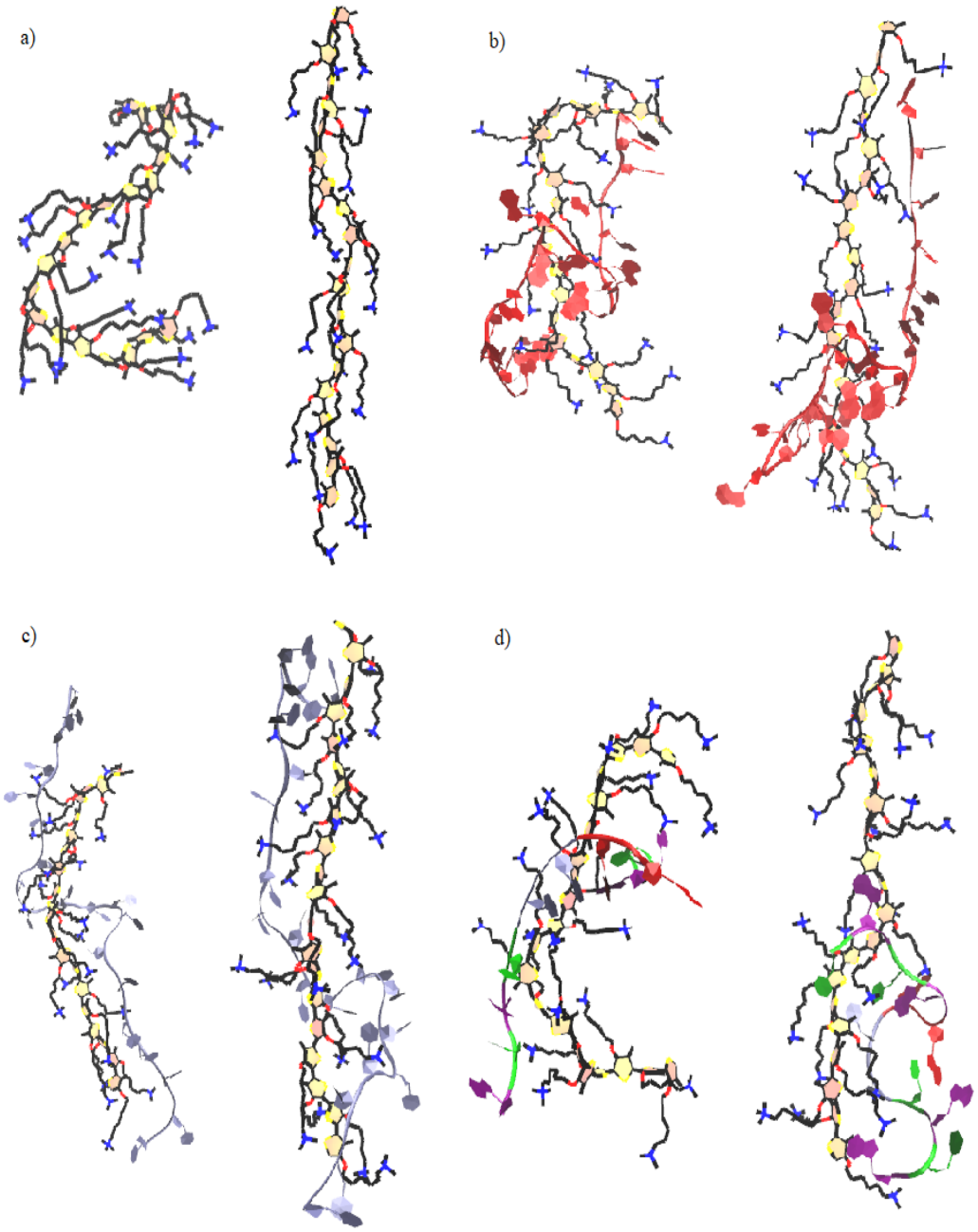


Figure 3.5 The structures of complexes at the minimum and maximum Ree; a) F0, b) F0-20A, c) F0-20T, d) F0-MIX

Average Ree values of F0 with different complexes were displayed in Figure 3.6. Adding ssDNAs to the system changes the Ree of F0; 20A decreases average Ree, whereas 20T increases. The decrease in Ree also can be seen in Rubio-Magnieto's article in 2015 (Rubio-Magnieto 2015). The addition of the MIX chain to the system has almost no effect on the structure of the oligomer backbone. Increasing DNA chain size lowers the average R_{ee} value for 40A and 40T duplexes. This might be explained by the rise of

the number of interactions in the complexes when the DNA sequences doubled. There is no significant difference between F0-MIX and F0-2MIX complexes (only 1.6 Å). This means that the MIX chain, which has twelve nucleotides could not interact enough to fold or unfold F0 (see Figure A2). Adding complementary chains to ssDNA removes the effects of F0-ssDNA duplexes. The average R_{ee} value is about 53 Å for all F0-dsDNA triplexes. It seems that the oligomer backbone recovers its free structure (Figure A2), as seen in some UV/Vis and colorimetric experimental studies of CPTs.

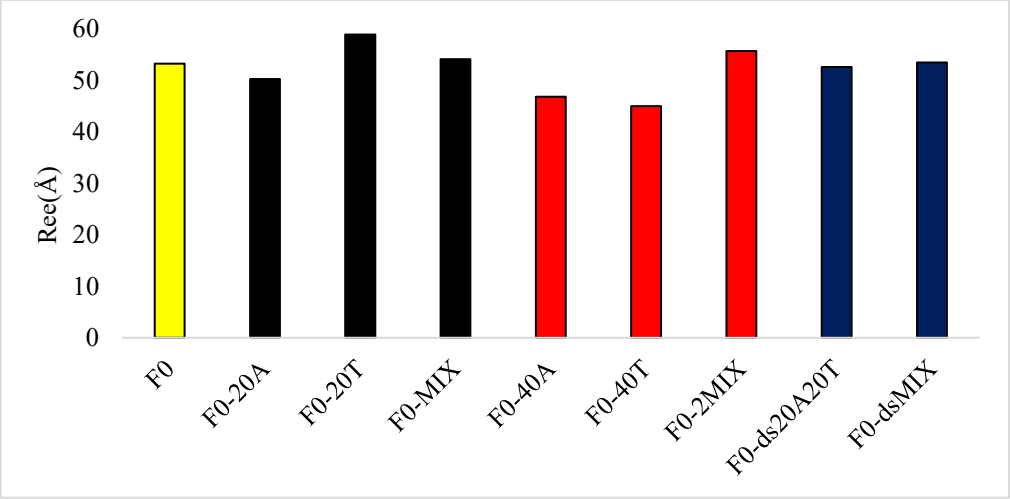


Figure 3.6 Average Ree of F0 with different complexes; only F0 (yellow), F0 with single-stranded 20mer DNAs (black), F0 with single-stranded 40mer DNAs (red), F0 with dsDNAs (blue).

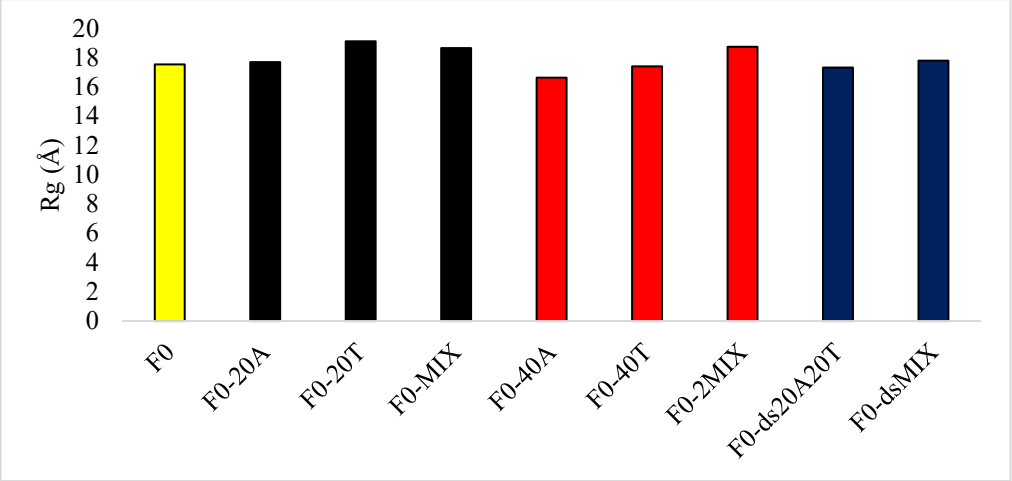


Figure 3.7 Average R_g of F0 with different complexes; only F0 (yellow), F0 with single-stranded DNAs (black), F0 with single-stranded and longer chain DNAs (red), F0 with dsDNAs (blue).

Average R_g values are shown in Figure 3.7 for F0 and complexes with DNAs. Based on this plot, increasing the DNA chain length from 20 mer to 40 mer decreases R_g, which means F0 becomes more compact except for MIX complexes. ssDNAs with only adenine chain are more random coil form than those with thymine chain similar to F0. When the complementary chain was added to ssDNA, the average R_g of the F0 was not dependent on the chain length or sequence of the DNA.

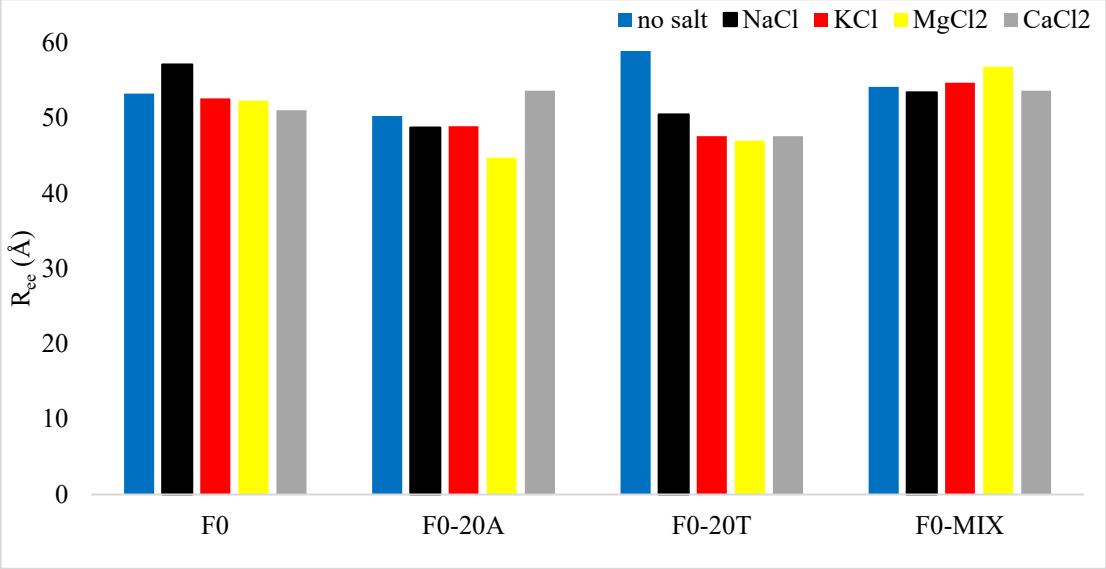


Figure 3.8 Average R_{ee} of F0 and its complexes with DNA with different salts; no salt (blue), NaCl (black), KCl (red), $MgCl_2$ (yellow), and $CaCl_2$ (grey).

The average R_{ee} values of F0 and its complexes in various salt solutions are displayed in Figure 3.8. It was observed that salts have almost no effect on the average R_{ee} of oligomer in F0, F0-20A, and F0-MIX duplexes, with some exceptions. NaCl increased the end-to-end distance of F0 in water. $MgCl_2$ lowered the R_{ee} of F0 in the F0-20A complex. On the other hand, with the addition of salts, the average end-to-end distance decreased significantly in the F0-20T complex (about 11 Å). As a summary, the most affected complex is the F0-20T by the additions of salts that cause a coiled backbone. However, the presence of sodium ions leads to elongate F0.

F0-20A is affected by divalent ions; the Mg^{2+} ion shortened the length of the oligomer. On the other hand, Ca^{2+} ions stretched the oligomer backbone.

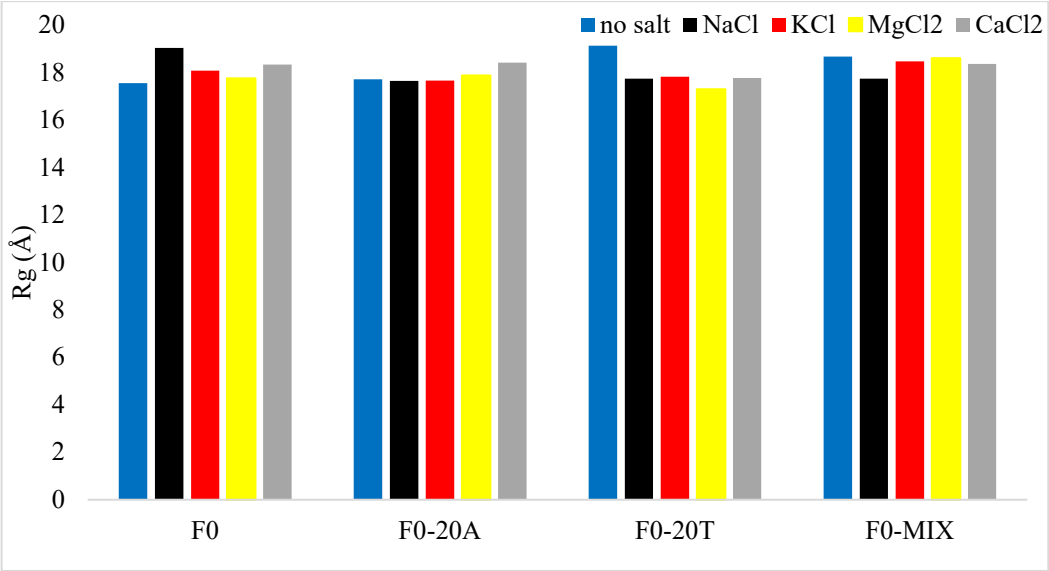


Figure 3.9 Average R_g of F0 and its complexes with DNA with different salts; no salt (blue), NaCl (black), KCl (red), $MgCl_2$ (yellow), and $CaCl_2$ (grey).

The average R_g of F0 and its complexes with salt ions are shown in Figure 3.9. There is almost no effect of the salts in Rg except in the F0-20T duplex. In the F0-20T duplex, the addition of salts to the system decreased the R_g of F0, which means the oligomer becomes more compact.

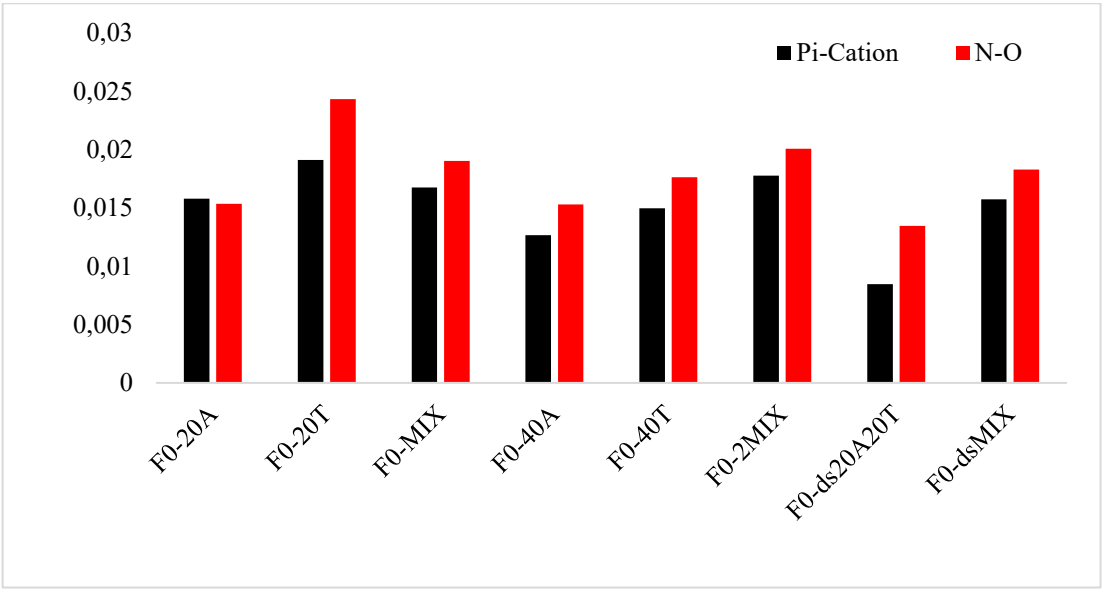


Figure 3.10 Normalized interactions between F0 and DNA strands; π-cation (Nitrogen of the F0 and center of the 6 membered ring of the nucleobases) (black) and electrostatic interactions (N-O) (Nitrogen atoms of F0 and oxygen atoms of the phosphate group of the DNA) (red).

Normalized π-cation (Nitrogen of the F0 and center of the 6 membered ring of the nucleobases) and electrostatic interactions (Nitrogen atoms of F0 and oxygen atoms of the phosphate group of the DNA) are plotted in Figure 3.10. It was detected that complexes exhibit higher electrostatic interactions than π-cation interactions except for the F0-20A duplex. Chains with only thymine have a larger number of interactions than chains with only adenine. Increasing the chain length of DNA from 20 mer to 40 mer decreased the number of interactions in thymine duplexes. On the other hand, increasing the chain length increased the number of interactions on the F0-MIX complex. When the DNA size doubled in adenine complexes, it did not alter the number of electrostatic interactions, but π-cation interactions decreased. The addition of a complementary ssDNA chain lowered both electrostatic and π-cations interactions. When the DNA size doubled in adenine and thymine complexes, some parts of DNA did not interact (did not come together) with F0 since it is much longer than F0. However, in the normalized number of interactions, all possible sides are counted. Thus, the interactions seem to be lowered. The same reason is true for dsDNAs. The complementary DNA interact both with other ssDNA and oligomer in some parts, but there are some pieces of DNA that look like free (see Figure A.2.)

As a summary, F0 interacted more strongly (both π-cation and electrostatic) with thymine DNA strand than with adenine and mix strands. That's why the oligomer responded more to the 20T ssDNA. This result agreed well with experimental and previous theoretical findings (Rubio-Magnieto 2015)

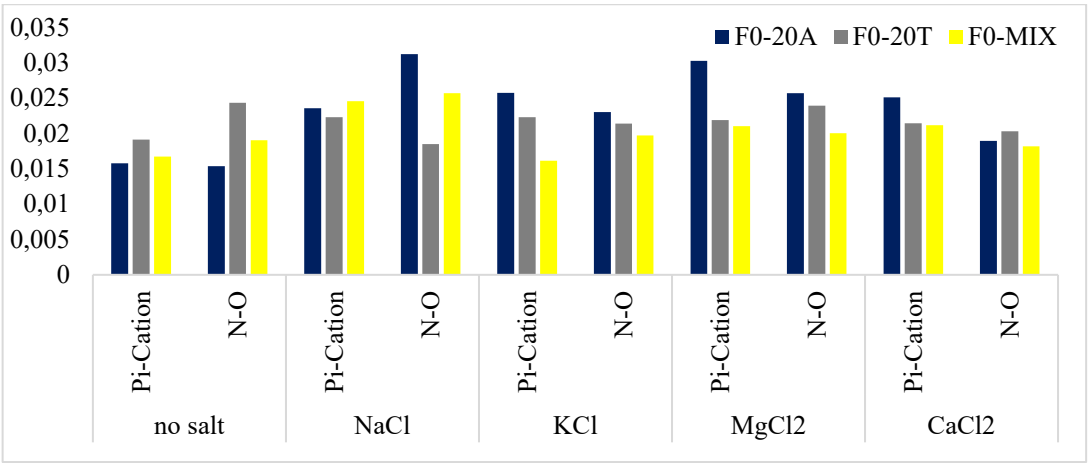


Figure 3.11 Normalized π-cation and N-O interactions between F0 and DNA strands in salts; F0-20A (blue), F0-20T (gray), and F0-MIX (yellow).

Normalized interaction numbers of π-cation and electrostatic interactions are given in Figure 3.11 for complexes in salt solutions. The addition of salt ions to the system increased the number of π-cation and electrostatic interactions in the F0-20A complex. In the F0-20T duplex, π-cation interactions rose with the addition of salt. On the other hand, N-O interactions were decreased. In the F0-MIX duplex, π-cation interactions were increased except for the addition of KCl to the system. The normalized number of N-O interactions of F0-MIX increased with the addition of NaCl, KCl, and $MgCl_2$. However, with the addition of $CaCl_2$ to the system, N-O interaction decreased for the F0-MIX duplex. For F0-20A, π-cation interactions were higher than N-O interactions except with NaCl salt. For the F0-20T duplex, unlike F0-20A, complex N-O interactions were higher than π-cation interactions except for $MgCl_2$. For the F0-MIX complex, π-cation interactions are higher than electrostatic interactions with $MgCl_2$ and $CaCl_2$. However, N-O interactions were higher than π-cation interactions with KCl and NaCl. This graph shows that different salts affect the number of interactions; unlike R_{ee} and Rg, no correlations are observed in the interaction numbers about the size and charges of the ions. The most affected duplex by the salts is the F0-20A complex. The addition of NaCl salt doubled N-O interactions while $MgCl_2$ addition doubled π-cation interaction.

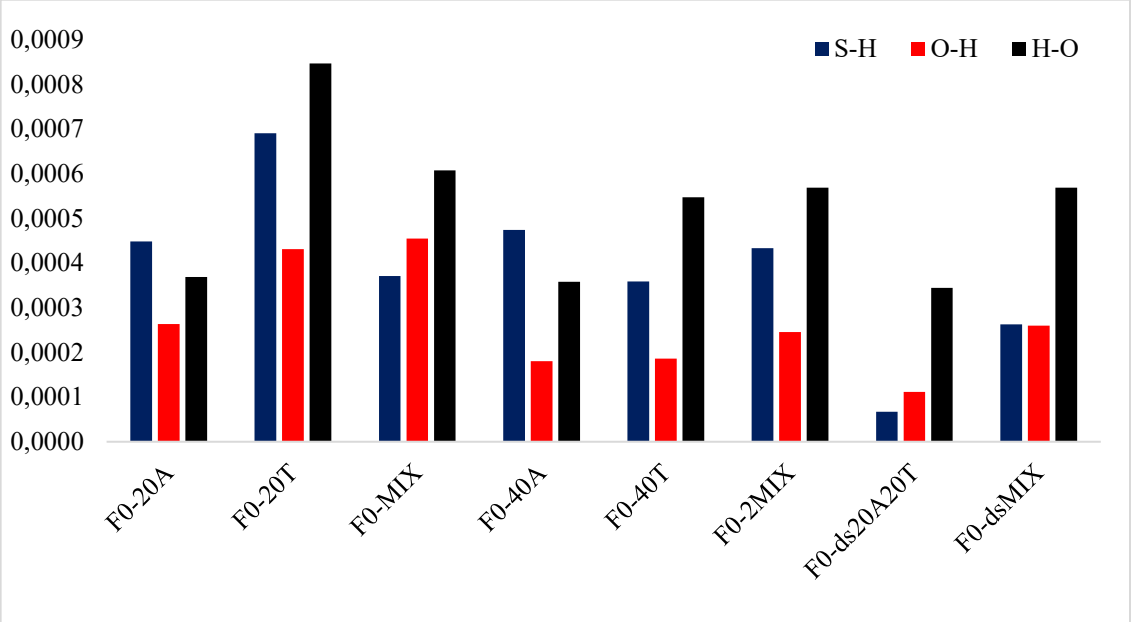


Figure 3.12 Normalized number of interactions between F0 and DNA complexes; S-H (sulfur atom of F0 and hydrogens of DNA) (Blue), O-H (Oxygens of F0 and hydrogens of F0) (red), and H-O (Hydrogens of F0 and oxygens of DNA) (black).

Normalized interaction numbers of hydrogen bonding (S-H, O-H, and H-O) are shown in Figure 3.12. Only adenine and thymine chains have similar trends within themselves. For the F0-20A duplex, S-H bonding is the highest, and O-H has the smallest value. On the other hand, the F0-20T complex has the highest H-O bonding interactions while having the lowest O-H interactions. When the chain length doubled, both adenine and thymine did not lose their trend, and the overall number of interactions decreased since ssDNA has a lot of non-interactive atoms caused by the size difference. For all MIX complexes, H-O bonding has the highest value. S-H bonding has the lowest number of interactions in the F0-MIX, but for the F0-2MIX, the O-H bonding is the lowest. Unlike adenine and thymine complexes, reduction of O-H interactions occurred in the MIX complex. The addition of the complementary ssDNA decreased the normalized number of interactions in the F0-20A20T triplex.

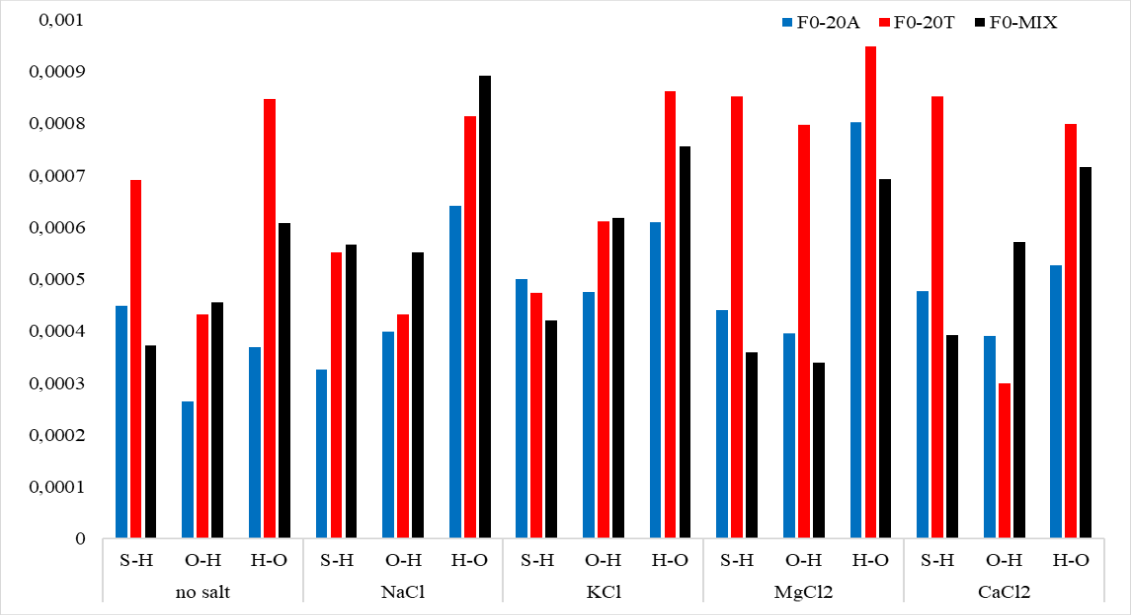


Figure 3.13 Normalized number of hydrogen bonding between F0 and ssDNA with different salts (NaCl, KCl, $MgCl_2$, and $CaCl_2$); F0-20A (blue), F0-20T (red), and F0-MIX (black).

The normalized number of hydrogen bonding in different salts is given in Figure 3.13. Each salt affected the number of hydrogen bonding differently. No correlation is observed between the size and valence of the ions.

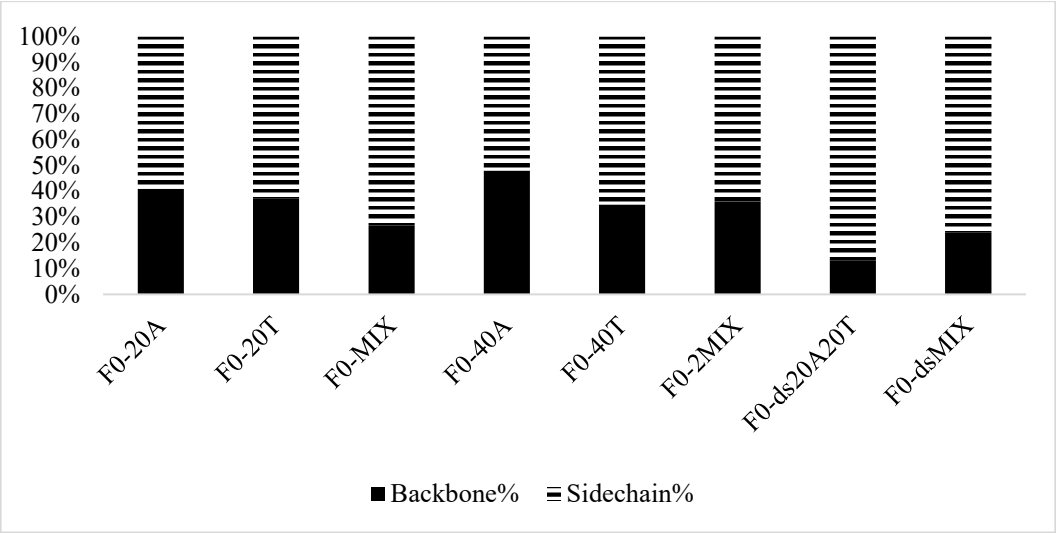


Figure 3.14 Percentages of the backbone (solid) and side chain (dashed) of the F0 interacted with the DNA.

Backbone and side chain of F0 interaction percentages were calculated as described in the method part to deduce which part of the oligomer participates more contacts with DNAs, and they are displayed in Figure 3.14. In all the complexes, F0 highly interacted with the side chain as expected because of the size of the side chain except F0-40A, in which both side and backbone contribute almost equally. F0-20A duplex interacted with backbone more than F0-20T duplex (about 4%). Doubling the chain length of the DNA increases the backbone percentage of 20A and MIX complexes by about 8% and 10%, respectively. On the other hand, for the F0-20T, doubling the chain length of the DNA decreases the backbone percentage by about 3%. The addition of the complementary DNA chain to the system increases the side chain percentage highly.

The interaction percentages from the backbone and side chain of oligomer for complexes in different salt solutions are shown in Figure 3.15. In all salts, F0-20T has the largest backbone interactions. F0-20A has a higher side chain percentage than F0-MIX in NaCl. On the other hand, F0-MIX has a higher side chain percentage in KCl, $MgCl_2$, and $CaCl_2$. The addition of all kinds of salts to the complex solution cause to more interactions from the backbone of the oligomer for the Adenine duplex though they almost have no influence on the MIX and 20T duplexes except the presence of Ca^{2+} ion in the 20T complex. Na^+ ions increased the interactions considerably from the side chain of F0 with 20A.

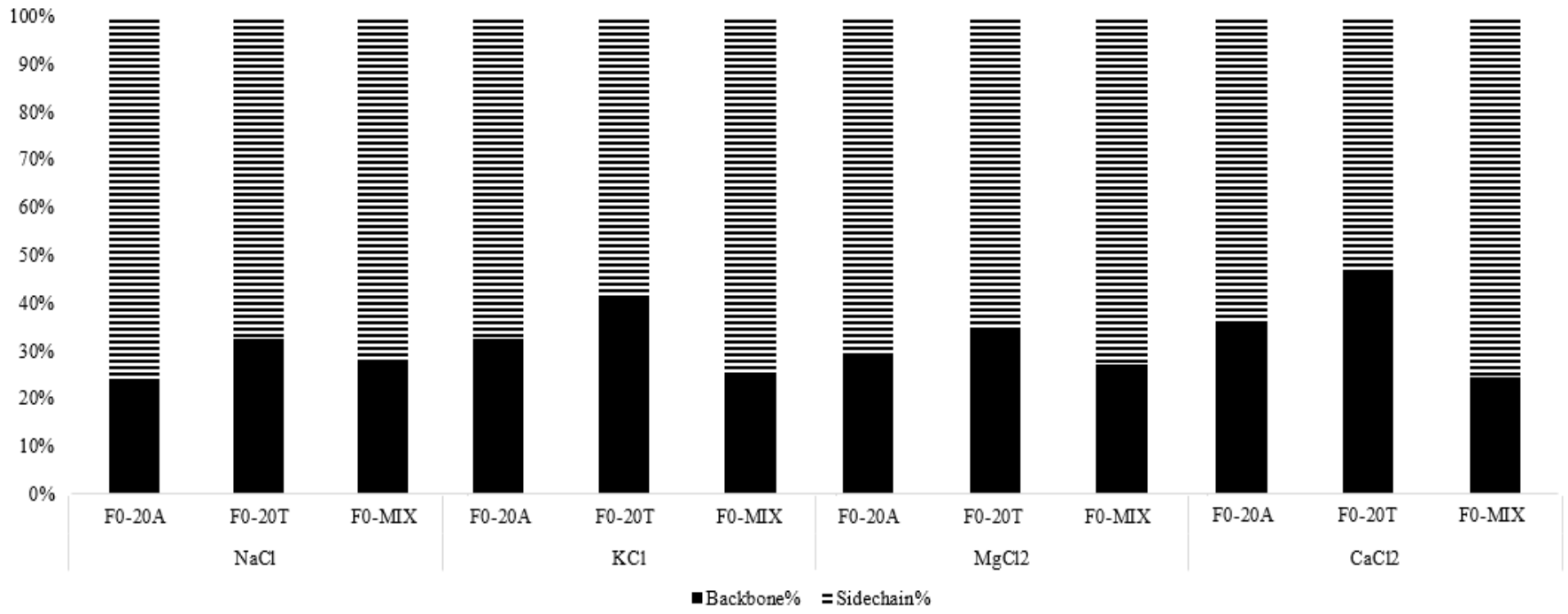


Figure 3.15 Percentages of the backbone (solid) and side chain (dashed) of the F0 interacted with the DNA in salt.

CHAPTER 4

CONCLUSION

In the first part of this thesis, parametrization of poly- N, N, N-trimethyl-3-(4-methylthiophen- 3-yl) oxy) hexan-1-aminium was done. Some parameters were taken from our research group's published article (Kıbrıs 2021). Missing parameters were calculated with ffTK via the VMD program. QM and MM validation was performed, and RMSE values were found as 0.02 Å and 2.49˚ for bonds and angles, respectively. For dihedrals, RMSE values are 0.4 kcal/mol and 0.8 kcal/mol. All the RMSE values except 0.8 kcal/mol are under the accepted values for CHARMM.

Using those parameters, MD simulations were carried out to understand the response of F0 upon the complexation with various DNA strands in different conditions (with/without salt).

The contraction and elongation of the oligomer backbone upon the addition of 20A and 20T were observed, respectively. Increasing the length of DNA decreased the R_{ee} of F0 with 20A and 20T, while the MIX chain was not affected significantly. F0-dsDNA triplexes have almost the same R_{ee} values as F0, and this suggests that the addition of complementary ssDNA causes back transformation of the oligomer backbone structure in free form. It seems that the oligomer backbone recovers its structure before complexation, as seen in some UV/Vis and colorimetric experimental studies of CPTs.

Salts addition almost did not affect the R_{ee}, except for the F0-20T complex, which decreased significantly. Based on the simulation results, the R_g of F0 is influenced by DNA chains. R_g of F0 decreased when the 20T chain was added instead of 20A. This finding supports the results from Rubio-Magnieto's article (Rubio-Magnieto 2015).

In the interaction analysis, it was observed that all the complexes except F0-20A favor electrostatic interaction instead π-cation interaction. F0-20T complexes made more interactions than F0-20A. When DNA size doubled, adding a complementary chain to ssDNA decreased the number of interatomic interactions. The addition of salts to the system leads to more interactions in the F0-20A duplex. All the complexes prefer reacting with the side chain more than the backbone.

REFERENCES

Alder, B. J., and Wainwright, T.E. "Phase Transition for a Hard Sphere System." *The Journal of Chemical Physics 27 (5):*, 1957: 1208-9.

Ammanath, G., Delachi, C.G., Karabacak, S., Ali, Y., Boehm, B.O., Yildiz, U.H., Alagappan, P., Liedberg, B. "Colorimetric and Fluorometric Profiling of Advanced Glycation End Products." *ACS Applied Materials & Interfaces*, 2022: 94-103.

Carreon, A.C., Santos, W.L., Matson, J.B., So, R.C. "Cationic polythiophenes as responsive DNA-binding polymers." *Polym. Chem. 5*, 2014: 314-317.

Dennington, R., Keith, T.A., Millam, J.M.,. "GaussView Version 5." *Semichem Inc. Shawnee Mission KS*, 2009.

Feng, F., He, F., An, L., Wang, S., Li, Y., Zhu, D. "Fluorescent Conjugated Polyelectrolytes for Biomacromolecule Detection." *Advanced Materials 20(15)*, 2008: 2959-2964.

Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Petersson, G.A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A.V., Bloino, J., Janesko, B.G., Gomperts, R., Mennucci, B., Hratchian, H.P. "Gaussian 09 Revision A.02." *Gaussian, Inc., Wallingford CT*, 2009.

Hanwell, M.D., Curtis, D.E., Lonie, D.C., Vandermeersch, T., Zurek, E., Hutchison, G.R. "Avogadro: An advanced semantic chemical editor, visualization, and analysis platform." *Journal of Cheminformatics*, 2012.

Humphrey, W., Dalke, A., Schulten, K. "VMD Visual Molecular Dynamics." *Journal of Molecular Graphics 14*, 1996: 33-38.

Kıbrıs, E., Barbak, N. N., Irmak N. E. "CHARMM force field generation for a cationic thiophene oligomer with ffTK." *Journal of Molecular Modeling 27*, 2021.

Leclerc, M., Faid, K. "Electrical and Optical Properties of Processable Polythiophene Derivatives: Structure-Property Relationships." *Advanced Materials 357 (14)*, 1997: 1087-1094.

Leclerc, M., Ho, H., Boissinot, M., Bergeron, M., Corbeil, G., Doré, K., Boudreau, D. "Colorimetric and Fluorometric Detection of Nucleic Acids Using Cationic Polythiophene Derivatives." *Angewandte Chemie International Edition 41 (9)*, 2002: 1548-1551.

Leclerc, M., Ho, H.-A., Bra-Abrem, M.,. "Optical sensors based on hybrid dna/conjugated polymer complexes." *Chemistry A European Journal 11(6)*, 2005: 1718-1724.

Li, C., Numata, M., Takeuchi, M. and Shinkai, S. "A Sensitive Colorimetric and Fluorescent Probe Based on a Polythiophene Derivative for the Detection of ATP." *Angewandte Chemie International Edition, 44*, 2005: 6371-6374.

Mayne, C.G., Saam, J., Schulten, K.,Tajkhorshid, E., Gumbart, J.C. "Rapid parameterization of small molecules using the Force Field Toolkit." *J. Comput. Chem. 34*, 2013: 2757-2770.

McCammon, J. A., Gelin, B. R. & Karplus, M. "Dynamics of folded proteins." *Nature 267*, 1977: 585-590.

McCullough, R. D. "The Chemistry of Conducting." *Adv. Mat. 2*, 1998: 93-116.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M. N., Teller, A. H., and Teller, E. "Equation of state calculations by fast computing machines." *J. Chem. Phys., 21*, 1953: 1087-1092.

Ozenler, S., Kaya, H., Elmaci, N., Yildiz, U.H. "Transition-Metal Free Direct C-H Arylation of Thiophene in Aqueous Media via Potassium Peroxymonosulfate." *ChemistrySelect*, 2019a: 8516-8521

Ozenler, S., Yucel, M., Tuncel, O., Kaya, H., Ozcelik, S., Yildiz U.H. "Single Chain Cationic Polymer Dot as a Fluorescent Probe for Cell Imaging and Selective

Determination of Hepatocellular Carcinoma Cells." *Analytical Chemistry*, 2019b: 10357-10360.

Patil, A. O., Y. Ikenoue, F. Wudl, and A. J. Heeger. "Water-Soluble Conducting Polymers." *Journal of the American Chemical Society 109 (6)*, 1987: 1858-1859.

Phillips, J.C., Hardy, D.J., Maia, J.D.C., Stone, J.E., Ribeiro, J.V., Bernardi, R.C., Buch, R., Fiorin, G., Hnin, J., Jiang, W., McGreevy, R., Melo, M.C.R., Radak, B.K., Skeel, R.D., Singharoy, A., Wang, Y., Roux, B., Aksimentiev, A., Luthey-Schulten, Z. "Scalable molecular dynamics on CPU and GPU architectures with NAMD." *Journal of Chemical Physics*, 2020.

Rahman, A. "Correlations in the Motion of Atoms in Liquid Argon." *Physical Review*, 1964: A405-11.

Rubio-Magnieto, Azene, E.G., Knoops, J., Knippenberg, S., Delcourt, C., Thomas, A., Richeter, S., Mehdi, A., Dubois, P., Lazzaroni, R., Beljonne, D., Clment, S., Surin, M. "Self-assembly and hybridization mechanisms of DNA with cationic polythiophene." *Soft Matter 11*, 2015: 6460-6471.

Rubio-Magnieto, Thomas, A., Richeter, S., Mehdi, A., Dubois, P., Lazzaroni, R., Clment, S., Surin, M. "Chirality in dna-conjugated polymer supramolecular structures: insights into the self-assembly." *Chemical Communications*, 2013: 5483-5485.

Thomas, S.W., Joly, G.D., Swager, T.M. "Chemical Sensors Based on Amplifying Fluorescent Conjugated Polymers." *Chem. Rev. 2007*, 2007: 1339-1386.

Vanommeslaeghe, K, Hatcher, E, Acharya, C, Kundu, S, Zhong, S, Shim, J, Darian, E, Guvench, O, Lopes, P, Vorobyov, I, Mackerell AD Jr. "Charmm general force field: a force field for drug-like molecules compatible with the charmm all atom additive biological force fields." *J Comput Chem 31(4)*, 2010: 671-690.

Wang, L., Feng, Q., Wang, X., Pei, M., Zhang, G. "A novel polythiophene derivative as a sensitive colorimetric and fluorescent sensor for anionic surfactants in water." *New J. Chem. 36*, 2012: 1897-1901.

Yildiz, U.H., Sheng, C.W., Mailepessov, D. "Real-time determination of the activity of ATPase by use of a water-soluble polythiophene." *Anal Bioanal Chem. 404*, 2012: 269-275.

Yucel, M., Koc, A., Ulgenalp, A., Akkoc, G.D., Ceyhan, M., Yildiz, U.H. "PCR-Free Methodology for Detection of Single-Nucleotide Polymorphism with a Cationic Polythiophene Reporter." *ACS Sensors*, 2021: 950-957

APPENDIX A

SNAPSHOTS OF F0 IN DIFFERENT COMPLEXES AT AVERAGE R_{ee}

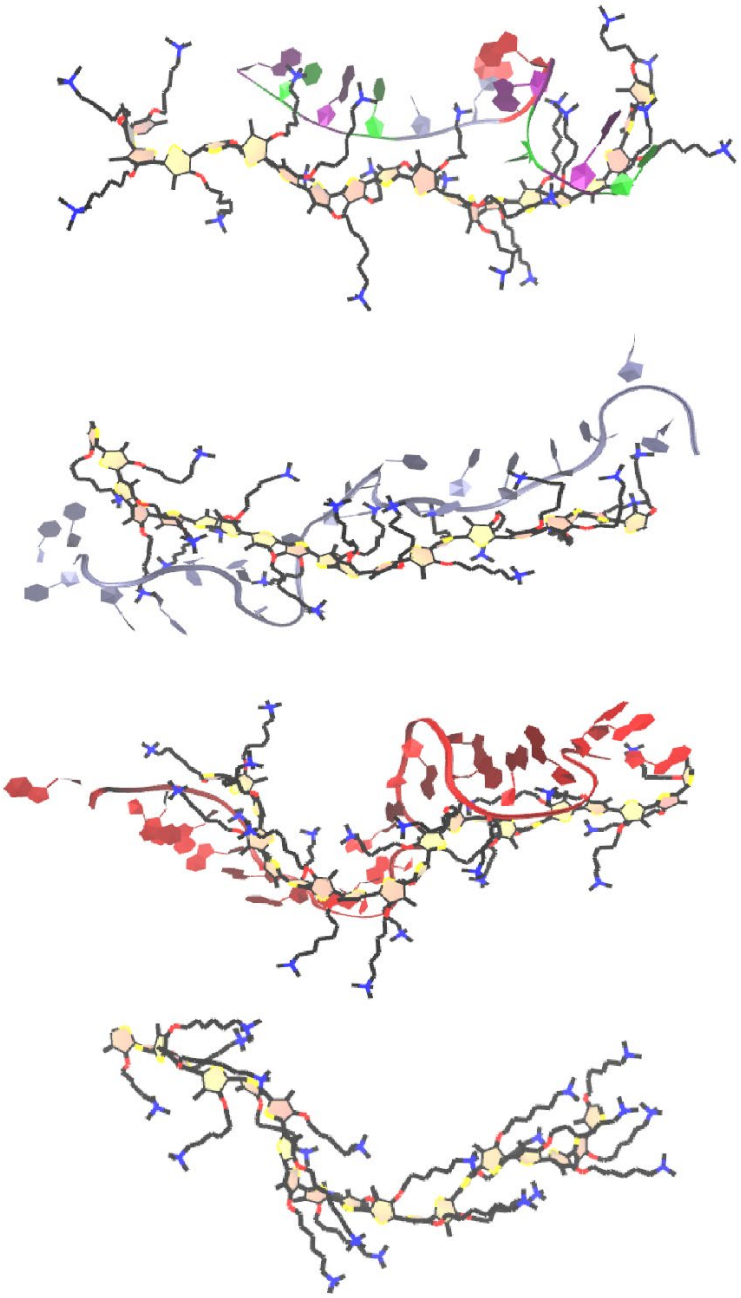


Figure A.1 Snapshots of F0, F0 in F0-20A, F0 in F0-20T, and F0 in F0-MIX at average R_{ee} respectively

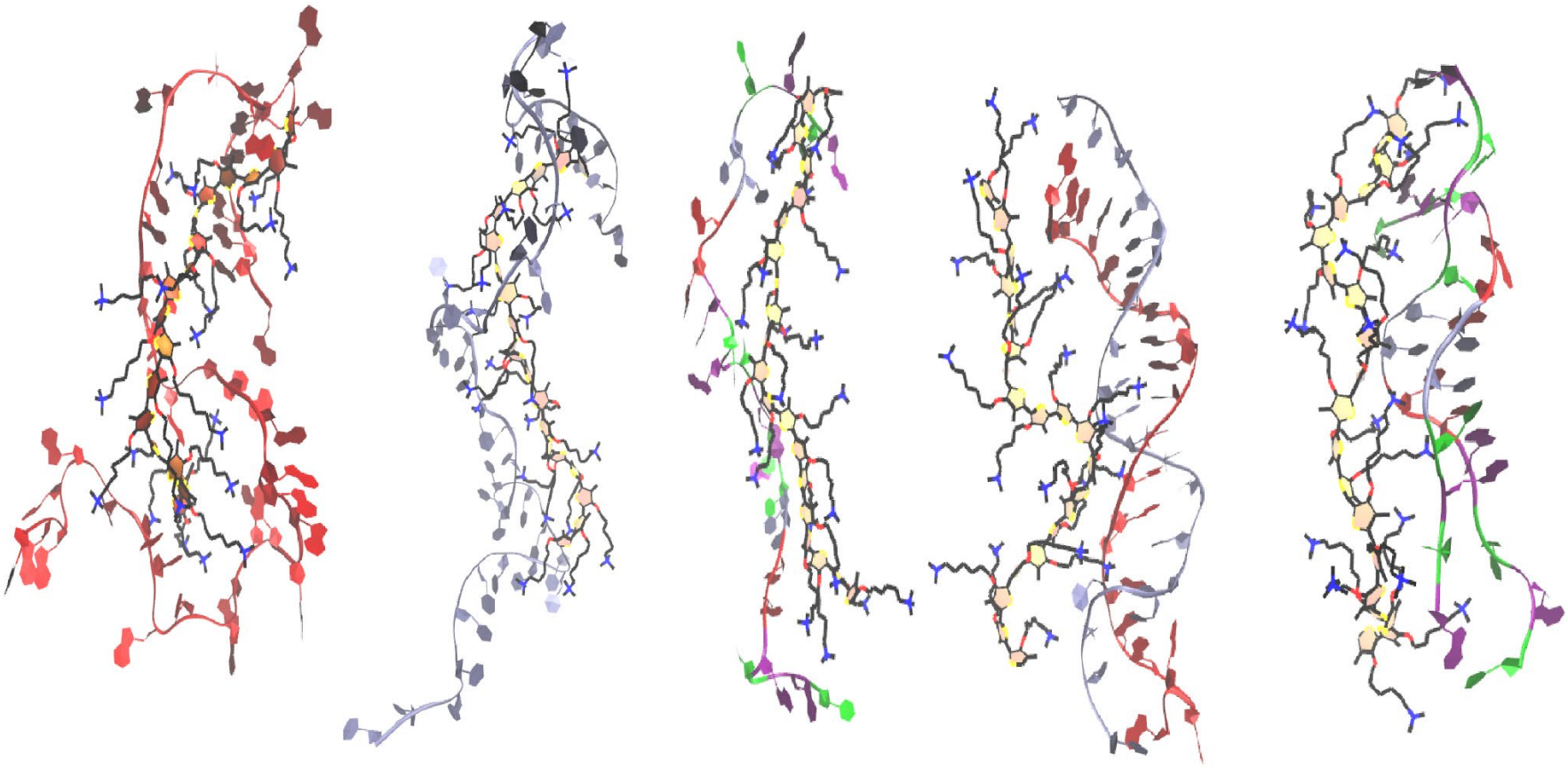


Figure A.2 Snapshots of F0 in F0-40A, F0 in F040T, F0 in F0-MIX, F0 in F0-ds20AT, F0 in F0-dsMIX at average R_{ee} respectively.

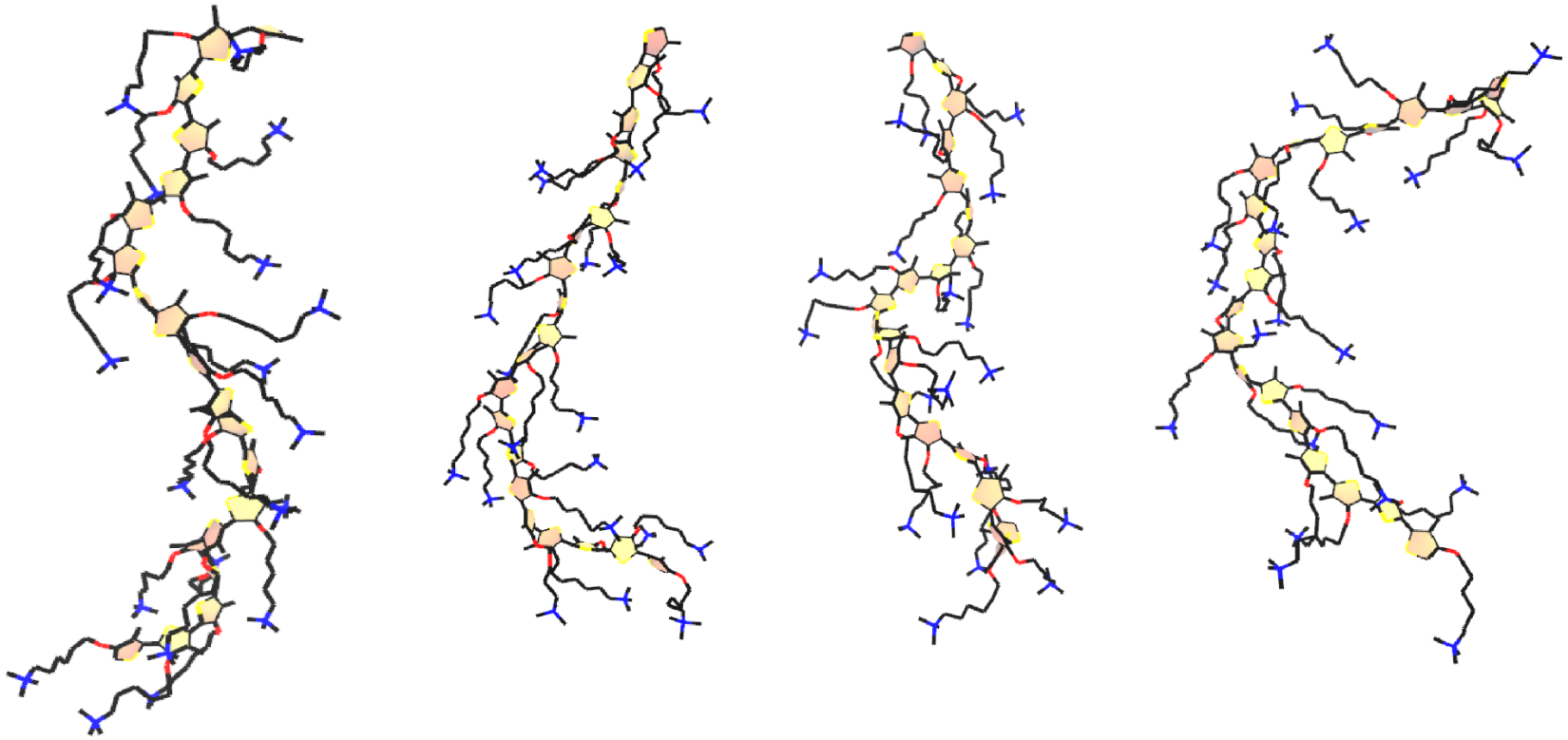


Figure A.3 Snapshots of F0 with different salts; NaCl, KCl, $MgCl_2$, and $CaCl_2$ respectively.

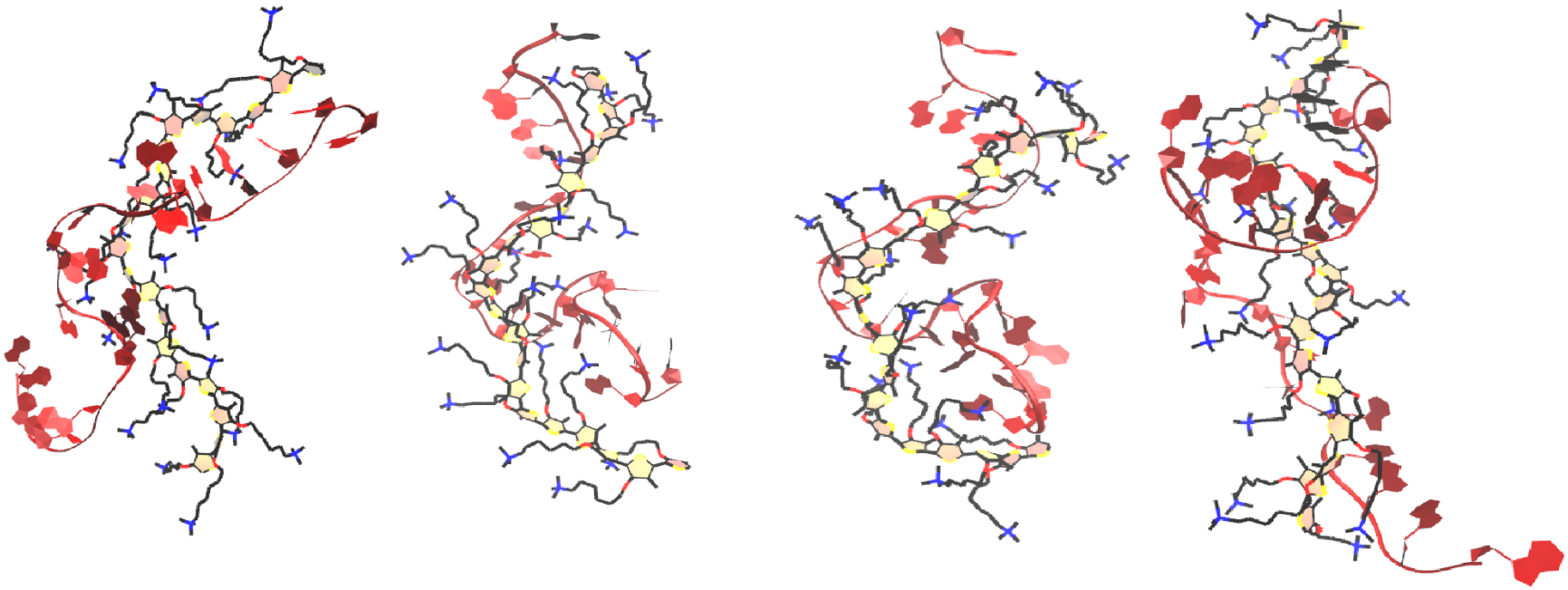


Figure A.4 Snapshots of F0 in F0-20A with different salts; NaCl, KCl, $MgCl_2$, and $CaCl_2$ respectively.

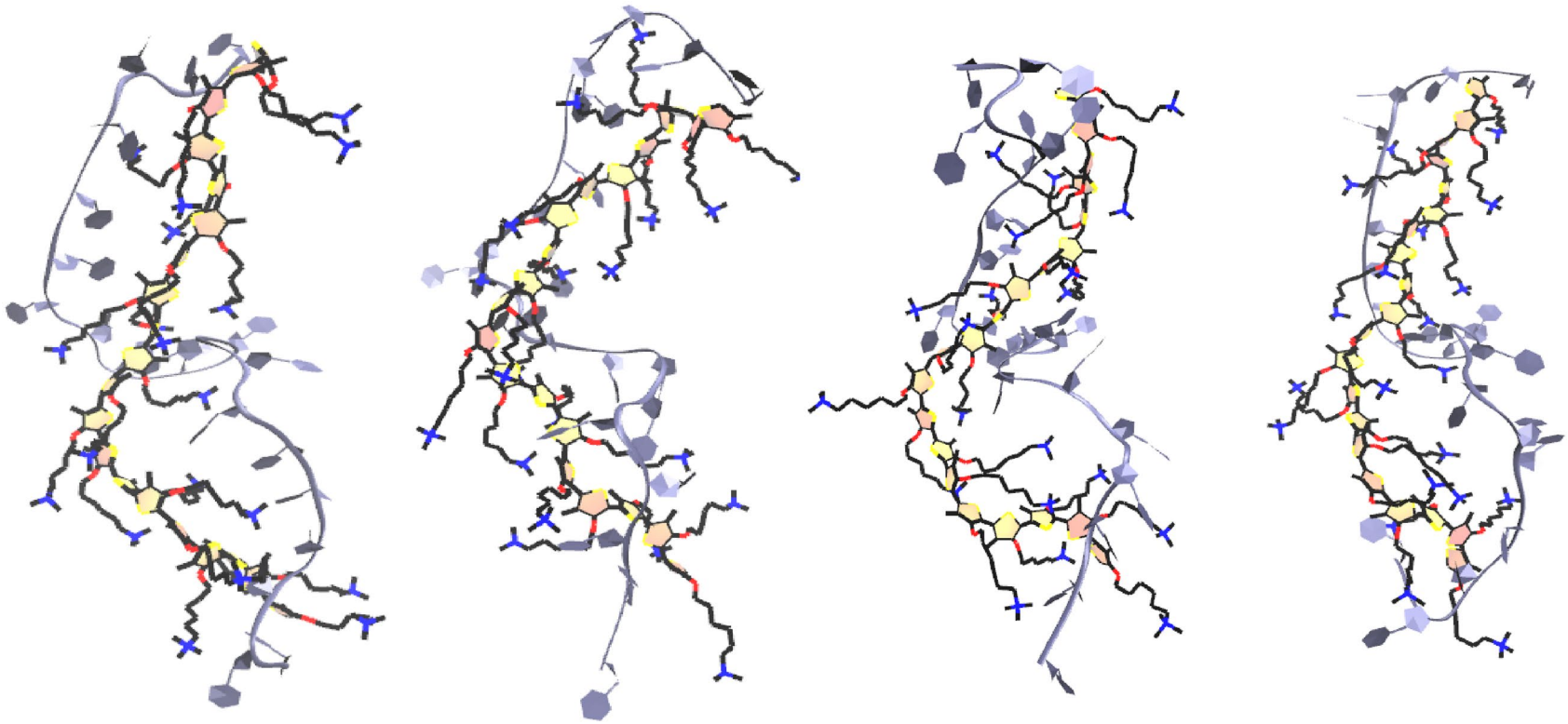


Figure A.5 Snapshots of F0 in F0-20T with different salts; NaCl, KCl, $MgCl_2$, and $CaCl_2$ respectively.

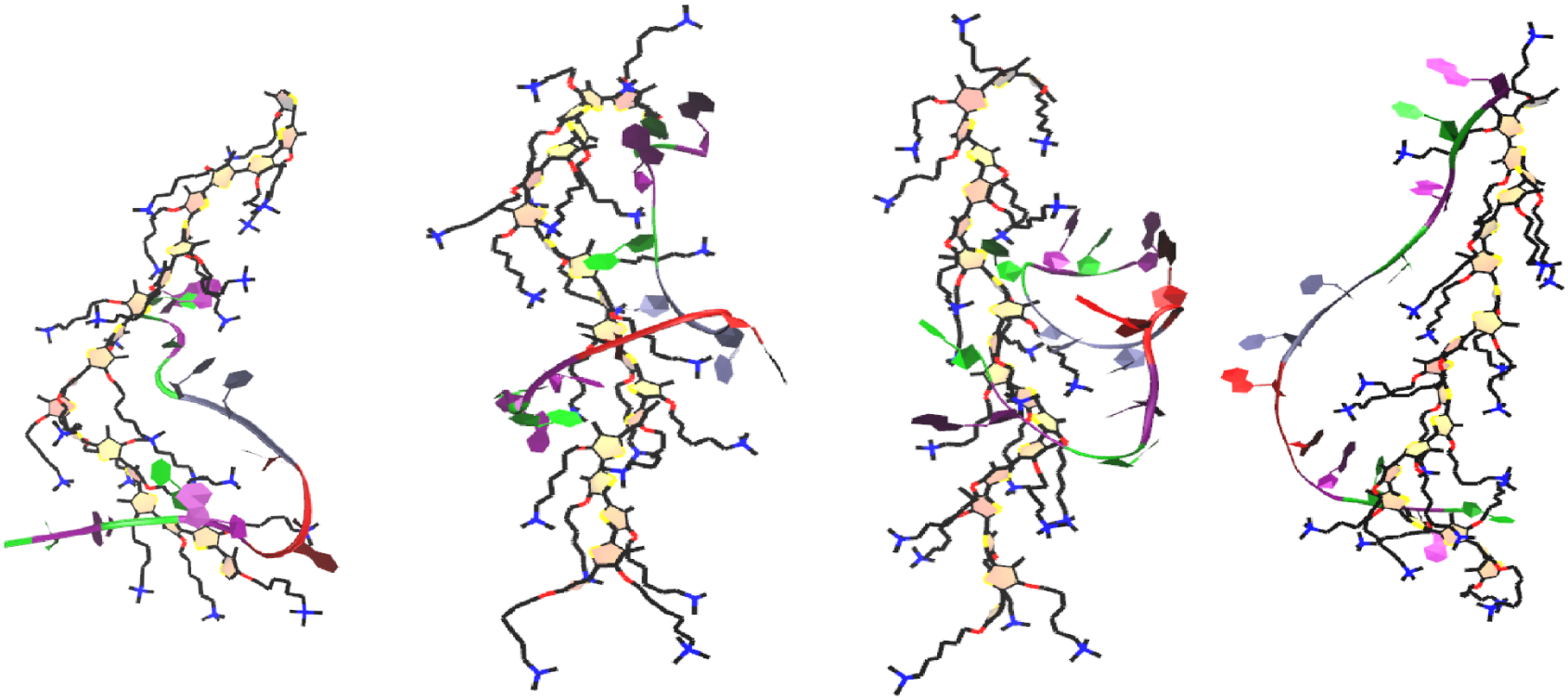


Figure A.6 Snapshots of F0 in F0-MIX with different salts; NaCl, KCl, $MgCl_2$, and $CaCl_2$ respectively.

APPENDIX B

PARAMETERS THAT WERE CREATED

ANGLES	**ATOM TYPES**	k_θ	θ_0	
	GC321 CG321 CG321	59,657	112,84	
DIHEDRALS	**ATOM TYPES**	**$k\Phi$**	**n**	**δ**
	CG321 CG321 CG321 CG321	0,005	2	0
	CG321 CG321 CG321 CG321	0,497	3	180
	CG321 CG321 CG321 CG321	0,086	4	0
	OG301 CG321 CG321 CG321	1,095	1	180
	OG301 CG321 CG321 CG321	0,744	2	0
	CG321 CG321 CG321 HGA2	0,284	3	0
	CG321 CG321 CG321 CG324	0,043	3	0
	CG2R51 CG2R57 CG2R57 SG2R50	2,996	2	180