ACTIVITY OF TOPOTECAN AS TOPOISOMERASE IB INHIBITOR

by

Semiha Kevser Bali

B.S., Molecular Biology and Genetics, Boğaziçi University, 2015

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Department of Chemistry
Boğaziçi University

2018

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Prof. Viktorya Aviyente, for her continuous support, motivation, guidence and patience. She has been a great inspiration for me, and it is always a pleasure to be her student. I would also like to thank my thesis supervisor, Assoc. Prof. Saron Catak, for her guidance and valuable comments.

I wish to extend my appreciation to the members of my committee: Prof. Ilknur Doğan and Prof. Aylin Konuklar, for agreeing to be in my committee on such a short notice, and their valuable advices and comments.

I would like to thank Dr. İlke Ugur and Assist. Prof. Antoine Marion for their help, guidance and patience. This work would not be completed without their insightful comments and advices.

It is quite hard to believe that four years have passed since my first day in KB332A. My sincere thanks go to the former and current members of CCBG for the warm and friendly environment. Especially my dear friend E. Birsen Boydas, thank you for being there for the good, the bad, the weird and The Office times. If we had Dundies, you would definitely get the Birsen Boydaş Award. I want to thank Oğuzhan Kucur, Başak Koca, Volkan Fındık and Haydar T. Turan for their support and friendship for all these years. It is always a pleasure to work and have fun with you all.

My biggest gratitude goes to Duygu Karademir. You are an awesome person and I am really lucky to have you as my friend. You know, maybe the real thesis is friendship.

I am thankful for the financial support from the TUBITAK (Project No: 215Z399) that funded this work.

Lastly, I wish to thank my greatest supporters in life, my parents. I wouldn't be here if it weren't for you. Thank you for supporting me all my life sentimentally. I am really lucky to have you all.

ABSTRACT

ACTIVITY OF TOPOTECAN AS TOPOISOMERASE IB INHIBITOR

Human Topoisomerase I (TopoI) is an enzyme that relaxes the supercoils on the DNA. First, it creates a nick only on one strand of DNA by covalently binding to it, and relaxes DNA by rotating the nicked strand around the intact one and re-ligating it back. Camptothecin (CPT) and its derivatives are known to target only Topoisomerase enzymes. When CPT or its analog Topotecan (TPT) is added to DNA-TopoI complex, the drug intercalates where the nick occurs and prevents relaxation by turning TopoI into a DNA-damaging agent. TPT has two forms: lactone and carboxylate forms, which are in equilibrium at neutral pH. It was shown that the lactone form is the "active" form of the drug, however, in the crystal structure of ternary complex (TPT-DNA-TopoI) both forms were intercalating at the same site. This finding raises the question regarding the cause of activity difference between two forms despite binding to the same site. In this study, using Molecular Dynamics (MD) and hybrid Quantum Mechanics/Molecular Mechanics (QM/MM) approaches the difference in activity is investigated. 450 ns long MD simulation of both forms was performed and interaction energies were calculated using QM/MM method. Decomposition of the interaction energies showed that the interaction with TopoI is the reason for their activity difference. When the contribution of residues on the interaction energy was investigated, it was found that the most striking effect was due to K532 residue, and the topological analysis has further supported this finding. In addition, when the hydrogen bond networks around the drug were investigated, it was observed that in lactone form K532 was interacting with the drug, which was lost in carboxylate form. In conclusion, lactone form of TPT is held by K532 residue strongly and N722 forms a hydrogen bond network around the binding pocket in the lactone form only.

ÖZET

TOPOİZOMERAZ IB İNHİBİTORÜ OLARAK TOPOTECAN

Topoizomeraz enzimleri, DNA'nın kendini eşlemesi/transkripsiyonu sırasında meydana gelen süpersarmalların rahatlatılması sürecinde önemli role sahiplerdir. İnsan Topoizomeraz enzimi IB (TopoI), DNA üzerinde tek bir zinciri kesip diğer zincirin etrafında dönüş yaparak supersarmalları düzeltmektedir. Bu mekanizmanın engellenmesi, çift zincir kırıklarına sebep olabilmekte ve hücreyi öldürebilmektedir. Bu rolünden dolayı, kanser tedavisinde önemli bir hedef olmuşlardır. Topotecan (TPT), Camptothecin (CPT)'den türetilmiş bir yarı-sentetik ilaç molekülüdür. DNA ve ona kovalent bağlı olarak bulunan TopoI'in olduğu ortama eklenmesi halinde, ikili kompleksin birbirlerine bağlı oldukları yerden araya girmekte ve süpersarmal rahatlatılmasını engellemektedir. TPT'nin iki şekli bulunmaktadır: lakton ve karbosil yapıları. Aralarından lakton yapısının aktif olduğu keşfedilmiş olmakla birlikte, Staker ve ark. tarafından 2002 yılında elde edilen kristal yapıda her iki yapı da DNA-TopoI kompleksinde aynı yere bağlanabilmektedir. Şimdiye kadar yapılmış olan çalışmalarla laktonun neden sadece DNA-TopoI kompleksine bağlandığı açıklanmış olsa da, TPT'nin engelleme mekanizması tam olarak aydınlatılmamıştır. Bu çalışmada, iki yapı arasındaki farkın sebebi, Moleküler Dinamik ve hibrit Quantum Mekanik/Moleküler Mekanik yöntemleri kullanarak açıklanmıştır. İki ayrı DNA-TPT-TopoI kompleksi için 450 ns boyunca MD benzetimi yapılmıştır. Yapılan QM/MM etkileşim enerji hesaplamaları sonucu, lakton ve karboksilat yapılarının arasındaki en büyük farkın K532 amino asitinden kaynaklandığı bulunmuştur. Yapıların çevreleriyle yaptıkları hidrojen bağları incelenmesi sonucu, N722 kalıntısının da DNA'nin stabilize edilmesinde rolü olduğu anlaşılmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

ΔE	Total interaction energy
$\Delta E_{residue}$	Total interaction energy without the residue
$\Delta^2 E_{DNA,TopoI}$	Interaction energy between DNA and TopoI
$\Delta^2 E_{DNA,TPT}$	Interaction energy between DNA and TPT
$\Delta^3 E_{DNA,TopoI,TPT}$	Energy due to the interction of DNA, TopoI and TPT

LIST OF ACRONYMS/ABBREVIATIONS

-1A	Adenine at position -1
-1T	Thymine at position -1
-2A	Adenine at position -2
+1C	Cysteine at position -1
+1G	Guanine at position +1
Ala	Alanine
CPT	Camptothecin
Cys	Cysteine
D533	Asparagine 533
dsDNA	Double-stranded DNA
DFT	Density functional theory
E356	Glutamate 356
E418	Glutamate 418
G363	Glycine 363
G503	Glycine 503
H632	Histidine 632
IQN	Indenoisoquinoline
K374	Lysine 374
K425	Lysine 425
K532	Lysine 532
MD	Molecular Dynamics
MM/GBSA	Molecular Mechanics/Generalized Born Surface Area
N722	Asparagine 722
PBC	Periodic boundary conditions
PDB	Protein Data Bank
QM	Quantum Mechanics
QM/MM	Quantum Mechanics/Molecular Mechanics
R364	Arginine 364

RMSD	Root mean square deviation
Ser	Serine
T718	Threonine 718
TopoI	Topoisomerase IB enzyme
TopoIcc	Topoisomerase IB/DNA cleavable complex
TOP1mt	Mitochondrial Topoisomerase 1
TOP2α	Topoisomerase 2α
TOP2β	Topoisomerase 2β
TOP3α	Topoisomerase 3α
TOP3β	Topoisomerase 3β
TPT	Topotecan
TPT-closed	Topotecan lactone form
TPT-open	Topotecan carboxyl form
vdw	van der Waals
Val	Valine
Y723	Tyrosine 723

1. INTRODUCTION

It was 1953 when the currently known structure of DNA molecule was first obtained by Watson, Crick and Franklin [1, 2]. The double helical structure provided an explanation of semi-conservative replication in which each antiparallel strand of DNA provides a guide for the newly synthesized daughter strands [2]. However, the structure of DNA molecule brings many challenges during the key cellular processes including transcription, replication and as well as repair mechanisms. For these events to occur, the cell needs to be able to reach the nucleotide sequence encoded in DNA duplex and opening of the double helix results in underwinded and overwinded regions, which can interfere with the gene expressions and cause strand breaks on DNA molecule if not treated by some machinery. The main cellular components involved in this "treatment" process are Topoisomerase enzymes that have the capacity to cut and religate the DNA.

1.1. Topological Features of DNA Molecule

Topoisomerase enzymes can be found in both eukaryotic and prokaryotic life forms. The first enzyme that is found to interfere with the topology of double-stranded DNA molecule was in *Escherichia coli* in 1970s by James C. Wang [3]. A couple of years later, Champoux and Dulbecco identified a similar enzyme in mouse embryonic cells [3]. Since then, different types of Topoisomerase enzymes have been discovered and many of them have been the focus of the research groups due to their pivotal roles in main cellular events. Although the main purposes of these enzymes are similar, their mechanism of action as well as their roles in cell differ. Here, for the sake of clarity, only the eukaryotic Topoisomerases will be briefly introduced and then the focus will be the human Topoisomerase IB (TopoI).

1.2. Topoisomerase Enzymes in Eukaryotes

Eukaryotic Topoisomerase enzymes can be divided into three sub-families: Type IA, Type IB and Type IIA [3]. The main distinction between the Type I and Type II families is the number of DNA strands cleaved: while Type I family only cuts a single strand, Type II family performs by cutting both strands. In addition, Type II family also require ATP to perform its function [3].

1.2.1. Type IA: The Strand-Passage Mechanism with Mg^{2+}

This sub-family uses a "strand-passage" mechanism to change the topology of DNA. First, one strand of DNA is cut, and the enzyme-DNA link is established through a 5'-phospho-tyrosyl bond. Then the intact strand passes through the nicked one [4]. A metal cofactor, Mg^{2+}, is also needed. This sub-family mainly prefers to relax hypernegative supercoils [4]. The mammalian form of this sub-group consists of TOP3α and TOP3β isoforms; TOP3β can function as RNA helicase on circular RNA molecules [4–6] and TOP3α can be found in mitochondria as well as in nucleus [7]. All eukaryotic Type IA Topoisomerases have critical roles during neural development which will be discussed later in this chapter [8].

1.2.2. Type IIA: The Strand-Passage Mechanism with ATP

Eukaryotic Topoisomerases in this sub-family act as homodimers, and along with ATP requirement for them to function, their mechanism involves cutting both strands of DNA at the same time [3, 4, 9, 10]. Via a mechanism called "strand passage", these enzymes cut both strand of one DNA molecule, and then an intact DNA is passed through the transient break [9, 10]. Positive and negative supercoiled DNA, and decatenation of entangled double-stranded DNA (dsDNA) molecules can be relaxed by this sub-family [9, 11].

The two isoforms of eukaryotic Type IIA family, TOP2α and TOP2β, have significant mechanistic and structural similarities, although their primary sequences and

arrangement of domains are different [12–16]. The former has an expression that is cycle-dependent: its expression increases during S phase and reaches a peak at G_2/M phase, assisting the cell proliferation and chromosome condensation and segregation/replication [3, 17–19]. On the other hand, the expression of TOP2β is constitutive in both dividing and dormant cells, and it has a crucial role during neuronal cell development [3, 20–23].

1.2.3. Type IB: Controlled Rotation Mechanism

This kind of Topoisomerases are particular due to their characteristic 3'-phosphotyrosyl bond with DNA, and it does not share any homology -structural, mechanistic or sequence- with other types of eukaryotic Topoisomerases [3]. In addition, it is believed that mitochondrial TOP1 enzyme (TOP1mt) shares a common ancestor with nucleic TOP1 enzyme due to their high sequence as well as mechanical similarities [24].

The mechanism of Type IB starts with the formation of single-stranded break on DNA, continues with the rotation of intact strand around the nicked one, and finishes with the religation of the nicked DNA backbone [25]. The interactions between the DNA and enzyme regulate the "controlled rotation" mechanism [26], which also assists keeping the nicked ends close for religation [25, 27]. Unlike other members of Type I family, eukaryotic Type IB Topoisomerases can relax both negative and positive supercoils, and they do not require Mg^{+2} as cofactor [28]. Human Topoisomerase Type IB enzyme will be referred as TopoI from now on.

1.3. Cellular Roles of Eukaryotic Topo Enzymes

There are many types of Topoisomerase enzymes with very specialized functions required for complex DNA activities such as replication, transcription and chromosome segregation. To have a general understanding of the activities mediated by Topoisomerases, their roles in various cellular processes will be mentioned briefly in this section.

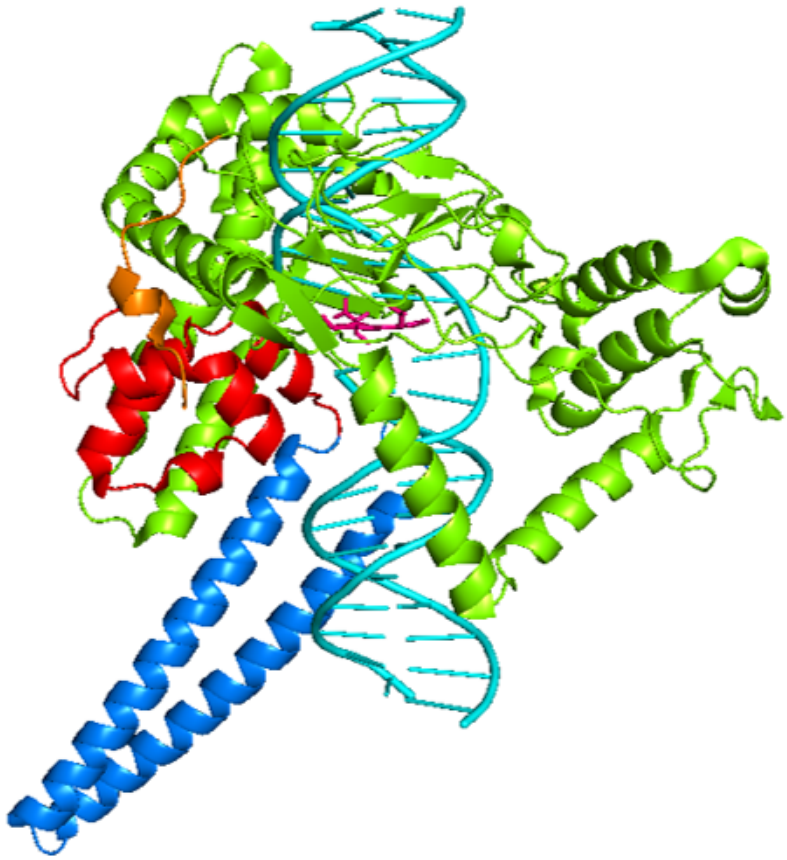

Figure 1.1. The crystal structure of TopoIB in complex with DNA and TPT (PDB ID: 1K4T). N-terminal, core, linker and C-terminal domains are shown with orange, yellow, blue and red color, respectively. DNA is colored as cyan and TPT is shown in magenta.

1.3.1. Replication and Chromosome Segregation

Replication mechanism of the genetic material mainly involves three steps: initiation, replication fork progression and termination [29]. The initiation step is where the duplex DNA molecule is separated at specific genomic loci, called "origins of replication", and the necessary protein complexes for the replication is employed [30]. A protein called Cdt1, a part of the pre-replication complex, is shown to have an interaction with both TopoI and TOP2α [31, 32]. In another study by Gonzalez *et al.*, the simultaneous knockdown of both TOP2α and TOP2β did not slow down the DNA

replication but it significantly affected the chromosome condensation as well as chromosome segregation, indicating that Type IIA Topoisomerases might have a role in chromosome stability rather than the replication complex stabilization [33]. Although the reason for employment of Topoisomerases during the early initiation is still not known completely, it is thought that it might be related with the preparation of DNA for assembly of the necessary complexes [29].

Due to the nature of the replication process, positive supercoils on DNA accumulate as replicative helicases unwind DNA and the synthesis of new strands progress. These supercoils mainly treated by Topo IB or Type II Topoisomerases; while the former is proposed to relax supercoils ahead of the replication fork, and the latter is thought to function behind of the fork [34, 35]. Positive supercoil formation is not the only topological change that occurs during the replication. The fork rotation causes the sister chromatids to interwind and form precatenanes. Although precatenane formation is a natural event, if not treated, it might form catenated or tangled DNA molecules, which might lead to unnatural chromosome segregation [4, 36]. In addition, another topological linkage, called "hemicatenane", can be also formed in-between two replication forks where the DNA is unreplicated, and if hemicatenanes are not treated by Type IA or Type IIA Topoisomerases, an incomplete resolution of DNA might happen [37].

After completion of DNA replication, the daughter chromosomes should be separated to opposite ends. The segregation of chromosomes can be achieved with the help of chromosome-partitioning complexes and cytoskeletal elements [29]. Type IA and Type IIA Topoisomerases are also found to be necessary during the separation [4]. Especially TOP2 was discovered to be present at decatenation sites and when its activity is blocked in yeast cells, it was observed that chromosome separation failed and ultimately led to cell death [38, 39]. In addition, some studies showed that decatenation of chromosomes could cause oncogenesis and that chromatin remodeling complexes strongly control the topoisomerases during chromosome segregation [40].

In addition to their roles in chromosome segregation, topoisomerases are associ-

ated with packed DNA structures as well. In human cancer cells, it has been found that the complexes related to the nucleosome deposition and histone methylation have modified activities [40, 41] . One of these complexes, known as SWI/SNF, has a catalytic subunit that recruits TopoI as well as TOP2α to keep the genome stable by preventing the formation of mutagenic DNA structures, namely G-quadruplexes and Z-form of DNA [40, 41].

1.3.2. Transcription and Gene Regulation

Transcription is the first step of converting the genomic data into three - dimensional protein structures. At the first glance, transcription process has three stages: initiation step where RNA polymerase binds to DNA at the promoter region, elongation step in which the free nucleotides are added to the new RNA chain, and finally the termination step where the termination sequence is read by RNA polymerase and finishes the transcription event.

Initiation is the far most important step since it involves the recognition of promoter site as well as the separation of DNA duplex. In previous studies, TopoI was thought to be required for the recruitment of the protein that binds to a specific promoter site (TATA-box binding protein, TBP) [42]. However, in recent studies it was shown that the transcription activation does not need a catalytically active TopoI protein [42, 43]. Experiments performed with cancer cells in which TopoI activity was inhibited using Camptothecin (CPT) or Topotecan (TPT) caused a halt in transcription elongation of long genes in particular [44, 45]. One of the most striking result is that the genes related to autism and synaptic functions have altered expression when treated with TPT to inhibit TopoI function or when TOP2β is depleted [46, 47], and this points out that the expression of long genes demands presence of both TopoI and TOP2β proteins [46, 48].

In addition, R-loops could be formed at the transcription bubbles and they might generate genomic disturbances if not removed. As defined before, R-loops are RNA-DNA hybrids and could be dangerous for genome integrity and potentially lead to

double stranded breaks on DNA [4,49]. It has been reported that reducing the activity of TopoI and TOP3β proteins in human cell induces R-loop accumulation, and TopoI cleavable complex (TopoIcc) formation due to CPT poisoning is also related with R-loop formations [50–54]. Additionally, TopoIcc could lead the formation of short R-loops [55, 56].

All in all, each enzyme in Topoisomerase family has its own distinct role in core cellular events. TopoI mainly assists during the initiation of DNA replication as well as elongation step. Although its function in transcription still is not clear, TopoI has a role in inhibiting R-loop formation while the transcription advances. Therefore, targeting TopoI protein for the treatment of cancer cells has been the focus of many research groups. In the next section, the molecules used for TopoI inhibition will be addressed.

1.4. TopoI Inhibitors

TopoI enzyme is quite vulnerable after the formation of the phosphotyrosine bond with DNA, and this covalently bonded DNA/TopoI complex is called cleavage complex (TopoIcc) [57]. If the religation of the TopoIcc is hindered by TopoI inhibitors, it transforms into a DNA-damaging agent due to possible collisions with replication and transcription factors. In fast dividing cells, such as tumor cells, trapping TopoIcc could even lead to cell death. The molecules that have been used to influence the enzyme activities could function by:

- Catalytically inhibiting the enzyme activity,
- Changing the activity toward other substrates rather than the original one,
- Transforming the enzyme into a cell-damaging agent [58].

The first known TopoI inhibitor, Camptothecin (CPT), and its derivatives use the third mechanism (Figure 1.2). Long before discovering that its target is the TopoI, anti-cancer activity of CPT has been known [59, 60]. Discovering that CPT binds to TopoI within the cell was a breakthrough since it provided scientists a scaffold for developing new molecules that could also target TopoI. However, this family of

inhibitors suffer from serious drawbacks despite their potential as anti-cancer drugs. To reduce the effects of these weaknesses, various semi-synthetic CPT analogues have been synthesized and only a few are available for clinical use as of now.

1.4.1. Camptothecin: The First TopoI Inhibitor

CPT is a natural alkaloid first obtained from the bark of Chinese tree, *Camptotheca acuminate*, by Wall and his co-workers [61]. As it can be seen from Figure 1.2, it has five fused rings, and the E ring is a lactone ring which can be hydrolyzed to carboxyl at physiological and high pHs [62].

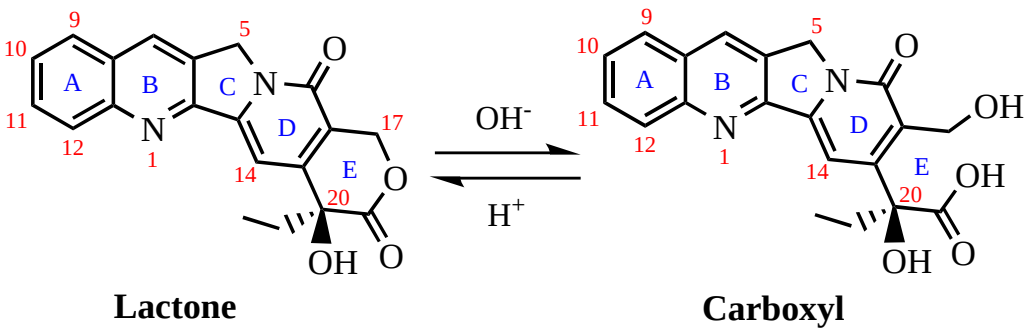

Figure 1.2. The equilibrium between lactone and carboxyl forms of Camptothecin (CPT). Atoms are numbered to show where possible substitutions could happen, and the rings are named from A through E.

The cytotoxic nature of CPT is due to its binding to DNA/TopoI complex and consequent inhibition the religation mechanism, which could introduce double-stranded breaks upon collision with other DNA-binding agents [63]. TOP1B gene deletion experiments in yeasts proved that CPT can only bind to DNA/TopoI complex, not TopoI alone [64, 65].

However, the hydrolysis of the lactone ring of CPT is a big problem. The carboxyl CPT is considered inactive and it has high affinity towards serum albumin [62]. Furthermore, due to its low solubility in water and adverse side effects, CPT could

not complete the clinical trials [66, 67]. Experiments with derivations of CPT showed that any substitution at position 12 resulted in decrease in CPT activity, whereas amino substitution at position 9 enhanced the drug acitivity (Figure 1.2) [68]. It was also shown that only natural S-isomer is effective [68]. Among many CPT-derivatives, currently only three (Topotecan, Irinotecan and Belotecan) are approved for clinical use [69, 70].

1.4.2. Topotecan: CPT with Better Solubility

Topotecan (9-[(dimethylamino)methyl]-10-hydroxy-camptothecin) is the first CPT-derivative that has been approved by FDA and is currently being used in the treatment of ovarian and small-cell lung cancers [71–73]. Compared to its parent molecule, TPT has higher water solubility and less adverse side effects. TPT also has a hydrolyzable lactone ring, however, compared to CPT, binding of TPT to DNA/TopoI complex increases the half-life of the complex more [61, 74, 75].

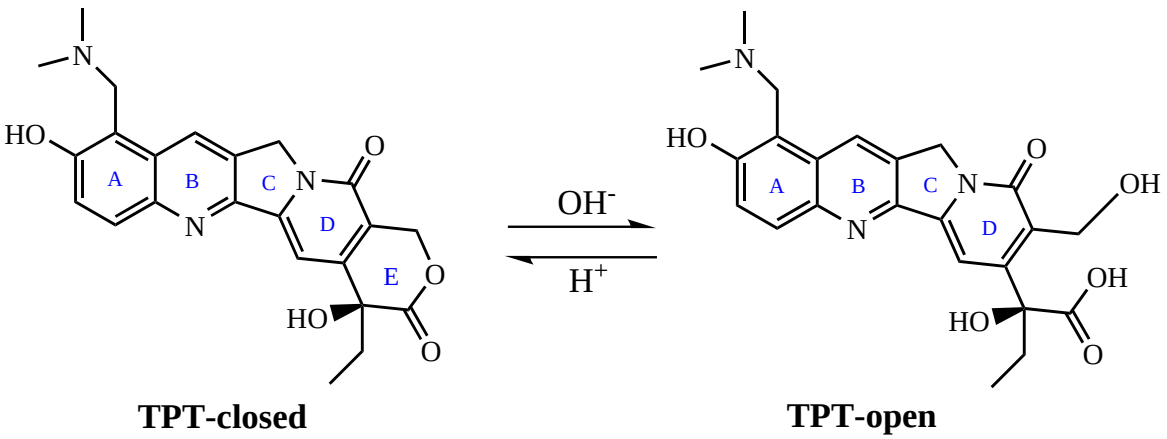

Figure 1.3. The equilibrium between lactone and carboxyl structures of Topotecan (TPT).

The E-ring of TPT is in 50:50 equilibrium between the lactone (TPT-closed) and carboxyl (TPT-open) forms at pH 6.28 (Figure 1.2) [76, 77]. While at high pH values (pH>10) TPT-open form predominates, acidic pH (pH<4) causes the equilibrium to

shift towards TPT-closed [78]. The binding affinity of carboxyl form of CPT for blood proteins was a prominent problem since it shifts the equilibrium toward carboxyl form and reduces the concentration of lactone form in the blood. This problem is not observed in TPT; the presence of HSA in the medium did not affected the lactone-carboxyl ratio of TPT [60, 76, 79].

Human Topoisomerase IB is a 765 amino acid long monomeric protein which acts on supercoils formed on dsDNA during replication and transcription processes. It has four sub-domains: core domain (420 aa), linker domain (76 aa), carboxylic terminal (52 aa) and amino terminal (214 aa) [80]. The domain of TopoI can be observed in Figure 1.1.The first crystal structure of TopoI was obtained in 1998 by Redinbo *et al.*, a study in which the mechanism of TopoI on DNA duplex was also clarified [80]. With the electrostatic interactions, as well as hydrogen bonds, the enzyme forms a clamp-like structure around DNA molecule, and a covalent link between TopoI and DNA formed by catalytic Tyrosine 723 (Y723) residue attacking the DNA backbone. After the nick formation on the phosphate backbone of DNA, the intact strand is rotated around the nicked one via the mechanism called "controlled-rotation" [26]. Following this rotation, the cut on the phosphate backbone of DNA is restored by an attack of 3'-OH end, and the supercoil is relaxed.

The covalently bonded DNA/TopoI (TopoIcc), however, is quite vulnarable to inhibitors such as CPT and Indenoisoquinoline (IQN) derivatives. These small molecules can intercalate between the bases where the nick occurs and prevent the last religation step from happening. The stabilization of the complex at this stage could result in permanent ds breaks on DNA upon collision with other DNA binding agents, and the cell dies; which is why, these TopoI inhibitors are considered as anti-cancer agents. Among them, Topotecan is currently available on the market; however, the exact mechanism of inhibition is still not completely understood. It was proven that TPT does not prefer to bind TopoI protein alone, but it does bind to ss or ds DNA molecules as proven with HPLC and NMR studies. The interactions and π-π stackings with DNA bases stabilize the lactone form, and hence shift the equilibrium towards TPT-closed [76,81]. From this, it can be inferred that TPT molecule indeed targets the binary DNA/TopoI

complex, and DNA-TPT interactions are as important as TopoI-TPT interactions.

1.5. Interaction of TPT with DNA/TopoI Complex

Similar to its parent molecule, TPT has a lactone ring which hydrolyses to carboxyl at the physiological and high pHs (Figure 1.3), and this lessens its activity since the carboxyl form (TPT-open) is proved to be inactive [81, 82]. Despite this, in the crystal structure (PDB ID: 1K4T) of DNA/TopoI with TPT obtained in 2002, the occupancy of TPT-open and TPT-closed was found to be 37% and 63%, respectively, which brings the question of activity difference between the two forms into mind [63].

The crystal structure of the ternary complex gives the interactions that are critical for the inhibiton mechanism of drug TPT. First, upon intercalation of TPT, regardless of the status of E-ring, the distance between two nicked ends increases from 3.5 Å to 11.5 Å, which renders the religation physically impossible [63]. Secondly, both TPT-open and TPT-closed form a hydrogen bond with Aspartate 533 (D533) residue.

1.6. Mutation Studies

Mutating the residues around the intercalation site can help to identify structure-wise and mechanism-wise important residues that might play a role in inhibition via small drug molecules and the mechanism of activity. The lip region, which is located on core domain, is the region that wraps around the DNA molecule, hence mutations in this location can interfere with the TopoI activity [83]. Upon mutation of Glycine 363 (G363) residue to Val, Ser and Cys, both the sensitivity for CPT and enzyme activity did not affected much [84]. Mutating Threonine 718 (T718) to Alanine also did not change the enzyme activity but this mutation mimicked the effect of CPT [84]. In another study, changing Arginine 364 (R364) residue, which is also located in the core domain, into Histidine enabled the protein to be resistant against CPT [85]. Glutamate 418 (E418) to Alanine mutation also caused CPT resistance in human nasopharyngeal carcinoma cell lines [86]. Studies with chinese hamster DNA Topoisomerase revealed that Glycine 503 (G503S) mutation at a highly conserved region grants the cell re-

sistance against CPT [87]. In addition, the change of the amino acid at position 533 (D533) to either Glycine or Aspartate also resulted in CPT resistance in human leukemia cells [88, 89]. The mutation at Asparagine 722 (N722), which is located very close to the catalytic site, also induced drug resistance against CPT [90]. However, it was discovered that the durg resistance was not due to the structural perturbations on Tyrosine 723 (Y723), but due to the loss of interaction with CPT [90]. Finally, Lysine 532 (K532) residue which interacts with the Thymine base at position -1 (-1T), is mutated to Alanine and Arginine, and both mutation resulted in mimicking of the effect of CPT, although K532R case was 400-fold faster in religation [91].

1.7. Computational Studies on the Interaction of CPT-like Drug Molecules

The residues located near the catalytic residue Y723 indeed have a role in either the controlled rotation mechanism or CPT resistance. MM/GBSA (Molecular Mechanics/Generalized Born Surface Area) interaction energies of three mutant TopoI proteins (E418K, G503S, D533G) revealed that D533G mutation has the lowest affinity for the TPT, followed by G503S and E418K [92]. The decomposition of MM/GBSA energies points out the residues that have important interactions with TPT. In the study, they showed that residues R364, D533 and K425 have the largest contributions to the interaction with TPT [92].

Among them, K532 and N722 were discovered to have direct interactions with the CPT-like drugs that intercalate the nicked site. In the crystal structure that has TPT co-crystallized with DNA/TopoI complex, it was observed that TPT-closed only makes direct interaction with these residues [63]. The Molecular Dynamics simulations of binary and ternary complexes supported the mutation studies. In the study performed by Mancini *et al.*, TPT interacts with R364 and K532, and intercalation of TPT to the binding site caused important topological changes [93]. R364 and TPT formed two hydrogen bonds through B and E rings of TPT. The physical distance between the ends of the nicked strand is increased up to 8.6 Å from 4.5 Å, and more importantly,

the flexibility of the linker domain(residues between 655-680) dropped drastically with the RMSF (root-mean-square-fluctuation) values of 6.3 Å for binary and 7.3 Å for the ternary complex. However, in the region between the linker and core domains (residues between 633-643) the observed behaviour was the opposite, which caused a decrease in the flexbility of the linker region [93].

1.8. Purpose of the Study

The main focus of this thesis is to clarify the pH-dependent activity of Topotecan by studying its two different forms, i.e. lactone and carboxyl, in the TopoI/DNA complex. As stated previously in this section, the role of TopoI enzyme in a cell is quite crucial, hence understanding the mechanism of inhibition enables one to construct better inhibitor molecules. For this reason, the investigation of activity difference between lactone and carboxyl forms of TPT is essential. Here, the study has been performed using classical Molecular Dynamics, hybrid Quantum Mechanical/Molecular Mechanism approch and topological analysis. The results have been published in Biochemistry early in this year [94].

2. METHODOLOGY

In this chapter, the theories behind the methodologies used in this study are explained. These methodologies are Molecular Dynamics, hybrid Quantum Mechanics/Molecular Mechanics method and docking.

2.1. Molecular Mechanics and the Concept of Force Field

Molecular Mechanics is an approach that makes use of classical mechanics, in which the atoms are represented as perfect spheres and the bonds are defined as strings. This approach is preferred for systems that cannot be modelled using quantum mechanics due to the computational cost of modelling huge number of atoms. That's why, classical approach that disposes this problem is preferred to study large systems including proteins.

A force field is a collection of parameters that is used to define the total energy of the system. Two main components of a force field are bonded and non-bonded terms. Bonded terms consist of bond stretching, angle bending, torsion (bond rotation), while non-bonded interactions include electrostatic and van der Waals terms. These parameters are derived from empirical data to predict the particular parameters of the molecules.

Two of the bonded terms, bond stretching and angle bending can be expressed using simple harmonic motion:

$$V_{bonded} = \sum_{bond} \frac{1}{2} k_{bond}(l_0 - l)^2 \tag{2.1}$$

$$V_{angle} = \sum_{angle} \frac{1}{2} k_{angle}(\theta_0 - \theta)^2 \tag{2.2}$$

where k_{bond} and k_{angle} are force constants for bonds and angles, respectively. l_0 and θ_0 are the reference values for the bond length and angle, respectively. The force needed to change the bond length between two bonded atoms would be massive, hence any significant changes in both bond lengths and bond angles are prevented.

Torsional term provides the largest contribution to the total energy, and is described as follows:

$$V_{torsion} = \sum_{n=0}^{N} \frac{1}{2} V_n [1 + cos(n\omega - \gamma)] \tag{2.3}$$

In Equation 2.3, the term V_n describes the depth of the potential energy surface of rotations about ω, over the periodicity of n, with the minimum angle of γ.

Improper torsions are used to define the out-of-plane bending motions, and the following equation shows its functional form:

$$V_{improper} = \frac{1}{2} k_\omega [1 - cos2\omega] \tag{2.4}$$

where the improper torsion is defined as the angle between four atoms that are not bonded in sequential order.

To define the electrostatic properties of atoms, they are considered as point charges in molecular mechanics calculations, hence the electrostatic contributions can be defined using Coulomb's law:

$$V_{el} = \sum_i \sum_j \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{2.5}$$

where q_i and q_j define the point charges on the atoms, and r_{ij} is the distance between them. There are a couple of methodologies that can be utilized to calculate the point charges including quantum mechanical methods and experimental approaches. AMBER force field, which is used in this study, uses the point charges derived from

electrostatic potentials [95].

The final term, which is the van der Waals interaction energy, is defined as the sum of all interactions of the molecules while considering their positions as well as the relative orientations. Lennard-Jones function is the most widely used representation to define vdW interactions:

$$V_{vdW} = 4\epsilon_{AB}\left[\left(\frac{\sigma_{AB}}{r}\right)^{12} - \left(2\frac{\sigma_{AB}}{r}\right)^{6}\right] \tag{2.6}$$

where ϵ_{AB} is the amplitude of the potential, and σ_{AB} is called as collision diameter, which is the arithmetic mean of individual diameters of pure species (σ_{A} and σ_{B}).

2.2. Molecular Dynamics

Molecular Dynamics (MD) uses the forcefield terms to calculate forces for each particle in the system, and with the help of statictical mechanical approach it models the dynamics of particles, hence gives a prediction for the position of the particles within a given system. This prediction is obtained by numerically solving Newton's equations of motion.

Classical Hamiltonian ($H(\mathrm{p_i}(t),\mathrm{r_i}(t))$) can be used to desribe the time evaluated motion of a system with N particles:

$$H(\mathbf{p}_i(t), \mathbf{r}_i(t)) = \sum_{i=1}^{N} \frac{1}{2m_i}\mathbf{p}_i^2 + V(\mathbf{r}_i) \tag{2.7}$$

in this expression $\mathrm{p_i}(t)$ is the momentum vector and $\mathrm{V(r_i)}$ is the potential energy. The Hamiltonian is the sum of potential and kinetic energies of all particles in the system, and its partial derivation yields the equations of motion, i.e. velocity of the particle and the force acting on the particle.

$$\frac{\partial \mathbf{r}_i}{\partial t} = \frac{\partial H}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i} = \mathbf{v}_i \tag{2.8}$$

$$\frac{\partial \mathbf{p}_i}{\partial t} = -\frac{\partial H}{\partial \mathbf{r}_i} = -\frac{\partial V}{\partial \mathbf{r}_i} = F \tag{2.9}$$

which leads us to the Newton's 2nd law:

$$\frac{\partial^2 \mathbf{r}_i}{\partial t^2} = \frac{F}{m_i} \tag{2.10}$$

for a particle with mass m_i to move along coordinate $\mathrm{r_i}$ under the influence of an external force F. Equation 2.10 is used to obtain the trajectories of the particles. When the position of a particle changes, the force acting on it also changes. F(t), the total force at time t, is obtained by the vector sum of all interactions between the individual paticles, and if the time step, ∂t, is small, it is assumed to be constant. There are many methods implemented in MD softwares to integrate the equations of motion, and *Velocity Verlet* is one of them [96]. In the *Velocity Verlet* algorithm, at time (t + ∂t), the velocity of a particle (v) and its position (r) are defined as follows:

$$\mathbf{r}(t+\partial t) = \mathbf{r}(t) + \partial t \mathbf{v} t + \frac{1}{2}\partial t^2 m^{-1} F(t) \tag{2.11}$$

$$\mathbf{v}(t+\partial t) = \mathbf{v}(t) + \frac{1}{2m}\partial t (F(t) + F(t+\partial t)) \tag{2.12}$$

Initial coordinates, r(0), for the system of interest can be obtained from various sources. For instance, the coordinates of biological systems can be obtained via X-Ray crystallography studies or NMR structures. Maxwell-Boltzman distribution is used to derive the initial velocities of the particles, v(0), and it is adjusted so that the total momentum would be zero for the whole system. Initial forces at t=0 are obtained using the Equation 2.9.

For a system that has all requirements (initial coordinates, r(0); initial velocities, v(0); initial forces, F(0)) that will be used in MD simulation, the simulation cycle follows these steps:

(i) Using Equation 2.11, the displacement of coordinates are calculated with respect to the initial positions within a time interval ∂t.

(ii) With the help of Equation 2.9, the forces on the particles are calculated using the positions form *step (i)*.

(iii) Using Equation 2.12, new velocities are calculated with the initial force and the new force that is obtained in *step (ii)*.

(iv) Steps above are repeated until reaching a proper amount of simulation time.

The question about simulation length does not have a certain answer yet and it is an ongoing debate [97]. According to *ergodic hypotesis*, within an infinite time of simulation, all possible states of a system can be obtained; however, an infiniely long simulation is practically impossible, and with the current computational limitations, there is a tradeoff between long simulation time and the cost, which needs to be considered. To be able to reach a sufficient simulation time without increasing the computational cost too much, the time step (∂t) should be taken as long as possible. But one should also bear in mind that ∂t is also limited with the integration algorithm that is used by the software.

In MD simulations, to mimic the bulk effect *Periodic Boundary Conditions* (PBC) has employed. When PBC has used, the simulation box is replicated inifinitely in each direction, and only the coordinates of the original simulation box is recorded. If a particle in the box moves to outside of the box, its periodic image from the other box moves in the same direction and hence, the total number of particles in the box is kept constant. In PBC, to treat the electrostatic interaction between particles, generally Ewald sum method is employed [98].

Statistical ensembles are used to obtain the thermodynamical properties of the system of interest. These ensembles are constructed based on the temperature (T), volume (V), number of paticles (N) and pressure (P). The most common ones are:

- NVE: microcanonical ensemble - has constant N, V and E;
- NVT: canonical ensemle - has constant N, V and T;

- NPT: isothermal isobaric ensemble - has constant N, P and T

For the biological systems, NVT ensemble is the most popular choice due to its computational efficiency. When using the canonical enseble in simulations, the exchange of energy is contained with different thermostat models by adjusting the temperature to the desired value. The classical definition of average kinetic energy in NVT ensemble is follows:

$$< E_k >_{NVT} = \frac{1}{2} \sum_{i=1}^{N} m_i v_i^2 \tag{2.13}$$

and the average kinetic energy also can be rewritten such that it is related to the temperature using the classical equipartition theory:

$$< E_k >_{NVT} = \frac{3N-6}{2} k_\beta T \tag{2.14}$$

in which k_β is the Boltzman constant, and T corresponds to the temperature. The velocities of each step are scaled as $v_{new}=\lambda v_i$, then the tempetature change ($\Delta T = Ti\text{-}T(t)$) can be calculated using the correlation between v_i and T from Equation 2.13 and 2.14 as follows:

$$\Delta T = \frac{1}{2} \sum_{i=1}^{N} \frac{2}{3} \frac{m_i (\lambda v_i)^2}{N k_\beta} - \frac{1}{2} \sum_{i=1}^{N} \frac{2}{3} \frac{m_i v_i^2}{N k_\beta} \tag{2.15}$$

which gives

$$\Delta T = (\lambda^2 - 1) T(t) \tag{2.16}$$

To obtain the value of λ with respect to the target temperature, T_{max}, and the instantenous temperature, $T(t)$:

$$\lambda = \sqrt{T_{new}/T(t)} \tag{2.17}$$

This is the procedure of the simplest thermostat [99], in which one could obtain the T_{max} by multiplying the velocity with λ and using the temperature obtained from the kinetic energy, *T(t)*. However, the drawback of this thermostat is that along the simulation, the temperature difference between the solute and solvent may occur.

To solve the problems with the previous thermostat model, Andersen thermostat, which is based on the stochastic collision model, was developed [100]. In this thermostat, the system is immersed in a heat bath, and the velocity of the particles in a random time interval is assigned using Maxwell-Boltzman distribution at temperature *T(t)*. At each step, the simulation is performed with constant energy so no thermal difference within the system occures. In addition, the calculated velocities follow the Gaussian distribution.

Langevin thermostat is another model in which all particles obtain a random force at each time step, and their velocities are lowered due to the use of a constant friction [101]. In order to obey the "fluctuation-dissipation" theorem, average strength of the random forces and the friction are related. In the Langevin method, it is assumed that big particles (solute) are swimming in a pool of smaller particles (solvent), and the smaller particles usually randomly collide with the solute molecules and influence their dynamics. Moreover, solvent molecules also have dampening effect on the solute molecules described as a fictional drag force. Langevin thermostat incorporates these two factors.

In order to mimic the natural conditions of biological systems, protein structures need to be solvated in suitable environments. For Molecular Dynamics, this solvent is usually explicit water molecules that are defined by certain models. The most common solvation model is the rigid TIP3P water model, in which Coloumb's law and Lennard-Jones potentials are used to describe the electrostatic interactions.

When all of these considered, one can summarize the steps for a standard MD simulation as follows:

(i) Preparing the initial coordinates of the systems - either immersed in explicit water or not,

(ii) Preparing the topology of the system, in which the number and the mass of atoms, interaction potentials and the connections between the particles are described,

(iii) Minimizing the energy of the system using PBC and MM,

(iv) Heating the system to a desired temperature (generally it is 300K that corresponds to the room temperature),

(v) Equilibrating the system,

(vi) Generating the trajectories by using appropriate ensembles, thermostats and/or barostats, using the Newtonian equations of motion integrated with an appropriate time step,

(vii) Analyzing the trajectories using statistical mechanical methods.

Some details regarding the equilibration time, time step, cutoff values, the choice of thermostat and barostat might change with respect to the system of interest. For this reason, computational details of the systems used in the studies here are given in their respective chapters.

2.3. Quantum Mechanics

Molecular Dynamics simulations provide a powerful tool to see the time-dependent change in the structure of systems, however, one cannot study the electronic properties of molecules using MD due to the limitations of MD. Quantum Mechanics is developed to describe the microscopic systems and it depends on the approximate solution to Schrödinger equation. The Schrödinger equation

$$\hat{H}\Psi = E\Psi \tag{2.18}$$

gives the energy of the system (E) as an eigenvalue of the Hamiltonian operator($\hat{H}$). Typically, $\hat{H}$ consists of potential and kinetic energy terms of electrons and nuclei:

$$\hat{H} = -\hbar^2 \sum_i [\frac{\nabla_i^2}{2m_e} + V_{e-e} + V_{e-N} + V_{N-N}] \tag{2.19}$$

in which m_e corresponds to the mass of the electorn, m_k is the mass of nuclei, ∇ is the Laplacian operator and $\hbar$ is the Planck's constant. If one defines the system in the cartesian coordinates, Laplacian becomes

$$\nabla_i^2 = \frac{d^2}{dx_i^2} + \frac{d^2}{dy_i^2} + \frac{d^2}{dz_i^2} \tag{2.20}$$

The terms $V_{e\text{ - }e}$, $V_{e\text{ - }N}$ and $V_{N\text{ - }N}$ define the electron-electron repulsion, electron-nuclei attaction and nuclei-nuclei repulsion behaviors, respectively.

Although it is relatively easy to find an exact solution in small systems, for bigger ones, to obtain an exact solution for Schrödiger's equation requires some approximations. Under physical conditions, the electrons move faster that the nuclei since they are lighter. Keeping this fact in mind, Born-Oppenheimer approximation proposes that electronic energies can be calculated by fixing the nuclear positions, which removes the $V_{e\text{ - }N}$ term and the kinetic energy term that is not depended on electrons. Also, the nuclear-nuclear potential energy term becomes constant. Therefore, the Schrödinger equation becomes

$$(H_{el} + V_{N-N})\psi_{el} = E_{el}\psi_{el} \tag{2.21}$$

Molecular orbitals (MO) and the linear combination of atomic orbitals (LCAO), which can be used to form MOs, are obtained from the eigenfunction of Equation 2.21, and atomic orbitals(AO) are usually Gaussian function and defined by basis sets.

2.3.1. ab initio Methods

If one wants to solve the Schrödinger equation, molecular orbitals are not provided as an input. The electronic structure calculations are performed with the guide of the variational theorem, in which the eigenvalue, the energy, of an approximate wavefunction will be higher than of the true wavefunction. Hence, first the minimum value of energy should be searched. For this purpose, an iterative process called "self-consistent field (SCF)" theory is being used. It uses Hartree-Fock (HF) equations and Coulomb and exchange integrals are considered.

$$f_i x_i = \epsilon_i x_i \tag{2.22}$$

in which x_i is the MO of the electron i, f_i is the Fock operator and ϵ_i corresponds to the energy of the system. SCF cycle can be divided into the following steps:

(i) preparing an input for the moelcular geometry,
(ii) obtaining an initial guess for MOs,
(iii) constructing HF secular determinant and solving ϵ_i,
(iv) obtaining a second set of solutions to x_i by using ϵ_i,
(v) repeating *step (iii)* until reaching to the minimum energy point at which the results are self-consistent.

This method, known as SCF theory, which is also called as HF method, is an *ab initio* method and it does not utilize the empirical parameters.

2.3.2. Density Functional Theory

Density Functional Theory (DFT) is derived from Hohenberg-Kohn Existence Theorem in which for a system with known density, Hamiltonian of that system can be obtained and Schrödinger equation can be solved to have the energy [102]. Hohemberg and Kohn proposed that only external potential ($V(r)$) affecting the electrons is the

potential due to the presence of nuclei, which is imposed by the electron density ($\rho(r)$):

$$\rho(r) = N \int ... \int \mid \psi(r_1, r_1, ...r_n) \mid^2 dr_1 dr_2 ... dr_n \tag{2.23}$$

in which coordinates of electrons are shown by r_i. According to Hohenberg-Kohn Variational Theorem, energies higher than the ground state values are obtained with any electron density except the real one. The ground state electronic energy is a function of electron density:

$$E[\rho(r)] = \int V(r)\rho(r)dr + F[\rho(r)] \tag{2.24}$$

and here, the electron-external potential interaction is described by the first term, and kinetic energy of electrons as well as the interatomic interactions are shown by $F[\rho(r)]$. The latter can be divided into individual terms according to Kohn-Sham theory [102]:

$$F[\rho(r)] = E_{KE}[\rho(r)] + E_H[\rho(r)] + E_{XC}[\rho(r)] \tag{2.25}$$

where $E_{KE}[\rho(r)]$, $E_H[\rho(r)]$, $E_{XC}[\rho(r)]$ terms correspond to the kinetic energy of non-interacting electons, Coulomb energy, and exchange-correlation enegry, respectively. The electron-electron Coulombic energy does not consider the correlation between motions of electrons, and it can be described as follows:

$$E_H[\rho(r)] = -\sum_{A=1}^{M} \int \frac{Z_A}{\mid r - R_A \mid} \rho(r)dr + \int \frac{\rho(r_1)\rho(r_2)}{\mid r_1 - r_2 \mid} dr_1 dr_2 \tag{2.26}$$

where M is the total number of nuclei. Then, the kinetic energy for non-interacting electrons becomes:

$$E_{KE}[\rho(r)] = \sum_{i=1}^{N} \int \psi_i(r) - \frac{\nabla^2}{2} \psi_i(r)dr \tag{2.27}$$

where N is the total number of electrons. Exchange-correlation term includes both the contributions of exchange and correlation, and a correction term that correlates the

real kinetic energy and $E_{KE}[\rho(r)]$ term. The "local density approach (LDA)" is based on the uniform gas approach, and it assumes that the electron density is uniformly scattered in all space. The electron correlation energy as a function of ϵ_{XC} is:

$$E_{XC}[\rho(r)] = \int \rho(r)\epsilon_{XC}(\rho(r))dr \tag{2.28}$$

And it can be arranged so that it provides the exchange correlation function:

$$V_{XC}[\rho(r)] = \frac{dE_{XC}[\rho(r)]}{d\rho} \tag{2.29}$$

hence the following equation is obtained:

$$V_{XC}[\rho(r)] = \rho(r)\frac{d\epsilon_{XC}(\rho(r))}{d\rho(r)} + \epsilon_{XC}(\rho(r)) \tag{2.30}$$

Then, the Hamiltonian of Kohn-Sham h_{KS} is:

$$h_{KS}\psi_i = \epsilon_i\psi_i \tag{2.31}$$

which can also be expressed as

$$h_{KS} = -\frac{\nabla^2}{2} - \sum_{A=1}^{N}\frac{Z_A}{\mid r - R_A \mid} + \int \frac{\rho(r_1)\rho(r_2)}{\mid r_1 - r_2 dr_1 dr_2} V_{XC}[\rho(r)] \tag{2.32}$$

That is, to solve Kohn Sham equation, solving E_{XC} should be enough. The exchange-correlation term is divided into E_X and E_C, which are described by gradient corrected methods. M06-2X is one of the fuctionals that use meta-generalized gradient approximation and its E_X term is as follows:

$$E_X = \sum_{\sigma}\int dr[F_{x\sigma}^{PBE}(\rho_\sigma, \delta\rho_\sigma)f(\omega_\sigma) + \epsilon_{LSDA}^{X\sigma}h_X(x_\sigma, z_\sigma)] \tag{2.33}$$

Basis sets are terms that used to describe the orbitals of a system, and mainly based on Gaussian functions. The most commonly used ones are developed by Pople and they are split valence double zeta basis sets [103–109]. In these basis sets, the inner shell is described by the contracted and diffuse parts, hence it is splitted. Depending on the atom types, one can either include diffusion and polarization functions can be added to the basis set.

2.4. Hybrid Quantum Mechanics/Molecular Mechanics Method

The effects of the environment on a reaction mechanism is undeniable. However, modelling of whole environment along with the molecules of interest at QM level is computationally demanding. That's why, hybrid QM/MM methods are developed [110–112].

The core idea of QM/MM method is that the number of atoms that are treated with QM descriptors are much less then the rest of the system, which affects the former only via non-bonded interactions. For QM/MM calculations, the system of interest is separated into two parts: QM region and MM region. The corresponding Hamiltonian is defined by the QM contributions (H_{QM}), MM contributions (H_{MM}) and their interaction ($H_{QM/MM}$):

$$H = H_{QM} + H_{MM} + H_{QM/MM} \tag{2.34}$$

in which the last term depicts the electrostatic interaction between atom in MM region and electrons in QM region, as well as their repulsion. The non-electrostatic term is defined as:

$$H_{QM/MM} = \sum_{k} \frac{q_M Z_k}{R_{kM}} - \sum_{iM} \frac{q_M}{r_{iM}} + V^{vdw}_{QM/MM} \tag{2.35}$$

where the point charge is shown with q_M, core charges in QM is shown with Z_k, R_{kM} and r_{iM} are distances between an atom k in QM and an atom MM in MM regions, and

the separation between QM electron and MM atoms. Then the partition for QM/MM Hamiltonian can be written as:

$$H = H_{elec} + H_{nonelec} \tag{2.36}$$

in which the non-electronic interaction terms are defined as:

$$H_{nonelec} = H_{MM} + V^{vdW}_{(QM/MM)} + V^{nuclei}_{(QM+QM/MM)} \tag{2.37}$$

Hence, the parts in which quantum treatment is required, H_{el} and $V^{nuclei}_{(QM+QM/MM)}$, calculated by QM tools at different level of theories such as *ab initio* or DFT, while the rest of the terms, H_{MM} and $V^{vdW}_{(QM/MM)}$, are treated with MM.

For the covalently bonded atoms that fall in between the QM and MM part, a special treatment is required since the MM atoms do not contain valence electrons. One method to overcome this obstacle is the "link atom" method in which a monovalent atom, most likely hydrogen, is aligned along the C_{QM}-C_{MM} bond, and it interacts with the both regions. In this study, the QM part does not covalently bonded to any other atom in the system, hence the link atom approach was not needed.

QM/MM method can also be used to calculate the single-point energies of the snapshots taken from the trajectory. Again, the system should be separated to two different parts, and for a given snapshot, the interaction energy between the QM region and the rest of the complex can be computed.

2.5. Computational Details

Systems for both TPT-closed and TPT-open are constructed using the ternary crystal structure (PDB ID: 1K4T, resolution: 2.1 rA) [63]. This structure contains both TPT-closed and TPT-open forms intercalated to DNA/TopoI complex. The -SH end of the nicked strand was modified to -OH, and all solvent molecules, except crystal

water molecules, were removed. Amber14/AmberTools15 software's tleap module was used to create the topology and the coordinate files [95].

2.5.1. System Preparations

Both systems, TPT-closed and TPT-open complexes, were solvated explicitly in TIP3P water box with 15 Å buffer region. Na^+ atoms were added to neutralize the systems. The final number of atoms in both systems were around 144000. By adjusting the box size gradually, the initial density, 1.0 $g.cm^{-3}$ has been reached. RESP procedure was used to obtain the partial charges on TPT-closed, TPT-open and phosphotyrosine residue (Y723) [113, 114]. ff14SB, ff99bsc0 and gaff parameters were used for TopoI, DNA and TPT molecules, respectively [115–118]. The hydrogen atoms of TopoI and DNA were added by tleap module of AmberTools15 as defined in their respective forcefields, and standard protanation states were applied for all residues.

2.5.2. Molecular Dynamics Simulations

Both complexes were first equilibrated starting from 0K up to 300K in ten steps with Langevin thermostat, and the collision frequency was chosen to be 4.0 ps^{-1} [119]. The steps for equilibration procedure were as follows:

(i) with all atoms were constrained and under the constant volume, within 160 ps time, the systems were heated up to 20K (3.0 kcal mol^{-1} $Å^{-2}$),
(ii) with only backbone atoms were constrained, within 1350 ps time, the temperature was increased up to 300K gradually,
(iii) without any restraint, the final equilibration was performed using NPT ensemble for 300 ps, and Langevin thermostat and Berendsen barostat were used with the default AMBER settings [119, 120].

SHAKE algorithm was employed during all calculations, and the time step was chosen to be 1 fs. For production runs, NPT ensemble was used. The runs were performed using AMBER14 software on CUDA-enabled GPUs with pmemd module

[119–123]. For both systems, the trajectories were 450 ns long, and at each 0.5 ns the coordinates were printed. Additional runs were also performed starting with different velocities to asses the convergence of the systems, and they are 350 ns long simulations. In addition, root-mean square deviations (RMSD), hydrogen bonds and distances were calculated with the cpptraj module of AmberTools15 with the default cutoff values [124]. For the following analysis, the first 30 ns of the simulations were excluded due to the fluctuations in RMSD.

2.5.3. Interaction Energy Calculations with Hybrid QM/MM Method

Quantum Mechanics/Molecular Mechanics (QM/MM) interaction energy calculations were performed between TPT molecules and DNA/TopoI binary complex. To calculate the QM/MM interaction energies, the complex must be separated into two components: the drug, which is the QM region, and the DNA/TopoI, which corresponds to the MM region. Before the calculations, all solvent water molecules and Na^+ ions were removed. TPT molecules were treated at M06-2X/6-31G(d) level of theory through Amber/Gaussian09 interface, as implemented in AmberTools15; and parameters for the MM site were kept as the same [114, 125, 126]. By taking one snapshot at every nanosecond, a total of 420 snapshots for each system were extracted and the non-relaxed interaction energies ΔE were calculated as follows:

$$\Delta E = E_{TPT:DNA/TopoI} - E_{DNA/TopoI} - E_{TPT} \tag{2.38}$$

where the first term on the right of the equation represents the single-point energy of the whole complex, and $E_{DNA/TopoI}$ and E_{TPT} terms correspond to the single-point energies of binary complex and TPT molecules, respectively. In order to obtain the individual contributions of DNA and TopoI to ΔE, the term can also be rewritten as:

$$\Delta E = \Delta^2 E_{DNA,TPT} + \Delta^2 E_{TopoI,TPT} + \Delta^3 E_{DNA,TopoI,TPT} \tag{2.39}$$

Then, the individual terms can be defined as follows:

$$\Delta^3 E_{DNA,TopoI,TPT} = E_{TPT:DNA/TopoI} - E_{TopoI} - E_{DNA} - E_{TPT} - \Delta^2 E_{DNA,TPT} - \Delta^2 E_{TopoI,TPT} - \Delta^2 E_{TopoI,DNA} \quad (2.40)$$

$$\Delta^2 E_{TopoI,TPT} = E_{TPT:TopoI} - E_{TPT} - E_{TopoI} \quad (2.41)$$

$$\Delta^2 E_{TopoI,DNA} = E_{DNA/TopoI} - E_{TopoI} - E_{DNA} \quad (2.42)$$

$$\Delta^2 E_{DNA,TPT} = E_{TPT:DNA} - E_{TPT} - E_{DNA} \quad (2.43)$$

In Equations 2.40-2.43, the $\Delta^2 E_{i,j}$ terms represents the interaction energy between only two components of the whole complex (e.g.,$\Delta^2 E_{TopoI,DNA}$ term gives the interaction between TopoI and DNA). The $\Delta^3 E_{DNA,TopoI,TPT}$ term describes the interaction energy which resulted from the effect of three components on each other. This term would be zero if the TPT is not treated with QM.

2.5.4. The Contributions from Residues

To understand the contributions of closeby as well as catalytically important residues, additional calculations were performed. The vdW parameters and the charge of selected residues were set to zero one by one, and the interaction energies were re-calculated. In addition to the residues, the immediate DNA bases where TPT molecule intercalated (T:A and G:C, at positions -1 and +1, respectively) were also investigated. The single-residue effect was calculated as following:

(i) set the charge of the residue to zero,

(ii) calculate the new interaction energy ($\Delta E_{residue}$),

(iii) to obtain the single-residue effect, $\Delta E_{residue}$ was subtracted from ΔE,

(iv) if the final energy is larger than zero, then the residue has an destabilizing effect, and vice versa.

Standard deviations are calculated as follows:

$$SD = \sqrt{\frac{\sum \Delta E_i^2}{N} - \frac{\sum \Delta E_i}{N}} \tag{2.44}$$

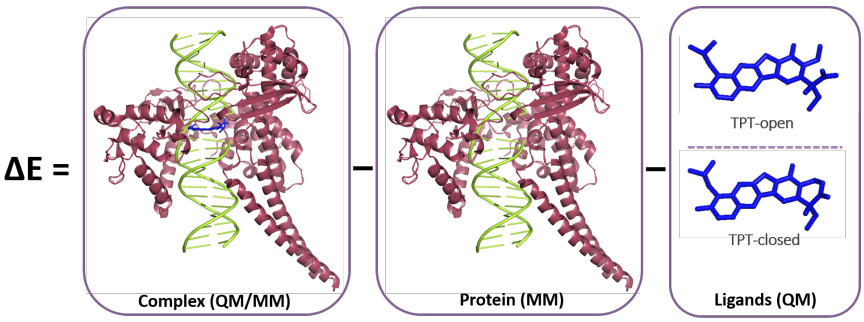

Figure 2.1. Separation of regions for QM/MM calculations.

3. RESULTS

3.1. Stability of the Simulations

Stability of a given simulation can be assessed using different approaches, however, the most common one is "Root-Mean-Square Deviation" (RMSD). RMSD of the main simulations for TPT-closed and TPT-open are obtained with respect to the crsytal structure (PDB ID: 1K4T) in Figure 3.1 and Figure 3.2, respectively. It can be immediately seen that absence of the linker domain stabilizes the system in both cases. For TPT-closed complex, the average RMSD for whole complex is calculated to be 2.88 Å and it is 1.95 Å if the linker domain is excluded. The trend is similar in TPT-open complex as well: RMSD for entire complex is 2.23 Å, and when the linker domain is not present, it drops to 1.72 Å.

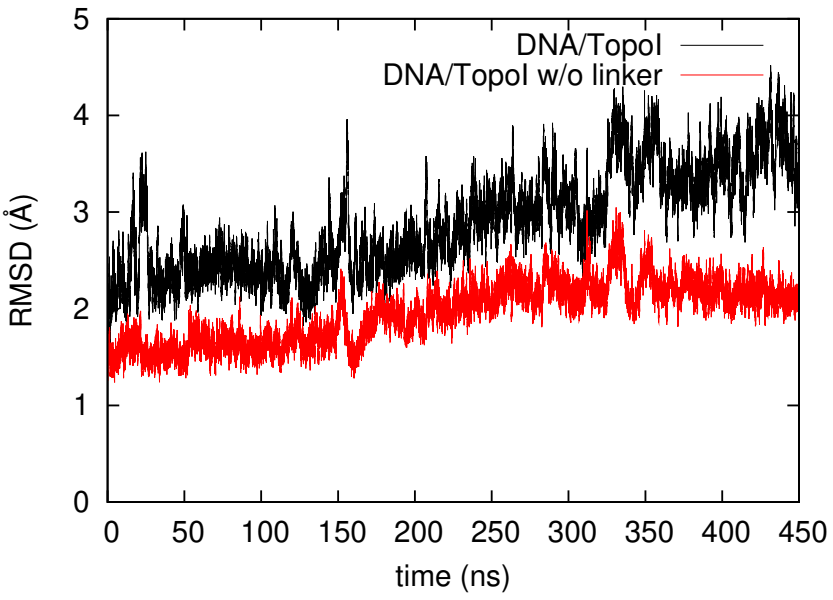

Figure 3.1. RMS deviations in Å for the main simulation TPT-closed complex. Black line: RMSD of backbone atoms of DNA/TopoI complex. Red line: RMSD of backbone atoms of DNA/TopoI without the linker domain [94].

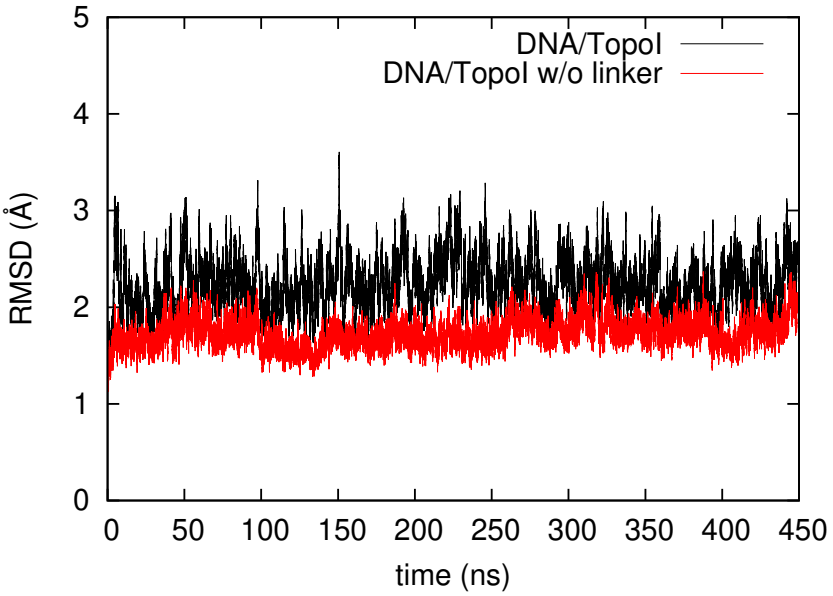

Figure 3.2. RMS deviations in Å for the main simulation TPT-open complex. Black line: RMSD of backbone atoms of DNA/TopoI complex. Red line: RMSD of backbone atoms of DNA/TopoI without the linker domain [94].

To determine the convergence of the main simulation set, a second set of simulations has been performed for 350 ns. These duplicate simulations have different initial velocities. Similar behavior in RMS deviations has been observed: the presence of the linker domain destabilizes the complex. The RMSD plots for the duplicate simulations can be found in Figure A.1 and Figure A.2 for TPT-closed and TPT-open, respectively.

During the interaction energy calculations, only the last 420 ns of the main simulations have been used, and the first 30 ns have been considered as equilibration time. For the topological analysis, main simulations as well as the duplicate simulations have been investigated.

3.2. QM/MM Interaction Energies

Interaction energies for TPT-closed and TPT-open have been calculated as described in Section 2.5.3. For snapshots taken at every 1 ns, with the total of 420 frames,

the average energy values can be found in Table 3.1. In addition, in Figure 3.3 and Figure 3.4, time series of the individual energy values can be tracked for TPT-closed and TPT-open, respectively. Average ΔE values tell us that while both ligands, TPT-closed and TPT-open, bind to the DNA/TopoI complex tightly, TPT-closed has more favorable interactions. This result is consistent with the experimental findings in which both forms have been found to intercalate at the same site in the crsytal structure [63]. The next question is which component of the complex causes this difference. When ΔE is decomposed as in Equation 2.39, one obtains the individual contributions of DNA ($\Delta^2 E_{DNA,TPT}$) and TopoI ($\Delta^2 E_{TopoI,TPT}$) seperately. The former does not change between two complexes. On the other hand, the latter differs quite a lot between TPT-closed and TPT-open complexes. Having the interaction of TPT-closed with TopoI 26 kcal.mol^{-1} more stable than of TPT-open points out that main contributor for ΔE difference is the interaction with the protein component, TopoI.

Table 3.1. The interaction energies (kcal.mol^{-1}) for TPT-closed and TPT-open complexes averaged over last 420 ns of the main simulations [94].

	TPT-closed	**TPT-open**
$\mathbf{\Delta E}$	-102.73±9.61	-79.94 ±9.43
$\mathbf{\Delta^2 E_{DNA,TPT}}$	-57.82±3.17	-59.19±6.56
$\mathbf{\Delta^2 E_{TopoI,TPT}}$	-53.54±10.40	-27.59±8.48
$\mathbf{\Delta^3 E_{DNA,TopoI,TPT}}$	8.62±1.71	6.84±2.51

3.3. Single Residue Contributions to the Interaction Energy

Identifying the interaction with TopoI as the main contributor of the interaction energy difference leads one to ask which of the amino acid residue/residues is/are key player. The residues in close proximity of the ligand as well as the ones that were shown to have a role in the catalytic cycle are considered for this analysis. Moreover, to see the individual contributions of the imediate stacking bases, they are also subjected to the same analysis. A total of twelve residues and four DNA bases have been analyzed

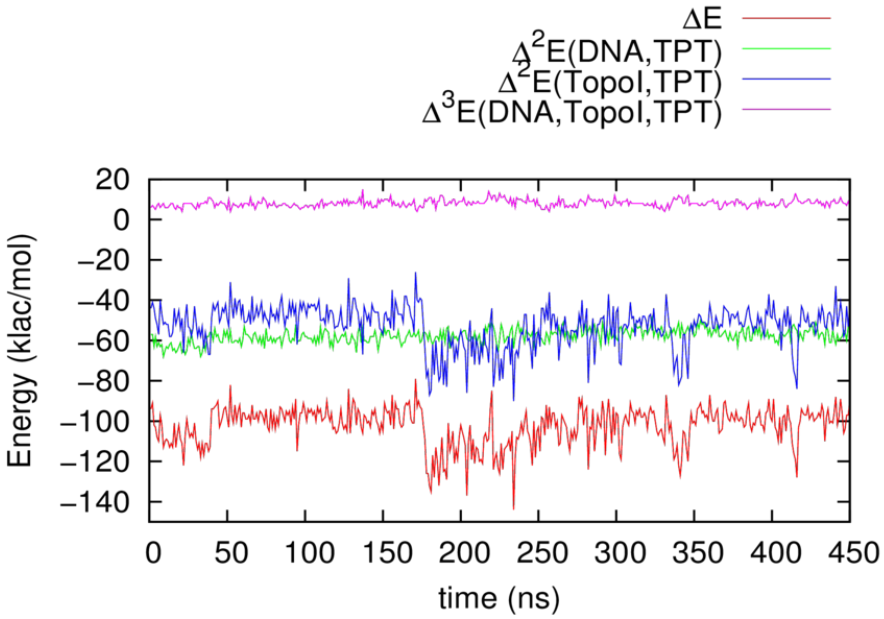

Figure 3.3. Time series of QM/MM interaction enegies (kcal.mol^{-1}) of TPT-closed complex throughout 450 ns simulation time [94].

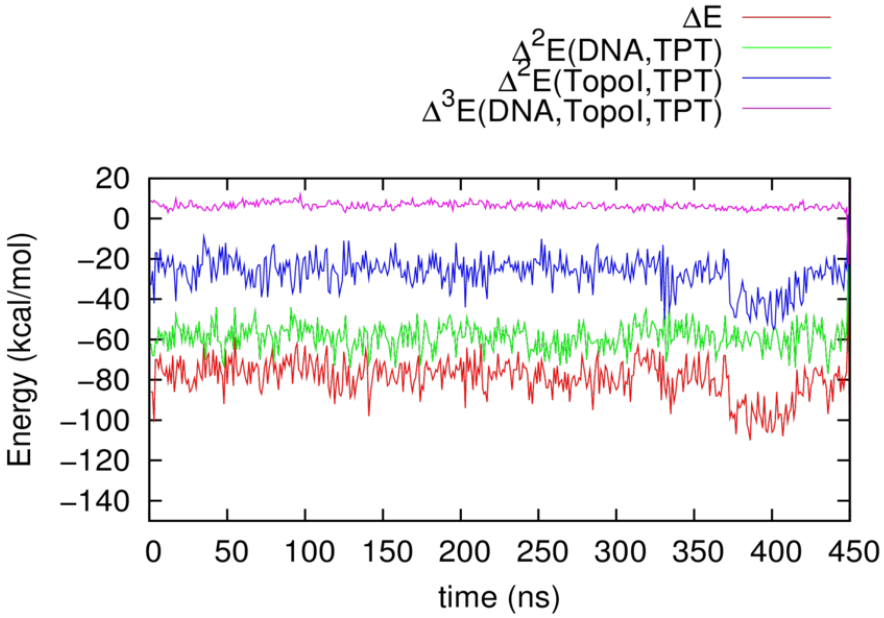

Figure 3.4. Time series of QM/MM interaction enegies (kcal.mol^{-1}) of TPT-open complex throughout 450 ns simulation time [94].

and the names of the residues and DNA bases can be found in the first column of Table 3.2. The standard deviations are calculated using Equation 2.44. The time series of individual residues can be seen in Figure 3.5 and Figure 3.6. In these Figures, the behaviours of the energies are quite consistent throughout the simulations.

Table 3.2. Average contribution of each residue to ΔE for TPT-closed and TPT-open [94].

	TPT-closed	**TPT-open**
E356	-7.29±6.94	-3.69±2.24
R364	-2.39±2.31	2.79±5.47
K374	6.08±2.53	4.81±2.20
E418	-1.31±1.10	-2.20±1.07
K425	2.97±5.25	5.94±2.42
G503	-0.1±0.02	-0.06±0.03
K532	-18.71±4.00	-1.90±5.81
D533	0.43±2.81	-0.18±4.51
H632	-1.51±0.36	-0.51±0.44
T718	-2.46±1.14	-1.15±2.07
N722	-0.65±0.68	-1.77±2.47
Y723	7.06±1.00	3.68±1.86
-1T	-3.36±1.38	-7.68±2.76
-1A	-15.33±2.01	-14.71±2.83
+1G	-18.17±2.22	-16.86±3.01
+1C	-8.82±1.55	-9.02±2.28

When Table 3.2 is examined, it can be seen that only three residues (K374, K435, Y723) have positive value, which expresses the repulsive interaction between the residue and the ligand. These two lysines are located very close to A-ring of TPT, and Y723 is the phosphotyrosine residue that links the TopoI with DNA and it is near the E-ring. K374, K425 and Y723 do not change considerably between TPT-closed and

TPT-open. The residues with negative contributions are E356, E418 and T718, and while the glumate residues are positioned near A-ring, T718 is around the E-ring of TPT. Between the two glumate residues, E356 is the one with a notable contribution value (-7.29 kcal.mol^{-1}). This can be due to a direct but rather trivial interaction with TPT-closed. For TPT-open, the contribution is not as significant as TPT-closed case.

Another set of three resiues (G503, H632, N722) does not have any significant contribution to the interaction energy in both TPT-closed and TPT-open. It is an expected result when one considers their distance to the binding site. Therefore, the impact of these residues might be more structural. One surprising result in Table 3.2 is the very small contribution of D533 residue. D533 is located quite close to E-ring of TPT, and it has two neighboring positively charged residues: K532 and R364. As a result, the effect of D533 might be shielded by these positively charged residues.

The two amino acid residues that are positioned very close to TPT, R364 and K532, show significant contributions. R364 has its side chain very close to the nitrogen atom on the B-ring (distance between N-N: 3.8 Å). Although slightly, it has a favorable contribution in TPT-closed results. On the other hand, the contribution in TPT-open case indicates repulsive interaction between R364 and TPT-open. At the end, having such small difference in absolute energy values and large deviations leads to a minor resulting effect for R364 residue.

K532 is the one residue that has contributions quite different for TPT-closed (-18.71 kcal.mol^{-1}) and TPT-open (-1.90 kcal.mol^{-1}). It is also known that K532 has an important role in the enzymatic acitivity of TopoI [91]. From these values, one can safely conclude that the difference in the total interaction energy is due to the interaction loss of TPT-open with K532 residue.

3.4. Hydrogen Bond Networks around the Binding Site

Interaction energy calculations have provided us the information on which residues might have a direct interaction with TPT. To identify the interactions that are not be-

ing captured by the previous analysis, hydrogen bonds and inter-atomic distances for the residues around the active site are investigated. This analysis has been performed for both main set of simulations and for the duplicate simulations. Hydrogen bond analysis results are summarized in Table 3.3-3.4 for the main set of simulations, and Table 3.5-3.6 for the duplicate simulations. The corresponding distances can be found in Tables 3.7-3.8 and Tables 3.9-3.10 for main set and duplicate simulations, respectively. The time series of inter-atomic distances are included in Figures 3.11-3.15 for the main set of simulations, and Figure A.3-A.6 for the duplicate simulation set. In general, the distances behave stable and do not fluctuate a lot in case of both sets.

In hydrogen bond analysis, the most striking difference in hydrogen bonding percentages is obtained for K532 residue. This residue has an interaction with O3 atom of TPT-closed for 53% and 49% of the simulation time in main and duplicate simulations, respectively. It can be seen from the times series in Figure 3.8 that this interaction is quite persistent and strong throughout the simulations. Moreover, there is another residue that interacts only with TPT-closed, but this interaction is not present in the duplicate simulations. The reason for this is that TPT-closed in the duplicate simulations has shifted 1 Å towards the scissile strand and caused an additional hydrogen bond between free -OH group on the +1G DNA base. One might think that this shift could have affected the interaction energy calculations; however, one should also keep in mind that a standard hydrogen bond has 5 kcal.mol^{-1} energy, which would not affect the overall trend. For TPT-open, there is no direct hydrogen bond with the surrounding residues, and this corralates with the findings in interaction energy calculations. As a result, the interaction of TPT-closed with K532 is causing the difference in the total interaction energy, resulting in more stable complex than of TPT-open.

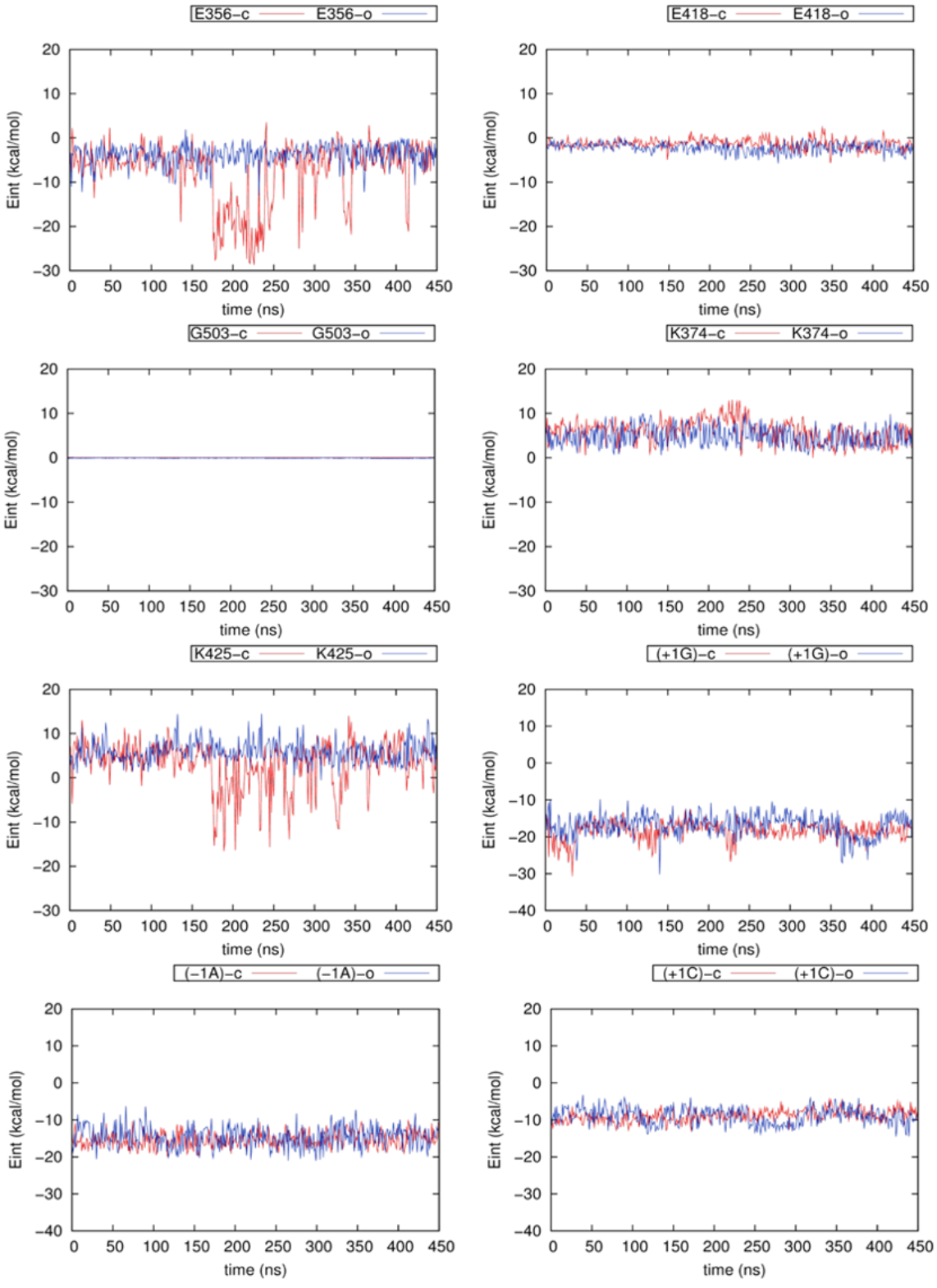

Figure 3.5. Time series of the single-residue contributions to the interaction energy around A-ring. "-c" an "-o" are used to imply TPT-closed and TPT-open, respectively [94].

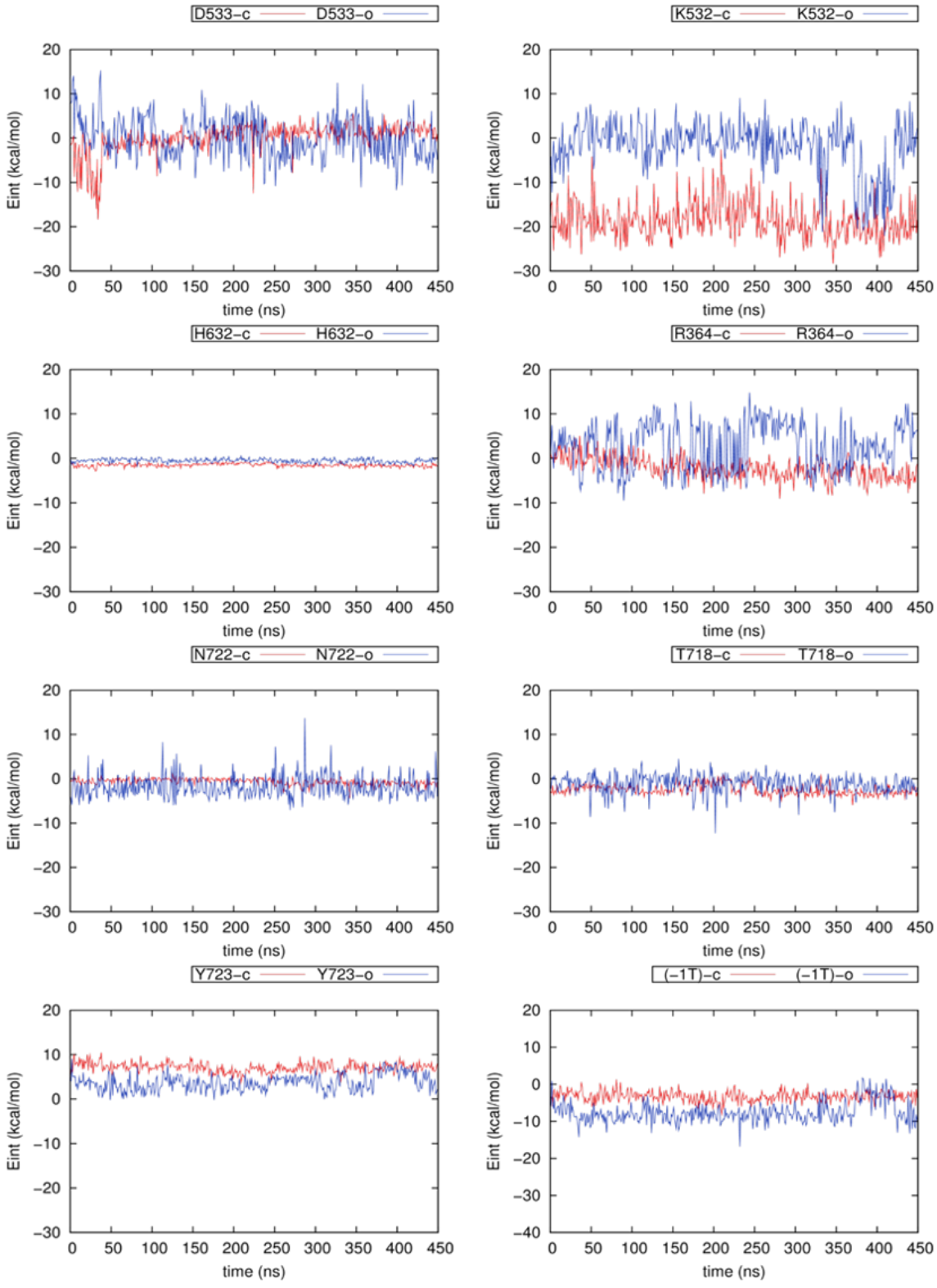

Figure 3.6. Time series of the single-residue contributions to the interaction energy around E-ring. "-c" an "-o" are used to imply TPT-closed and TPT-open, respectively [94].

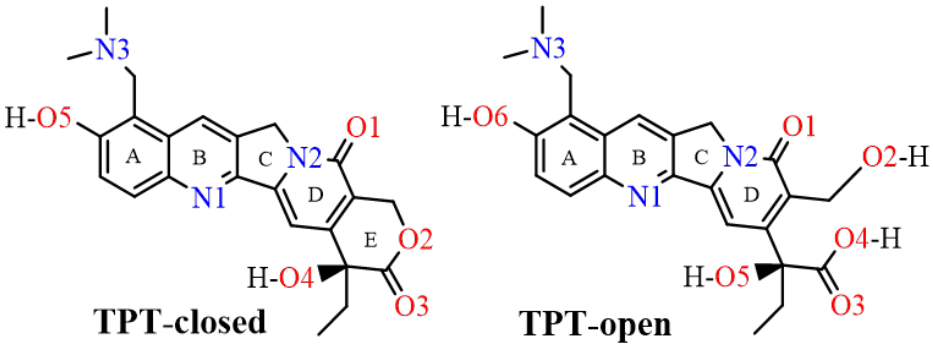

Figure 3.7. Atoms of TPT-closed and TPT-open structures are labeled.

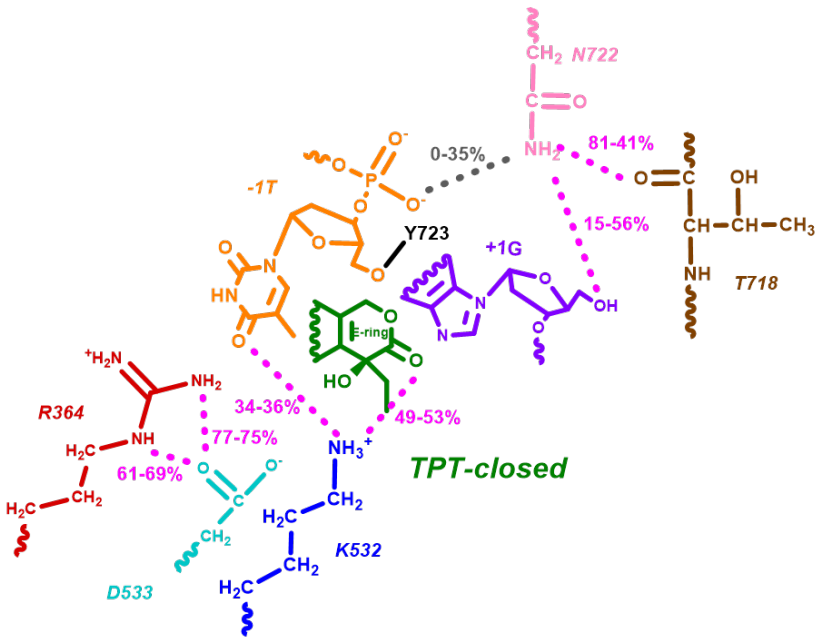

Figure 3.8. Hydrogen bond network and percentages around A-ring of TPT-closed molecule [94].

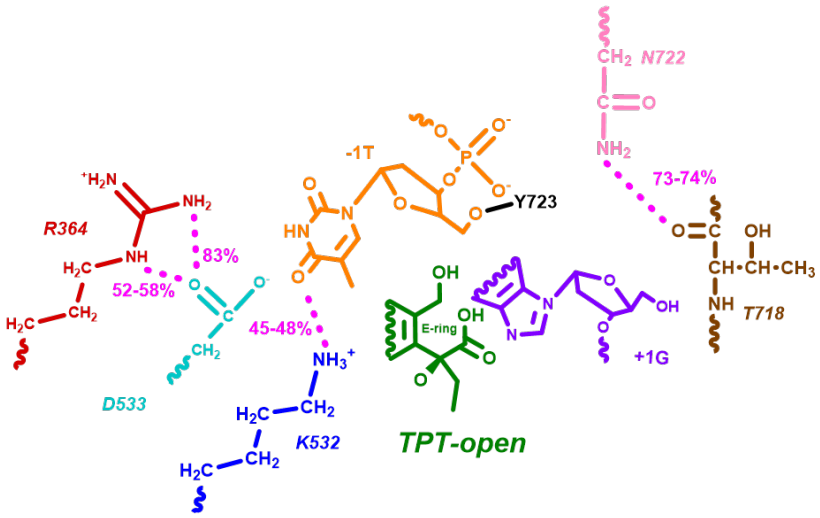

Figure 3.9. Hydrogen bond network and percentages around E-ring of TPT-open molecule [94].

Table 3.3. Hydrogen bond percentages between selected residues for main simulations around E-ring [94].

Residues	TPT-closed	TPT-open
K532-NZ/TPT	0.53 (O3)	0.00 (O3)
E356-OE2/TPT	0.15 (O5)	—
41G-O5'/TPT	0.04 (O1)	0.00 (O2)
N722-ND2/TPT	—	0.03 (O2)
K532-NZ/-1T-O2	0.36	0.48
N722-ND2/T718-O	0.41	0.73
N722-N/T718-O	0.28	0.12
N722-ND2/-1T-OP2	0.35	0.0
N722-ND2/+1G-O5'	0.15	0.0
Y723-O1P/H632-NE2	0.74	0.74
Y723-O1P/r488-NH2	0.83	0.85
Y723-O1P/R590-NH1	0.68	0.73
Y723-O2P/R488-NE	0.87	0.90

Table 3.4. Hydrogen bond percentages between selected residues for main simulations around A-ring [94].

Residues	**TPT-closed**	**TPT-open**
E356-OE1/K374-NZ	0.37	0.27
E356-OE2/K374-NZ	0.33	0.28
E356-OE1/K425-NZ	0.25	0.25
E356-OE2/K425-NZ	0.25	0.24
R362-N/-1A-OP1	0.94	0.78
R362-NH1/-1A-OP1	0.59	0.03
G363-N/-1A-OP1	0.50	0.40
K374-NZ/-1A-OP2	0.84	0.71
R364-NE/D533-OD1	0.61	0.52
R364-NE/D533-OD2	0.33	0.47
R364-NH2/D533-OD1	0.77	0.83
R364-NH2/-1A-N3	0.51	0.33
R364-NH1/-1A-N3	0.22	0.41
R364-N/-2A-OP1	0.82	0.77
E418-OE1/K425-NZ	0.17	0.13
E418-OE2/K425-NZ	0.17	0.18

Table 3.5. Hydrogen bond percentages between selected residues for duplicate simulations around A-ring [94].

Residues	**TPT-closed-2**	**TPT-open-2**
E356-OE1/K374-NZ	0.36	0.44
E356-OE2/K374-NZ	0.41	0.32
E356-OE1/K425-NZ	0.25	0.28
E356-OE2/K425-NZ	0.22	0.39
R362-N/-1A-OP1	0.90	0.95
R362-NH1/-1A-OP1	0.0	0.0
G363-N/-1A-OP1	0.44	0.34
K374-NZ/-1A-OP2	0.93	0.93
R364-NE/D533-OD1	0.69	0.58
R364-NE/D533-OD2	0.29	0.47
R364-NH2/D533-OD1	0.75	0.83
R364-NH2/-1A-N3	0.53	0.33
R364-NH1/-1A-N3	0.20	0.34
R364-N/-2A-OP1	0.82	0.82
E418-OE1/K425-NZ	0.06	0.0
E418-OE2/K425-NZ	0.04	0.0

Table 3.6. Hydrogen bond percentages between selected residues for duplicate simulations around E-ring [94].

Residues	TPT-closed-2	TPT-open-2
K532-NZ/TPT	0.49 (O3)	0.00 (O3)
E356-OE2/TPT	0.01 (O5)	—
41G-O5'/TPT	0.67 (O1)	0.01 (O2)
N722-ND2/TPT	—	0.02 (O2)
K532-NZ/-1T-O2	0.34	0.45
N722-ND2/T718-O	0.81	0.74
N722-N/T718-O	0.12	0.11
N722-ND2/-1T-OP2	0.0	0.0
N722-ND2/+1G-O5'	0.56	0.0
Y723-O1P/H632-NE2	0.66	0.68
Y723-O1P/r488-NH2	0.89	0.90
Y723-O1P/R590-NH1	0.79	0.56
Y723-O2P/R488-NE	0.90	0.89

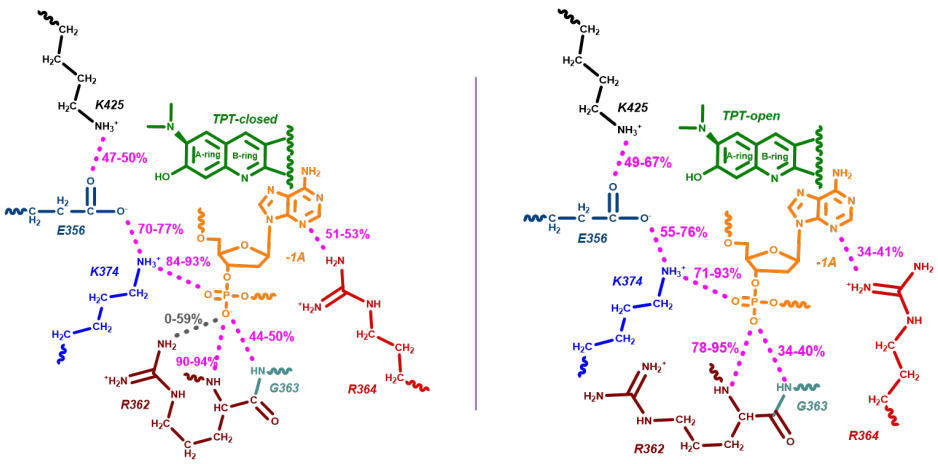

Figure 3.10. Hydrogen bond network and percentages around A-ring of TPT-closed molecule [94].

The residues around the intercalation site are also investigated. In Tables 3.3 and 3.6, and Figures 3.10 and 3.11, it can be seen that K532 forms very stable hydrogen bond with -1T DNA base for 34-36% and 45-48% of the simulation time for TPT-closed and TPT-open, respecively. Another interaction is detected between +1G DNA base and the side chain of N722 residue. This interaction has been observed for 15-56% of simulation time in only complexes with TPT-closed molecule. Only in one of the runs for TPT-closed complex, an interaction between the side chain of N722 and -1T DNA base has been observed for 35% of the time.

When the hydrogen bond network around A-ring is examined, it can be easily seen that both TPT-closed and TPT-open have the same motifs. The whole interaction scheme can be seen in Figure 3.12. K374 residue interacts with the backbone phosphate group of -1A DNA base, which is coordinated by the backbone of G363 and R364 resdiues at the same time. Another interaction motif observed at this region is between the side groups of R364 and D533 residues, and it persists most of the simuation time in both complexes. R364 has contacts with two DNA bases (-1A and -2A) as well. All these interactions have been reported in the literature also [63, 91, 93]. Interaction formed between catalytic residue Y723 and H632 is found to be stable in all simulations.

In conclusion, it was observed that the hydrogen bond motifs around A-ring are well-preserved regardless of which form of TPT is present in the complex. This also supports the fact that the chemical changes happen around E-ring of the TPT molecule. The energy contributions of these residues near A-ring are mostly positive, and this suggests that they could have a role in stabilizing the interaction between TopoI and DNA. The most striking result between TPT-closed and TPT-open forms was observed in the interactions of K532 and N722 residues. It was shown previously that K532A mutation causes extremely stable cleavable complex, which is unable to religate the scissile DNA strand back [91]. An MD study performed by Mancini *et al.* demonstrated that K532 interacts with +1G DNA base in the presence of TPT-closed only [93]. This interaction was not observed in the simulation performed in this thesis, however, the hydrogen bond between K532 and TPT-closed suggests that the role of K532 in religation step might be mechanical only, by assisting that the nicked ends

of the DNA backbone are at the right position. Another experimental study shows that N722 mutation inhibits the CPT binding [90]. The interaction between N722 and +1G, which is observed by Mancini *et al.*, has been observed only in the complex with TPT-closed ligand. This proposes a role for N722 residue in stabilizing the DNA base such that π-stackings with the active form of the drug is favored. The positions of K532 and N722 can be seen in Figure 3.13.

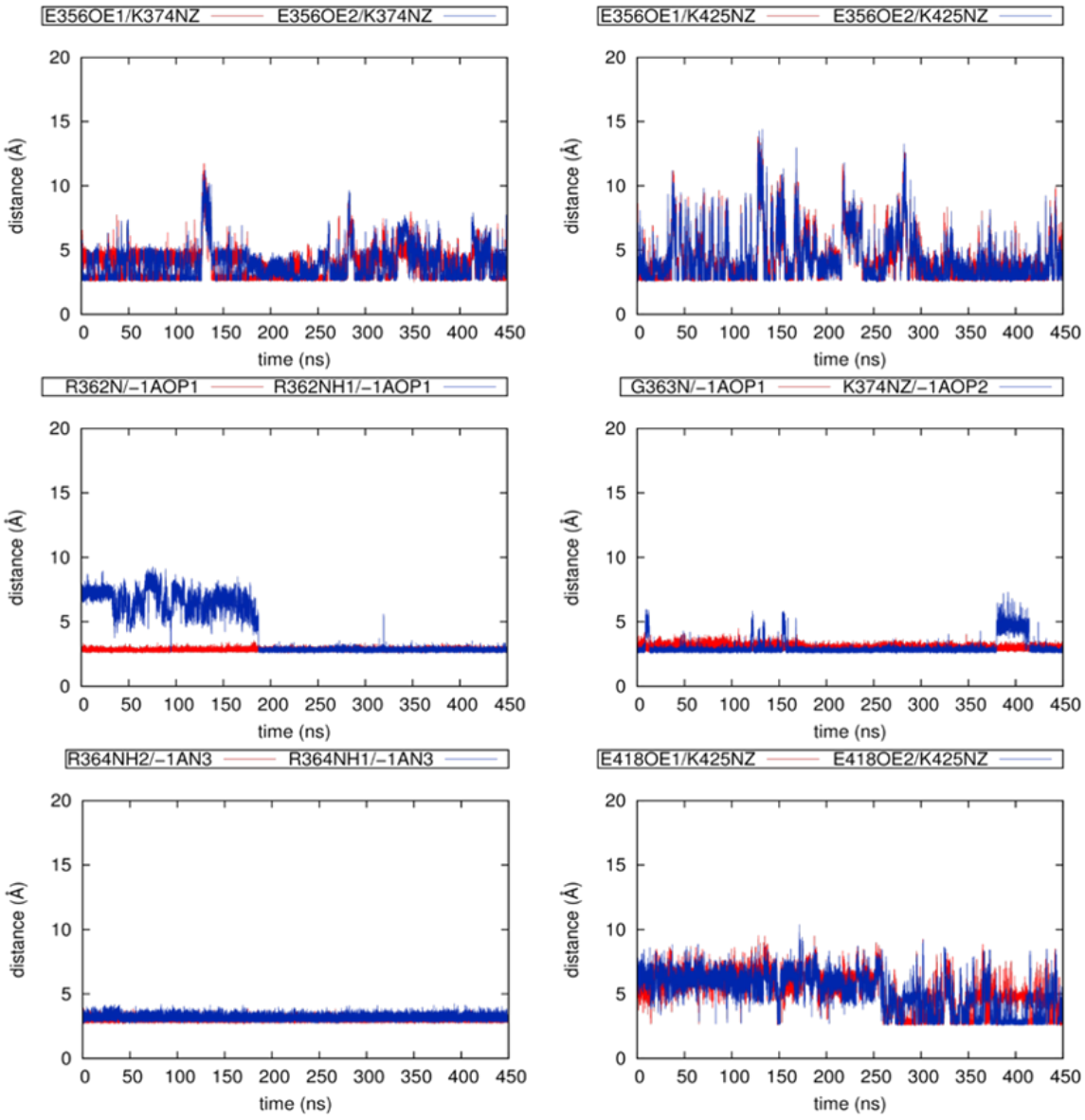

Figure 3.11. Time series for inter-atomic distances around A-ring for main simulation of TPT-closed [94].

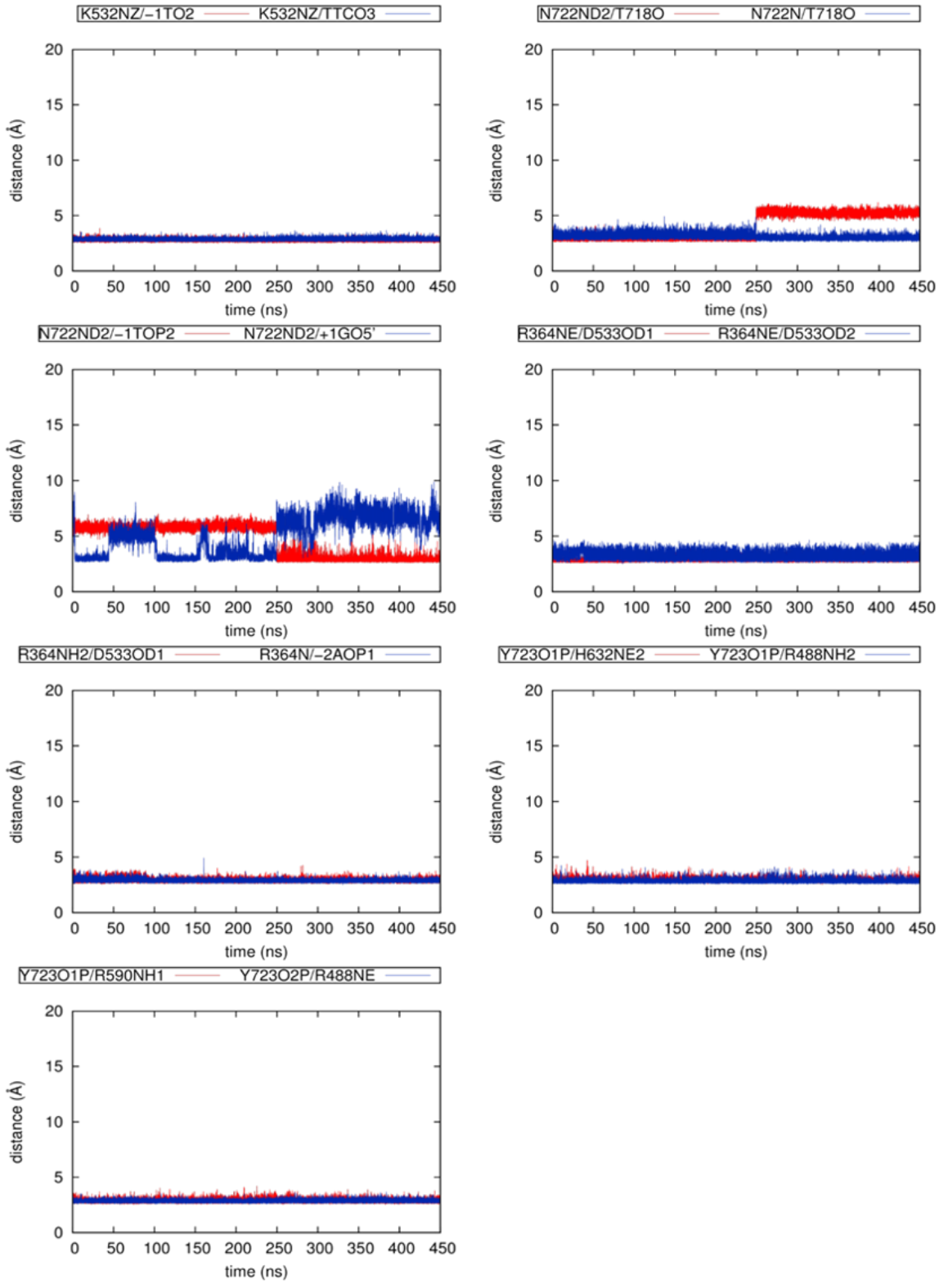

Figure 3.12. Time series for inter-atomic distances around E-ring for main simulation of TPT-closed [94].

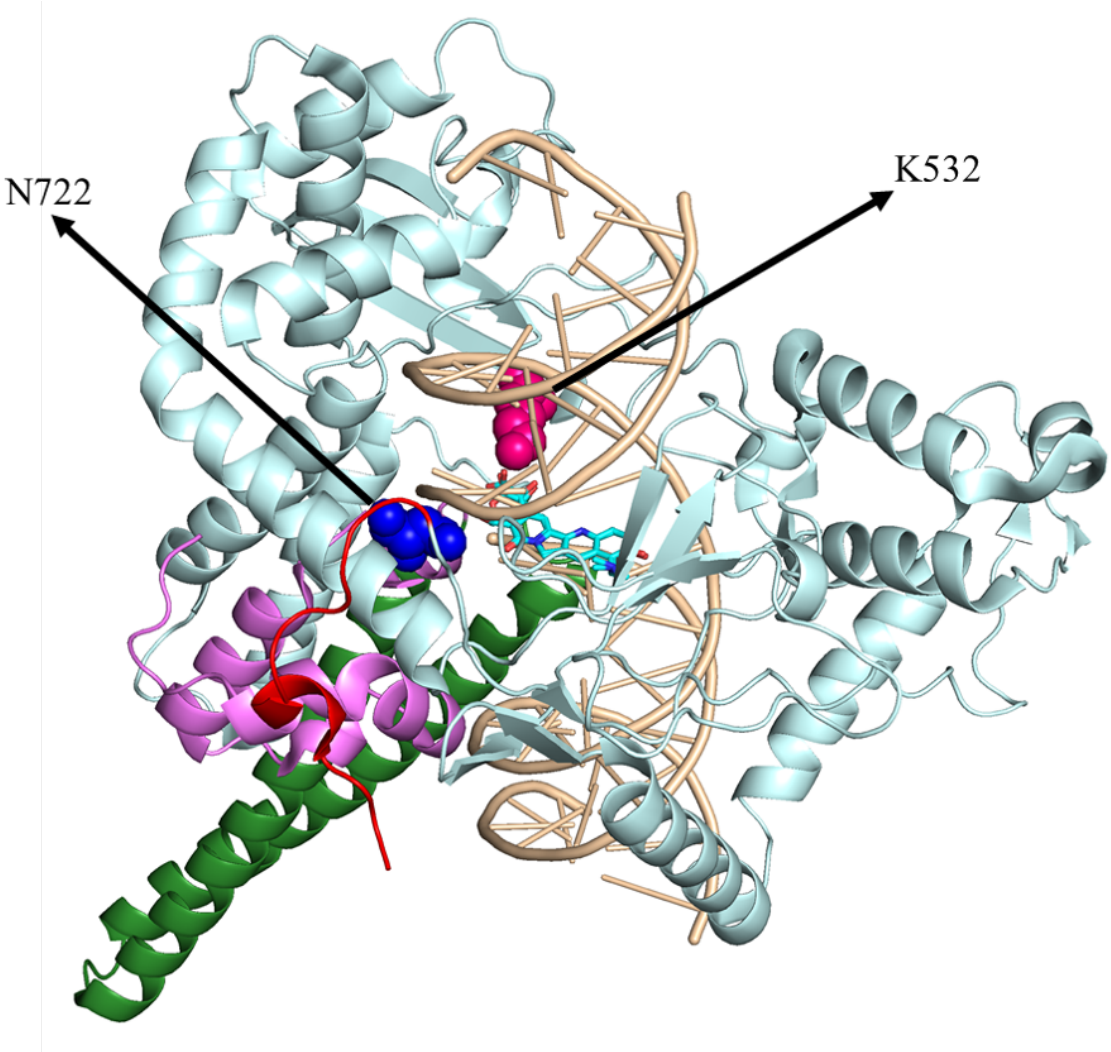

Figure 3.13. The positions of two important residues, K532 and N722 are displayed with respect to TPT (in cyan color) [94].

Table 3.7. Inter-atomic distances (in Å) between selected residues for main simulations around A-ring [94].

Residue	**TPT-closed**	**TPT-open**
E356-OE1/K374-NZ	3.83±1.22	4.76±1.91
E356-OE2/K374-NZ	3.90±1.27	4.76±1.99
E356-OE1/K425-NZ	4.23±1.81	5.30±3.18
E356-OE2/K4225-NZ	4.24±1.79	5.32±3.17
R362-N/-1A-OP1	2.82±0.10	2.94±0.37
R362-NH1/-1A-OP1	4.40±1.98	7.04±0.83
G363-N/-1A-OP1	3.03±0.18	3.12±0.39
K374-NZ/-1A-OP2	2.98±0.61	3.24±0.86
R364-NE/D533-OD1	2.92±0.17	2.94±0.17
R364-NE/D533-OD2	3.23±0.36	3.17±0.43
R364-NH2/D533-OD1	2.90±0.16	2.89±0.24
R364-NH2/-1A-N3	3.01±0.14	3.08±0.18
R364-NH1/-1A-N3	3.16±0.20	3.11±0.33
R364-N/-2A-OP1	2.89±0.13	2.90±0.13
E418-OE1/K425-NZ	5.34±1.43	5.12±1.47
E418-OE2/K425-NZ	5.17±1.48	5.17±1.40

Table 3.8. Inter-atomic distances (in Å) between selected residues for duplicate simulations around A-ring [94].

Residue	**TPT-closed**	**TPT-open**
E356-OE1/K374-NZ	4.063±1.39	3.68±1.03
E356-OE2/K374-NZ	4.05±1.53	3.92±1.01
E356-OE1/K425-NZ	5.37±2.72	3.84±1.55
E356-OE2/K4225-NZ	5.40±2.70	3.75±1.55
R362-N/-1A-OP1	2.83±0.12	2.80±0.10
R362-NH1/-1A-OP1	6.89±0.87	7.16±0.60
G363-N/-1A-OP1	3.06±0.21	3.11±0.22
K374-NZ/-1A-OP2	2.81±0.25	2.80±0.18
R364-NE/D533-OD1	2.90±0.16	2.93±0.16
R364-NE/D533-OD2	3.30±0.42	3.12±0.39
R364-NH2/D533-OD1	2.91±0.17	2.87±0.15
R364-NH2/-1A-N3	3.00±0.15	3.07±0.15
R364-NH1/-1A-N3	3.21±0.25	3.10±0.19
R364-N/-2A-OP1	2.89±0.14	2.88±0.12
E418-OE1/K425-NZ	5.73±1.17	6.17±0.78
E418-OE2/K425-NZ	5.46±1.10	6.20±0.27

Table 3.9. Inter-atomic distances (in Å) around E-ring between selected residues for main simulations around E-ring [94].

Residue	**TPT-closed**	**TPT-open**
K532-NZ/TPT-O3	3.06±0.41	6.32±1.16
K532-NZ/-1T-O2	2.79±0.11	2.86±0.16
N722-ND2/T718-O	3.95±1.19	2.92±0.16
N722-N/T718-O	3.12±0.25	3.21±0.22
N722-ND2/-1T-OP2	4.54±1.44	5.68±0.27
N722-ND2/+1G-O5'	4.99±1.71	6.90±0.58
Y723-O1P/H632-NE2	2.92±0.16	2.91±0.15
Y723-O1P/R488-NH2	2.83±0.15	2.85±0.14
Y723-O1P/R590-NH1	2.88±0.15	2.85±0.13
Y723-O2P/R488-NE	2.86±0.11	1.84±0.11

Table 3.10. Inter-atomic distances (in Å) between selected residues for duplicate simulations around E-ring [94].

Residue	**TPT-closed**	**TPT-open**
K532-NZ/TPT-O3	3.17±0.70	6.58±0.63
K532-NZ/-1T-O2	2.80±0.12	2.84±0.16
N722-ND2/T718-O	2.89±0.15	2.91±0.17
N722-N/T718-O	3.23±0.25	3.25±0.25
N722-ND2/-1T-OP2	5.76±0.29	5.61±0.27
N722-ND2/+1G-O5'	3.40±0.94	6.24±0.68
Y723-O1P/H632-NE2	2.96±0.18	2.97±0.27
Y723-O1P/R488-NH2	2.82±0.13	2.82±0.12
Y723-O1P/R590-NH1	2.86±0.14	2.92±0.19
Y723-O2P/R488-NE	2.84±0.11	2.85±0.11

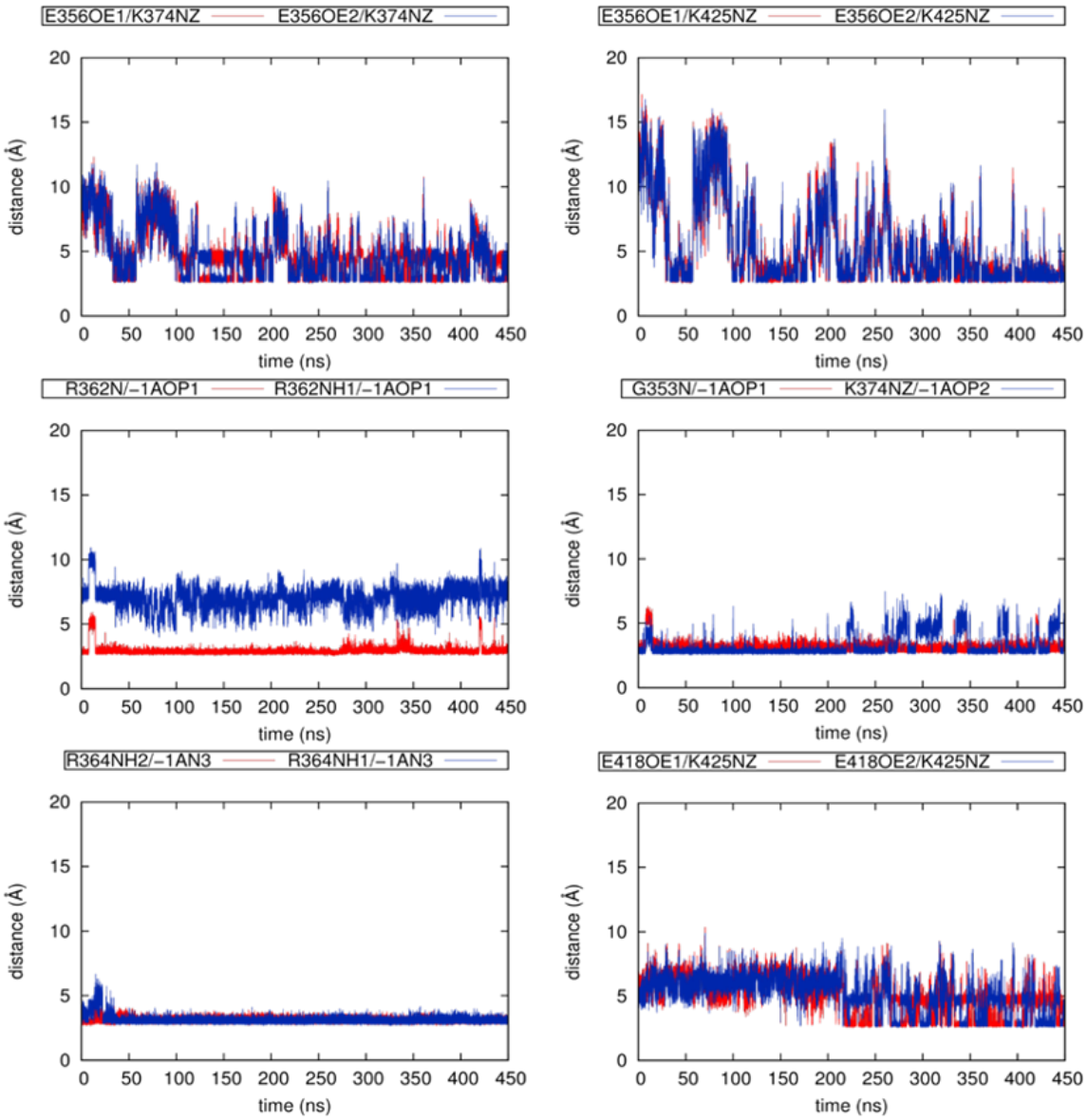

Figure 3.14. Time series for inter-atomic distances around A-ring for main simulation of TPT-open [94].

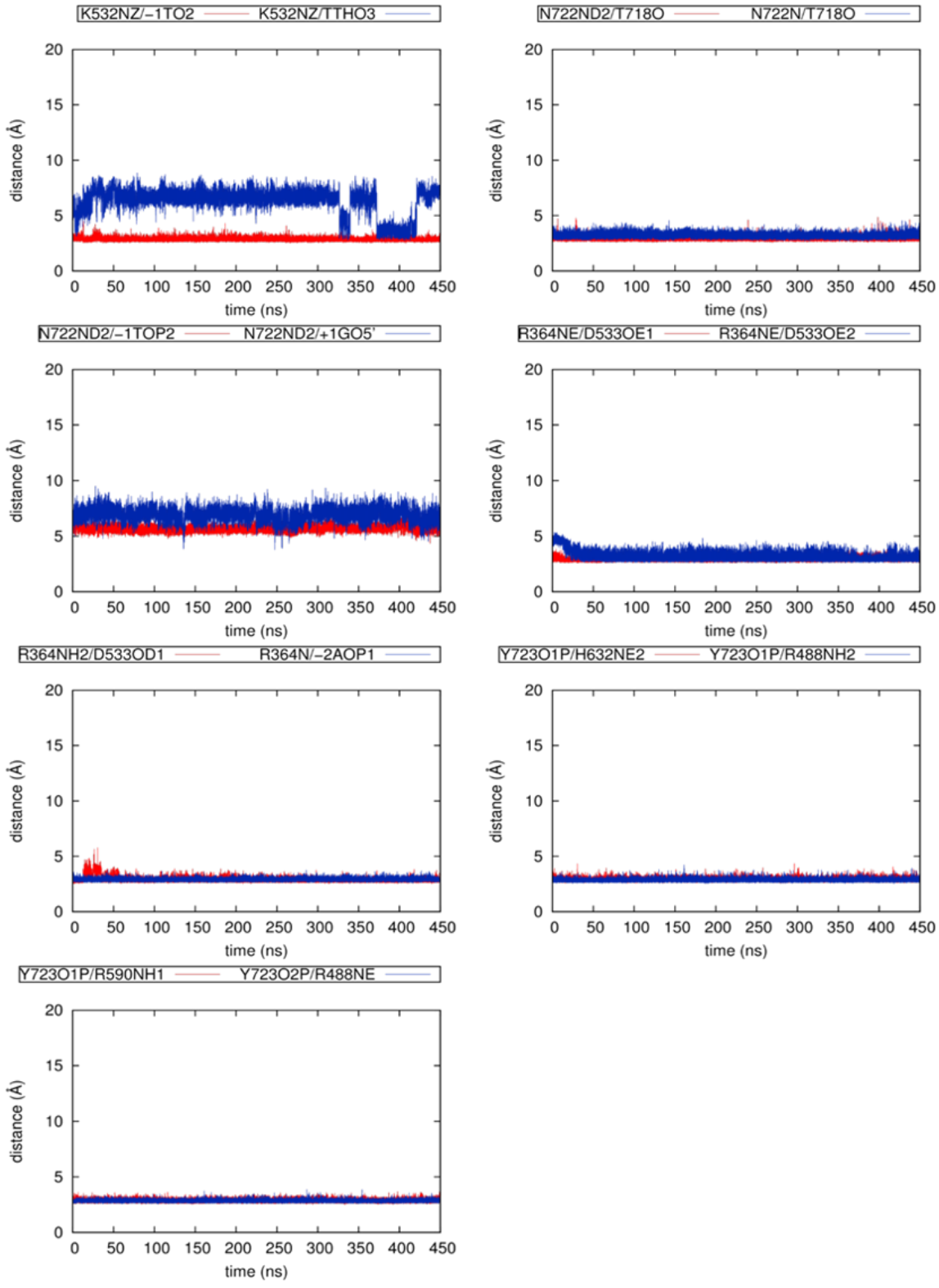

Figure 3.15. Time series for inter-atomic distances around E-ring for main simulation of TPT-open [94].

4. CONCLUSIONS

In this study, different chemical forms of drug Topotecan, i.e. lactone (TPT-closed) and carboxyl (TPT-open), have been investigated to shed a light on their activity difference by performing Molecular Dynamics simulations and calculating the interaction energies. Interaction energy calculations using hybrid QM/MM approach revealed that TPT-closed forms more stable interactions with DNA/TopoI complex than TPT-open. To understand which component is responsible for this difference, total interaction energy has been decomposed to individual terms, and TopoI has been identified as the main contributor to the interaction energy difference between TPT-closed and TPT-open. With a more detailed view, residue K532 is identified as the only residue that has a direct hydrogen bond with TPT, and this interaction is found to be only present in TPT-closed. Moreover, hydrogen bond networks around the binding site helped to identify another important residue, i.e. N722, which forms stable interaction with +1G only in the presence of TPT-closed. In conclusion, two residues, i.e. K532 and N722, have essential roles in inhibiton of TopoI activity. The results of the calculations imply that TPT-closed causes the active site to be more rigid, hence inhibits the religation process.

5. FUTURE WORK

This study clarifies the contributions from the TopoI when TPT is the bound ligand molecule. However, another set of newly synthesized synthetic TopoI inhibitor molecules are currenly under clinical trial, and they belong to the "Indenoisoquinolines (IQN)". IQNs have unique features. They do not have a lactone ring, and this makes them more stable than CPT derivatives. They also have less affinity for blood proteins, and they are more potent. One important feature of IQNs is that they prefer to intercalate between DNA base pairs that CPT does not bind to. Hence, they can target different gene groups/locations on the genome.

As future work, this study can be extended by adding IQN derivatives which are under clinical trial, and by investigating their choice of DNA base pairs to intercalate. Also, again performing the same analysis applied in this thesis, a comparison between the mechanism of action between IQNs and TPT can also be part of the future work.

REFERENCES

1. Watson, J. D. and F. H. C. Crick, "Genetical Implications of the Structure of Deoxyribonucleic Acid", *Nature*, Vol. 171, No. 4361, pp. 964–967, 1953.

2. Watson, J. D. and F. H. C. Crick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid", *Nature*, Vol. 171, No. 4356, pp. 737–738, 1953.

3. Cuya, S. M., M.-A. Bjornsti and R. C. van Waardenburg, "DNA topoisomerase-targeting chemotherapeutics: what's new?", *Cancer Chemotherapy and Pharmacology*, Vol. 80, No. 1, pp. 1–14, 2017.

4. Pommier, Y., Y. Sun, S.-y. N. Huang and J. L. Nitiss, "Roles of eukaryotic topoisomerases in transcription, replication and genomic stability", *Nature Reviews Molecular Cell Biology*, Vol. 17, No. 11, pp. 703–721, 2016.

5. Swuec, P. and A. Costa, "Molecular mechanism of double Holliday junction dissolution", *Cell & Bioscience*, Vol. 4, No. 1, p. 36, 2014.

6. Chaudhury, I., A. Sareen, M. Raghunandan and A. Sobeck, "FANCD2 regulates BLM complex functions independently of FANCI to promote replication fork recovery", *Nucleic Acids Research*, Vol. 41, No. 13, pp. 6444–6459, 2013.

7. Wang, Y., Y. L. Lyu and J. C. Wang, "Dual localization of human DNA topoisomerase IIIalpha to mitochondria and nucleus.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 99, No. 19, pp. 12114–9, 2002.

8. Nott, A. and L.-H. Tsai, "The Top3β way to untangle RNA", *Nature Neuroscience*, Vol. 16, No. 9, pp. 1163–1164, 2013.

9. Liu, L. F., C. C. Liu and B. M. Alberts, "Type II DNA topoisomerases: enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break.", *Cell*, Vol. 19, No. 3, pp. 697–707, 1980.

10. Gellert, M., K. Mizuuchi, M. H. O'Dea and H. A. Nash, "DNA gyrase: an enzyme that introduces superhelical turns into DNA.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 73, No. 11, pp. 3872–6, 1976.

11. Hsieh, T.-s. and D. Brutlag, "ATP-dependent DNA topoisomerase from D. melanogaster reversibly catenates duplex DNA rings", *Cell*, Vol. 21, No. 1, pp. 115–125, 1980.

12. Bergerat, A., B. de Massy, D. Gadelle *et al*, "An atypical topoisomerase II from archaea with implications for meiotic recombination", *Nature*, Vol. 386, No. 6623, pp. 414–417, 1997.

13. Bergerat, A., D. Gadelle and P. Forterre, "Purification of a DNA topoisomerase II from the hyperthermophilic archaeon Sulfolobus shibatae. A thermostable enzyme with both bacterial and eucaryal features.", *The Journal of biological chemistry*, Vol. 269, No. 44, pp. 27663–9, 1994.

14. Corbett, K. D. and J. M. Berger, "Structure of the topoisomerase VI-B subunit: implications for type II topoisomerase mechanism and evolution", *The EMBO Journal*, Vol. 22, No. 1, pp. 151–163, 2003.

15. Nichols, M. D., K. DeAngelis, J. L. Keck and J. M. Berger, "Structure and function of an archaeal topoisomerase VI subunit with homology to the meiotic recombination factor Spo11", *The EMBO Journal*, Vol. 18, No. 21, pp. 6177–6188, 1999.

16. Corbett, K. D., P. Benedetti and J. M. Berger, "Holoenzyme assembly and ATP-mediated conformational dynamics of topoisomerase VI.", *Nature structural &*

molecular biology, Vol. 14, No. 7, pp. 611–9, 2007.

17. Ye, J., C. Lenain, S. Bauwens *et al*, "TRF2 and Apollo Cooperate with Topoisomerase 2α to Protect Human Telomeres from Replicative Damage", *Cell*, Vol. 142, No. 2, pp. 230–242, 2010.

18. Niimi, A., N. Suka, M. Harata *et al*, "Co-localization of chicken DNA topoisomerase IIα, but not β, with sites of DNA replication and possible involvement of a C-terminal region of α through its binding to PCNA", *Chromosoma*, Vol. 110, No. 2, pp. 102–114, 2001.

19. Grue, P., A. Grässer, M. Sehested *et al*, "Essential mitotic functions of DNA topoisomerase IIalpha are not adopted by topoisomerase IIbeta in human H69 cells.", *The Journal of biological chemistry*, Vol. 273, No. 50, pp. 33660–6, 1998.

20. Fortune, J. M. and N. Osheroff, "Topoisomerase II as a target for anticancer drugs: when enzymes stop being nice.", *Progress in nucleic acid research and molecular biology*, Vol. 64, pp. 221–53, 2000.

21. McClendon, A. K. and N. Osheroff, "DNA topoisomerase II, genotoxicity, and cancer.", *Mutation research*, Vol. 623, No. 1-2, pp. 83–97, 2007.

22. Berger, J. M., S. J. Gamblin, S. C. Harrison and J. C. Wang, "Structure and mechanism of DNA topoisomerase II", *Nature*, Vol. 379, No. 6562, pp. 225–232, 1996.

23. Wang, J. C., "Moving one DNA double helix through another by a type II DNA topoisomerase: the story of a simple molecular machine", *Quarterly Reviews of Biophysics*, Vol. 31, No. 2, pp. 107–144, 1998.

24. Zhang, H., L. Meng, D. B. Zimonjic *et al*, "Thirteen-exon-motif signature for vertebrate nuclear and mitochondrial type IB topoisomerases", *Nucleic Acids Research*, Vol. 32, No. 7, pp. 2087–2092, 2004.

25. Koster, D. A., V. Croquette, C. Dekker *et al*, "Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB", *Nature*, Vol. 434, No. 7033, pp. 671–674, 2005.

26. Stewart, L., M. R. Redinbo, X. Qiu *et al*, "A model for the mechanism of human topoisomerase I.", *Science*, Vol. 279, No. 5356, pp. 1534–41, 1998.

27. Champoux, J. J. and R. Dulbecco, "An activity from mammalian cells that untwists superhelical DNA–a possible swivel for DNA replication (polyoma-ethidium bromide-mouse-embryo cells-dye binding assay).", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 69, No. 1, pp. 143–6, 1972.

28. Champoux, J. J., "DNA Topoisomerases: Structure, Function, and Mechanism", *Annual Review of Biochemistry*, Vol. 70, No. 1, pp. 369–413, 2001.

29. Vos, S. M., E. M. Tretter, B. H. Schmidt and J. M. Berger, "All tangled up: how cells direct, manage and exploit topoisomerase function", *Nature Reviews Molecular Cell Biology*, Vol. 12, No. 12, pp. 827–841, 2011.

30. Parker, M. W., M. R. Botchan and J. M. Berger, "Mechanisms and regulation of DNA replication initiation in eukaryotes", *Critical Reviews in Biochemistry and Molecular Biology*, Vol. 52, No. 2, pp. 107–144.

31. Abdurashidova, G., S. Radulescu, O. Sandoval *et al*, "Functional interactions of DNA topoisomerases with a human replication origin", *The EMBO Journal*, Vol. 26, No. 4, pp. 998–1009, 2007.

32. Sugimoto, N., I. Kitabayashi, S. Osano *et al*, "Mechanisms and regulation of DNA replication initiation in eukaryotes", *Molecular Biology of the Cell*, Vol. 52, No. 2, pp. 107–144, 2017.

33. Gonzalez, R. E., C.-U. Lim, K. Cole *et al*, "Effects of conditional depletion of

topoisomerase II on cell cycle progression in mammalian cells", *Cell Cycle*, Vol. 10, No. 20, pp. 3505–3514, 2011.

34. Brill, S. J., S. DiNardo, K. Voelkel-Meiman and R. Sternglanz, "Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA", *Nature*, Vol. 326, No. 6111, pp. 414–416, 1987.

35. Postow, L., N. J. Crisona, B. J. Peter *et al*, "Topological challenges to DNA replication: conformations at the fork.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 98, No. 15, pp. 8219–26, 2001.

36. Baxter, J. and J. F. X. Diffley, "Topoisomerase II inactivation prevents the completion of DNA replication in budding yeast.", *Molecular cell*, Vol. 30, No. 6, pp. 790–802, 2008.

37. Lee, S.-H., G. E.-L. Siaw, S. Willcox *et al*, "Synthesis and dissolution of hemicatenanes by type IA DNA topoisomerases.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 110, No. 38, pp. E3587–94, 2013.

38. Ramamoorthy, M., T. Tadokoro, I. Rybanska *et al*, "RECQL5 cooperates with Topoisomerase II alpha in DNA decatenation and cell cycle progression", *Nucleic Acids Research*, Vol. 40, No. 4, pp. 1621–1635, 2012.

39. Rouzeau, S., F. P. Cordelières, G. Buhagiar-Labarchède *et al*, "Bloom's Syndrome and PICH Helicases Cooperate with Topoisomerase IIα in Centromere Disjunction before Anaphase", *PLoS ONE*, Vol. 7, No. 4, p. e33905, 2012.

40. Dykhuizen, E. C., D. C. Hargreaves, E. L. Miller *et al*, "BAF complexes facilitate decatenation of DNA by topoisomerase IIα", *Nature*, Vol. 497, No. 7451, pp. 624–627, 2013.

41. Husain, A., N. A. Begum, T. Taniguchi *et al*, "Chromatin remodeller SMARCA4 recruits topoisomerase 1 and suppresses transcription-associated genomic insta-

bility", *Nature Communications*, Vol. 7, p. 10549, 2016.

42. Shykind, B. M., J. Kim, L. Stewart *et al*, "Topoisomerase I enhances TFIID-TFIIA complex assembly during activation of transcription.", *Genes & development*, Vol. 11, No. 3, pp. 397–407, 1997.

43. Merino, A., K. R. Madden, W. S. Lane *et al*, "DNA topoisomerase I is involved in both repression and activation of transcription", *Nature*, Vol. 365, No. 6443, pp. 227–232, 1993.

44. Ljungman, M. and P. C. Hanawalt, "The anti-cancer drug camptothecin inhibits elongation but stimulates initiation of RNA polymerase II transcription", *Carcinogenesis*, Vol. 17, No. 1, pp. 31–36, 1996.

45. Solier, S., M. C. Ryan, S. E. Martin *et al*, "Transcription poisoning by Topoisomerase I is controlled by gene length, splice sites, and miR-142-3p.", *Cancer research*, Vol. 73, No. 15, pp. 4830–9, 2013.

46. King, I. F., C. N. Yandava, A. M. Mabb *et al*, "Topoisomerases facilitate transcription of long genes linked to autism", *Nature*, Vol. 501, No. 7465, pp. 58–62, 2013.

47. Mabb, A. M., P. H. M. Kullmann, M. A. Twomey *et al*, "Topoisomerase 1 inhibition reversibly impairs synaptic function.", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 111, No. 48, pp. 17290–5, 2014.

48. Lyu, Y. L., C.-P. Lin, A. M. Azarova *et al*, "Role of topoisomerase IIbeta in the expression of developmentally regulated genes.", *Molecular and cellular biology*, Vol. 26, No. 21, pp. 7929–41, 2006.

49. Drolet, M., "Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology", *Molecular Microbiology*, Vol. 59, No. 3, pp. 723–730, 2006.

50. Yang, Y., K. M. McBride, S. Hensley *et al*, "Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation.", *Molecular cell*, Vol. 53, No. 3, pp. 484–97, 2014.

51. Sollier, J., C. T. Stork, M. L. García-Rubio *et al*, "Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability", *Molecular Cell*, Vol. 56, No. 6, pp. 777–785, 2014.

52. Tuduri, S., L. Crabbé, C. Conti *et al*, "Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription", *Nature Cell Biology*, Vol. 11, No. 11, pp. 1315–1324, 2009.

53. Sordet, O., A. J. Nakamura, C. E. Redon and Y. Pommier, "DNA double-strand breaks and ATM activation by transcription-blocking DNA lesions", *Cell Cycle*, Vol. 9, No. 2, pp. 274–278, 2010.

54. Sordet, O., C. E. Redon, J. Guirouilh-Barbat *et al*, "Ataxia telangiectasia mutated activation by transcription- and topoisomerase I-induced DNA double-strand breaks", *EMBO reports*, Vol. 10, No. 8, pp. 887–893, 2009.

55. Marinello, J., S. Bertoncini, I. Aloisi *et al*, "Dynamic Effects of Topoisomerase I Inhibition on R-Loops and Short Transcripts at Active Promoters", *PLOS ONE*, Vol. 11, No. 1, p. e0147053, 2016.

56. Marinello, J., G. Chillemi, S. Bueno *et al*, "Antisense transcripts enhanced by camptothecin at divergent CpG-island promoters associated with bursts of topoisomerase I-DNA cleavage complex and R-loop formation", *Nucleic Acids Research*, Vol. 41, No. 22, pp. 10110–10123, 2013.

57. Pommier, Y., "Topoisomerase I inhibitors: camptothecins and beyond", *Nature Reviews Cancer*, Vol. 6, No. 10, pp. 789–802, 2006.

58. Delgado, J. L., C.-M. Hsieh, N.-L. Chan and H. Hiasa, "Topoisomerases as anti-

cancer targets.", *The Biochemical journal*, Vol. 475, No. 2, pp. 373–398, 2018.

59. Wall, M. E. and M. C. Wani, "Camptothecin and taxol: discovery to clinic–thirteenth Bruce F. Cain Memorial Award Lecture.", *Cancer research*, Vol. 55, No. 4, pp. 753–60, 1995.

60. Hsiang, Y. H., R. Hertzberg, S. Hecht and L. F. Liu, "Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I.", *The Journal of biological chemistry*, Vol. 260, No. 27, pp. 14873–8, 1985.

61. Wall, M. E., M. C. Wani, C. E. Cook *et al*, "Plant Antitumor Agents. I. The Isolation and Structure of Camptothecin, a Novel Alkaloidal Leukemia and Tumor Inhibitor from Camptotheca acuminata", *Journal of the American Chemical Society*, Vol. 88, No. 16, pp. 3888–3890, 1966.

62. Mi, Z. and T. G. Burke, "Differential interactions of camptothecin lactone and carboxylate forms with human blood components.", *Biochemistry*, Vol. 33, No. 34, pp. 10325–36, 1994.

63. Staker, B. L., K. Hjerrild, M. D. Feese *et al*, "The mechanism of topoisomerase I poisoning by a camptothecin analog", *Proceedings of the National Academy of Sciences*, Vol. 99, No. 24, pp. 15387–15392, 2002.

64. Hertzberg, R. P., M. J. Caranfa and S. M. Hecht, "On the mechanism of topoisomerase I inhibition by camptothecin: evidence for binding to an enzyme-DNA complex", *Biochemistry*, Vol. 28, No. 11, pp. 4629–4638, 1989.

65. Eng, W. K., L. Faucette, R. K. Johnson and R. Sternglanz, "Evidence that DNA topoisomerase I is necessary for the cytotoxic effects of camptothecin.", *Molecular pharmacology*, Vol. 34, No. 6, pp. 755–60, 1988.

66. Thomas, C. J., N. J. Rahier and S. M. Hecht, "Camptothecin: current perspectives", *Bioorganic & Medicinal Chemistry*, Vol. 12, No. 7, pp. 1585–1604, 2004.

67. Garcia-Carbonero, R. and J. G. Supko, "Current perspectives on the clinical experience, pharmacology, and continued development of the camptothecins.", *Clinical cancer research : an official journal of the American Association for Cancer Research*, Vol. 8, No. 3, pp. 641–61, 2002.

68. Jaxel, C., K. W. Kohn, M. C. Wani *et al*, "Structure-activity study of the actions of camptothecin derivatives on mammalian topoisomerase I: evidence for a specific receptor site and a relation to antitumor activity.", *Cancer research*, Vol. 49, No. 6, pp. 1465–9, 1989.

69. Wani, M. C., A. W. Nicholas, G. Manikumar and M. E. Wall, "Plant antitumor agents. 25. Total synthesis and antileukemic activity of ring A substituted camptothecin analogues. Structure-activity correlations.", *Journal of medicinal chemistry*, Vol. 30, No. 10, pp. 1774–9, 1987.

70. Pommier, Y., E. Leo, H. Zhang and C. Marchand, "DNA topoisomerases and their poisoning by anticancer and antibacterial drugs", *Cell Chemical Biology*, Vol. 17, No. 5, pp. 421–433, 2010.

71. Stewart, D. J., "Topotecan in the First-Line Treatment of Small Cell Lung Cancer", *The Oncologist*, Vol. 9, No. suppl_6, pp. 33–42, 2004.

72. Rodriguez-Galindo, C., K. Radomski, C. F. Stewart *et al*, "Clinical use of topoisomerase I inhibitors in anticancer treatment.", *Medical and pediatric oncology*, Vol. 35, No. 4, pp. 385–402, 2000.

73. Minsky, B. D., "Combined-modality therapy of rectal cancer with oxaliplatin-based regimens.", *Clinical colorectal cancer*, Vol. 4 Suppl 1, pp. S29–36, 2004.

74. Bailly, C., "Homocamptothecins: potent topoisomerase I inhibitors and promising anticancer drugs.", *Critical reviews in oncology/hematology*, Vol. 45, No. 1, pp. 91–108, 2003.

75. Holden, J. A., M. E. Wall, M. C. Wani *et al*, "Human DNA Topoisomerase I: Quantitative Analysis of the Effects of Camptothecin Analogs and the Benzophenanthridine Alkaloids Nitidine and 6-Ethoxydihydronitidine on DNA Topoisomerase I-Induced DNA Strand Breakage", *Archives of Biochemistry and Biophysics*, Vol. 370, No. 1, pp. 66–76, 1999.

76. Yang, D., J. Thompson Strode, H. Peter Spielmann *et al*, "DNA interactions of two clinical camptothecin drugs stabilize their active lactone forms", *Journal of the American Chemical Society*, Vol. 120, No. 12, pp. 2979–2980, 1998.

77. Warner, D. L. and T. G. Burke, "Simple and versatile high-performance liquid chromatographic method for the simultaneous quantitation of the lactone and carboxylate forms of camptothecin anticancer drugs.", *Journal of chromatography. B, Biomedical sciences and applications*, Vol. 691, No. 1, pp. 161–71, 1997.

78. Underberg, W. J., R. M. Goossen, B. R. Smith *et al*, "Equilibrium kinetics of the new experimental anti-tumour compound SK&F 104864-A in aqueous solution.", *Journal of pharmaceutical and biomedical analysis*, Vol. 8, No. 8-12, pp. 681–3, 1990.

79. Hsiang, Y. H. and L. F. Liu, "Identification of mammalian DNA topoisomerase I as an intracellular target of the anticancer drug camptothecin.", *Cancer research*, Vol. 48, No. 7, pp. 1722–6, 1988.

80. Redinbo, M. R., L. Stewart, P. Kuhn *et al*, "Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA.", *Science*, Vol. 279, No. 5356, pp. 1504–13, 1998.

81. Yao, S., D. Murali, P. Seetharamulu *et al*, "Topotecan lactone selectively binds to double- and single-stranded DNA in the absence of topoisomerase I.", *Cancer research*, Vol. 58, No. 17, pp. 3782–6, 1998.

82. Gallo, J. M., P. B. Laub, E. K. Rowinsky *et al*, "Population Pharmacokinetic

Model for Topotecan Derived From Phase I Clinical Trials", *Journal of Clinical Oncology*, Vol. 18, No. 12, pp. 2459–2467, 2000.

83. Li, X. G., P. Haluska, Y. H. Hsiang *et al*, "Involvement of amino acids 361 to 364 of human topoisomerase I in camptothecin resistance and enzyme catalysis.", *Biochemical pharmacology*, Vol. 53, No. 7, pp. 1019–27, 1997.

84. Fiorani, P., J. F. Amatruda, A. Silvestri *et al*, "Domain interactions affecting human DNA topoisomerase I catalysis and camptothecin sensitivity.", *Molecular pharmacology*, Vol. 56, No. 6, pp. 1105–15, 1999.

85. Urasaki, Y., G. S. Laco, P. Pourquier *et al*, "Characterization of a novel topoisomerase I mutation from a camptothecin-resistant human prostate cancer cell line.", *Cancer research*, Vol. 61, No. 5, pp. 1964–9, 2001.

86. Chang, J.-Y., J.-F. Liu, S.-H. Juang *et al*, "Novel mutation of topoisomerase I in rendering cells resistant to camptothecin.", *Cancer research*, Vol. 62, No. 13, pp. 3716–21, 2002.

87. Tanizawa, A., R. Beitrand, G. Kohlhagen *et al*, "Cloning of Chinese hamster DNA topoisomerase I cDNA and identification of a single point mutation responsible for camptothecin resistance.", *The Journal of biological chemistry*, Vol. 268, No. 34, pp. 25463–8, 1993.

88. Saleem, A., N. Ibrahim, M. Patel *et al*, "Mechanisms of resistance in a human cell line exposed to sequential topoisomerase poisoning.", *Cancer research*, Vol. 57, No. 22, pp. 5100–6, 1997.

89. Tamura, H., C. Kohchi, R. Yamada *et al*, "Molecular cloning of a cDNA of a camptothecin-resistant human DNA topoisomerase I and identification of mutation sites.", *Nucleic acids research*, Vol. 19, No. 1, pp. 69–75, 1991.

90. Fujimori, A., W. G. Harker, G. Kohlhagen *et al*, "Mutation at the catalytic site

of topoisomerase I in CEM/C2, a human leukemia cell line resistant to camptothecin.", *Cancer research*, Vol. 55, No. 6, pp. 1339–46, 1995.

91. Interthal, H., P. M. Quigley, W. G. J. Hol *et al*, "The Role of Lysine 532 in the Catalytic Mechanism of Human Topoisomerase I", *Journal of Biological Chemistry*, Vol. 279, No. 4, pp. 2984–2992, 2004.

92. Pan, P., Y. Li, H. Yu *et al*, "Molecular Principle of Topotecan Resistance by Topoisomerase I Mutations through Molecular Modeling Approaches", *Journal of Chemical Information and Modeling*, Vol. 53, No. 4, pp. 997–1006, 2013.

93. Mancini, G., I. D'Annessa, A. Coletta *et al*, "Structural and Dynamical Effects Induced by the Anticancer Drug Topotecan on the Human Topoisomerase I – DNA Complex", *PLoS ONE*, Vol. 5, No. 6, p. e10934, 2010.

94. Bali, S. K., A. Marion, I. Ugur *et al*, "Activity of Topotecan toward the DNA/Topoisomerase I Complex: A Theoretical Rationalization", *Biochemistry*, Vol. 57, No. 9, pp. 1542–1551, 2018.

95. Case, D., J. Berryman, R. Betz *et al*, "AMBER 2015", 2015.

96. Verlet, L., "Computer Experiments on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules", *Physical Review*, Vol. 159, No. 1, pp. 98–103, 1967.

97. Genheden, S. and U. Ryde, "Will molecular dynamics simulations of proteins ever reach equilibrium?", *Physical Chemistry Chemical Physics*, Vol. 14, No. 24, p. 8662, 2012.

98. Cerutti, D. S. and D. A. Case, "Multi-Level Ewald: A hybrid multigrid / Fast Fourier Transform approach to the electrostatic particle-mesh problem.", *Journal of chemical theory and computation*, Vol. 6, No. 2, pp. 443–58, 2010.

99. Woodcock, L., "Isothermal molecular dynamics calculations for liquid salts", *Chemical Physics Letters*, Vol. 10, No. 3, pp. 257–261, 1971.

100. Andersen, H. C., "Molecular dynamics simulations at constant pressure and/or temperature", *The Journal of Chemical Physics*, Vol. 72, No. 4, pp. 2384–2393, 1980.

101. Grest and Kremer, "Molecular dynamics simulation for polymers in the presence of a heat bath.", *Physical review. A, General physics*, Vol. 33, No. 5, pp. 3628–3631, 1986.

102. Hohenberg, P. and W. Kohn, "Inhomogeneous Electron Gas", *Physical Review*, Vol. 136, No. 3B, pp. B864–B871, 1964.

103. Hehre, W. J., R. F. Stewart and J. A. Pople, "Self Consistent Molecular Orbital Methods. I. Use of Gaussian Expansions of Slater Type Atomic Orbitals", *The Journal of Chemical Physics*, Vol. 51, No. 6, pp. 2657–2664, 1969.

104. Collins, J. B., P. von R. Schleyer, J. S. Binkley and J. A. Pople, "Self consistent molecular orbital methods. XVII. Geometries and binding energies of second row molecules. A comparison of three basis sets", *The Journal of Chemical Physics*, Vol. 64, No. 12, pp. 5142–5151, 1976.

105. Gordon, M. S., J. S. Binkley, J. A. Pople *et al*, "Self consistent molecular orbital methods. 22. Small split valence basis sets for second row elements", *Journal of the American Chemical Society*, Vol. 104, No. 10, pp. 2797–2803, 1982.

106. Binkley, J. S., J. A. Pople and W. J. Hehre, "Self consistent molecular orbital methods. 21. Small split valence basis sets for first row elements", *Journal of the American Chemical Society*, Vol. 102, No. 3, pp. 939–947, 1980.

107. Ditchfield, R., W. J. Hehre and J. A. Pople, "Self Consistent Molecular Orbital Methods. IX. An Extended Gaussian Type Basis for Molecular Orbital Studies

of Organic Molecules", *The Journal of Chemical Physics*, Vol. 54, No. 2, pp. 724–728, 1971.

108. Hehre, W. J., R. Ditchfield and J. A. Pople, "Self Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules", *The Journal of Chemical Physics*, Vol. 56, No. 5, pp. 2257–2261, 1972.

109. Dobbs, K. D. and W. J. Hehre, "Molecular orbital theory of the properties of inorganic and organometallic compounds 4. Extended basis sets for third and fourth row, main group elements", *Journal of Computational Chemistry*, Vol. 7, No. 3, pp. 359–378, 1986.

110. Warshel, A. and A. Lappicirella, "Calculations of ground and excited state potential surfaces for conjugated heteroatomic molecules", *Journal of the American Chemical Society*, Vol. 103, No. 16, pp. 4664–4673, 1981.

111. Warshel, A. and M. Levitt, "Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme.", *Journal of molecular biology*, Vol. 103, No. 2, pp. 227–49, 1976.

112. Levitt, M. and A. Warshel, "Computer simulation of protein folding", *Nature*, Vol. 253, No. 5494, pp. 694–698, 1975.

113. Bayly, C. I., P. Cieplak, W. Cornell and P. A. Kollman, "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model", *The Journal of Physical Chemistry*, Vol. 97, No. 40, pp. 10269–10280, 1993.

114. M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al*, "Gaussian 09", 2016.

115. Maier, J. A., C. Martinez, K. Kasavajhala *et al*, "ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB", *Journal of*

Chemical Theory and Computation, Vol. 11, No. 8, pp. 3696–3713, 2015.

116. Wang, J., R. M. Wolf, J. W. Caldwell *et al*, "Development and testing of a general amber force field", *Journal of Computational Chemistry*, Vol. 25, No. 9, pp. 1157–1174, 2004.

117. Wang, J., W. Wang, P. A. Kollman and D. A. Case, "Automatic atom type and bond type perception in molecular mechanical calculations", *Journal of Molecular Graphics and Modelling*, Vol. 25, No. 2, pp. 247–260, 2006.

118. Pérez, A., I. Marchán, D. Svozil *et al*, "Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers", *Biophysical Journal*, Vol. 92, No. 11, pp. 3817–3829, 2007.

119. Wu, X. and B. R. Brooks, "Self-guided Langevin dynamics simulation method", *Chemical Physics Letters*, Vol. 381, No. 3-4, pp. 512–518, 2003.

120. Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren *et al*, "Molecular dynamics with coupling to an external bath", *The Journal of Chemical Physics*, Vol. 81, No. 8, pp. 3684–3690, 1984.

121. Götz, A. W., M. J. Williamson, D. Xu *et al*, "Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born", *Journal of Chemical Theory and Computation*, Vol. 8, No. 5, pp. 1542–1555, 2012.

122. Salomon-Ferrer, R., A. W. Götz, D. Poole *et al*, "Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald", *Journal of Chemical Theory and Computation*, Vol. 9, No. 9, pp. 3878–3888, 2013.

123. Le Grand, S., A. W. Götz and R. C. Walker, "SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations", *Computer Physics Communications*, Vol. 184, No. 2, pp. 374–380, 2013.

124. Roe, D. R. and T. E. Cheatham, "PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data", *Journal of Chemical Theory and Computation*, Vol. 9, No. 7, pp. 3084–3095, 2013.

125. Zhao, Y. and D. G. Truhlar, "The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other function", *Theoretical Chemistry Accounts*, Vol. 120, No. 1-3, pp. 215–241, 2008.

126. Götz, A. W., M. A. Clark and R. C. Walker, "An extensible interface for QM/MM molecular dynamics simulations with AMBER", *Journal of Computational Chemistry*, Vol. 35, No. 2, pp. 95–108, 2014.

APPENDIX A: RMSD AND DISTANCE PLOTS FOR DUPLICATE SIMULATIONS

RMSD plots for duplicate simulations can be found in Figure A.1 and Figure A.2 The time series for the inter-atomic distances of duplicate simulations can be found in Figures A3-A6.

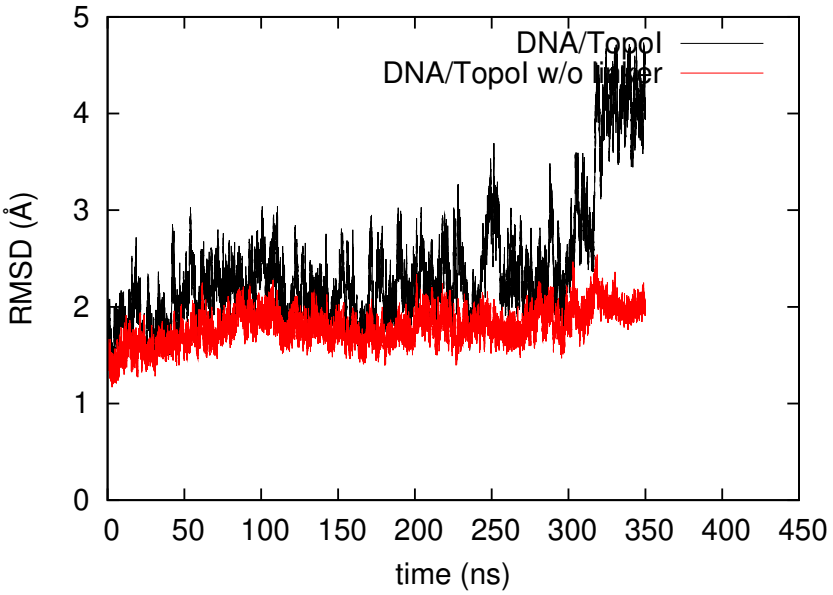

Figure A.1. RMS deviations in Å for the duplicate simulation TPT-closed complex. Black line. RMSD of backbone atoms of DNA/TopoI complex. Red line. RMSD of backbone atoms of DNA/TopoI without the linker domain [94].

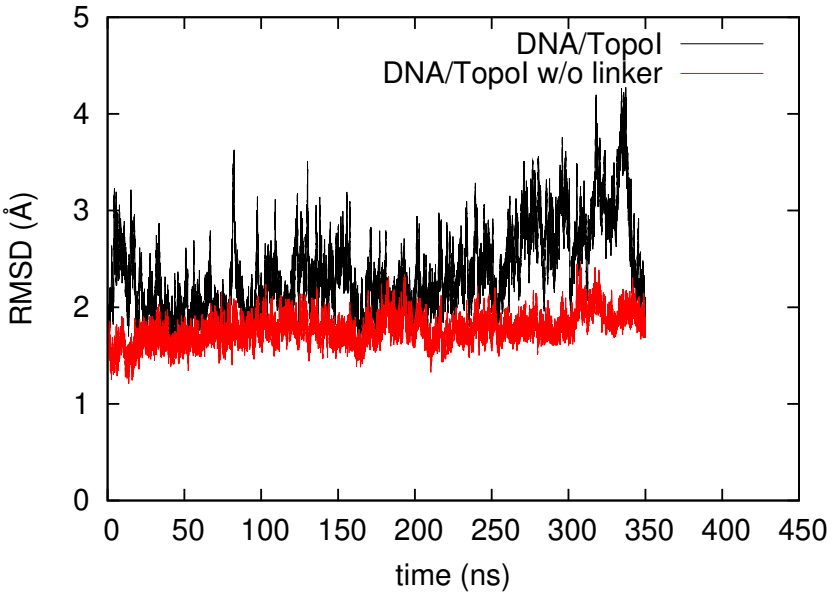

Figure A.2. RMS deviations in Å for the duplicate simulation TPT-open complex. Black line: RMSD of backbone atoms of DNA/TopoI complex. Red line: RMSD of backbone atoms of DNA/TopoI without the linker domain [94].

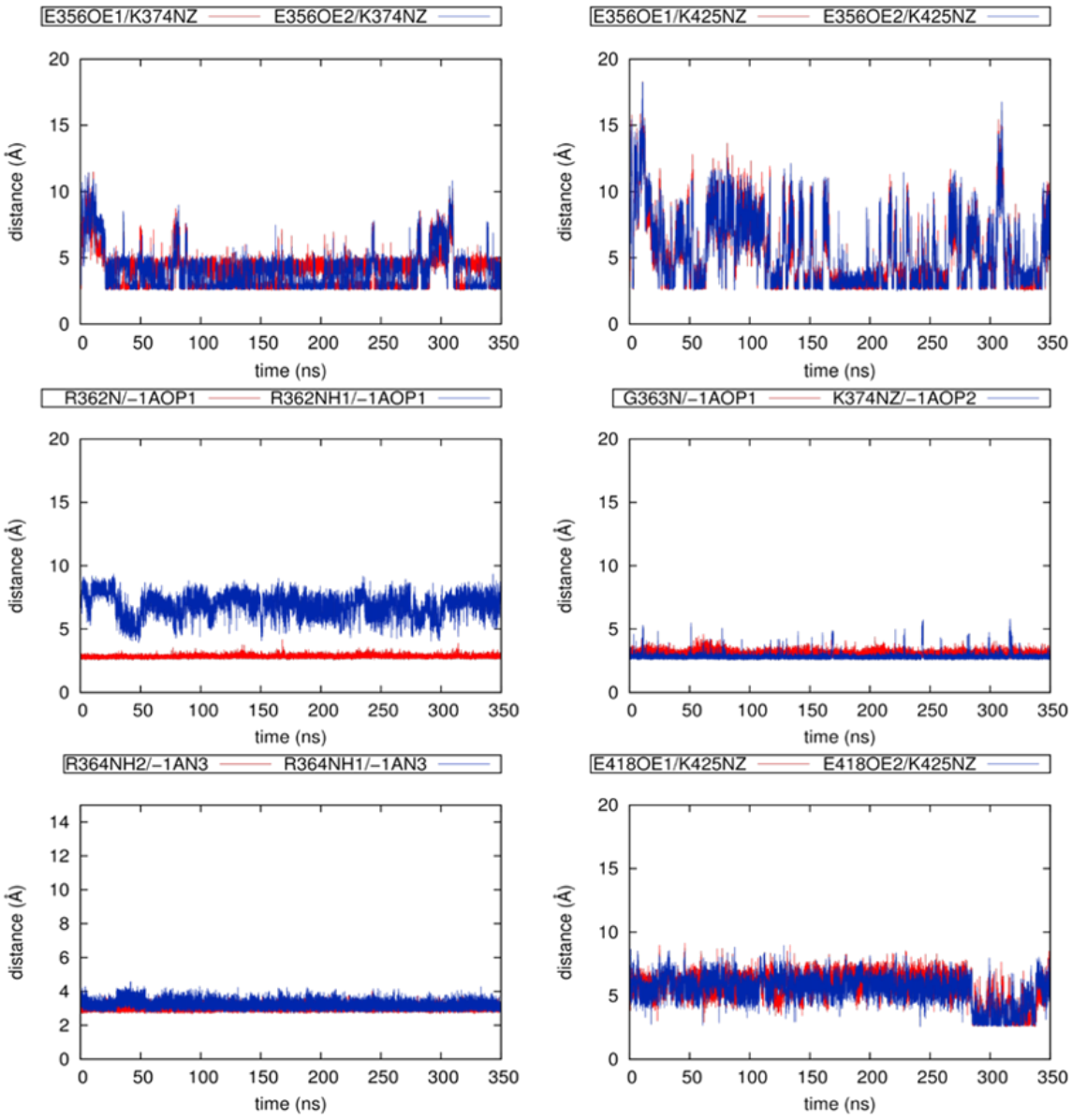

Figure A.3. Time series for inter-atomic distances around A-ring for duplicate simulation of TPT-closed [94].

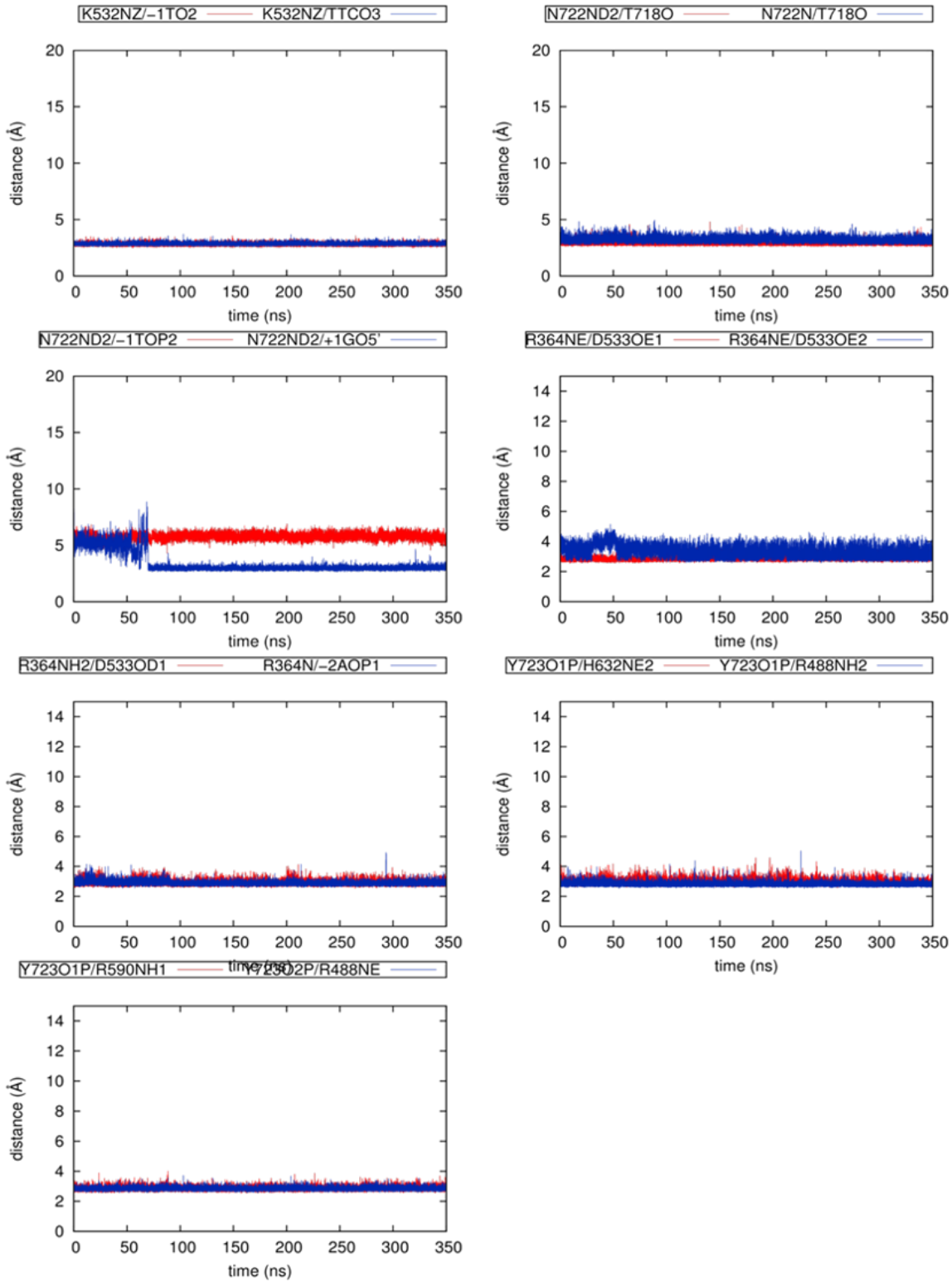

Figure A.4. Time series for inter-atomic distances around E-ring for duplicate simulation of TPT-closed [94].

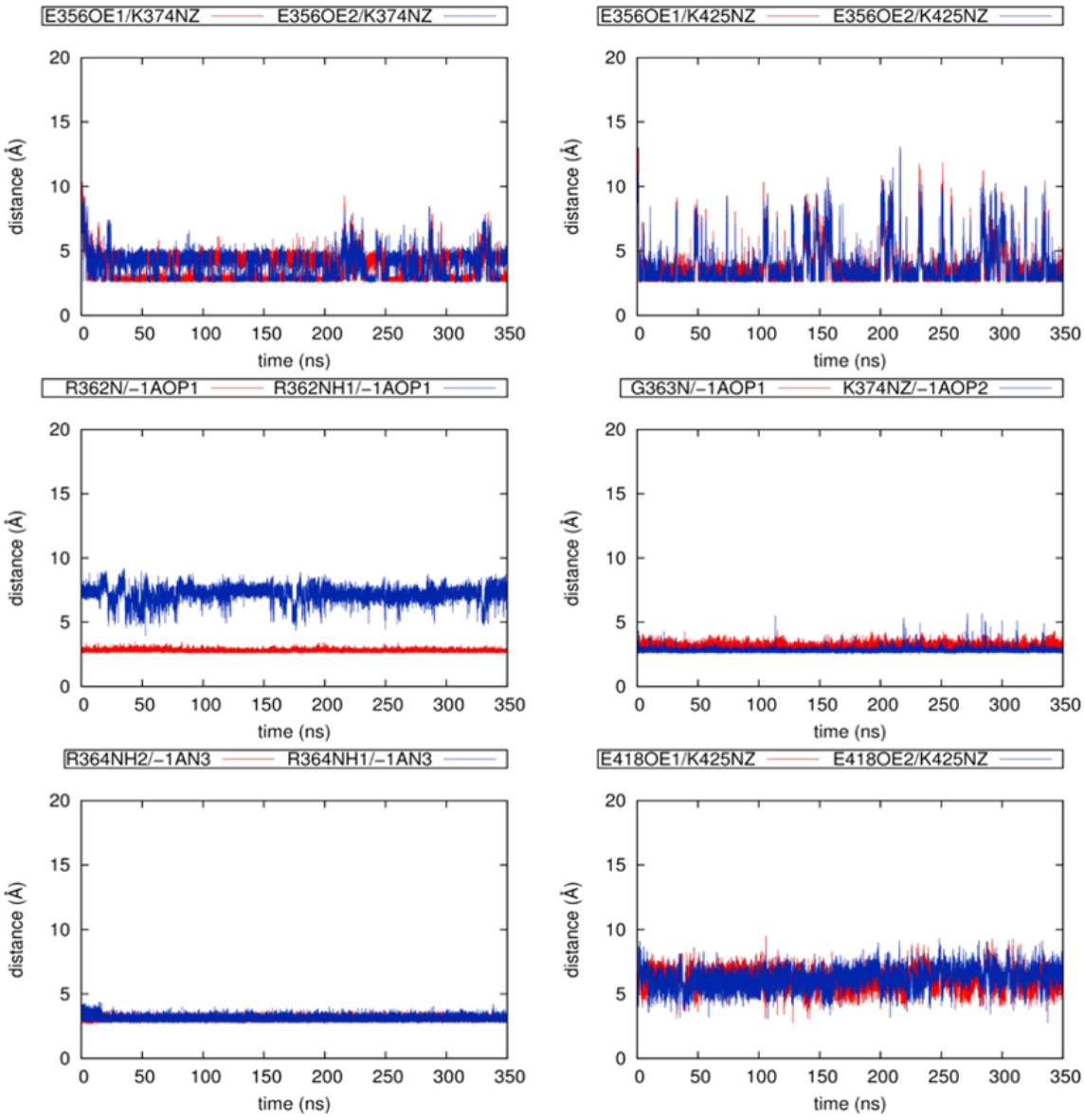

Figure A.5. Time series for inter-atomic distances around A-ring for duplicate simulation of TPT-open [94].

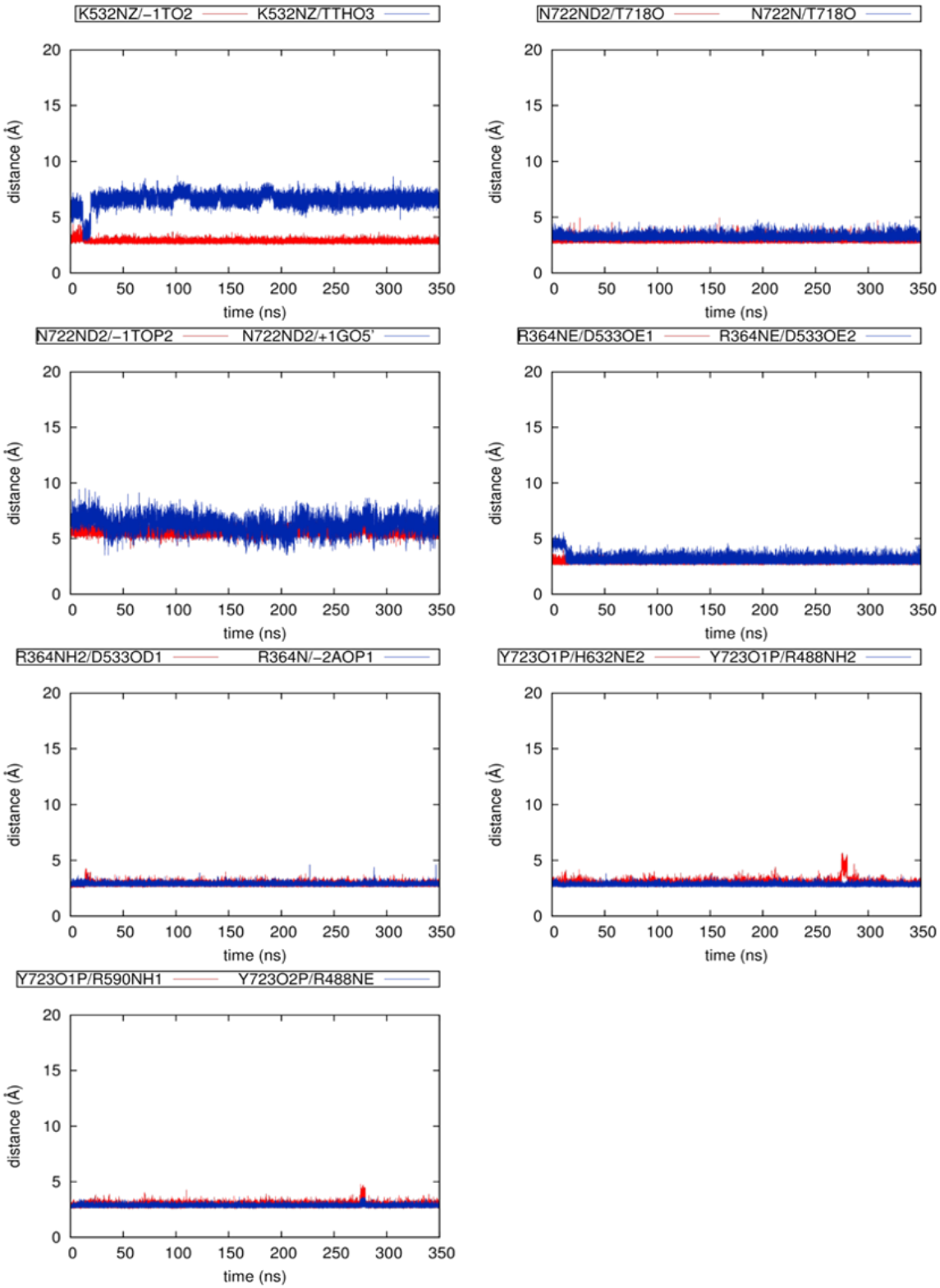

Figure A.6. Time series for inter-atomic distances around E-ring for duplicate simulation of TPT-open [94].