



T.R.
EGE UNIVERSITY
Graduate School of Applied and Natural Science



**COMPUTATIONAL INVESTIGATION OF
INTERACTION OF SEVERAL
ANTIINFLAMMATORY AND ANTI-HYPERTENSIVE
DRUG MOLECULES WITH INDIVIDUAL DNA
NUCLEOBASES AND THEIR BASE PAIRS BY
SEVERAL DFT AND QM/MM MD METHODS**

PhD Thesis

Işıl ÖZTÜRK

Department of Chemistry

İzmir
2022

T.R.
EGE UNIVERSITY
Graduate School of Applied and Natural Science

**COMPUTATIONAL INVESTIGATION OF
INTERACTION OF SEVERAL
ANTIINFLAMMATORY AND ANTI-HYPERTENSIVE
DRUG MOLECULES WITH INDIVIDUAL DNA
NUCLEOBASES AND THEIR BASE PAIRS BY
SEVERAL DFT AND QM/MM MD METHODS**

İşıl ÖZTÜRK

Supervisor : Prof. Dr. Armağan KINAL
Co-supervisor: Prof. Dr. Toomas TAMM

Department of Chemistry
Physical Chemistry Third Cycle Programme

İzmir
2022



Işılay ÖZTÜRK tarafından doktora tezi olarak sunulan “**Computational investigation of interaction of several antiinflammatory and anti-hypertensive drug molecules with individual DNA nucleobases and their base pairs by several DFT and QM/MM MD methods (Çeşitli iltihap önleyici ve anti-hipertansif ilaç moleküllerinin DNA nükleobazları ve bu nükleobazların baz çiftleri ile olan etkileşimlerinin çeşitli DFT ve QM/MM MD yöntemleri ile incelenmesi)**” başlıklı bu çalışma EÜ Lisansüstü Eğitim ve Öğretim Yönetmeliği ile EÜ Fen Bilimleri Enstitüsü Eğitim ve Öğretim Yönergesi’nin ilgili hükümleri uyarınca tarafımızdan değerlendirilerek savunmaya değer bulunmuş ve 18/08/2022 tarihinde yapılan tez savunma sınavında aday oybirliği ile başarılı bulunmuştur.

Jüri Üyeleri:**İmza**

Jüri Başkanı	: Prof. Dr. Armağan KINAL
Raportör Üye	: Prof. Dr. Nursel Acar SELÇUKİ
Üye	: Prof. Dr. Ali ÇAĞIR
	: Doç. Dr. Kamil ŞİRİN
	: Doç. Dr. Burcu OKUTUCU



EGE ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

ETİK KURALLARA UYGUNLUK BEYANI

EÜ Lisansüstü Eğitim ve Öğretim Yönetmeliğinin ilgili hükümleri uyarınca Doktora Tezi olarak sunduğum **“Computational investigation of interaction of several antiinflammatory and anti-hypertensive drug molecules with individual DNA nucleobases and their base pairs by several DFT and QM/MM MD methods (Çeşitli iltihap önleyici ve anti-hipertansif ilaç moleküllerinin DNA nükleobazları ve bu nükleobazların baz çiftleri ile olan etkileşimlerinin çeşitli DFT ve QM/MM MD yöntemleri ile incelenmesi)”** başlıklı bu tezin kendi çalışmam olduğunu, sunduğum tüm sonuç, doküman, bilgi ve belgeleri bizzat ve bu tez çalışması kapsamında elde ettiğimi, bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara atıf yaptığımı ve bunları kaynaklar listesinde usulüne uygun olarak verdiğimi, tez çalışması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını, bu tezin herhangi bir bölümünü bu üniversite veya diğer bir üniversitede başka bir tez çalışması içinde sunmadığımı, bu tezin planlanmasından yazımına kadar bütün safhalarda bilimsel etik kurallarına uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul edeceğimi beyan ederim.

18/ 08/2022

İmzası

Işıl ÖZTÜRK



ÖZET**ÇEŞİTLİ İLTİHAP ÖNLEYİCİ VE ANTİ-HİPERTANSİF İLAÇ
MOLEKÜLLERİNİN DNA NÜKLEOBAZLARI VE BU
NÜKLEOBAZLARIN BAZ ÇİFTLERİ İLE OLAN
ETKİLEŞİMLERİNİN ÇEŞİTLİ DFT VE QM/MM MD
YÖNTEMLERİ İLE İNCELENMESİ**

ÖZTÜRK, Işıl

Doktora tezi, Kimya Anabilim Dalı

Tez Danışmanı: Prof. Dr. Armağan KINAL

Yardımcı Danışman: Prof. Dr. Toomas TAMM

18 Ağustos 2022, 141 sayfa

Bu tez, seçilmiş çeşitli anti-inflamatuar ve anti-hipertansif ilaç moleküllerinin DNA nükleobazları ve baz çiftleri ile etkileşimlerini DFT ve QM/MM MD hesaplamalı kimya yöntemlerini kullanarak araştırmaktadır. İlk olarak, CCSD(T) sonuçlarıyla yapılan karşılaştırma, ω B97 ailesinin (ω B97X-D, ω B97M-V ve ω B97X-V) ilaç-DNA etkileşimleri için en iyi yöntemler olduğunu ortaya koydu. ω B97X-D hesaplamaları, ilaç-tek baz kompleksleri arasında en negatif bağlanma enerjisinin celiprolol ve guanin arasında olduğunu göstermiştir. Bununla birlikte, tek başına etkileşim enerjisinin değerlendirilmesi, ilaç-nükleobaz etkileşimini anlamak için yetersizdir. İlaç moleküllerinin hidrojen bağları üzerindeki etkileri açı ve mesafe açısından incelendiğinde, DFT hesaplamaları mefenamik asit ve oksaprozinin hem A-T hem de G-C baz çiftleri ile etkili bir şekilde etkileşen iki ilaç olduğuna işaret etmektedir. İlaçlar ve nükleobazlar için moleküler elektrostatik potansiyel (MEP) haritaları belirlenmiş ve bunların donör-alıcı ilişkisini açıklamada içgörü sağladıkları ve potansiyel etkileşimlere ışık tuttuğu gösterilmiştir. Ayrıca mefenamik asit ve oksaprozinin sulu ortamda DNA ile etkileşimleri üzerinde QM/MM MD hesaplamaları yapıldı. Model kuantum hesaplamaları (DFT), mefenamik asidin, QM/MM MD hesaplamalarında olduğu gibi, karboksilik grup aracılığıyla nükleobaz çiftleri ile etkileşim sağladığını ortaya çıkardı. Bu nedenle, burada bağlı olmayan etkileşimler için kullanılan DFT yöntemlerinin oldukça güvenilir olduğu söylenebilir.

Anahtar sözcükler: NSAIDs ilaçları, beta bloker ilaçlar, DNA-nükleobaz ve baz çiftleri, kovalent olmayan etkileşimler, QM/MM MD simülasyonları, DFT.



ABSTRACT**COMPUTATIONAL INVESTIGATION OF INTERACTION OF SEVERAL ANTIINFLAMMATORY AND ANTI-HYPERTENSIVE DRUG MOLECULES WITH INDIVIDUAL DNA NUCLEOBASES AND THEIR BASE PAIRS BY SEVERAL DFT AND QM/MM MD METHODS**

ÖZTÜRK, Işıl

PhD in Chemistry.

Supervisor: Prof. Dr. Armağan KINAL

Co-supervisor: Toomas TAMM

18 August 2022, 141 pages

This thesis investigates the interactions of various selected anti-inflammatory and anti-hypertensive drug molecules with DNA nucleobases and their base pairs using the DFT and QM/MM MD computational chemistry methods. Firstly, the comparison with the CCSD(T) results revealed that the ω B97 family (ω B97X-D, ω B97M-V, and ω B97X-V) are the best methods for drug-DNA interactions. The ω B97X-D calculations indicated that the most negative binding energy among the drug-single base complexes is between celiprolol and guanine. However, evaluation of interaction energy alone is insufficient to understand drug-nucleobase interaction. When the effects of drug molecules on hydrogen bonds were examined in terms of angle and distance the DFT calculations point out that mefenamic acid and oxaprozin are the two drugs interacting effectively with both A-T and G-C base pairs. The molecular electrostatic potential (MEP) maps were determined for the drugs and nucleobases and it has been shown that they provide insight into explaining the donor-acceptor relationship and shed light on potential interactions. In addition, QM/MM MD calculations were performed on interactions of mefenamic acid and oxaprozin with DNA in the aqueous medium. The model quantum calculations (DFT) revealed that mefenamic acid gives interaction with nucleobase pairs through the carboxylic group, as did the QM/MM MD calculations. Hence, the DFT methods employed here for non-bonded interactions can be said to be rather reliable.

Keywords: DFT, non-covalent interactions, NSAIDs, beta blockers, DNA-nucleobase; binding energies, docking, QM/MM MD simulation model.



PREFACE

In this thesis, I present the computational investigation of the interaction of several anti-inflammatory and anti-hypertensive drug molecules with individual DNA nucleobases and their base pairs by quantum chemical methods. Due to my special interest in computational chemistry, all the results of this thesis include theoretical research. Non-steroidal anti-inflammatory drugs and blood pressure medications are widely used today. It has many side effects in long-term use. Therefore, we investigated their interactions with DNA-nucleobase and base pairs. It is aimed to find their interaction sites, since it is not known how interaction sites behave in DNA. Therefore, the interactions of some non-steroidal anti-inflammatory drugs and beta-blocker drugs with DNA nucleobase and base pairs were compared using computational chemistry methods. The interaction of the detected drug in DNA was examined in MD simulation with the QM/MM method. This thesis aims to provide information to individuals who will be working on these issues.

İZMİR

18/08/2022

Işıl ÖZTÜRK



*to Nuran Nişli ÖZTÜRK, my lovely mum
and
to Ahmet Zeki ÖZTÜRK, my lovely dad*



CONTENTS

	<u>Page</u>
ÖZET.....	VII
ABSTRACT	IX
PREFACE	XI
CONTENTS	XV
LIST OF FIGURES.....	XIX
LIST OF TABLES	XXV
SYMBOLS AND ABBREVIATIONS	XXVIII
1. INTRODUCTION.....	1
1.1 Drug.....	3
1.2 Drug Discovery	4
1.2.1 Chemical Modification	5
1.2.2 Rational Drug Design	5
1.3 Deoxyribonucleic Acid (DNA)	6
1.4 DNA Drug Interaction.....	7
1.5 Driving Forces Involved in Drug-DNA Interactions	8
1.6 Non-Bonded Interactions	8
1.7 Aromatic Molecules	9
1.8 Electrostatic Potential MAPS.....	9
1.9 Drugs Used in The Thesis Project.....	10
1.9.1 Antihypertensive Drugs.....	11
1.9.2 Beta Bloklers	11
1.9.2 Non-Steroidal Anti-Inflammatory Drugs (NSAIDs).....	12
1.9.3 Mefenamic Acid Toxic Effects	12

CONTENTS (Continued)

	<u>Page</u>
1.10 The Main Goal of The Thesis Project	13
1.11 The Significance of The Thesis Project.....	13
2. COMPUTATIONAL METHODS.....	15
2.1 Computational Chemistry Methods.....	15
2.2 Schrödinger's Equation	16
2.3 Molecular Mechanics	18
2.4 Molecular Dynamics.....	19
2.5 Quantum Chemical Methods.....	20
2.5.1 Ab Initio Methods.....	20
2.5.2 Semi-Empirical Methods.....	20
2.5.3 Density Functional Theory (DFT).....	21
2.6 Basis Sets.....	22
2.6.1 Slater Type Orbital	22
2.6.2 Gaussian Type Orbital	23
2.6.3 Pople Style Basis Sets.....	23
2.7 Polarizable Continuum Model (PCM).....	24
2.8 Couple Cluster (CC) Theory.....	24
2.9 QM/MM Methods.....	26
2.9.1 Briefly QM, MM and QM/MM methods	28
2.10 The Computational Chemistry Suites Employed	28
3. RESULTS AND DISCUSSION.....	30
3.1 Geometry Optimization of All Drugs, Nucleobases and Base Pairs	30
3.2 Preliminary Study	33
3.2.1 Binding Enthalpies of Drugs to Nucleobases	33

CONTENTS (Continued)

	<u>Page</u>
3.2.2 Drug-Nucleobase-Pair Binding Enthalpies	38
3.4 The Performance of The Methods for The Non-Covalently Interacted Systems.....	44
3.5 Computational Investigation Drug-Nucleobase and Nucleobase Pair Interaction with ω b97x-D Method.....	56
3.5.1 Drug – Nucleobase Binding Energies.....	56
3.5.2 Drug-Nucleobase-Pair Binding Energies	64
3.5.3 H-Bonds Distance of Drug–Nucleobase Pair	72
3.5.4 H-Bonds Angles of Drug–Nucleobase Pair	77
3.6 Computational Investigation Drug-Nucleobase and Drug-Nucleobase Pair Interaction with ω b97M-V And ω b97X-V Methods.....	81
3.6.1 Binding Energies of Drugs to Nucleobases	81
3.6.2 Binding Energies of Drugs – Nucleobase pairs.....	83
3.6.3 Molecular Electrostatic Potentials (Meps) Of Drug-Nucleobase Pair	86
3.7 Introduction to QM/MM Part	94
3.7.1 What is CP2K?	94
3.7.2 Which Levels of Theory Used?	94
3.7.3 How Did We Define the QM, MM Regions and The Border?	95
3.8 MD Simulation Of Mefenamic Acid-DNA Interaction with QM/MM Method.....	95
3.8.1 The Process of Setting Up a Simulation System Containing DNA in Complex with Mefenamic Acid	98
3.8.2. Energy Minimization (EM).....	100
3.8.3 NVT Equilibration	101
3.8.4 NPT Equilibration.....	103
3.8.5 Production.....	104

CONTENTS (Continued)

	<u>Page</u>
3.8.6 Total 100 ps of Production Part.....	111
3.8.7 Experimental Researchs About MEF.....	111
3.9 MEF Interaction with A-T Nucleobase Pair in DNA with MD Simulation of QM/MM Method	112
3.9.1 The Process of Setting up A Simulation System of MEF Interaction with A-T Nucleobase Pair In DNA	113
3.9.2 Energy Minimization (EM)	113
3.9.3 NVT Equilibration.....	114
3.9.4 NPT Equilibration.....	115
3.9.5 Production.....	116
3.10 DNA-Oxaprozin Drug Simulation with PM3/MM Methods (QM/MM).....	117
3.11 MD Simulation Of OXA-DNA Interaction with QM/MM Method.....	118
3.11.1 The Process of Setting Up A Simulation System Of OXA Interaction with DNA	120
3.11.2 Energy Minimization (EM)	120
3.11.3 NVT Equilibration.....	122
3.11.4 NPT Equilibration.....	123
3.11.5 Production.....	124
4. CONCLUSION.....	125
ACKNOWLEDGMENT	128
ACKNOWLEDGMENT (CONTINUED)	129
REFERENCES	130
CURRICULUM VITAE.....	138

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
Figure 1. 1 Conformers designed for Indole-Guanine complexes (Toupkanloo and Rahmani, 2018).....	2
Figure 1. 2 Cost of developing a Drug	3
Figure 1. 3 Schematic of drug discovery.....	4
Figure 1. 4 Structure of chemical modification.....	5
Figure 1. 5 Types of rational drug design	5
Figure 1. 6 Showing of general information about DNA.....	6
Figure 1. 7 Showing of information about DNA Drug interaction schematically	7
Figure 1. 8 Showing of schematical driving forces involved in Drug-DNA interactions (Intechopen, 2019), (Longdom,2019).....	8
Figure 1. 9 Showing of schematical electrostatic potential MAPS.....	9
Figure 1. 10 Showing the properties of drugs used.....	10
Figure 1. 11 Classification of antihypertensive drug molecules	11
Figure 1. 12 Used of beta blockers drugs in this thesi	11
Figure 1. 13 Used of NSAIDs drugs in this thesis	12
Figure 2. 1 Schematical computational chemistry methods.....	16
Figure 2. 2 Energy partitioning in the molecular mechanics methods.....	18
Figure 2. 3 Approximations of exchange-correlation sheme	21
Figure 2. 4 Showing of QM, MM and QM/MM methods schematically	28
Figure 2. 5 Showing of the general protocol we will use for the calculation.. ..	29
Figure 3. 1 Stable geometries of Non-steroidal anti-inflammatory drugs molecules.....	30
Figure 3. 2 Stable geometries of Non-steroidal anti-inflammatory drugs molecules.....	31
Figure 3. 3 Stable geometries of anti-hypertensive/ Beta- Blockers drugs molecules.....	31

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure 3. 4 Stable geometries of Anti-hypertensive/ Beta- Blockers drugs molecules	32
Figure 3. 5 Stable geometries of nucleobase pairs structures	32
Figure 3. 6 Highly interacting drug-base structures	34
Figure 3. 7 ω B97X-D optimized structures of drug–nucleobase complexes.....	35
Figure 3. 8 Highly interacting drug-base structures	36
Figure 3. 9 ω B97X-D/ 6-31+G (d) and ω B97X-D/ 6-311++G(d,p) level optimized structures of drug–nucleobase complexes	37
Figure 3. 10 Orientations of drug–nucleobase pair complexes optimized by ω B97X-D (T: T shaped)	38
Figure 3. 11 General showing that orientations of drug–nucleobase pair complexes.....	38
Figure 3. 12 The highly interacting drug-base pair structures.....	40
Figure 3. 13 ω B97X-D optimized structures of drug–nucleobase pair complexes	40
Figure 3. 14 Optimized structures of drug–nucleobase pair complexes.....	43
Figure 3. 15 Intermolecular interaction of drug–nucleobase pair complexes	44
Figure 3. 16 Structures and definition of edges of the aromatic amino acids (histidine (HIS), phenylalanine (PHE), tyrosine (TYR), tryptophan (TRP)) and the natural nucleobases (cytosine (C), adenine (A), guanine (G), thymine (T)	45
Figure 3. 17 Binding energy (kcal/mol) values rank of the top strong 20 drug–nucleobase complexes optimized with the IEF-PCM/ ω B97X-D /6-31+G(d,p) methods.....	58
Figure 3. 18 Structures of DNA nucleobases	58

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure 3. 19 The strongest binding energy of celiprolol, acebutolol, and oxaprozin – nucleobase complexes with the IEF-PCM/ ω B97X-D /6-31+G(d,p) methods	59
Figure 3. 20 Crystal structure of oxaprozin.....	61
Figure 3. 21 Oxaprozin- nucleobase complexes optimized with the IEF-PCM/ ω B97X-D /6-31+G(d,p) method	62
Figure 3. 22 Mefenamic acid- nucleobase complexes optimized with the IEF-PCM/ ω B97X-D /6-31+G(d,p) method	63
Figure 3. 23 Symbolic of parallel stacking and T-stacking interactions between all drugs and nucleobase pairs using.....	66
Figure 3. 24 Orientations of celiprolol–nucleobase pair (A-T) complexes optimized by ω B97X-D /6-31+G(d,p)	66
Figure 3. 25 Orientations of drug molecules- nucleobase pairs ((A-T), (G-C)) complexes optimized by ω B97X-D /6-31+G(d,p)	67
Figure 3. 26 Orientations of mefenamic acid–nucleobase pairs ((A-T), (G-C)) complexes optimized by ω B97X-D /6-31+G(d,p)	68
Figure 3. 27 Orientations of 3_at_mefenamic acid (nucleobase pairs ((A-T)) complexes optimized by ω B97X-D /6-31+G(d,p)	69
Figure 3. 28 Orientations of oxaprozin –nucleobase pairs ((A-T), (G-C)) complexes optimized by ω B97X-D /6-31+G(d,p).....	70
Figure 3. 29 Intermolecular interaction of drug–nucleobase pair complexes.....	73
Figure 3. 30 Intermolecular interaction of drug–nucleobase pair complexes.....	74
Figure 3. 31 Intermolecular interaction of drug–nucleobase pair complexes.....	76
Figure 3. 32 H-bonds angles changes of drug– A-T nucleobase pair complexes.....	78
Figure 3. 33 H-bonds angles changes of drug– A-T nucleobase pair complexes.....	79

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure 3. 34 H-bonds angles changes of drug– G-C nucleobase pair complexes	81
Figure 3. 35 Electronic energy values in kcal/mol of drug-nucleobase calculated with ω B97M-V/def2-TZVPP, ω B97X-V/def2-TZVPP and ω B97X-D/6-31+G(d,p).....	82
Figure 3. 36 Electronic energy values in kcal/mol of drug-nucleobase calculated with ω B97X-V, ω B97M-V, and ω B97X-D	83
Figure 3. 37 Electronic energy values in kcal/mol of drug-nucleobase pairs calculated with ω B97X-V, ω B97M-V, and ω B97X-D	84
Figure 3. 38 Electronic energy values in kcal/mol of drug-nucleobase pair are calculated with ω B97X-V, ω B97M-V, and ω B97X-D	85
Figure 3. 39 Molecular electrostatic potential maps of A-T nucleobase pair and drug oxaprozin	87
Figure 3. 40 Interaction of Oxaprozin and A-T nucleobase pair result.	88
Figure 3. 41 Molecular electrostatic potential maps of A-T nucleobase pair and diflunisal drug molecule	88
Figure 3. 42 Interaction of diflunisal and A-T nucleobase pair result	89
Figure 3. 43 Molecular electrostatic potential maps of A-T nucleobase pair and drug atenolol	89
Figure 3. 44 Interaction of atenolol and A-T nucleobase pair result .	90
Figure 3. 45 Molecular electrostatic potential maps of A-T nucleobase pair and drug celecoxib.....	90
Figure 3. 46 Molecular electrostatic potential maps of A-T nucleobase pair and drug celecoxib.....	91
Figure 3. 47 Molecular electrostatic potential maps of A-T nucleobase pair and drug nepafenac	91
Figure 3. 48 Molecular electrostatic potential maps of A-T nucleobase pair and drug nepafenac	92

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure 3. 49 Molecular electrostatic potential maps of G-C nucleobase pair and drug mefenamic acid	92
Figure 3. 50 Molecular electrostatic potential maps of A-T nucleobase pair and drug mefenamic acid	93
Figure 3. 51 Orientations of mefenamic acid–nucleobase pairs ((A-T), (G-C)) complexes optimized by ω B97X-D /6-31+G(d,p)	96
Figure 3. 52 MD simulation steps of mefenamic acid- DNA interactions with Gromacs+CP2K interface of QM/MM method.....	97
Figure 3. 53 Haddock docking result of the MEF with DNA	97
Figure 3. 54 Total system of MEF-DNA	99
Figure 3. 55 Energy minimization of geometry result	100
Figure 3. 56 Potential energy from EM.....	101
Figure 3. 57 Temperature from NVT equilibrium	102
Figure 3. 58 Pressure from NPT equilibrium.....	103
Figure 3. 59 Density from NPT equilibrium	104
Figure 3. 60 Graphs of the RMSD relative to the structure of crystal DNA	105
Figure 3. 61 Graphs of the radius of gyration to the structure present in the minimized, equilibrated system.....	106
Figure 3. 62 Total hydrogen bonds numbers of MEF-DNA.....	107
Figure 3. 63 MEF-DNA interactions at the end of production simulation with QM/MM method (Discovery studio was used for drug-DNA interactions) (distance unit is ångstrom).....	108
Figure 3. 64 Finding the interaction energy of the highly interacting regions on the DNA chain with ω B97X-D /6-31+G(d,p).....	109
Figure 3. 65 Comparing of QM model result and MD simulation with QM/MM result	110
Figure 3. 66 Haddock docking result	112
Figure 3. 67 QM model of MEF interaction with A-T nucleobase pair and their H-bonds distance and angles changes	112

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure 3. 68 Potential energy from EM	113
Figure 3. 69 Temperature from NVT equilibrium.....	114
Figure 3. 70 Pressure from NPT equilibrium	115
Figure 3. 71 MEF-DNA interactions at the end of production simulation with QM/MM method (Discovery studio was used for drug-DNA interactions) (distance unit is ångstrom)	116
Figure 3. 72 Schematic of MD simulation steps of oxaprozin–DNA interactions with QM/MM method using the Amber and CP2K.....	117
Figure 3. 73 DNA -oxaprozin drug simulation with PM3/MM methods (QM/MM)	118
Figure 3. 74 Haddock docking result of OXA interaction with DNA... ..	119
Figure 3. 75 Figure QM model of OXA interaction with A-T nucleobase pair and their H-bonds distance and angles changes	119
Figure 3. 76 Energy minimization of geometry result.....	120
Figure 3. 77 Potential energy from EM	121
Figure 3. 78 Temperature from NVT equilibrium.....	122
Figure 3. 79 Pressure from NPT equilibrium	123
Figure 3. 80 OXA-DNA interactions at the end of production simulation with QM/MM method (Discovery studio was used for drug-DNA interactions) (distance unit is ångstrom)	124

LIST OF TABLES

<u>Tables</u>	<u>Page</u>
Table 3. 1 Binding enthalpy (kcal/mol) values of the drug–nucleobase complexes optimized with the B3LYP, B3LYP-D2 and ω B97X-D methods	33
Table 3. 2 Binding enthalpy (kcal/mol) values of the drug–nucleobase complexes optimized with the B3LYP, B3LYP-D2 and ω B97X-D methods	35
Table 3. 3 ΔH_{bind} (kcal/mol) values of B3LYP and ω B97X-D optimized drug–nucleobase pair complexes	39
Table 3. 4 ΔH_{bind} (kcal/mol) values of ω B97X-D and B3LYP, B3LYP-D2 optimized the parallel-oriented drug–nucleobase pair complexes	41
Table 3. 5 ΔH_{bind} (kcal/mol) values of ω B97X-D and B3LYP, B3LYP-D2 optimized the parallel stacking -oriented drug–nucleobase pair complexes	42
Table 3. 6 Distances between hydrogen bonds of ω B97X-D optimized drug–G-C basepair complexes	43
Table 3. 7 Distances between hydrogen bonds of ω B97X-D optimized drug– A-T base-pair complexes	43
Table 3. 8 B3LYP, MP2/6-31G*(0.25) and CCSD(T)/CBS Parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases	46
Table 3. 9 B3LYP-D2, MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases	47
Table 3. 10 ω B97X-D, MP2/6-31G*(0.25) and CCSD(T)/CBS Parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases	48
Table 3.11 ω B97, MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases	49

LIST OF TABLES (Continued)

<u>Tables</u>	<u>Page</u>
Table 3.12 B97-D, MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases	50
Table 3. 13 ω B97X-D ω B97M-V, ω B97X-V, MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases	51
Table 3. 14 B3LYP/6-31+G(d,p), B3LYP-D2/6-31+G(d,p), ω B97X-D/6-31+G(d,p), MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases	53
Table 3. 15 B3LYP/6-311++G(d,p), B3LYP-D2/6-311++G(d,p), ω B97X-D/6-311++G(d,p), ω B97M-V/def2-TZVPP, ω B97X-V/def2-TZVPP, MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases	55
Table 3. 16 Binding energies (kcal/mol) for drug-nucleobase complexes with ω B97X-D /6-31+G(d,p) levels	57
Table 3. 17 Experimental and optimized geometrical parameters of oxaprozin [bond length (Å)]	60
Table 3. 18 ΔE_{bind} (kcal/mol) values of IEF-PCM/ ω B97X-D /6-31+G(d,p) optimized drug–nucleobase pair complexes.....	65
Table 3. 19 Estimated Dug effect on H-bonding energy in kcal/mol for adenine–thymine (A–T) and guanine-cytosine (C–G) DNA base pairs with mefenamic acid.....	69
Table 3. 20 Estimated Drug effect on H-bonding energy in kcal/mol for adenine–thymine (A–T) and guanine-cytosine (C–G) DNA base pairs with oxaprozin	70
Table 3. 21 H-Bonds Distances (Ångstrom) of the ω B97X-D optimized drug– A-T base pair complexes.....	72

LIST OF TABLES (Continued)

<u>Tables</u>	<u>Page</u>
Table 3. 22 Distances between H-bonds (Ångstrom) of ω B97X-D optimized drug– G-C basepair complexes	74
Table 3. 23 Differences in angles of H-bonds (degrees) of ω B97X-D optimized drug-A-T base pair complexes	77
Table 3. 24 Differences Between H-bonds Angles (degrees) of ω B97X-D Optimized Drug– G-T Basepair Complexes	80
Table 3. 25 Comparison of interaction values in large system with previous quantum model study	109

SYMBOLS AND ABBREVIATIONS

<u>Symbols</u>	<u>Explanations</u>
ξ	Exponent of the Slater type orbital (STO)
$\hbar$	Planck constant
χ_i	i^{th} Atomic Orbital
φ_λ	One-electron wave function
ψ	Wave function
c_{ij}	Coefficient of linear combination
E_o	Ground state energy
f_l and E_{xc}	Functional of electron density
H_o	Hamiltonian operator of one electron
$I(i)$	Hartree Operator
$J(i)$	Fock Operator
M_a	Mass of nucleus a
R_{ab}	Distance between the a^{th} and b^{th} nucleus
r_{ia}	Distance between the i^{th} and nucleus a
r_{ij}	Distance between the i^{th} and j^{th} electron

SYMBOLS AND ABBREVIATIONS (Continued)

<u>Symbols</u>	<u>Explanations</u>
S	Overlap matrix
T_e	Kinetic energy operator of the electrons
T_n	Kinetic energy operator of the nuclei
U	Electron – Electron coulomb interaction Unperturbed energy
V_{ee}	Repulsion energy operator of the electron-electron
V_{ne}	Electron – nucleus attraction energy operator
V_{nn}	Repulsion energy operator of the nucleus-nucleus
$W_1, W_2, \text{etc.}$	First order, second order, and etc. order corrections to the Unperturbed energy
Y_{lm}	Angular shape of wave function
Z_a	Atomic number of nucleus a
Z_j	Nucleus charge of the j th atom
α	Exponent of the Gaussian Type Orbital (GTO)
ϵ_λ	Lagrange multiplier
ρ	Molecular density

SYMBOLS AND ABBREVIATIONS (Continued)

<u>Symbols</u>	<u>Explanations</u>
Ψ_1, Ψ_2 , etc.	Wave functions
B3LYP	Becke's Three-parameter Hybrid Methods with Exchange Functional of Lee et al.
ω B97X-D	The latest functional from Head-Gordon and coworkers, which includes empirical dispersion
PBE	Hybrid Functional of Perdew, Burke and Ernserhof
DFT	Density Functional Theory
GTO's	Gaussian Type Orbitals.
MM	Molecular Mechanics
MO	Molecular Orbital
MD	Molecular Dynamics
PCM	Polarized Continuum Model
PES	Potential Energy Surface
QM	Quantum Mechanics
STO's	Slater Type Orbitals
MEF	Mefenamic acid

SYMBOLS AND ABBREVIATIONS (Continued)

<u>Symbols</u>	<u>Explanations</u>
OXA	Oxaprozin
AT	Adenine-Thymine nucleobase pair
GC	Guanine-Cytosine nucleobase pair



1. INTRODUCTION

In recent years, the rapid development of computer technologies, the increase in their capacities, the innovations in computational chemistry/physics, and the use of these approaches have opened up the possibility of finding answers to many questions in chemistry and materials science. These approaches have led to new research areas, not only by revealing interactions between molecules but also by reducing the cost of experimental research. Drug-DNA interactions are one of these areas. Drug-DNA interactions are important in pharmacology and play a critical role in determining drug mechanisms. There are non-covalent interactions between these two. Non-covalent interactions are important for biological systems (Toupkanloo and Rahmani, 2018). They include many interactions of different natures, such as hydrogen bonding, CH π - and π - π -interactions, dipole-dipole interactions and others. The binding of drug molecules and nucleobases is usually a van der Waals interaction. Non-covalent DNA interactants can alter DNA conformation. DNA consists of four nucleotides, adenine (A), thymine (T), guanine (G) and cytosine (C), and the combination of genetic information. Therefore, the interactions between drugs and nucleobases play an important role in determining the side effects (Abbas, 2017), (Kennard, 1993). In this thesis examines non-steroidal anti-inflammatory drugs (NSAIDs) and the β -blockers of antihypertensive drugs that are widely used in our daily lives. Non-steroidal anti-inflammatory drugs (NSAIDs) are an important class of agents used to relieve inflammation and pain associated with illness or injury. However, NSAIDs taken long term have various side effects such as kidney failure, gastrointestinal problems, colon, lung and breast cancer (Azam et al., 2018). β -Blockers are a class of blood pressure lowering drugs. Beta-blockers are mainly used to treat high blood pressure (hypertension), to treat cardiac arrhythmias and to reduce the risk of cardiac complications after a heart attack. However, the β -blocker toxicity of the degradation products/intermediates of the individual compounds is still unknown (Armakovic, Šetrajcic, and Šetrajcic, 2012). In recent years, studies have been conducted to understand how drugs interact with DNA/RNA nucleobases. In 2003, Baik, M. H. and colleagues (Baik, Friesner, and Lippard, 2003) studied how two possible hydrolysis products of cisplatin, the nucleobase guanine and adenine, bind. Their results show that guanine is the preferred reactant for platinisation. Platinisation of guanine is both

thermodynamically and kinetically more advantageous. In 2018, Toupkanloo and Rahmani (Toupkanloo and Rahmani, 2018) investigated the interaction between bicyclic fragments and nucleobases that may have genotoxic effects. They focused on π - π - parallel stacking interactions.

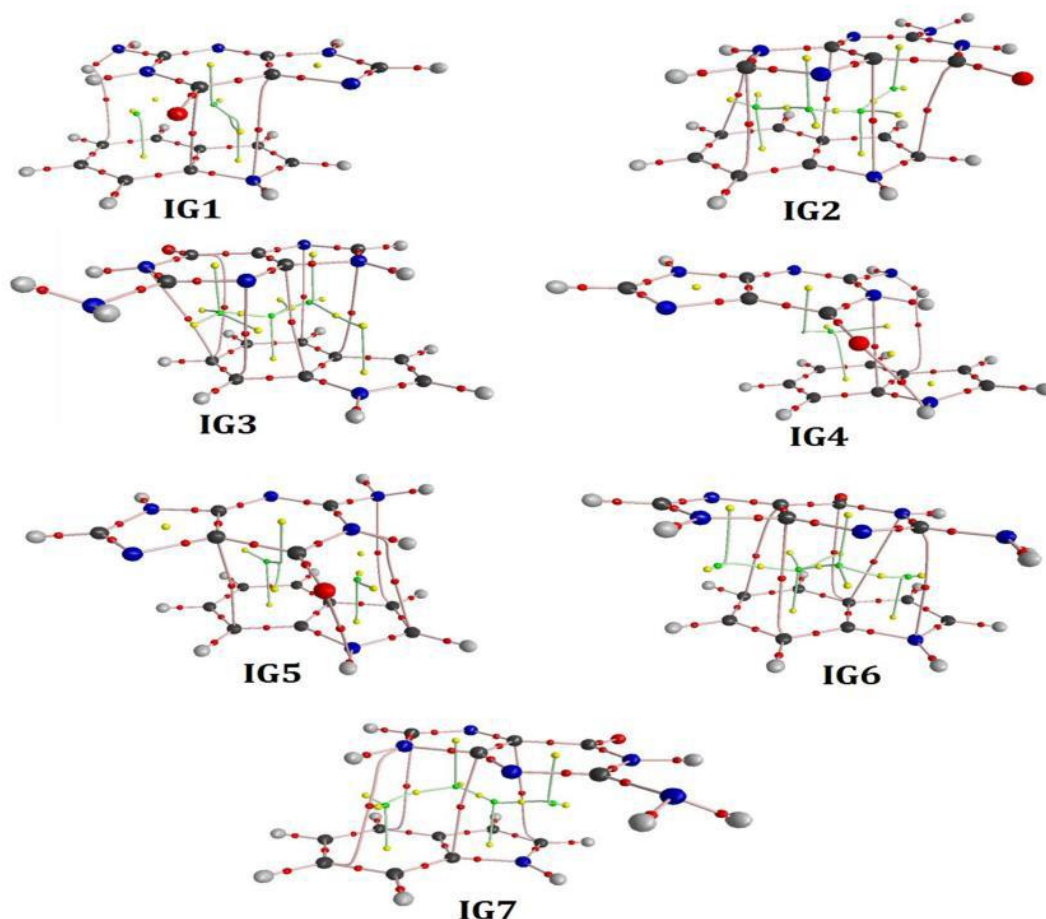


Figure 1. 1 Conformers designed for Indole-Guanine complexes (Toupkanloo and Rahmani, 2018)

Their results show that parallel stacking interactions lead to an increase and decrease in the length of hydrogen bonds involved in the interactions between nucleobase and drug fragment. There are computational studies in 1993 and 2017 studies indicating that interactions between drugs and nucleobases play an important role in determining side effects. Although there are experimental studies on the cancer-promoting effects of NSAIDs as well as other fatal side effects of these drugs, the interaction sites of NSAIDs and beta-blockers with DNA

nucleobases and base pairs are unknown, and there is no common computational study on this.

This section contains subheadings, relevant information about our drug molecules and non-bonded interactions.

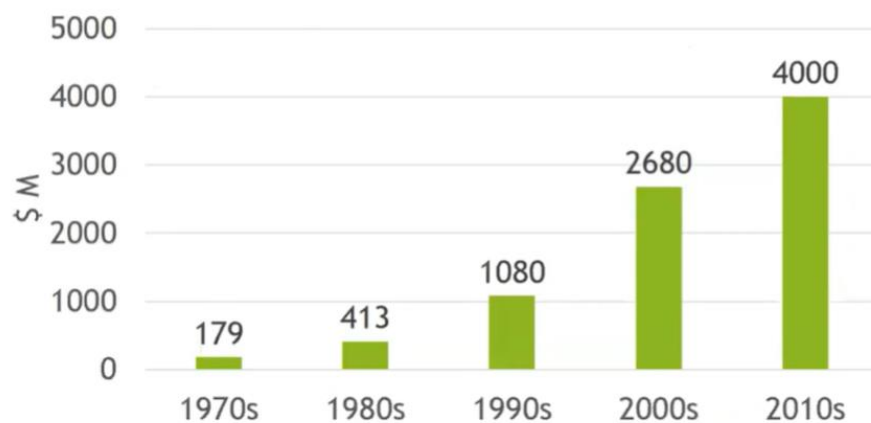
1.1 Drug

Drug cure that affects biological process and used to prevent , diagnose or treat disease.

Requirements for the ideal drug:

- Specific action
- Safe to use, no side effects or less side effects
- Chemically and metabolically stable
- Soluble in water and lipids

Cost of Developing a Drug



Source: Statista 2021 September

Figure 1. 2 Cost of developing a Drug

The cost of developing a drug increase exponentially it doubles every ten years. The cost of a new drug for ten years is estimated average in 4 billion dollars. As we can see understand from figure 1.2 that drugs are expensive products and has expensive procedure.

1.2 Drug Discovery

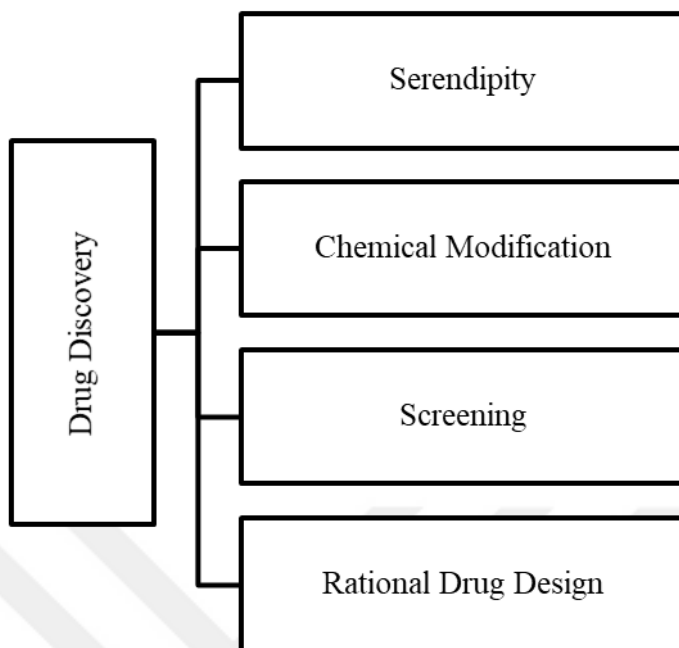


Figure 1. 3 Schematic of drug discovery

Drug discovery in the past with serendipity, generally. It means discovery by chance or trial. Other one is that biochemical modification of known drugs. Other one is that screening from data bases. Last and the most advanced method is that rational drug design. This is the smartest and cheapest way.

1.2.1 Chemical Modification

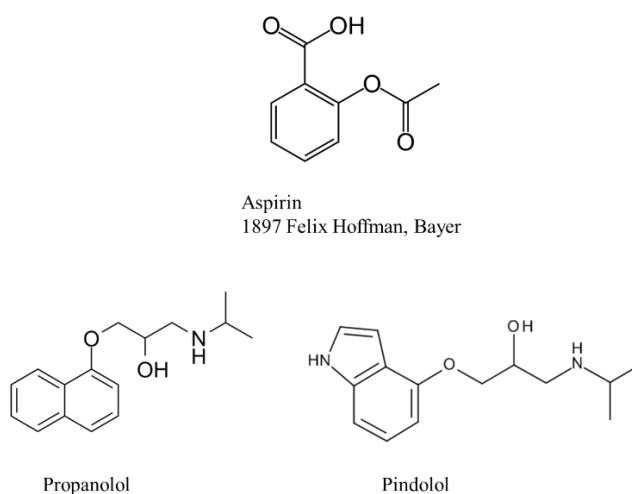


Figure 1. 4 Structure of chemical modification

Chemical modification was discovered aspirin is the acetyl stereotcyclic acid. It was circulated to the salicylic acid which is the natural product. The aim of this modification was to increase the stability and reduce irritating effects of salicylic acid. Modification is the still cheapest way to design drug.

1.2.2 Rational Drug Design

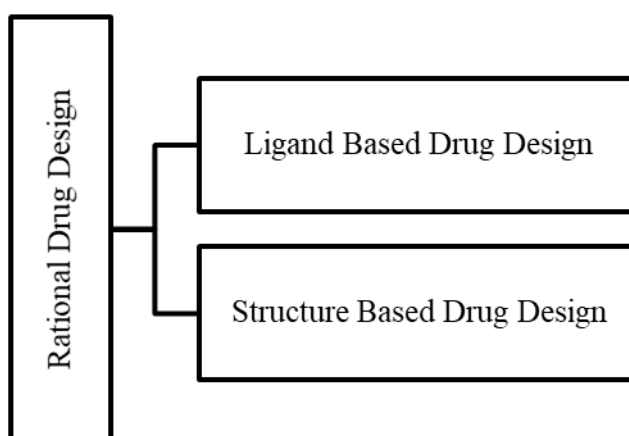


Figure 1. 5 Types of rational drug design

Drug design when the structure of the target molecule is unknown the structure of the ligand is designed and optimized based on the relationship between structure and activity. This is ligand-based drug design. Nevertheless, if the macromolecule

is known then ligand is design to be complementary to the binding site on the macromolecule (Wasko, Pellegrine, Madura, and Surratt, 2015).

1.3 Deoxyribonucleic Acid (DNA)

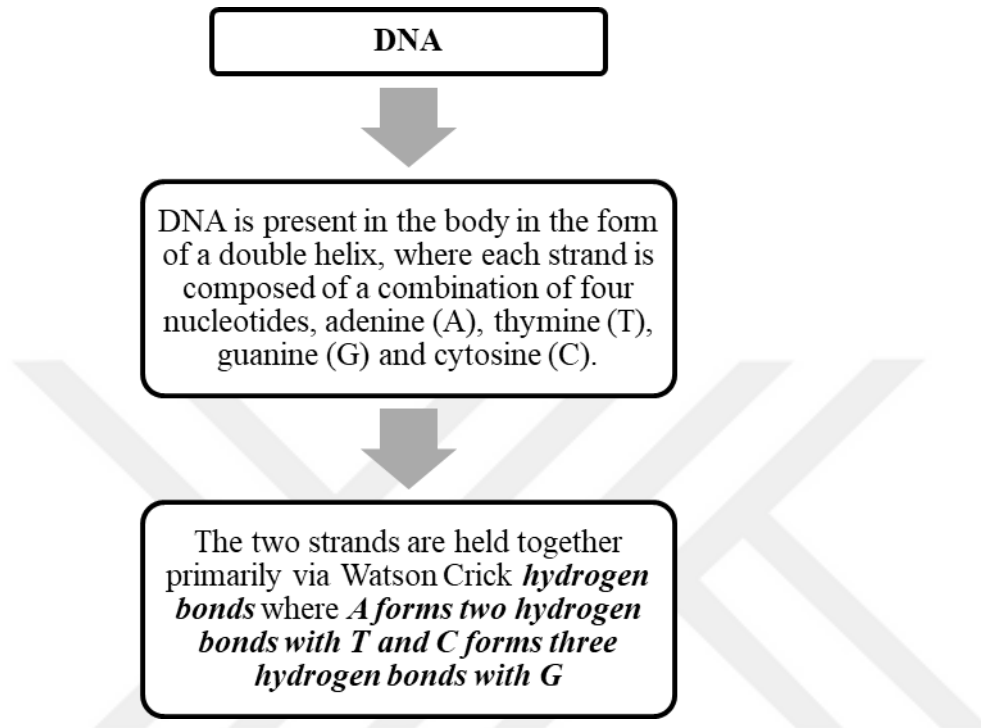


Figure 1. 6 Showing of general information about DNA (Kennard, 1993)

The main nucleobases of DNA are summarized in the figure above.

1.4 DNA Drug Interaction

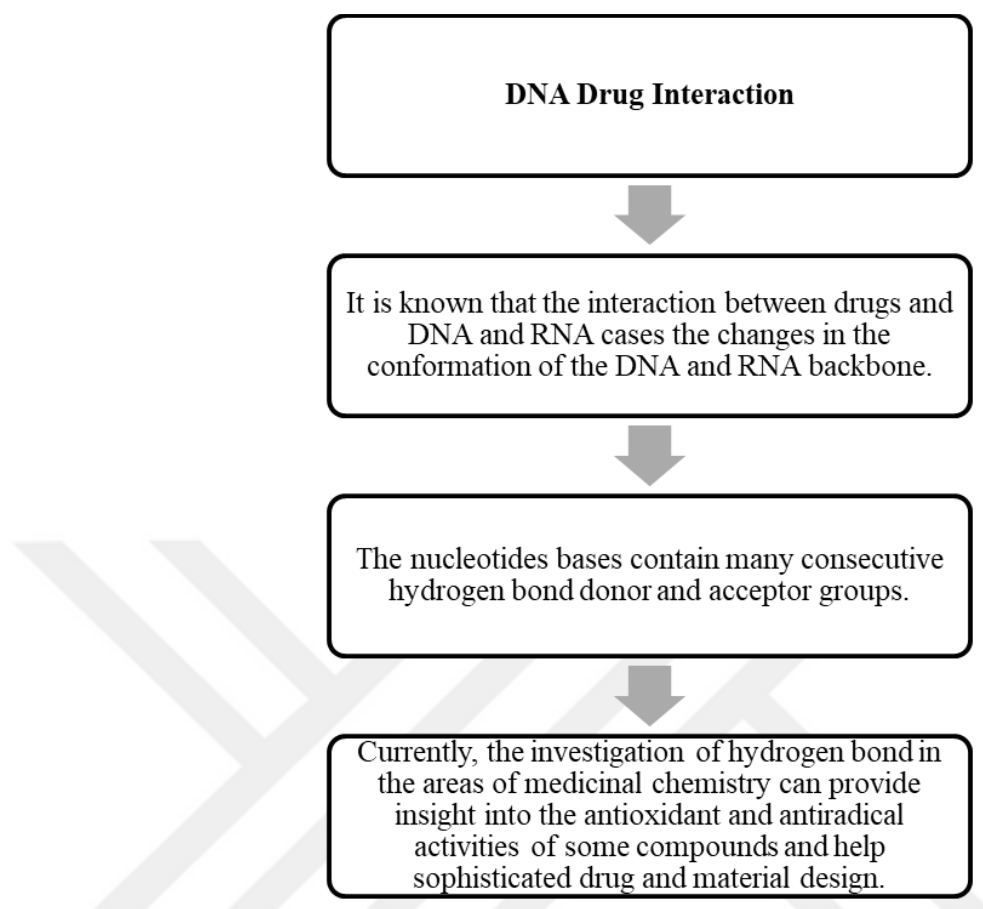


Figure 1. 7 Showing of information about DNA Drug interaction schematically

DNA – Drug interaction are summarized in the Figure 1.7.

1.5 Driving Forces Involved in Drug-DNA Interactions

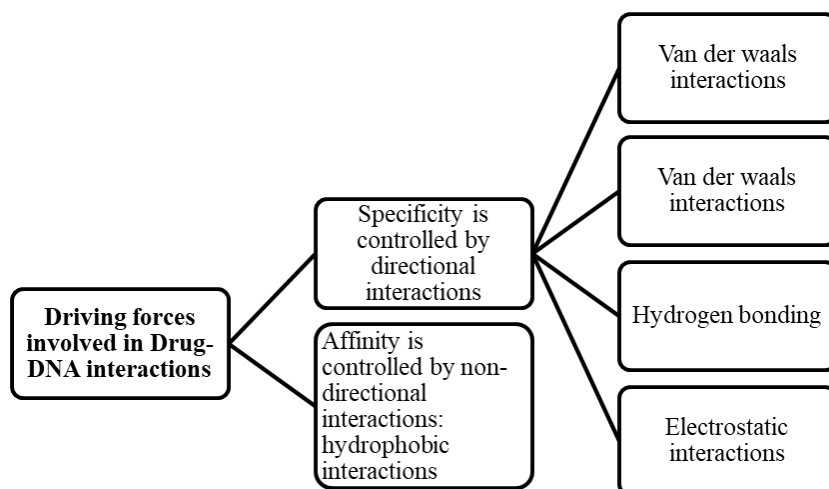


Figure 1. 8 Showing of schematical driving forces involved in Drug-DNA interactions (Intechopen, 2019), (Longdom,2019)

In the figure above, the interactions that are important in DNA-drug interactions are schematized.

1.6 Non-Bonded Interactions

Non-bonded interactions act between atoms in the same molecule and atoms in other molecules. Force fields are generally non-bonded interactions, which are divided into two categories: electrostatic interactions and Van der Waals interactions. (cmt, 2022). Non-bonded interaction has critical role to understanding to drug- DNA interactions. Non-covalent interactions include weak interactions such as hydrogen bond interactions, London dispersion effect, and π - π interactions. These play an important role in the identification of DNA-ligand interactions. Determination of these weak interactions are cheaper than experimental methods and can be quickly studied by quantum mechanical calculations, which give more reliable results than atomic-level force field calculations such as docking, MD, MM.

1.7 Aromatic Molecules

The benzene molecule is the simplest representative of a very large family of cyclic conjugated molecules. Aromatic molecules are formed essentially by the substitution, linkage, and fusion of the benzene ring, which consists of six carbon atoms linked together in a hexagonal cycle. Since there are six p-orbitals that interact with each other to form the molecular orbitals, there are three bonding π orbitals and three antibonding π^* orbitals.

1.8 Electrostatic Potential MAPS

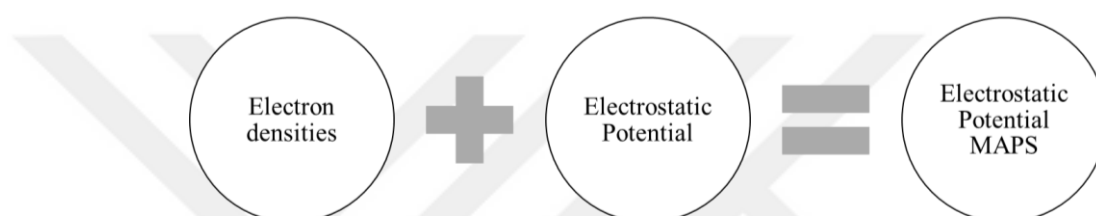


Figure 1. 9 Showing of schematical electrostatic potential MAPS

The locations of electrons or the chance that an electron will be in a specific position are displayed by electron densities. The energy of the interaction between a positive point charge and a molecule's nuclei and electrons is known as the electrostatic potential. The electrostatic potential MAPS is created by combining these two characteristics. Potential Electrostatic An effective technique for examining intra- and intermolecular bonding as well as interactions between bonds is MAPS analysis. Additionally, it offers a useful framework for the investigation of conjugative interactions or charge transfer in molecular systems. (Gangadharana and Krishnanb, 2014). It can also provide information about chemical reactivity or biological activity. It is also defined as the attack of electrophilic or nucleophilic substances. The potential electrostatic MAPS sheds light on molecular recognition procedures such interactions between drugs and receptors. However, the electrostatic potential MAPS ignores the polarisation of the molecule by the incoming species (Levine, 2014). Therefore, we have examined this feature in more detail in the Results section to control for the drug-nucleobase interactions that occurred in our study.

1.9 Drugs Used in The Thesis Project

In this thesis we employed nonsteroidal anti-inflammatory and antihypertensive drugs. The common point of these drugs is that they are widely used in our daily life. The properties of these drugs will be briefly mentioned in this section.

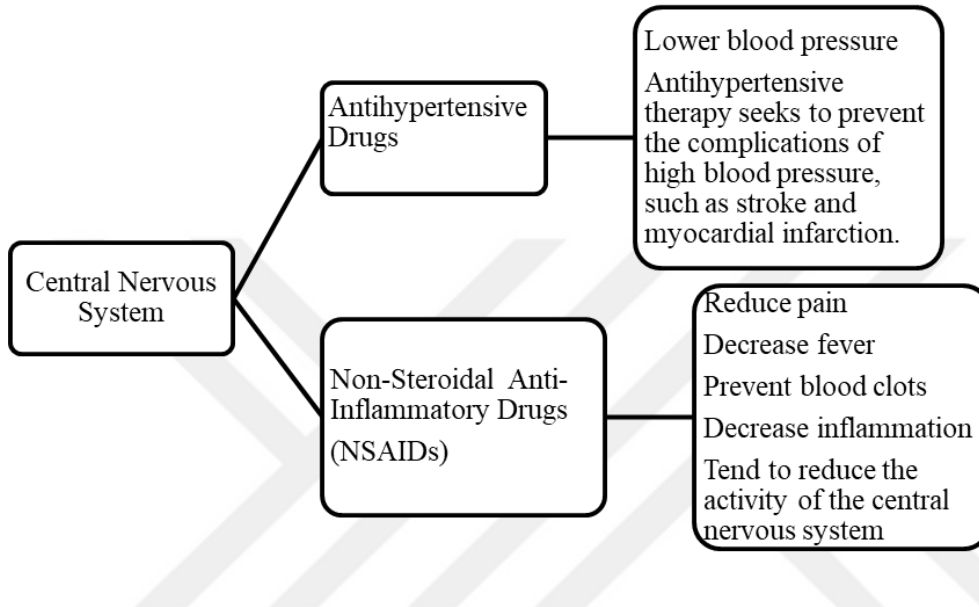


Figure 1. 10 Showing the properties of drugs used

1.9.1 Antihypertensive Drugs

Antihypertensives are a class of drugs that are used to treat hypertension (high blood pressure). They are schematized in figure 1.11.

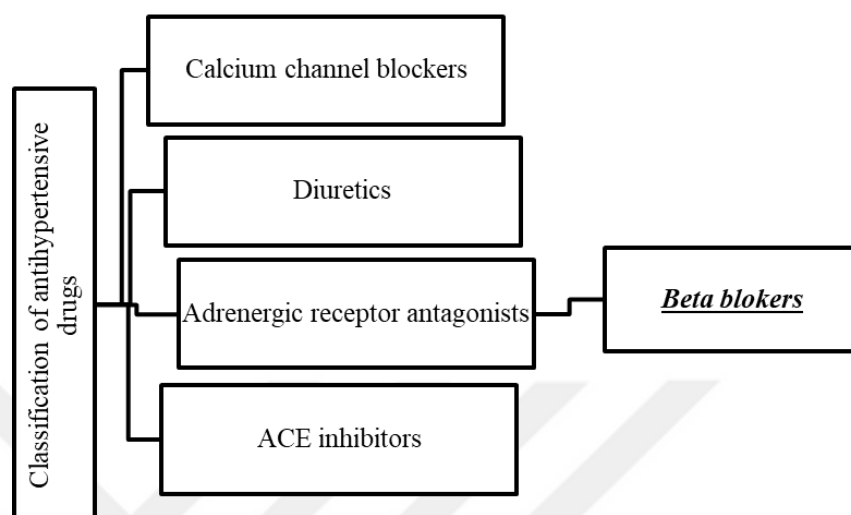


Figure 1. 11 Classification of antihypertensive drug molecules

1.9.2 Beta Blockers

Beta (β)-Blockers are used worldwide for the treatment of several diseases. The most common indications for β -blockers are: angina pectoris, atrial fibrillation, cardiac arrhythmias, congestive heart fail-ure, *hypertension*, myocardial infarction, glaucoma, migraine prophylaxis or anxiety. In figure 1.12, β -blockers function by preventing the effects of specific natural substances on the heart and blood arteries, such as epinephrine.

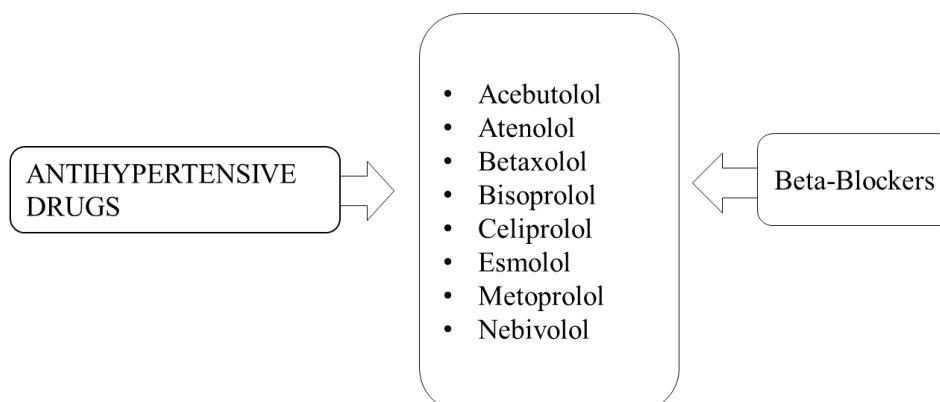


Figure 1. 12 Used of beta blockers drugs in this thesis

Bradycardia, congestive heart failure, hypotension, and hypoglycemia are among the overdose signs and symptoms that are anticipated. Symptoms of overdose include electromechanical dissociation and loss of consciousness. When beta-blockers were administered at the highest dose for a prolonged period of time, there was evidence of small benign lung tumours (Stankiewicz et al., 2015).

1.9.2 Non-Steroidal Anti-Inflammatory Drugs (NSAIDs)

Non-steroidal anti-inflammatory drugs (NSAIDs) are prescription medications used to treat inflammation, discomfort, and high body temperatures. NSAIDs are frequently used to treat osteoarthritis and rheumatoid arthritis pain and inflammation. Depending on the medication, side effects can include increased risk of stomach ulcers and bleeding, heart attack, and kidney illness. Taking NSAIDs can lead to high blood pressure. (Drugbank, 2019).

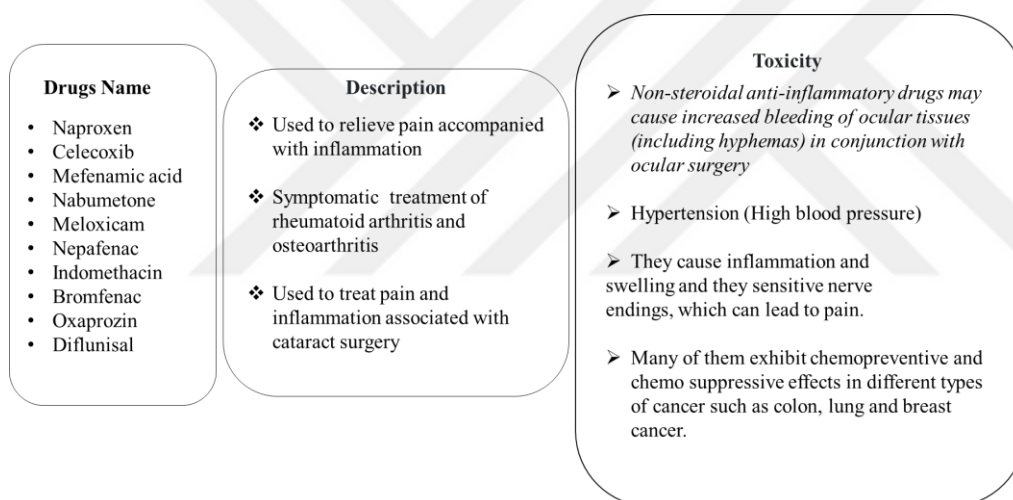


Figure 1. 13 Used of NSAIDs drugs in this thesis

In the figure 1.13, the drugs used in this study and their properties are summarized.

1.9.3 Mefenamic Acid Toxic Effects

Here we will take the drug mefenamic acid in a little more detail. Because, as a result of our calculations and rankings, we know that this drug is more prone to possible side effects or damage to DNA. Mefenamic acid is an anthranilic acid that is an aminobenzoic acid in which a 2,3-dimethylphenyl group has been substituted for one of the hydrogens connected to the nitrogen. It has hydrogen bond donor

count 2 , hydrogen bond acceptor count 3, rotatable bond count 3, lastly melting point is 230 – 231 °C (PubChem, 2022).

“This is the most serious warning from the Food and Drug Administration (FDA). A black box warning alerts doctors and patients to potentially dangerous effects. Potentially fatal heart risks. Mefenamic acid may increase your risk of heart problems, including heart attack, stroke, heart failure, or blood clot.” (Healthline,2020)

In clinical trials, the risk of cardiovascular (CV) thrombotic events has been shown to increase with prolonged and high-dose use of mefenamic acid. It has been found that it can lead to heart attacks or strokes, which can be fatal. In addition, It could potentially result in the development of hypertension or a worsening of pre-existing hypertension, both of which could raise the risk of cardiovascular disease. It can also lead to serious gastrointestinal adverse events. There are side effects such as gastritis, small bowel inflammation and colonic bleeding, which can be fatal if high doses are taken long-term (Accessdata, 2022).

Based on our findings and possible side effects, the interaction of mefenamic acid with DNA was studied in detail in the last section.

1.10 The Main Goal of The Thesis Project

The main objective of this study was to quantify and scale the interactions of some nonsteroidal anti-inflammatory and antihypertensive drugs with nucleobases using quantum chemical methods, especially DFT methods. We expect that knowledge of the strength of these interactions will provide insight into drug side effects.

1.11 The Significance of The Thesis Project

Inflammation, discomfort, and hypertension have long been treated with non-steroidal anti-inflammatory medicines (NSAIDs) and anti-hypertensive medications. In addition to these features, they display various other types of activity. Many of them have chemopreventive and chemosuppressive effects on various cancers, including breast, colon, and lung cancer. Therefore, a prerequisite step understanding cancer pathogenesis and targeted antimetabolite therapy will involve a detailed study of atomic-level interaction of a drug with DNA

nucleobases. Furthermore, to our knowledge, there are not many studies on the nucleobase's actions for the majority of these compounds.



2. COMPUTATIONAL METHODS

“Computational chemistry simulates chemical structures and reactions numerically, based in full or in part on the fundamental laws of physics.”

(Foresman and Frisch in Exploring Chemistry with Electronic Structure Methods, 1996)

Computational chemistry is a branch of chemistry. It goes hand in hand with theoretical chemistry and computer science techniques to enable the solution of chemical problems using computers. Computational modeling techniques can be used to make correct estimates about the geometric structures of the molecules under study. In addition, the properties of a reaction process can be predicted with chemical accuracy if we have sufficient computational capabilities. Computational chemistry plays two important roles in the chemical and physical sciences: One is to predict the properties of chemical systems before experiments are performed. The other is to interpret the complicated or ambiguous results of experiments. Both of these features make computational chemistry studies extremely significant because they provide plausible explanations and save time and money.

2.1 Computational Chemistry Methods

Computational chemistry methods include efficient mathematical approximations and computer programs. Computational chemistry uses fundamental physical laws, especially quantum mechanics, to calculate the properties of physical and chemical systems.

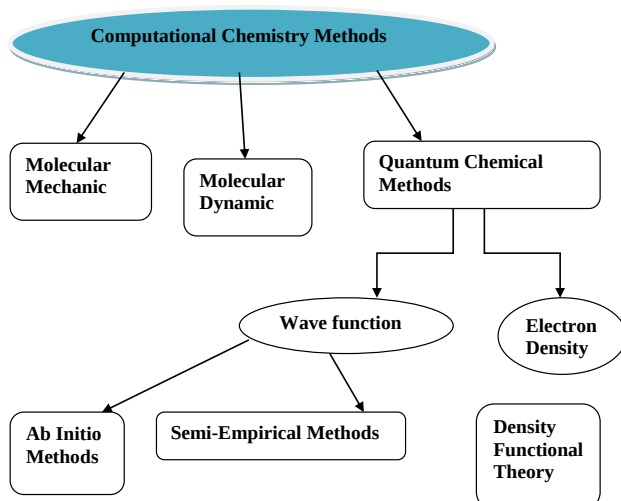


Figure 2. 1 Schematical computational chemistry methods

2.2 Schrödinger's Equation

In quantum mechanics, there is a wave function, represented as $\Psi(r,t)$, which describes all the properties of any system, such as the stable energy, the electron distribution of a particle in a given volume, the dipole moment, etc. The Schrödinger equation below is the basis of quantum mechanics and completely describes the electronic structure of an atomic or molecular system. The wave function has two main variables, namely r and t . While r is represented as the spatial coordinates of the particle, t is the time that makes the wave function dynamic. The understanding of the wave function including the time variables has been fully explained by Schroedinger's time-dependent equation:

$$i\hbar \frac{\partial \psi(r,t)}{\partial t} = \hat{H}\psi(r,t) = E\psi(r,t) \quad (2.1)$$

In this formula, the H is called the Hamilton operator, which gives the total energy of a particle in terms of kinetic and potential energy. In most situations, the wave functions of many particles do not contain a time variable. In other words, the time variable of the wave function is neglected when many properties of a particle are studied, such as finding stable coordinates of a particle in space. Therefore, the Schrödinger equation is divided into two separate equations, namely, time-independent and time-dependent equations, using the variable separation method.

Thus, the time-independent Schrödinger equation can be summarized by the following equation

$$H\psi = E\psi \quad (2.2)$$

In this formula, the H is given as:

$$\hat{H} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{a=1}^M \frac{1}{2M_a} \nabla_a^2 - \sum_{i=1}^N \sum_{a=1}^M \frac{Z_a}{r_{ia}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{a=1}^M \sum_{b>a}^M \frac{Z_a Z_b}{R_{ab}} \quad (2.3)$$

where i and j are represented as lower bounds of N electrons, while a and b are represented as lower bounds of M atomic nuclei. M_a is defined as the ratio of the mass of the nucleus a to the mass of the electron, Z_a is the atomic number of the nucleus a . While r_{ij} is the distance between the i_{th} and j_{th} electrons, r_{ia} is known as the distance between the i_{th} electron and the nucleus a . The first two terms give the kinetic energy of the electrons T_e and the nuclei T_n , respectively. The third term is an attractive operator, giving the energy of the Coulombic interactions between the electrons and the nuclei, V_{ne} , while the last two terms give the repulsive operator, giving the Coulombic interactions between electron and electron, V_{ee} , and nucleus-nucleus, V_{nn} , respectively. H is summarized in the following equation:

$$\hat{H} = T_e + T_n + V_{ne} + V_{ee} + V_{nn} \quad (2.4)$$

Using computational chemistry software one can perform:

- ✓ Electronic structure determinations
- ✓ Geometry optimizations
- ✓ Frequency calculations
- ✓ Protein calculations, i.e. docking
- ✓ Electron and charge distributions calculations
- ✓ Calculations of potential energy surfaces (PES)
- ✓ Calculations of rate constants for chemical reactions (kinetics) etc.

2.3 Molecular Mechanics

The method of molecular mechanics (MM) uses classical (Newtonian) physics and assumes the forces between atoms according to Hooke's law. It is based on simple, empirically derived relationships between energy and bond angles, dihedral angles, and distances, as shown in Figure 2.1.

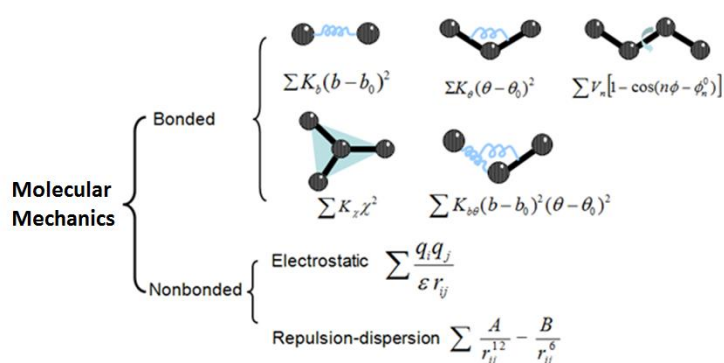


Figure 2. 2 Energy partitioning in the molecular mechanics methods

Molecular mechanics is concerned with the changes in the electronic energy of a molecule due to bond stretching (V_{str}), bond angle bending (V_{bend}), out-of-plane bending (V_{oop}), internal rotation (torsion) about bonds (V_{tors}), interactions

between these types of motion (producing transverse terms V_{cross}), van der Waals attraction and repulsion between unbonded atoms (V_{vdW}), and electrostatic interactions between atoms (V_{es}). For the motion of the atoms in the molecule (or molecules, if the system being computed contains more than one molecule), the aggregate of these contributions results in the molecular mechanical potential energy V , also known as the steric energy:

$$V = V_{\text{str}} + V_{\text{bend}} + V_{\text{oop}} + V_{\text{tors}} + V_{\text{cross}} + V_{\text{vdW}} + V_{\text{es}} \quad (2.5)$$

Due to the fact that the forces acting on the atoms are determined by the derivatives of the potential-energy function, the explicit equations used for each of the variables in (2.5) define what is known as a molecular mechanical force field. Analytical formulations for the terms in (2.5) and values for every parameter that appears in these formulas are contained in a force field. The empirical force field method is the name given to the MM method. Empirical force fields are used in molecular dynamics simulations of liquids and solutions, where Newton's second law is integrated to track the motions of atoms with time in systems with thousands of atoms, as well as in single-molecule molecular mechanics calculations of energy differences, geometries, and vibrational frequencies.

Although the methods of molecular mechanics provide only rough information about a calculated system, in most cases the results are reasonable but limited. Since they do not take into account the electrons of the system and the equations are easy to solve, they are mostly used for large systems.

2.4 Molecular Dynamics

By determining Newton's equation of motion, the molecular dynamics method (MD) determines a molecular system's time-dependent behavior. Comprehensive information on the fluctuations and structural changes of proteins and nucleic acids has been made available by MD methods. Additionally, the structure, dynamics, and thermodynamics of biological molecules and their complexes are typically studied using these techniques.

2.5 Quantum Chemical Methods

2.5.1 Ab Initio Methods

Ab initio methods attempt to solve the Schrödinger equation with a series of approximations. Ab initio methods do not use experimental data and the calculations are based on fundamental quantum mechanics. Ab initio methods allow the electronic structure of solids, surfaces, or clusters to be studied as accurately as possible with moderate computational effort. The calculations are more accurate than molecular mechanics methods, but they require much more time. Therefore, this method is applicable to systems with several hundred atoms.

In ab initio methods, Schrödinger equation is solved with Born-Oppenheimer and basis set approximations. They provide the energy and the wave function of the system.

2.5.2 Semi-Empirical Methods

Semi-empirical methods use experimental data to parameterize equations to solve the Schrödinger equation. Thus, in the semiempirical method, there is both theory and experiment.

Semi-empirical methods are less accurate than ab initio methods but are also much faster because these methods do not solve time-consuming two-electron integrals. The equations are parameterized to obtain certain results, usually geometry and heat of formation, but these methods can also be used to obtain other data. Nevertheless, it is not suitable for unusual molecules because of the parameterization. It can handle fairly large molecules. Unlike molecular mechanics methods, bond breaks can be calculated.

2.5.3 Density Functional Theory (DFT)

Density functional theory is considered an ab initio method, but it differs from other ab initio methods because the electron density, rather than the wave function, is used to describe a molecule.

Wave Function Theory

$$H\psi = E\psi \quad \psi \equiv \text{Trial wave function} \quad (2.6)$$

$$E = \langle \psi | H | \psi \rangle$$

Density Functional Theory

$$E = \min_n \left\{ \int V_{\text{nuclei}}(\vec{r}) n(\vec{r}) d^3\vec{r} + F[n(\vec{r})] \right\} \quad (2.7)$$

$\downarrow$ $n \equiv \text{Trial density}$ $\downarrow$ $F \equiv \text{Universal functional}$

The difference between wave functional theory and density functional theory is that wave functional theory calculates energy as a function of wavelength, while DFT calculates energy as a function of electron density. The universal functional is not fully known for DFT, so exchange-correlation approximations are made.

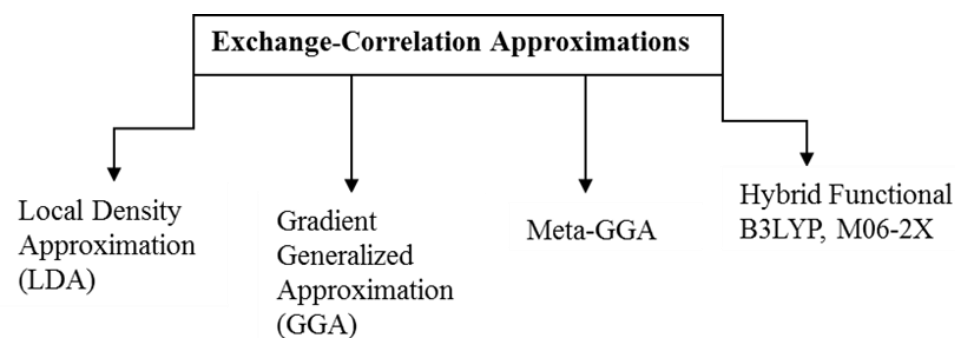


Figure 2. 3 Approximations of exchange-correlation scheme

Density functional theory is a very successful approach to calculating the properties of many electron systems. It must be ensured that it correctly contains all necessary information about the system. Many properties such as total energy, density, forces, eigenvalues, oscillations, frequencies, magnetic properties, etc. are calculated. DFT Methods provide errors of less than a few % and are therefore more reliable than molecular mechanics method. They can be interpreted as experimental results. This method has many other tricks to speed up the calculations. DFT Methods take electron correlation into account and are therefore superior. It gives more accurate results and more accurate geometries. DFT methods take less computational time than CCSD(T). It takes into account other interactions for larger systems such as electric and magnetic fields, relativistic corrections (spin orbit), etc. DFT method has some limitations, such as up to 103 atoms, band gap too small, complicated correction terms, ground state, no simple description of excited states, etc.

2.6 Basis Sets

A basis set is linear combination of atomic orbitals including a set of exponents and contraction coefficients for a set of atoms. Each basis set is a different set of constants used in the wave function of the Schrödinger equation in computational chemistry.

2.6.1 Slater Type Orbital

Slater type functions given in equation 2.8 are used for atomic calculations. They are solutions of hydrogen atom Schrödinger equation. They provide fairly fast convergence as the number of functions increases. Three- and four-center integrals cannot be performed analytically. They provide calculations with very high accuracy for atomic and diatomic systems.

$$\chi_{\alpha,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi) \underbrace{r^{n-1}}_{\substack{\text{Spherical harmonic} \\ \text{function for angular} \\ \text{distribution}}} \underbrace{e^{-\alpha r}}_{\substack{\text{Exponential function for} \\ \text{radial distribution}}} \quad (2.8)$$

2.6.2 Gaussian Type Orbital

Gaussian type orbitals (GTO) contains $e^{-\alpha r^2}$ function instead of $e^{-\alpha r}$ given in eqn 2.9. has a slope of zero; the STO has a cusp. The behavior near the nucleus is poorly represented. The GTOs decrease too rapidly with distance. Small α (exponent) values are used for preventing this rapid decay. The 'tail' behavior is poorly represented. The main advantage of GTOs is that they have analytical solutions that they linear combination of GTOs are used to overcome previously mentioned deficiencies of GTOs.

$$\chi_{\alpha, l_x, l_y, l_z}(x, y, z) = N x^{l_x} y^{l_y} z^{l_z} e^{-\alpha r^2} \quad (2.9)$$

Spherical harmonic function for angular distribution Exponential function for radial distribution

2.6.3 Pople Style Basis Sets

Name of Pople type basis sets comes from Prof. John Pople who won the Nobel Prize in Chemistry for his work in quantum chemistry in 1998. This basis set notation looks like **k-nlm++G**** or **k-nlm++G(idf,jpd)**.

k primitive GTOs for core electrons

n primitive GTOs for inner valence orbitals

l primitive GTOs for medium valence orbitals

m primitive GTOs for outer valence orbitals

+ means 1 p diffuse function added to heavy atoms.

++ means 1 p diffuse function added to heavy atoms and 1 s diffuse functions added to H atom.

***** means 1 d polarization functions added to heavy atoms.

****** means 1 d polarization functions added to heavy atoms and 1 p polarization functions added to H atom.

idf means i d and l f polarization functions added to heavy atoms.

E.g., 6-31G*, 6-31+G, 6-31+G(d,p)

2.7 Polarizable Continuum Model (PCM)

The dielectric continuum approach in quantum chemistry is the most widely used class of methods for modelling solvent effects in calculations, and apparent surface charge (ASC) models are the most popular subclass of it. Their first representative is the polarised continuum model (PCM), (Miertuš et al., 1981), which originally used the full dielectric boundary condition (EDBC) to calculate the vector of polarisation charge densities (σ) on the surface portions of the dissolution cavity (Γ). Later, this original version was called DPCM. In 1993, Klamt and Schueuermann presented a completely independently derived ladder-type shielding model (COSMO) that uses the much simpler boundary conditions of a ladder and accounts for the reduction in polarisation charge densities that occur at the finite dielectric constant (Klamt and Schueuermann, 1993). Solvent calculations at the DFT level provide reliable results using the well-established PCM (Mennucci et al., 2002; Cossi et al., 2003; Tomasi et al., 2005; Hiratani et al., 2005; Wang et al., 2000, Brooks et al., 1983, Wang et al., 2004).

2.8 Couple Cluster (CC) Theory

Perturbation theories with electron correlations deal with all excitations (S,D,T...etc) added to the reference wave function at a given order (2,3,4,...), but the Couple Cluster (CC) theory deals with all corrections added to the reference wave function from the given order to infinite order. The wave function CC is summarised in the following set of formulas

Ground state Slater determinations of HF theory

$$\psi = e^{\hat{T}}\Phi_0 \quad (2.10)$$

$$e^{\hat{T}} \equiv 1 + \hat{T} + \frac{\hat{T}^2}{2!} + \frac{\hat{T}^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{\hat{T}^k}{k!} \quad (2.11)$$

In the formula, exponential operator T is given as follow

$$\hat{T} \equiv \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_n \quad (2.12)$$

The HF reference wave function applied by using the T_i operator creates excited Slater determinants

$$\hat{T}_1 \Phi_0 \equiv \sum_{a=n+1}^{\infty} \sum_{i=1}^n t_i^a \Phi_i^a, \quad \hat{T}_2 \Phi_0 \equiv \sum_{b=a+1}^{\infty} \sum_{a=n+1}^{\infty} \sum_{j=i+1}^n \sum_{i=1}^{n-1} t_{ij}^{ab} \Phi_{ij}^{ab} \quad (2.13)$$

The expansion coefficients are represented as amplitudes. For the exponential operator can be rewritten by substituting equation 2.12 into eqn 2.11

$$e^{\hat{T}} = 1 + \hat{T}_1 + (\hat{T}_2 + \frac{1}{2} \hat{T}_1^2) + (\hat{T}_3 + \hat{T}_2 \hat{T}_1 + \frac{1}{6} \hat{T}_1^3) + \dots \quad (2.14)$$

In the formula, the first two terms (1 and $\hat{T}_1$) are the reference wave function HF and the individual excitations, respectively. The first terms in the parenthesis stand for the double excitations and the other terms in the second parenthesis for the triple excitations.

If the wave function CC is inserted into the Schrödinger equation, the following equation results:

$$\hat{H} e^{\hat{T}} \Phi_0 = E e^{\hat{T}} \Phi_0 \quad (2.15)$$

In the formula, the energy, E , becomes as follow when this equation is multiplied from the left side with Φ_0^*

$$E_{\text{CCD}} = \langle \Phi_0 | \hat{H} | e^{\hat{T}_2} \Phi_0 \rangle \quad (2.16)$$

In the formula, the Hamiltonian obtained by using the expandable e^T operator contains only one- and two-electron operators. Moreover, the matrix elements with singly excited Slater determinants are zero, since the HF orbital is used. E_{CC} is obtained as follows

$$E_{cc} = E_0 + \sum_{i \in j} \sum_{a \in b}^{occu \text{ virtual}} (t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a) (\langle \phi_i \phi_j | \phi_a \phi_b \rangle - \langle \phi_i \phi_j | \phi_b \phi_a \rangle) \quad (2.17)$$

Truncated Couple Cluster methods are used for cost reducing reasons. Note that adding the energy correction of the $\check{T}_1$ operator to the reference wave function HF does not give an accurate result. Therefore, the lowest level of approximation should be used when $\check{T} = \check{T}_2$ is used. Such an approximation is called a CCD (Pople et al., 1978). When $\check{T} = \check{T}_1 + \check{T}_2$ is used to calculate the reference wave function HF, this is referred to as the CCSD model (Purvis and Bartlett, 1982; Scuseria et al., 1988, 1989). Furthermore, the contribution of the triple excitation can be added to the CCSD method after it has been determined using perturbation theory. This method is called CCSD(T) (Pople et al., 1987) and is known as the "gold standard". The CCSD(T) method takes the full weak interaction of a system into account. The pair-cluster wave function can match the full CI if all cluster operators are added to the reference wave function HF. In this work, we have used the CCSD(T) method as a standard to check the quality of other methods since it provides the best recovery of the electron correlation.

2.9 QM/MM Methods

Although it can handle very big molecules, chemical reactions are not well-suited for molecular mechanics to handle. In contrast, quantum mechanical methods (ab initio, density functionals, and semiempirical) can handle chemical reactions but are not well suited to handling very large molecules. A combined quantum mechanical/molecular mechanical (QM/MM) technique can be used to tackle chemical reactions in very large systems, where quantum mechanics is used to treat the portion of the system that is most impacted by the reaction and molecular mechanics is used to treat the rest. When a reaction takes place in solution, the surrounding solvent molecules are treated with MM and the reacting solute molecules are treated with QM. Any quantum mechanical method can be applied to the QM. The majority of research use semi-empirical approaches like AM1,

PM3, PM6, and SCC-DFTB because of their speed, but more and more are turning to DFT-D. The QM/MM system's effective Hamiltonian is expressed as:

$$\hat{H}_{\text{eff}} = \hat{H}_{\text{QM}} + \hat{H}_{\text{QM/MM}} + \hat{H}_{\text{MM}} \quad (2.18)$$

where H_{QM} is the Hamiltonian of the quantum-mechanical region in a vacuum, $H_{\text{QM/MM}}$ is the Hamiltonian for the interaction between the QM and MM regions, and H_{MM} is taken as equal to the steric energy V_{MM} [Eq. (2.5)] of the MM region. If Ψ_{QM} is the wave function of the QM region, the energy E of the system is taken as

$$E = \langle \psi_{\text{QM}} | \hat{H}_{\text{QM}} + \hat{H}_{\text{QM/MM}} | \psi_{\text{QM}} \rangle + V_{\text{MM}} \quad (2.19)$$

Thus, in the QM region, i and a denote for the electrons and nuclei, respectively, and M denotes for an atom with a partial atomic charge Q_M in the MM region.

The polarization of the solvent molecules by the charges in the area QM is ignored by this approach. All terms in $H_{\text{QM/MM}}$ are constants for a fixed configuration (Born-Oppenheimer approximation), with the exception of the first double sum. The first double sum in $H_{\text{QM/MM}}$ results in the appearance of extra nuclei with charges Q_M when the electronic Schroedinger equation is solved with the Hamiltonian eff . As a result, the Fock matrix gains new one-electron integrals.. Both Ψ_{QM} and E are parametric functions of the nuclear coordinates of the QM and the coordinates of the atoms in the MM. Ψ_{QM} is a function of the electronic coordinates of the QM. Consider E 's partial derivatives in relation to nuclear coordinates. This generates the forces required for the molecular dynamics simulation.

2.9.1 Briefly QM, MM and QM/MM methods

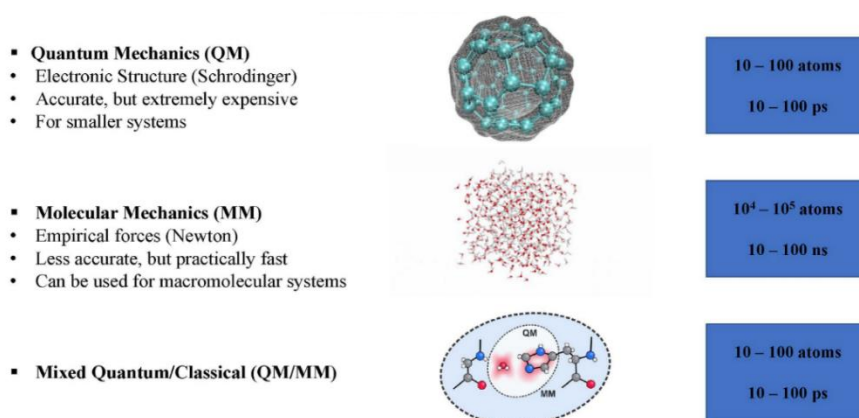


Figure 2. 4 Showing of QM, MM and QM/MM methods schematically

The basic property of the QM/MM method without going into details, it divides the system into sections and makes the calculations. The chemically active region calculated by quantum mechanics. The interface region is used quantum mechanics or molecular mechanics and The widest region is calculated by classical force field, that is, molecular mechanics. Besides, max QM/MM simulations time 100ps and the maximum atom number of the quantum mechanical region is 100 atoms.

2.10 The Computational Chemistry Suites Employed

Most of the calculations in this thesis was performed by utilizing GAUSSIAN 09 (Frisch et al., 2010) chemistry suit package program. Polarized continuum model (PCM) was used in all the calculations in water as solvent. In addition, the geometries of drug molecules and nucleobases have been obtained employing conformation analysis with molecular mechanics by the Spartan14 and geometry optimization and frequency were calculated with density functional theory (DFT) calculations by Gaussian 09. Moreover, ORCA software suite (version 4.2.1) chemistry suit programs also served in some calculations namely, DFT including dispersive interactions. Besides, CP2K chemistry suit package program was used for QM/MM calculations.

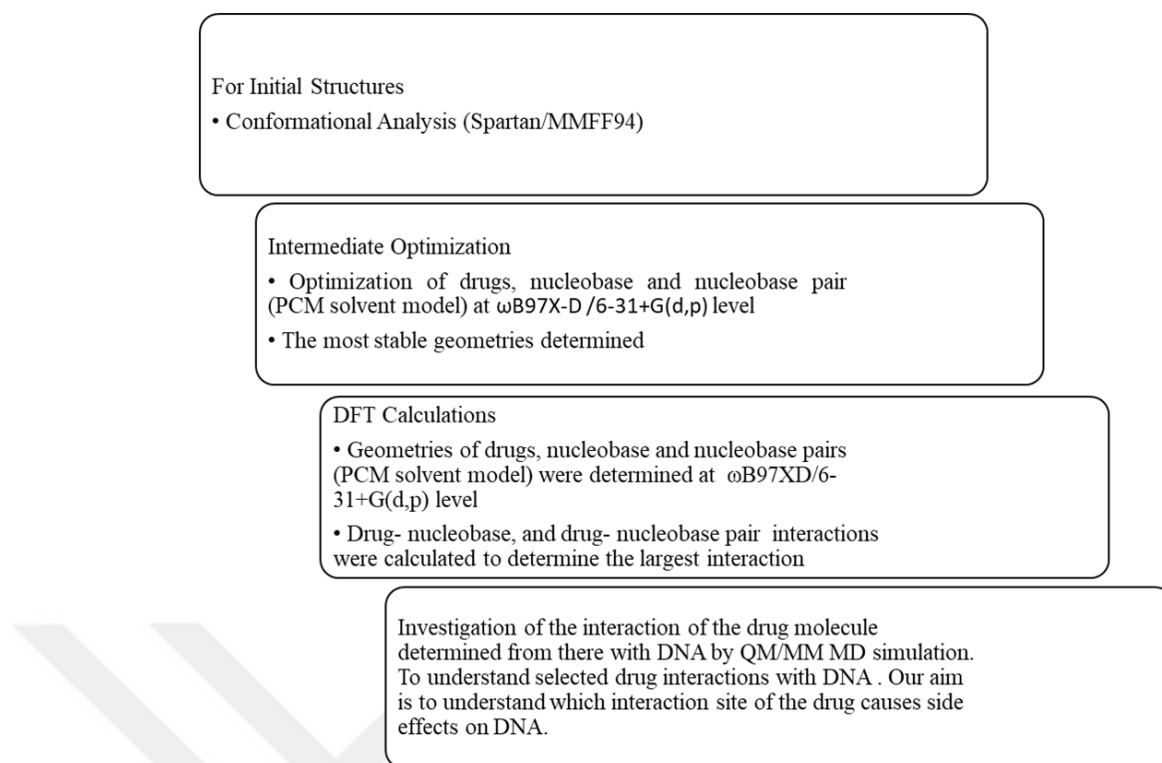


Figure 2. 5 Showing of the general protocol we will use for the calculation

This study consists of three main parts: 1) The performance of the methods for the non-covalently interacted systems 2) Interaction of drugs with nucleobases 3) Interaction of drugs with nucleobase pairs. All structures have been optimized at the ω B97X-D /6-31+G(d,p) levels of DFT. That each optimized geometry corresponds to a local minimum on the potential energy surface was identified via harmonic vibrational analyses. The effect of implicit water solvation on complex formation was calculated by employing integral equation formalism of the Polarized Continuum Model (IEF-PCM). The determination of selected drug molecules' interaction with DNA was investigated with QM/MM method by MD simulations.

3. RESULTS AND DISCUSSION

The findings and interpretations of the calculations conducted with the methods described in Chapter 1, and Chapter 2 are discussed in this chapter.

3.1 Geometry Optimization of All Drugs, Nucleobases and Base Pairs

Initially, the conformers of the drugs, nucleobases, and nucleobase pairs were determined using Spartan 14 and molecular mechanics calculations were performed using the MMFF94 force field. Then, their optimization and frequency calculations were performed at DFT B3LYP/6-31+G(d) in the Gaussian 09 molecular modeling suite. Each optimized geometry corresponds to a local minimum on the potential energy surface. The effect of implicit water solvation on complex formation was studied using the Polarized Continuum Model (IEF-PCM). The most stable conformer was determined, and calculations were continued with this structure.

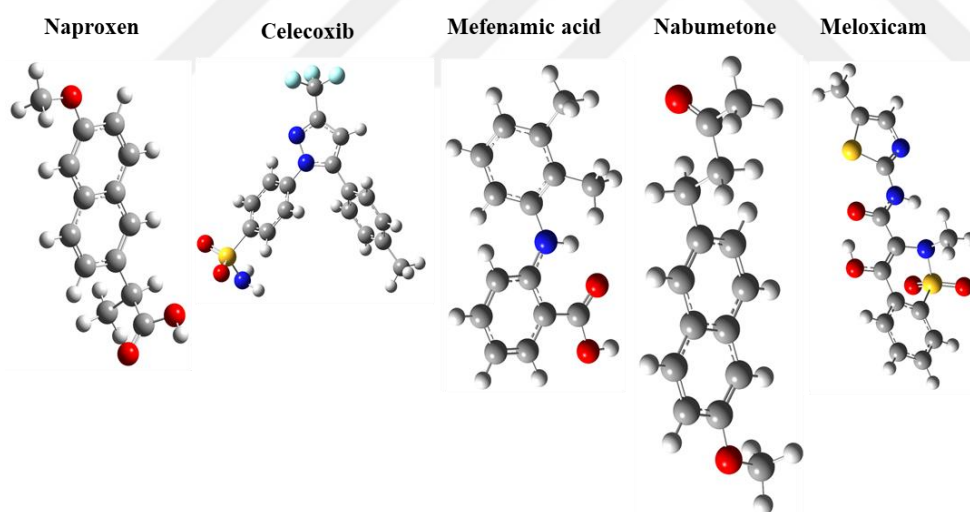


Figure 3. 1 Stable geometries of Non-steroidal anti-inflammatory drugs molecules

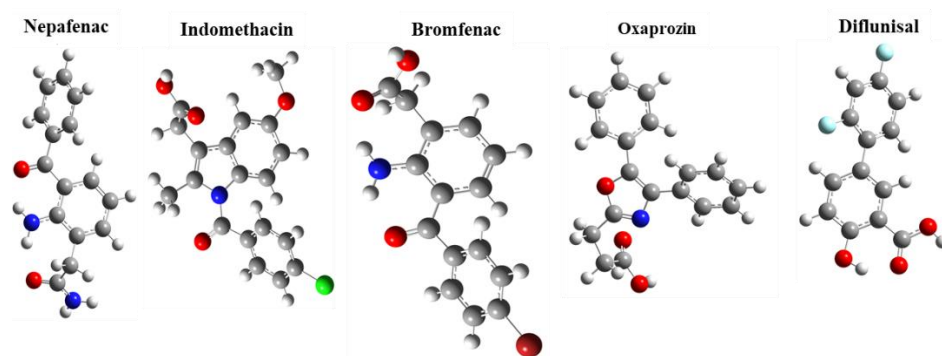


Figure 3. 2 Stable geometries of Non-steroidal anti-inflammatory drugs molecules

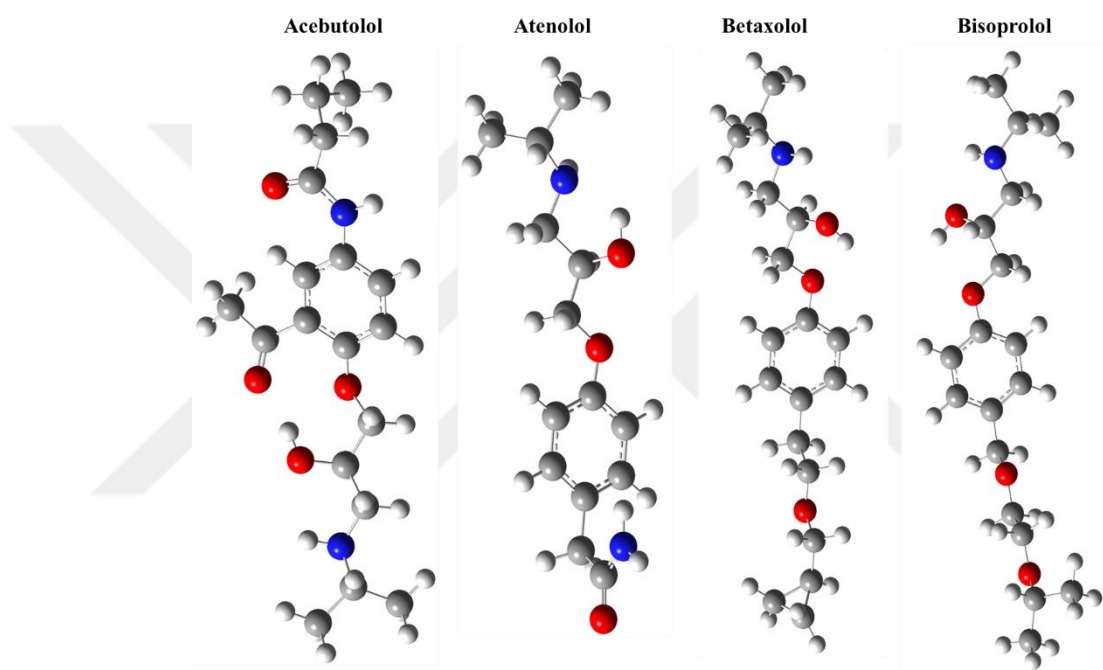


Figure 3. 3 Stable geometries of anti-hypertensive/ Beta- Blockers drugs molecules

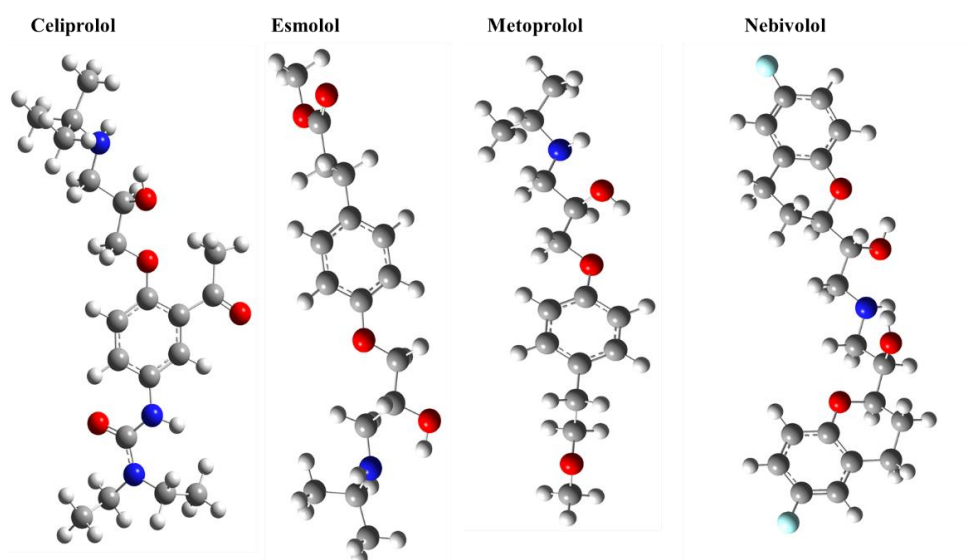


Figure 3. 4 Stable geometries of Anti-hypertensive/ Beta- Blockers drugs molecules

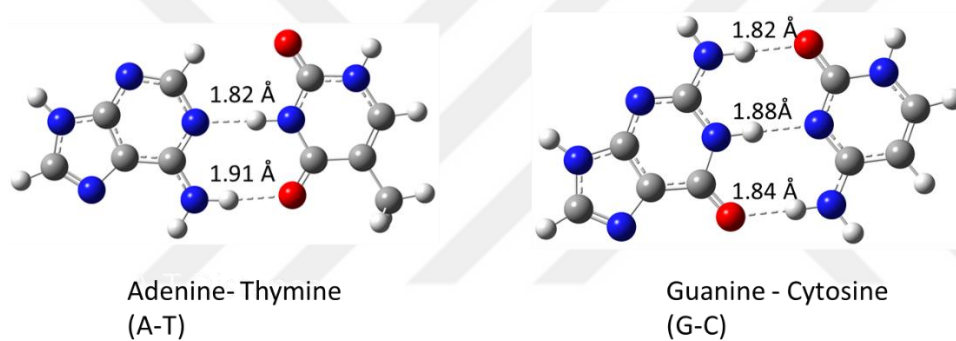


Figure 3. 5 Stable geometries of nucleobase pairs structures

Figures 3.1- 3.5 show the most stable geometries of the drug molecules optimized at B3LYP/6-31+G(d) level of theory. These geometries were employed for the rest of the study.

3.2 Preliminary Study

3.2.1 Binding Enthalpies of Drugs to Nucleobases

In this part, we investigated the effects of some selected drugs namely bisoprolol, acebutolol, naproxen, diflunisal on a single nucleobase with different DFT methods. In order to determine drug-nucleobase interactions, the drug-nucleobase complexes were optimized at the B3LYP/6-31G(d), B3LYP-D2/6-31G(d) and ω B97X-D/6-31G(d) levels of DFT. The binding enthalpies were calculated exploiting the following eqn 3.1.

$$\Delta H_{bind} = H_{298}(complex) - (H_{298}(drug) + H_{298}(base)) \quad (3.1)$$

Table 3.1 shows the binding enthalpy values of the drug–nucleobase complexes optimized with the B3LYP, B3LYP-D2 and ω B97X-D methods.

Table 3. 1 Binding enthalpy (kcal/mol) values of the drug–nucleobase complexes optimized with the B3LYP, B3LYP-D2 and ω B97X-D methods

Drug name	Base	B3LYP	B3LYP-D2	ω B97X-D
Bisoprolol	A	2.02	8.3	-8.17
	T	1.7	3.33	-11.33
	G	-2.39	5.72	-11.78
	C	0.17	5.25	-7.22
Acebutolol	A	1.44	9.03	-12.56
	T	1.02	7.74	-7.9
	G	1.57	8.29	-9.51
	C	-2.83	*	-8.32
Naproxen	A	0.91	6.24	-9.7
	T	1.03	7.07	-8.45
	G	*	3.86	-11.11
	C	0.94	7.36	-8.37
Diflunisal	A	-5.4	7.62	-7.45
	T	-6.05	7.22	-8.28
	G	-8.59	6.15	-7.59
	C	*	6.37	-6.91

* Calculations did not converge to a stable geometry.

As shown in the table, the most negative binding energy is found between adenine and acebutolol, while the lowest is between diflunisal and guanine. The ω B97X-D method predicted that all drug-nucleobase complexes are quite stable. We expected that the B3LYP-D2 method would give similar results because the D2 part includes the contribution of dispersion interactions that are not included in the B3LYP method. However, this is not the case. Although the B3LYP methods give reasonable results, they were degraded by the D2 correction. Therefore, it can be said that the ω B97X-D method gives more reliable results than the B3LYP-based methods.

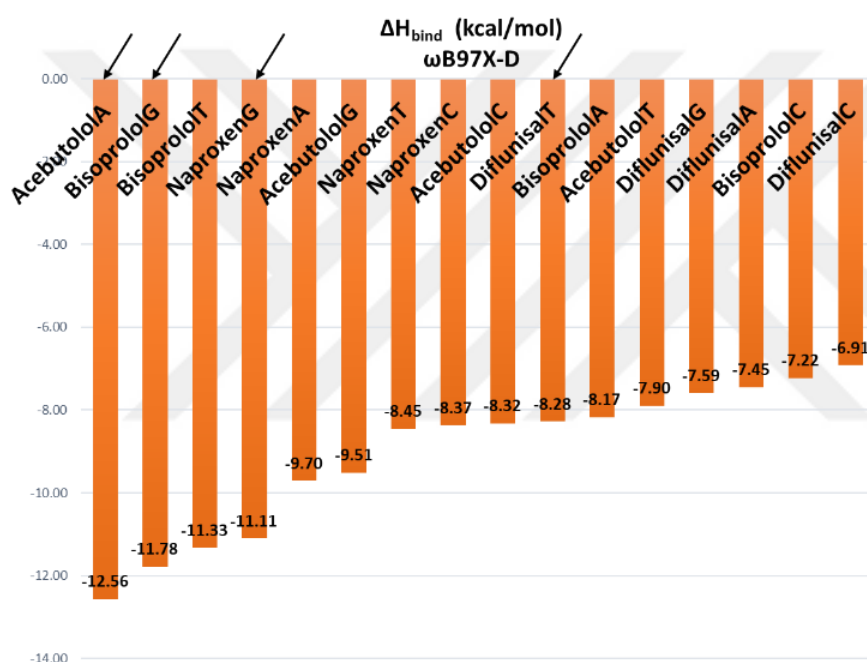


Figure 3. 6 Highly interacting drug-base structures

Figure 3.6 shows the structures of strongly interacting drugs and nucleobases obtained by the ω B97X-D method, which incorporates long-range interactions and, as mentioned earlier, gives good results for weakly interacting systems. Accordingly, acebutolol and adenine have the most negative binding energy among all individual nucleobase-drug complexes. The second strongest interacting complex is the bisoprolol guanine complex. The next two complexes are the naproxen-guanine and diflunisal-thymine complexes.

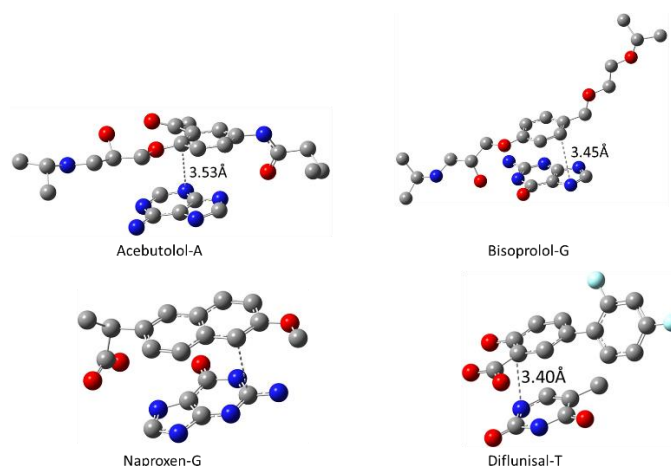


Figure 3. 7 ω B97X-D optimized structures of drug–nucleobase complexes

Table 3. 2 Binding enthalpy (kcal/mol) values of the drug–nucleobase complexes optimized with the B3LYP, B3LYP-D2 and ω B97X-D methods

Drug name	Base	B3LYP		B3LYP-D2		ω B97X-D	
		6-31-G(d)	6-311++G(d,p)	6-31-G(d)	6-311++G(d,p)	6-31-G(d)	6-311++G(d,p)
Bisoprolol	A	2.02	0.85	8.3	8.11	-8.17	-9.18
	T	1.7	-2.71	3.33	4.29	-11.33	-11.57
	G	-2.39	-2.68	5.72	6.10	-11.78	-12.65
	C	0.17	1.02	5.25	5.28	-7.22	-7.68
Acebutolol	A	1.44	-2.11	9.03	8.60	-12.56	-13.67
	T	1.02	0.96	7.74	7.53	-7.9	-8.69
	G	1.57	1.55	8.29	7.57	-9.51	-10.80
	C	-2.83	-2.89	*	6.89	-8.32	-9.07
Naproxen	A	0.91	-8.88	6.24	6.20	-9.7	-11.50
	T	1.03	1.04	7.07	6.99	-8.45	-9.08
	G	*	-6.50	3.86	4.83	-11.11	-12.61
	C	0.94	-0.28	7.36	7.18	-8.37	-4.99
Diflunisal	A	-5.4	1.25	7.62	7.39	-7.45	-8.03
	T	-6.05	0.75	7.22	6.97	-8.28	-9.09
	G	-8.59	1.40	6.15	7.29	-7.59	-8.93
	C	*	0.83	6.37	6.25	-6.91	-7.36

* Calculations did not converge to a stable geometry.

As can be seen from the Table 3.2, three density functional methods and two basis sets were used in the present calculations to obtain the interaction energies of these structures. Small basis sets are not flexible enough for accurate representation, which requires the use of much more functions to represent each atomic orbital (Chemistry Libretexts, 2020). Therefore, two different basis functions are used here. The 6-311++G(d,p) basis set contains both polarization and diffuse functions (Frisch et al., 2010). We have used two different basis functions to determine the errors that can arise from the basis sets. The ω B97X-D method makes all drug–nucleobase complexes quite stable. We expected that the B3LYP-D2 method would

give similar results because of the the empirical dispersion correction D2. Although the B3LYP methods give rather reasonable results, the D2 correction made the results worse. The ω B97X-D method gives negative binding energies, therefore, we can consider that the ω B97X-D method gives more reasonable results than the B3LYP-based methods. In addition, the results did not change much with the different basis functions.

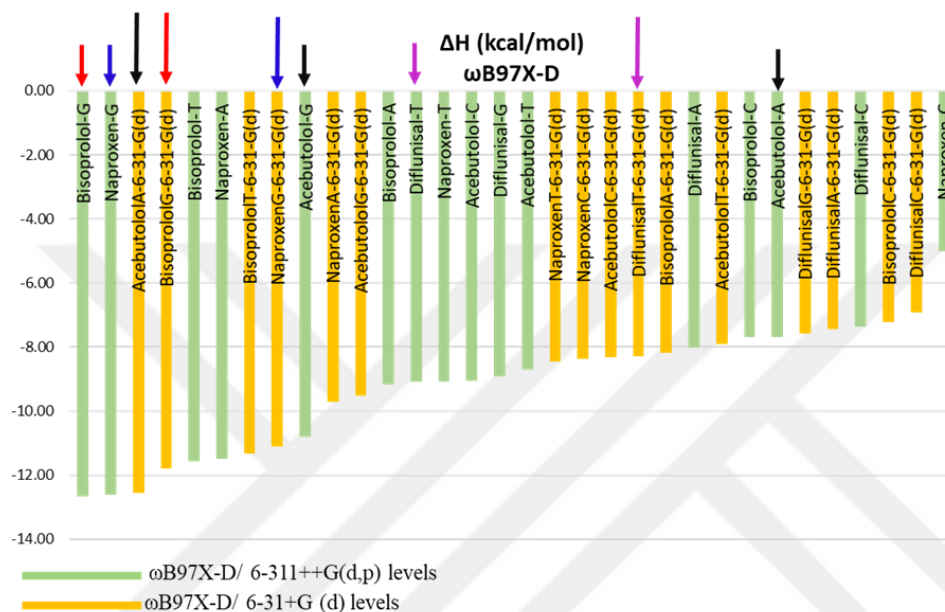


Figure 3. 8 Highly interacting drug-base structures

Figure 3.8 shows the structures of strongly interacting drugs and nucleobases obtained by the ω B97X-D method, which, as mentioned above, takes into account long-range interactions and gives good results in weakly interacting systems. The results slightly improved with the larger basis set, this means that the long-range of interactions are somewhat sensitive to basis set. As we have seen above, the most negative binding energy is between guanine and bisoprolol, while the lowest is between naproxen and cytosine. Accordingly, bisoprolol and guanine have the most negative binding energy among all individual nucleobase-drug complexes. The complex with the second strongest interaction is the naproxen-guanine complex. The next two complexes are bisoprolol-thymine and naproxen-adenine complexes by the ω B97X-D/6-311++G(d,p) method.

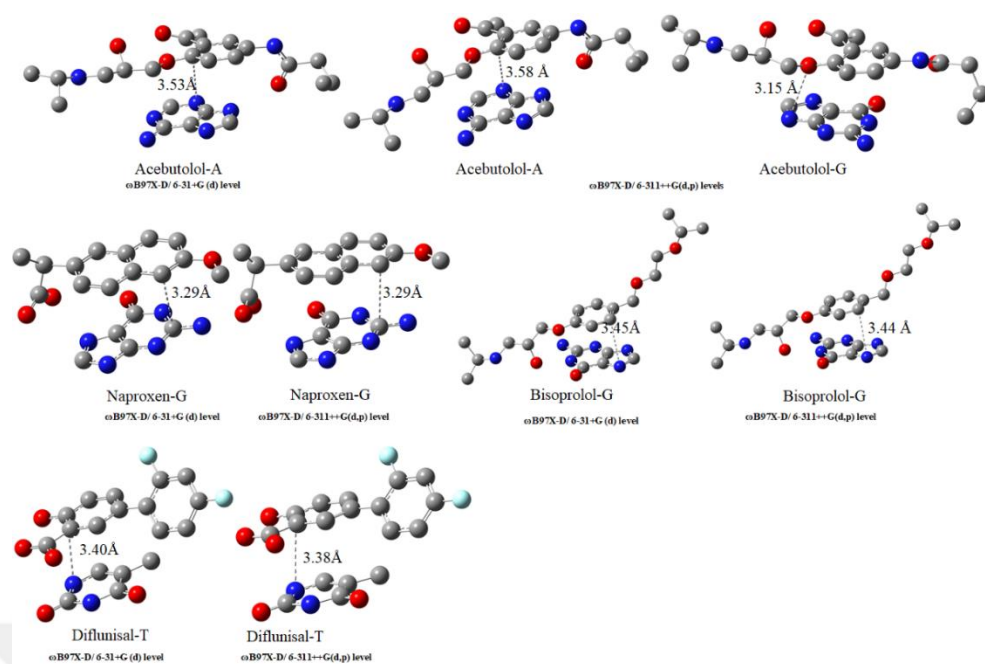


Figure 3.9 ω B97X-D/6-31+G(d) and ω B97X-D/6-311++G(d,p) level optimized structures of drug–nucleobase complexes

Figure 3.9 shows the structures obtained by the ω B97X-D method with strong drug-nucleobase interaction. The use of the large basis set does not have a significant change on the geometry. With a small change of 0.03 Å, we do not need to optimize the structures with a large basis set. While naproxen and bisoprolol have the largest interaction with guanine, diflunisal gives the largest interaction with thymine for both basis functions. However, acetobutolol gives the largest interaction with adenine for the ω B97X-D/6-31-G(d) method, while acetobutolol gives the largest interaction with guanine for the ω B97X-D/6-311++G(d,p) method. The performance of ω B97X-D in non-covalent interactions is significantly better than that of B3LYP-based methods(Mardirossian and Head-Gordon, 2014).

3.2.2 Drug-Nucleobase-Pair Binding Enthalpies

In this part, we examined the interactions of these selected drugs in different orientations with the nucleobase pairs (G-C and A-T). All complexes of drugs and nucleobase pairs were optimized at the B3LYP/6-31G(d) and ω B97X-D/6-31G(d) levels of theory to determine these interactions. The binding energies were calculated using eqn 3.2.

$$\Delta H_{bind} = H_{298}(complex) - (H_{298}(drug) + H_{298}(base\ pair)) \quad (3.2)$$

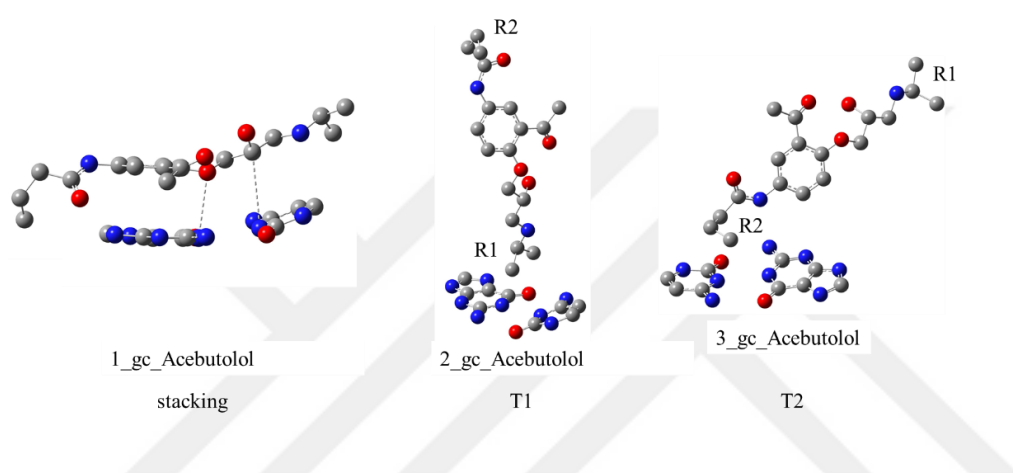


Figure 3. 10 Orientations of drug–nucleobase pair complexes optimized by ω B97X-D (T: T shaped)

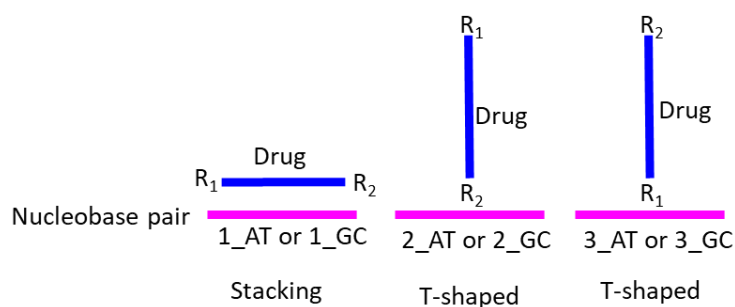


Figure 3. 11 General showing that orientations of drug–nucleobase pair complexes

The drug molecules were placed to interact with the nucleobase pairs in the orientations shown in Figure 3.10. Table 3.3 shows the binding enthalpy values for the drugs interacting with the nucleobase pairs in these three orientations. The calculations predict that the parallel oriented drug complexes have more negative binding enthalpies than the vertically oriented ones.

Table 3. 3 ΔH_{bind} (kcal/mol) values of B3LYP and ω B97X-D optimized drug–nucleobase pair complexes

Interreaction Site	Drug Name	Basepair	B3LYP	ω B97X-D
1_gc	Bisoprolol	G-C	1.26	-13.8
2_gc			1.74	-4.67
3_gc			1.66	-4.82
1_at	Bisoprolol	A-T	*	-13.38
2_at			1.67	-4.56
3_at			1.68	-4.66
1_gc	Acebutolol	G-C	0.63	-17.56
2_gc			0.59	-3.73
3_gc			1.13	-9.62
1_at	Acebutolol	A-T	-45.02	-11.87
2_at			-44.59	-3.67
3_at			-45.34	-6.63
1_gc	Naproxen	G-C	0.19	-9.99
2_gc			0.85	-12.79
3_gc			-1.83	-12.69
1_at	Naproxen	A-T	0.74	-10.31
2_at			1.02	-11.3
3_at			-2.29	-12.97

* Calculations did not converge to a stable geometry

As you can see from the table, the strongest interaction is between acebutolol and the G-C base pair. The ω B97X-D method predicted all drug-nucleobase

complexes to be more stable than those predicted by B3LYP. However, in only one case did the B3LYP method predict the π - π parallel stacking interaction of the drug acebutolol with the A-T base pair to be about four times stronger than that predicted by ω B97X-D. This is most likely a false stability due to the inability of the B3LYP method in this calculation. The reason is that this high binding energy (-45 kcal/mol) is larger than the covalent binding in the F_2 molecule, whose binding dissociation energy is about 37 kcal/mol (Wenthold and Squires, 2005). This clearly shows that B3LYP does not work for this particular case. Figure 3.12 shows the order of the binding enthalpies of the drug-nucleobase pairs calculated by the ω B97X-D method.

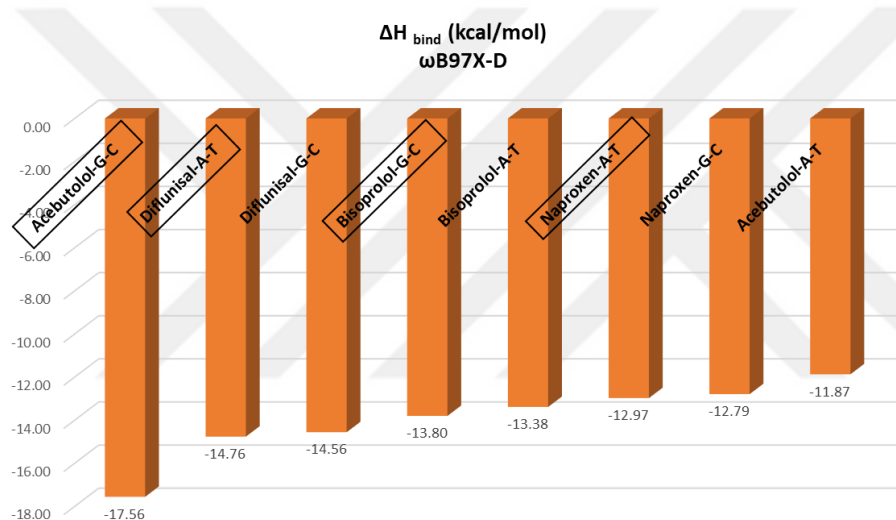


Figure 3. 12 The highly interacting drug-base pair structures

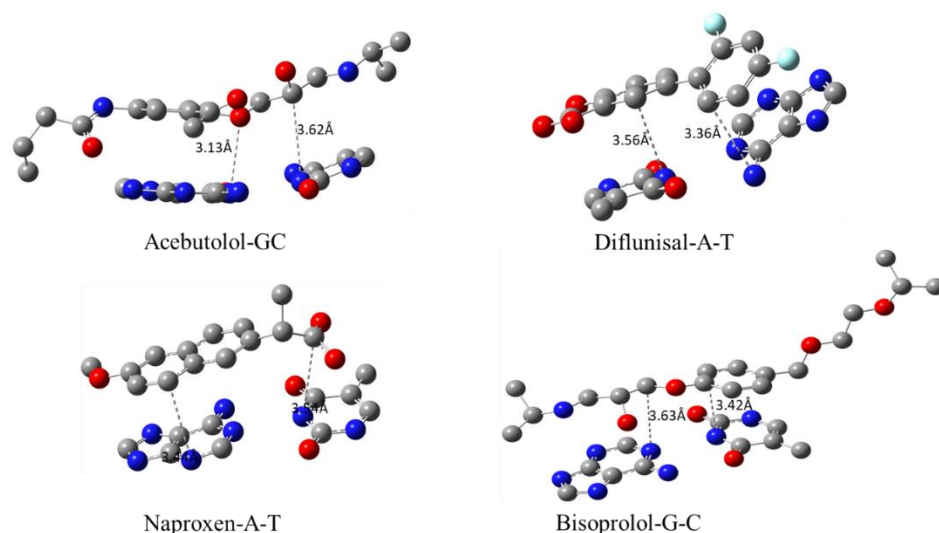


Figure 3. 13 ω B97X-D optimized structures of drug–nucleobase pair complexes

The Figure 3.13 shows the structures of the highly interacting drug nucleobases obtained by the ω B97X-D method. While acebutolol and bisoprolol, which are antihypertensive drugs, show the greatest interaction with the guanine-cytosine base pair, diflunisal and naproxen, which are nonsteroidal anti-inflammatory drugs, show the greatest interaction with the adenine-thymine base pair.

Table 3. 4 ΔH_{bind} (kcal/mol) values of ω B97X-D and B3LYP, B3LYP-D2 optimized the parallel-oriented drug–nucleobase pair complexes

Interreaction site	B3LYP	B3LYP-D2	ω B97X-D
Bisoprolol-G-C	*	11.39	-13.8
Bisoprolol-A-T	1.26	11.77	-13.38
Acebutolol-A-T	-45.02	11.35	-11.87
Acebutolol-G-C	0.63	*	-17.56
Naproxen-A-T	-2.29	3.68	-12.97
Naproxen-G-C	0.85	13.29	-12.79
Diflunisal-A-T	-5.9	13.39	-14.76
Diflunisal-G-C	-6.15	12.6	-14.56

* Calculations did not converge to a stable geometry

After deciding that the interactions were parallel, we repeated the calculations with the B3LYP-D2 method, as we thought that the B3LYP-D2 method would

improve the results. Table 3.4 shows the parallel oriented drug-nucleobase pair complexes optimized at the B3LYP/6-31G(d), B3LYP-D2/6-31G(d), and ω B97X-D/6-31G(d) levels of DFT. All of these complexes are π - π parallel stacking complexes. Here we show that B3LYP underestimates the stability of the complexes as expected, with the exception of acebutolol A-T, as explained above. However, the B3LYP-D2 results are very surprising. Although we performed D2 calculations to improve the B3LYP results, B3LYP-D2 made them worse than in the case of the interaction between nucleobase pair and a drug.

Table 3. 5 ΔH_{bind} (kcal/mol) values of ω B97X-D and B3LYP, B3LYP-D2 optimized the parallel stacking -oriented drug–nucleobase pair complexes

Interreaction site	B3LYP		B3LYP-D2		ω B97X-D	
	6-31-G(d)	6-311++G(d,p)	6-31-G(d)	6-311++G(d,p)	6-31-G(d)	6-311++G(d,p)
Bisoprolol-A-T	1.26	0.71	11.77	11.04	-13.38	*
Bisoprolol-G-C	*	0.73	11.39	11.75	-13.80	*
Acebutolol-A-T	-45.02	*	11.35	*	-11.87	*
Acebutolol-G-C	0.63	0.64	*	*	-17.56	*
Naproxen-A-T	-2.29	2.35	3.68	14.94	-12.97	-12.72
Naproxen-G-C	0.85	0.78	13.29	13.25	-12.79	-12.87
Diflunisal-A-T	-5.90	-0.66	13.39	12.99	-14.76	-15.97
Diflunisal-G-C	-6.15	0.50	12.60	12.55	-14.56	-14.90

*Calculations did not converge to a stable geometry.

Table 3.5 shows the parallel stacking-oriented drug-nucleobase pair complexes optimized at the B3LYP, B3LYP-D2, and ω B97X-D/6-31G(d) levels of DFT. All of these complexes are π - π parallel stacking complexes. This is most likely a false stability due to the inability of the B3LYP method in this calculation. This is because this high binding energy (-45 kcal/mol) is larger than the covalent bond in the F2 molecule, whose bond dissociation energy is about 37 kcal/mol (Wenthold and Squires, 2005). This shows that B3LYP underestimates the stability of the complexes. However, the B3LYP-D2 results are very surprising. Although we performed D2 calculations to improve the B3LYP results, B3LYP-D2 made them worse than in the case of the interaction between a single base and a drug. The B3LYP-D2 results are consistent, but the results are worse. Moreover, we can conclude that the ω B97X-D method gives better results for such non-covalently

interacting systems and also gives more consistent results when the base set is changed. However, to be sure, we also carried out another step.

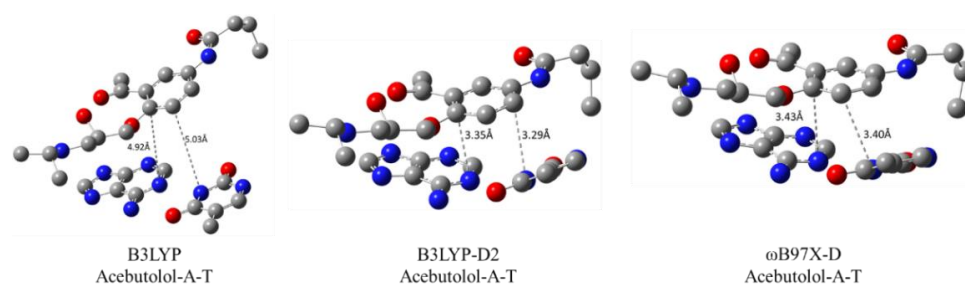


Figure 3. 14 Optimized structures of drug–nucleobase pair complexes

Table 3. 6 Distances between hydrogen bonds of ω B97X-D optimized drug–G-C basepair complexes

Angstrom	O-H	N-H	O-H	difference	difference	difference
G-C	1.840	1.906	1.858	0.000	0.000	0.000
G-C- Acebutolol	1.849	1.905	1.854	0.009	-0.001	-0.004
G-C-Bisoprolol	1.857	1.899	1.866	0.018	-0.007	0.009

Table 3. 7 Distances between hydrogen bonds of ω B97X-D optimized drug– A-T base-pair complexes

Angstrom	O-H	N-H	difference	difference
A-T	1.909	1.861	0.000	0.000
A-T-Diflunisal	1.967	1.857	0.058	-0.004
A-T-Naproxen	1.863	1.896	-0.045	0.034

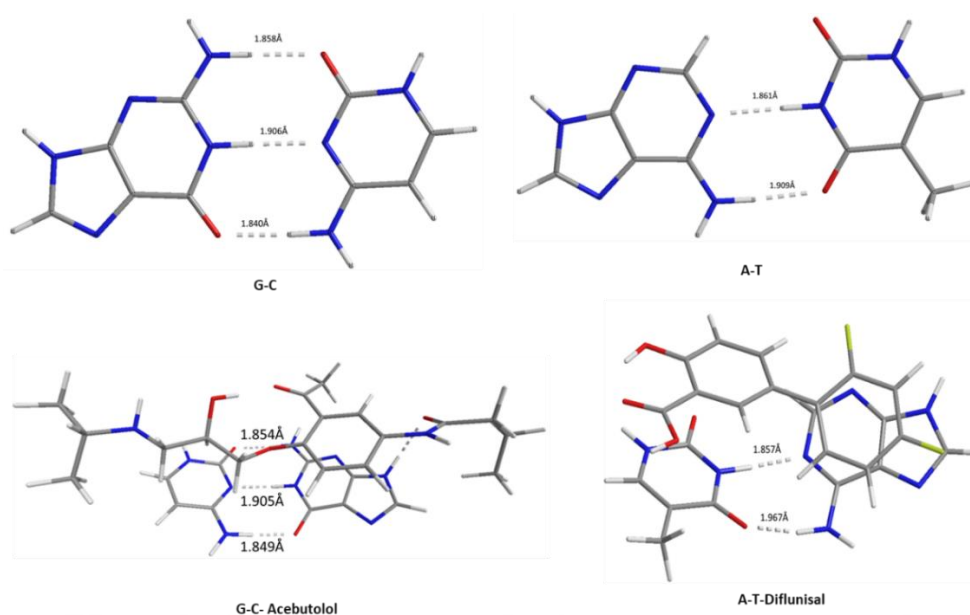


Figure 3. 15 Intermolecular interaction of drug–nucleobase pair complexes

Table 3.6 shows that the acebutolol affects the bond lengths of the hydrogen bonds between the G-C base pair. On the other hand, the diflunisal affects the bond lengths of the hydrogen bonds between the adenine-thymine base pairs, shown in Table 3.7. Accordingly, these drugs might cause DNA breakage in long-term usage (Öztürk et al., 2020).

3.4 The Performance of The Methods for The Non-Covalently Interacted Systems

In 2009, Rutledge et al. studied that parallel stacking and T-shaped interactions between natural DNA or RNA nucleobases (adenine, cytosine, guanine, thymine, uracil) and all aromatic amino acids (histidine, phenylalanine, tyrosine, tryptophan) (Rutledge et al., 2009). They characterized the potential energy surface of nucleobase amino acid dimers using the MP2/6-31G*(0.25) method and the CCSD(T)/CBS level of theory. Since their systems are similar to ours, we used their structures aromatic amino acids (histidine, phenylalanine, tyrosine, tryptophan) with DNA nucleobases to determine which method is more appropriate for our system of aromatic-aromatic interactions. Therefore, the selected examples in this section were optimized using B3LYP, B3LYP-D2, B97-D, ω B97, and ω B97X-D methods. We performed a comparison based on the structures studied by Rutledge et al.) Our structures were optimized with B3LYP,

B3LYP-D2, B97-D, ω B97, and ω B97X-D methods. We compared the results of the methods mentioned above with those of the CCSD(T) method considered the gold standard for these types of comparisons. In order to quantitatively determine the closeness of the methods used in this study to the performance of the CCSD(T) method, we calculated the root mean square error (RMSE). The equation below shows eqn 3.3 the formula of the RMSE value

$$\text{RMSE} = \sqrt{\sum \left(\frac{\text{current value} - \text{reference value}}{\text{reference value}} \right)^2} \quad (3.3)$$

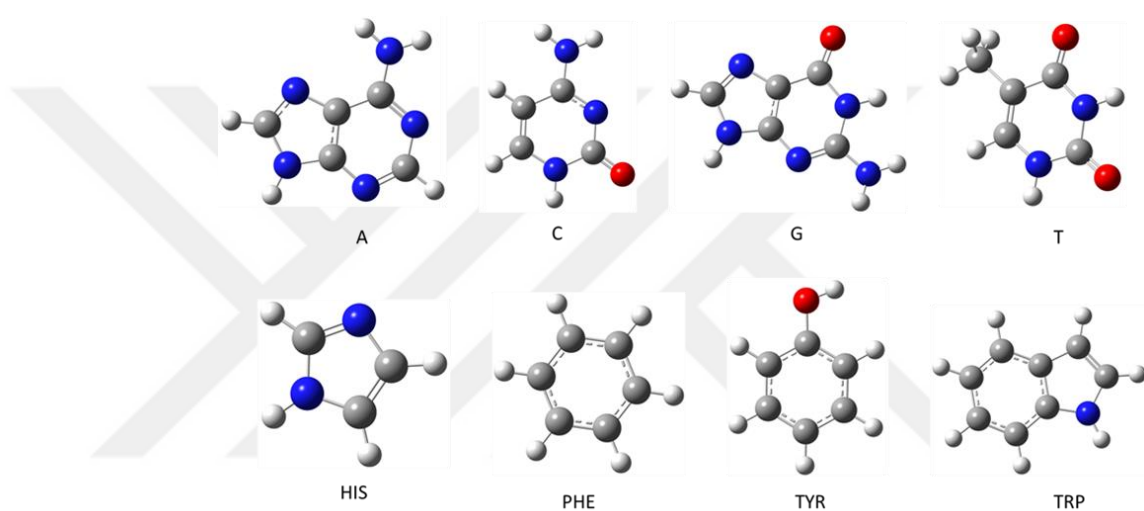


Figure 3. 16 Structures and definition of edges of the aromatic amino acids (histidine (HIS), phenylalanine (PHE), tyrosine (TYR), tryptophan (TRP)) and the natural nucleobases (cytosine (C), adenine (A), guanine (G), thymine (T))

Table 3. 8 B3LYP, MP2/6-31G*(0.25) and CCSD(T)/CBS Parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases

Adenine							
	B3LYP/ 6- 31G(d)	B3LYP/ 6- 31+G(d)	B3LYP/ 6- 31+G(d,p)	B3LYP/ 6- 311++G(d,p)	MP2/ 6- 31G*(0.25)	CCSD(T)/ CBS	Percent Error 6-31+G(d,p)
HIS	0.70	0.27	0.01	0.30	-7.10	-7.12	100.14
PHE	0.99	0.33	0.34	0.12	-5.81	-5.57	106.04
TYR	1.07	0.40	0.44	0.48	-6.91	-6.79	106.50
TRP	1.05	0.31	0.17	0.28	-7.65	-7.31	102.36
Cytosine							
	B3LYP/ 6- 31G(d)	B3LYP/ 6- 31+G(d)	B3LYP/ 6- 31+G(d,p)	B3LYP/ 6- 311++G(d,p)	MP2/ 6- 31G*(0.25)	CCSD(T)/ CBS	
HIS	0.89	0.23	0.35	0.19	-6.43	-7.05	105.02
PHE	0.91	0.30	0.34	0.40	-4.40	-4.92	106.83
TYR	0.77	0.60	0.55	0.44	-5.78	-6.48	108.51
TRP	1.00	0.42	0.33	0.30	-7.98	-8.46	103.88
Guanine							
	B3LYP/ 6- 31G(d)	B3LYP/ 6- 31+G(d)	B3LYP/ 6- 31+G(d,p)	B3LYP/ 6- 311++G(d,p)	MP2/ 6- 31G*(0.25)	CCSD(T)/ CBS	
HIS	0.90	0.41	0.33	0.25	-8.44	-9.03	103.69
PHE	1.06	0.37	0.25	0.41	-6.05	-6.05	104.14
TYR	1.17	0.49	0.45	0.57	-7.84	-8.41	105.31
TRP	1.25	0.30	0.45	0.32	-10.13	-10.06	104.50
Thymine							
	B3LYP/ 6- 31G(d)	B3LYP/ 6- 31+G(d)	B3LYP/ 6- 31+G(d,p)	B3LYP/ 6- 311++G(d,p)	MP2/ 6- 31G*(0.25)	CCSD(T)/ CBS	
HIS	1.01	0.26	0.24	0.34	-5.98	-6.48	103.71
PHE	1.17	0.25	0.22	0.38	-5.35	-5.86	103.70
TYR	1.47	0.44	0.34	0.46	-6.24	-6.76	104.99
TRP	1.19	0.29	0.26	0.32	-8.60	-9.03	102.88
						RMS	104.53

As you can see from the table 3.8, we analyzed histidine, phenylalanine, tyrosine, and tryptophan with DNA nucleobases using the DFT-B3LYP method for four basis sets. B3LYP has been seen to give always positively predict interaction energies, even for large basis sets. However, the CCSD (T) method we use as a reference show that all these interaction energies must be negative. In addition, the RMSE value of the B3LYP method has a very large value of 104.53%. Considering all this, it can be concluded that the B3LYP method does not perform well in systems involving remote interactions.

Table 3. 9 B3LYP-D2, MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases

Adenine						
B3LYP-D2/ 6-31G(d)	B3LYP-D2/ 6-31+G(d)	B3LYP-D2/ 6-31+G(d,p)	B3LYP/D2 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	Percent Error 6-31+G(d,p)
5.66	4.96	4.97	5.17	-7.10	-7.12	169.77
8.16	6.68	6.84	6.78	-5.81	-5.57	222.77
9.17	7.50	7.55	7.87	-6.91	-6.79	211.25
9.55	8.39	8.56	8.56	-7.65	-7.31	217.08
Cytosine						
B3LYP-D2/ 6-31G(d)	B3LYP-D2/ 6-31+G(d)	B3LYP-D2/ 6-31+G(d,p)	B3LYP/D2 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	
5.86	4.29	4.28	4.20	-6.43	-7.05	160.65
6.69	5.37	4.63	5.36	-4.40	-4.92	194.07
0.82	2.85	6.04	6.31	-5.78	-6.48	193.26
8.02	7.26	6.92	8.23	-7.98	-8.46	181.80
Guanine						
B3LYP-D2/ 6-31G(d)	B3LYP-D2/ 6-31+G(d)	B3LYP-D2/ 6-31+G(d,p)	B3LYP/D2 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	
5.25	5.48	5.53	5.69	-8.44	-9.03	161.24
6.86	5.77	6.22	6.39	-6.05	-6.05	202.81
8.30	6.55	6.82	7.18	-7.84	-8.41	181.05
5.98	7.05	7.13	8.10	-10.13	-10.06	170.84
Thymine						
B3LYP-D2/ 6-31G(d)	B3LYP-D2/ 6-31+G(d)	B3LYP-D2/ 6-31+G(d,p)	B3LYP/D2 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	
5.04	4.41	4.41	5.28	-5.98	-6.48	168.07
6.87	5.65	5.63	5.73	-5.35	-5.86	196.18
7.34	6.25	6.23	6.48	-6.24	-6.76	192.09
8.34	7.21	7.21	7.42	-8.60	-9.03	179.81
RMSE						188.60

The dispersion correction models significantly improve the original B3LYP results, so we expect B3LYP-D2 to give a relatively better result than B3LYP(Chen, Young, and Hsu, 2009)(Burns, et al.,2011) In our preliminary study, we showed that the B3LYP-D2 method gave worse results than the B3LYP. In this part, the RMSE value being 188.60% is a proof that the B3LYP-D2 method cannot model nucleobase interactions with aromatic amino acids well.

Table 3. 10 ω B97X-D, MP2/6-31G*(0.25) and CCSD(T)/CBS Parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases

Adenine							
	ω B97X-D/ 6-31G(d)	ω B97X-D/ 6-31+G(d)	ω B97X-D/ 6-31+G(d,p)	ω B97X-D/ 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	Percent Error 6-31+G(d,p)
HIS	-3.29	-3.98	-3.72	-3.93	-7.10	-7.12	47.77
PHE	-3.38	-4.30	-4.25	-4.73	-5.81	-5.57	23.67
TYR	-3.63	-4.30	-4.40	-4.52	-6.91	-6.79	35.21
TRP	-4.85	-5.34	-5.14	-5.73	-7.65	-7.31	29.71
Cytosine							
	ω B97X-D/ 6-31G(d)	ω B97X-D/ 6-31+G(d)	ω B97X-D/ 6-31+G(d,p)	ω B97X-D/ 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	
HIS	-2.61	-3.04	-2.99	-3.22	-6.43	-7.05	57.57
PHE	-2.50	-3.44	-3.43	-3.84	-4.40	-4.92	30.42
TYR	-2.56	-4.03	-4.09	-3.96	-5.78	-6.48	36.88
TRP	-3.41	-4.65	-4.60	-5.16	-7.98	-8.46	45.61
Guanine							
	ω B97X-D/ 6-31G(d)	ω B97X-D/ 6-31+G(d)	ω B97X-D/ 6-31+G(d,p)	ω B97X-D/ 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	
HIS	-3.19	-3.69	-3.71	-3.89	-8.44	-9.03	58.90
PHE	-3.72	-4.63	-4.63	-5.19	-6.05	-6.05	23.40
TYR	-3.65	-4.67	-4.80	-4.69	-7.84	-8.41	42.93
TRP	-4.48	-5.74	-5.77	-5.90	-10.13	-10.06	42.61
Thymine							
	ω B97X-D/ 6-31G(d)	ω B97X-D/ 6-31+G(d)	ω B97X-D/ 6-31+G(d,p)	ω B97X-D/ 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	
HIS	-2.90	-3.84	-3.85	-3.73	-5.98	-6.48	40.57
PHE	-3.82	-4.57	-4.60	-5.04	-5.35	-5.86	21.39
TYR	-4.10	-4.88	-4.80	-4.81	-6.24	-6.76	29.01
TRP	-5.44	-7.04	-7.14	-6.57	-8.60	-9.03	21.01
RMSE							38.49

The ω B97X-D function approaches the performance of MP2/6-31G*(0.25) and CCSD(T)/CBS. The ω B97X-D hybrid functionals are dispersion and long-range corrected and are better for non-covalent interactions. The RMSE value of the ω B97X-D method is 38.49%, which is closest to the result of the CCSD(T)/CBS method. We can conclude that ω B97X-D has better performance for dispersion-bound systems among the methods that can perform geometry optimization and its results are more reliable than those of the previous two methods.

Table 3.11 ω B97, MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases

	Adenine						Percent Error
	ω B97	ω B97	ω B97	ω B97	MP2/	CCSD(T)/	
	6-31G(d)	6-31+G(d)	6-31+G(d,p)	6-311++G(d,p)	6-31G*(0.25)	CBS	
HIS	-1.46	-2.90	-2.70	-3.20	-7.10	-7.12	62.10
PHE	-1.45	-3.07	-3.00	-3.04	-5.81	-5.57	46.05
TYR	-1.45	-3.08	-3.12	-3.22	-6.91	-6.79	53.96
TRP	-2.17	-3.75	-3.61	-3.70	-7.65	-7.31	50.66
	Cytosine						Percent Error
	ω B97	ω B97	ω B97	ω B97	MP2/	CCSD(T)/	
	6-31G(d)	6-31+G(d)	6-31+G(d,p)	6-311++G(d,p)	6-31G*(0.25)	CBS	
HIS	-1.42	-2.29	-2.30	-2.29	-6.43	-7.05	67.44
PHE	-1.03	-2.57	-2.56	-2.68	-4.40	-4.92	47.92
TYR	-0.87	-2.98	-3.00	-2.50	-5.78	-6.48	53.62
TRP	-1.42	-3.19	-3.24	-3.57	-7.98	-8.46	61.71
	Guanine						Percent Error
	ω B97	ω B97	ω B97	ω B97	MP2/	CCSD(T)/	
	6-31G(d)	6-31+G(d)	6-31+G(d,p)	6-311++G(d,p)	6-31G*(0.25)	CBS	
HIS	-1.61	-2.98	-2.89	-12.60	-3.01	-9.03	68.06
PHE	-1.69	-3.14	-3.20	-3.32	-6.05	-6.05	47.09
TYR	-4.44	-3.36	-3.11	-3.13	-7.84	-8.41	63.00
TRP	-1.70	-3.73	-3.61	-3.82	-10.13	-10.06	64.17
	Thymine						Percent Error
	ω B97	ω B97	ω B97	ω B97	MP2/	CCSD(T)/	
	6-31G(d)	6-31+G(d)	6-31+G(d,p)	6-311++G(d,p)	6-31G*(0.25)	CBS	
HIS	-1.61	-2.53	-2.50	-10.32	-2.47	-6.48	61.45
PHE	-2.12	-3.35	-3.40	-3.45	-5.35	-5.86	41.90
TYR	-2.29	-3.79	-3.76	-3.68	-6.24	-6.76	44.44
TRP	-3.23	-4.65	-4.70	-4.84	-8.60	-9.03	47.98
						RMSE	55.75

The 2008 study by M. (Chai and Head-Gordon, 2008) showed that the optimized parameters of the ω B97X-D method give slightly better results than those of ω B97. Thus, one of the methods that can predict weak interactions well is also ω B97. The RMSE value of the ω B97 method is 55.75%, so the ω B97 method has a better performance compared to B3LYP and B3LYP-D2.

Table 3.12 B97-D, MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases

	Adenine				MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	Percent Error 6-31+G(d,p)
	B97D 6-31G(d)	B97D 6-31+G(d)	B97D 6-31+G(d,p)	B97D 6-311++G(d,p)			
HIS	-3.43	-3.30	-3.46	0.97	-7.10	-7.12	51.38
PHE	-2.74	-3.55	-3.48	1.09	-5.81	-5.57	37.56
TYR	-2.79	-3.44	-3.44	1.43	-6.91	-6.79	49.33
TRP	-4.31	-4.95	-4.84	1.56	-7.65	-7.31	33.83
	Cytosine				MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	
	B97D 6-31G(d)	B97D 6-31+G(d)	B97D 6-31+G(d,p)	B97D 6-311++G(d,p)			
HIS	*	-2.78	*	0.88	-6.43	-7.05	
PHE	-1.97	-2.60	-2.66	0.95	-4.40	-4.92	45.89
TYR	-2.28	-2.85	-2.79	1.31	-5.78	-6.48	56.91
TRP	-3.03	-3.77	-4.01	1.33	-7.98	-8.46	52.57
	Guanine				MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	
	B97D 6-31G(d)	B97D 6-31+G(d)	B97D 6-31+G(d,p)	B97D 6-311++G(d,p)			
HIS	-3.55	-3.39	-3.44	1.15	-8.44	-9.03	61.88
PHE	-2.96	-3.61	-3.67	1.19	-6.05	-6.05	39.24
TYR	-3.18	-3.85	-3.86	1.49	-7.84	-8.41	54.15
TRP	-6.38	-6.40	-6.25	1.59	-10.13	-10.06	37.89
	Thymine				MP2/ 6-31G*(0.25)	CCSD(T)/ CBS	
	B97D 6-31G(d)	B97D 6-31+G(d)	B97D 6-31+G(d,p)	B97D 6-311++G(d,p)			
HIS	-2.80	-2.96	-2.99	0.98	-5.98	-6.48	53.86
PHE	-3.17	-3.70	-3.71	1.05	-5.35	-5.86	36.61
TYR	-3.49	-3.86	-3.79	1.45	-6.24	-6.76	44.03
TRP	-5.22	-6.13	-6.04	1.41	-8.60	-9.03	33.10
RMSE							45.24

As you can see from Table 3.12, B97D also describes the dispersion well. ω B97-families and B97D provide a relatively good estimate of these weak interactions. The RMSE value of the B97D method with 6-31+G(d,p) basis set is 45.24%. However, the B97D method gives positive interaction energies with the large basis set, 6-311++G(d,p), making the results worse. Increasing the basis set does not improve interaction energy for B97D (Chai and Head-Gordon, 2008). Therefore, employing this functional in weakly interacting systems may cause severe errors.

Table 3. 13 ω B97X-D ω B97M-V, ω B97X-V, MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases

	Adenine							CCSD(T)/ CBS
	ω B97X-D/ 6-31+G(d,p)	ω B97X-D/ 6-311++G(d,p)	wb97M-V def2-SVP	wb97M-V def2-TZVPP	wB97X-V def2-SVP	wB97X-V def2-TZVPP	MP2/ 6- 31G*(0.25)	
HIS	-3.72	-3.93	-9.66	-7.59	-9.06	-7.32	-7.10	-7.12
PHE	-4.25	-4.73	-7.64	-5.85	-6.90	-5.49	-5.81	-5.57
TYR	-4.40	-4.52	-8.25	-6.14	-7.42	-5.72	-6.91	-6.79
TRP	-5.14	-5.73	-11.77	-9.39	-10.80	-8.81	-7.65	-7.31
	Cytosine							CCSD(T)/ CBS
	ω B97X-D/ 6-31+G(d,p)	ω B97X-D/ 6-311++G(d,p)	wb97M-V def2-SVP	wb97M-V def2-TZVPP	wB97X-V def2-SVP	wB97X-V def2-TZVPP	MP2/ 6- 31G*(0.25)	
HIS	-2.99	-3.22	-9.87	-8.09	-9.36	-7.82	-6.43	-7.05
PHE	-3.43	-3.84	-7.34	-5.93	-6.77	-5.62	-4.40	-4.92
TYR	-4.09	-3.96	-11.55	-8.74	-10.83	-8.33	-5.78	-6.48
TRP	-4.60	-5.16	-12.17	-10.02	-11.34	-9.56	-7.98	-8.46
	Guanine							CCSD(T)/ CBS
	ω B97X-D/ 6-31+G(d,p)	ω B97X-D/ 6-311++G(d,p)	wb97M-V def2-SVP	wb97M-V def2-TZVPP	wB97X-V def2-SVP	wB97X-V def2-TZVPP	MP2/ 6- 31G*(0.25)	
HIS	-3.71	-3.89	-12.11	-10.17	-11.49	-9.85	-8.44	-9.03
PHE	-4.63	-5.19	-8.25	-6.42	-7.53	-6.09	-6.05	-6.05
TYR	-4.80	-4.69	-10.22	-8.06	-9.48	-7.66	-7.84	-8.41
TRP	-5.77	-5.90	-10.33	-7.87	-9.41	-7.46	-10.13	-10.06
	Thymine							CCSD(T)/ CBS
	ω B97X-D/ 6-31+G(d,p)	ω B97X-D/ 6-311++G(d,p)	wb97M-V def2-SVP	wb97M-V def2-TZVPP	wB97X-V def2-SVP	wB97X-V def2-TZVPP	MP2/ 6- 31G*(0.25)	
HIS	-3.85	-3.73	-9.39	-7.84	-8.91	-7.58	-5.98	-6.48
PHE	-4.60	-5.04	-8.90	-7.33	-8.29	-6.99	-5.35	-5.86
TYR	-4.80	-4.81	-10.50	-8.35	-9.77	-7.94	-6.24	-6.76
TRP	-7.14	-6.57	-14.96	-12.45	-14.02	-11.89	-8.60	-9.03
RMSE	38.49	36.21	46.05	20.79	36.67	17.28		

We can conclude that the ω B97X-D functional performs significantly better than B3LYP and B3LYP -D2. Our study constitutes another indicative of the fact that ω B97X-D works well in dispersion bond systems. The 2008 study by M. Head-Gordon et al. showed that the optimized parameters (S22 data set) of ω B97X-D were slightly larger for non-covalent interactions than for ω B97(Chai and Head-Gordon, 2008). Therefore, one of the methods that can predict weak interactions well is ω B97. In addition, B97-D also describes the dispersion well. However, the overall performance of the B97-D method does not seem to be able to provide reliable results. This result indicates that B97 is not good for non-covalent interactions. In 2017, N. Mardirossian and M. Head-Gordon investigated non-covalent interactions, isomerization energies, thermochemistry, and barrier heights in 84 data sets. Their results classify density functionals based on a series of S66 data sets.(Mardirossian and Head-Gordon, 2017) The ω B97M-V and ω B97X-V

methods rank 12th and 24th, respectively, in the S66 database, which includes a range of bioinspired molecular structures (e.g., amino acids from proteins or nucleic acids from DNA) that have the most common non-covalent interactions (including dispersion, electrostatic, and mixed systems). Their research shows that the ω B97M-V and ω B97X-V methods give better results with the S66 and S66x8 database pools. Using this knowledge, we decided using these functionals in our study. Unfortunately, Gaussian09 does not include these methods. The ω B97M-V and ω B97X-V methods have been coded in ORCA software without gradients. This means that only single point energy calculations can be performed with these methods.

Therefore, the single-point energy calculations of amino acids and nucleobases were performed. We used the optimized geometries obtained from ω B97X-D/6-31+G(d,p) for single-point energy calculations with ω B97M-V and ω B97X-V. As seen from Table 3.13, ω B97M-V and ω B97X-V describe the dispersion well, also. The RMSE value of the ω B97M-V/def2-TZVPP method is 20.79%, while the RMSE value of ω B97X-V is 17.28, therefore it is clear that the ω B97M-V and ω B97X-V methods are the best DFT methods describing weakly interacting systems. However, as mentioned before, the gradients of these two methods have not been coded in ORCA, yet. Hence, the ω B97X-D method, which produces energies closer to the CCSD(T) method, is the best choice that can perform geometry optimization.

Table 3. 14 B3LYP/6-31+G(d,p), B3LYP-D2/6-31+G(d,p), ω B97X-D/6-31+G(d,p), MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases

	Adenine						
	B3LYP/ 6-31+G(d,p)	B3LYP-D2/ 6-31+G(d,p)	B97D 6-31+G(d,p)	ω B97 6-31+G(d,p)	ω B97X-D/ 6-31+G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS
HIS	0.01	4.97	-3.46	-2.70	-3.72	-7.10	-7.12
PHE	0.34	6.84	-3.48	-3.00	-4.25	-5.81	-5.57
TYR	0.44	7.55	-3.44	-3.12	-4.40	-6.91	-6.79
TRP	0.17	8.56	-4.84	-3.61	-5.14	-7.65	-7.31
	Cytosine						
	B3LYP/ 6-31+G(d,p)	B3LYP-D2/ 6-31+G(d,p)	B97D 6-31+G(d,p)	ω B97 6-31+G(d,p)	ω B97X-D/ 6-31+G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS
HIS	0.35	4.28	*	-2.30	-2.99	-6.43	-7.05
PHE	0.34	4.63	-2.66	-2.56	-3.43	-4.40	-4.92
TYR	0.55	6.04	-2.79	-3.00	-4.09	-5.78	-6.48
TRP	0.33	6.92	-4.01	-3.24	-4.60	-7.98	-8.46
	Guanine						
	B3LYP/ 6-31+G(d,p)	B3LYP-D2/ 6-31+G(d,p)	B97D 6-31+G(d,p)	ω B97 6-31+G(d,p)	ω B97X-D/ 6-31+G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS
HIS	0.33	5.53	-3.44	-2.89	-3.71	-8.44	-9.03
PHE	0.25	6.22	-3.67	-3.20	-4.63	-6.05	-6.05
TYR	0.45	6.82	-3.86	-3.11	-4.80	-7.84	-8.41
TRP	0.45	7.13	-6.25	-3.61	-5.77	-10.13	-10.06
	Thymine						
	B3LYP/ 6-31+G(d,p)	B3LYP-D2/ 6-31+G(d,p)	B97D 6-31+G(d,p)	ω B97 6-31+G(d,p)	ω B97X-D/ 6-31+G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS
HIS	0.24	4.41	-2.99	-2.50	-3.85	-5.98	-6.48
PHE	0.22	5.63	-3.71	-3.40	-4.60	-5.35	-5.86
TYR	0.34	6.23	-3.79	-3.76	-4.80	-6.24	-6.76
TRP	0.26	7.21	-6.04	-4.70	-7.14	-8.60	-9.03
RMSE	104.53	188.60	45.24	55.75	38.49		

Table 3.14 shows that the B3LYP functional cannot evaluate dispersion interactions appropriately, and the errors of the B3LYP method are closely related to its inadequacy in describing van der Waals interactions. The dispersion correction models are expected to significantly improve the original B3LYP results. Therefore, we expected that B3LYP-D2 would give a relatively better result than B3LYP. However, the B3LYP-D2 method did not improve the results. The RMSE value is 188.60% for B3LYP-D2. This failure can be explained by the fact that the B3LYP and B3LYP-D2 methods systematically underestimate the electron binding energies of complexes of DNA bases (Chen et al., 2009).

On the other hand, the ω B97X-D functional whose results approaches to the results of MP2/6-31G*(0.25) and CCSD(T)/CBS are more powerful than previously mentioned two methods. In other words, the ω B97X-D hybrid functional is

dispersion and long-range corrected, so it is considerably better for non-covalent interactions compared to the B3LYP and B3LYP-D2 methods. The RMSE value of the ω B97X-D method is 38.49%, meaning that its results are closest to the result of the CCSD(T)/CBS method. We can conclude that ω B97X-D performs better in dispersion-bound systems and its results are more reliable. Besides, ω B97 is one of the methods that can predict weak interactions well. This method, whose results are better than those of B3LYP, gives relatively close results to the CCSD(T)/CBS method (RMSE value 55.75%). Although B97-D with 6-31+G(d,p) basis set also describes the dispersion features good since its RMSE value is 45.24%, as the basis set is enlarged its results become rather unexpectedly worse (it will be explained in detail in the next paragraph).



Table 3. 15 B3LYP/6-311++G(d,p), B3LYP-D2/6-311++G(d,p), ω B97X-D/6-311++G(d,p), ω B97M-V/def2-TZVPP, ω B97X-V/def2-TZVPP, MP2/6-31G*(0.25) and CCSD(T)/CBS parallel Stacking Interaction Energies (kcal/mol) between the Aromatic Amino Acids and the Natural DNA Nucleobases

	Adenine						
	B3LYP/ 6-311++G(d,p)	B3LYP/D2 6-311++G(d,p)	B97D 6-311++G(d,p)	ω B97 6-311++G(d,p)	ω B97X-D/ 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS
HIS	0.30	5.17	0.97	-3.20	-3.93	-7.10	-7.12
PHE	0.12	6.78	1.09	-3.04	-4.73	-5.81	-5.57
TYR	0.48	7.87	1.43	-3.22	-4.52	-6.91	-6.79
TRP	0.28	8.56	1.56	-3.70	-5.73	-7.65	-7.31

	Adenine						
	B3LYP/ 6-311++G(d,p)	B3LYP/D2 6-311++G(d,p)	B97D 6-311++G(d,p)	ω B97 6-311++G(d,p)	ω B97X-D/ 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS
HIS	0.19	4.20	0.88	-2.29	-3.22	-6.43	-7.05
PHE	0.40	5.36	0.95	-2.68	-3.84	-4.40	-4.92
TYR	0.44	6.31	1.31	-2.50	-3.96	-5.78	-6.48
TRP	0.30	8.23	1.33	-3.57	-5.16	-7.98	-8.46

	Adenine						
	B3LYP/ 6-311++G(d,p)	B3LYP/D2 6-311++G(d,p)	B97D 6-311++G(d,p)	ω B97 6-311++G(d,p)	ω B97X-D/ 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS
HIS	0.25	5.69	1.15	-12.60	-3.89	-8.44	-9.03
PHE	0.41	6.39	1.19	-3.32	-5.19	-6.05	-6.05
TYR	0.57	7.18	1.49	-3.13	-4.69	-7.84	-8.41
TRP	0.32	8.10	1.59	-3.82	-5.90	-10.13	-10.06

	Adenine						
	B3LYP/ 6-311++G(d,p)	B3LYP/D2 6-311++G(d,p)	B97D 6-311++G(d,p)	ω B97 6-311++G(d,p)	ω B97X-D/ 6-311++G(d,p)	MP2/ 6-31G*(0.25)	CCSD(T)/ CBS
HIS	0.34	5.28	0.98	-10.32	-3.73	-5.98	-6.48
PHE	0.38	5.73	1.05	-3.45	-5.04	-5.35	-5.86
TYR	0.46	6.48	1.45	-3.68	-4.81	-6.24	-6.76
TRP	0.32	7.42	1.41	-4.84	-6.57	-8.60	-9.03
RMSE	104.98	193.49	117.51	52.98	36.21		

All reported interaction energies include corrections for basis set superposition error (BSSE) (Boys and Bernardi, 1970). BSSE improves interaction energy, in this way it affects the DFT results. When we compare the basis sets using the same methods, there is not much difference between the results of the B3LYP, ω B97 and ω B97X-D. Besides, their results improved. However, the B3LYP-D2 and B97-D perform worse with large basis set. When the two method with larger basis sets are examined, the RMSE value of B97D/6-311++G(d,p) is 117.51% and the RMSE value of B97D/6-31+G(d,p) is 45.24%. The difference in RMSE values is almost 70%. This result suggests that the B97-D method is not suitable for non-covalent interactions.

These results make it clear that the ω B97X-D method is the best method that gives energies rather close to the CCSD(T)/CBS method among the methods that we can perform optimization. In addition, the functional which has the second-best performance is ω B97. Many complex systems require economic calculations. Optimizations with the larger basis sets generally give quantitatively better results, but in our calculations, they did not affect the results considerably. However, the larger basis sets take much longer processing time. Thus, we continued our calculations with smaller basis set, 6-31+G(d,p) employing ω B97X-D method.

3.5 Computational Investigation Drug-Nucleobase and Nucleobase Pair Interaction with ω B97X-D Method

3.5.1 Drug – Nucleobase Binding Energies

In this part, we have studied the effects of each drug on a single nucleobase to determine the drug-nucleobase interaction. Because π - π parallel stacking interactions play a fundamental role in chemical and biological recognition (Meyer, Castellano, and Diederich, 2003). Therefore, all drug-nucleobase complexes were optimized at the ω B97X-D /6-31+G(d,p) levels of DFT. The binding energies were calculated using eqn 3.4 as in the previous section. The effect of implicit water solvation on complex formation is studied using the integral equation formalism of the Polarized Continuum Model (IEF-PCM) with the Gaussian program.

$$\Delta E_{bind} = E_0(complex) - (E_0(drug) + E_0(base)) \quad (3.4)$$

Table 3. 16 Binding energies (kcal/mol) for drug-nucleobase complexes with ω B97X-D /6-31+G(d,p) levels

	ΔE_0	$\Delta E_0 + \text{BSSE}$	ΔE_0	$\Delta E_0 + \text{BSSE}$	ΔE_0	$\Delta E_0 + \text{BSSE}$	ΔE_0	$\Delta E_0 + \text{BSSE}$
Drug Name	A	A	C	C	G	G	T	T
Acebutolol	-13.34	-11.38	-9.04	-7.56	-10.56	-8.50	-8.59	-7.05
Atenolol	-8.23	-6.94	-8.04	-6.71	-9.57	-7.99	-8.12	-6.86
Betaxolol	-7.65	-6.35	-6.09	-4.97	-8.50	-6.98	-8.31	-7.05
Bisoprolol	-8.07	-6.67	-7.49	-6.02	-12.33	-10.32	-11.71	-10.10
Bromfenac	-9.08	-7.22	-6.44	-4.76	-9.86	-7.82	-9.71	-7.80
Celecoxib	-8.87	-7.02	-8.20	-6.38	-9.14	-7.08	-8.89	-7.16
Celiprolol	-10.03	-8.32	-11.90	-9.97	-16.03	-13.72	-8.94	-7.38
Diflunisal	-7.94	-6.65	-7.66	-6.30	-8.41	-6.82	-8.87	-7.39
Esmolol	-8.19	-6.83	-7.01	-5.71	-8.80	-7.26	-8.08	-6.74
Indometacin	-12.27	-10.36	-9.95	-8.22	-11.51	-9.41	-10.66	-8.93
Mefenamic Acid	-10.02	-8.56	-8.29	-6.81	-10.56	-8.90	-8.22	-6.58
Meloxicam	-7.10	-5.07	-10.08	-8.17	-13.05	-10.67	-9.54	-7.65
Metoprolol	-8.25	-6.83	-7.55	-5.84	-12.23	-10.18	-8.10	-6.79
Nabumetone	-8.89	-7.47	-7.81	-6.56	-9.64	-8.03	-9.02	-7.67
Naproxen	-10.46	-8.47	-4.51	-3.14	-9.48	-7.79	-9.10	-7.80
Nebivolol	-12.57	-11.00	-10.49	-8.27	-10.46	-8.69	-9.04	-7.62
Nepafenac	-9.26	-7.63	-8.57	-7.04	-9.57	-7.78	-8.95	-7.26
Oxaprozin	-13.17	-11.19	-7.58	-5.87	-14.06	-11.79	-12.77	-10.79

The basis set superposition error (BSSE) was applied to the optimized structures for compensating inconsistency of the energies. The basis functions of molecules that interact with one another or that are different components of the same molecule (intramolecular BSSE) overlap when the atoms move toward one another. Each monomer effectively expands its basis set and enhances the computation of derived attributes like energy by "borrowing" functions from other adjacent components. This leads to an error. In order to determine BSSE, all calculations are redone using mixed basis sets, and the error is then removed a posteriori from the uncorrected energy (Copeland, et al., 2013). Table 3.16 shows the effect of BSSE corrections on the interaction energies. As expected, there was a decrease in all binding energies due to the correction of over stabilization originated from basis sets. In general, it is seen that the BSSE corrections did not cause significant differences in energies of all drug-nucleobase interactions compared to magnitude of interaction energies.

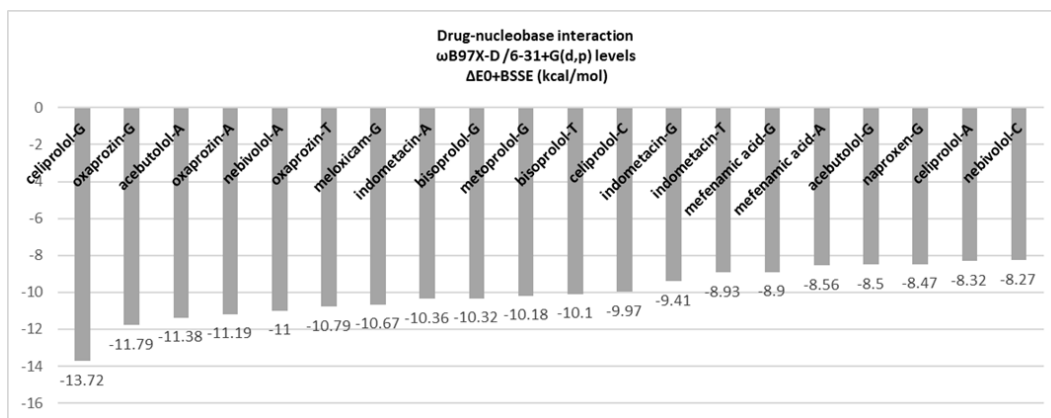


Figure 3. 17 Binding energy (kcal/mol) values rank of the top strong 20 drug–nucleobase complexes optimized with the IEF-PCM/ ω B97X-D /6-31+G(d,p) methods

Celiprolol, oxaprozin, acebutolol, celiprolol, nebivolol, and meloxicam have the highest interaction energy of all drug molecules, as shown in the Figure 3.17. Accordingly, celiprolol and guanine have the most negative binding energy among all individual nucleobase–drug complexes. The complex with the second strongest interaction is the oxaprozin–guanine complex. The following complexes are acebutolol–adenine, nebivolol–adenine, and meloxicam–guanine complexes.

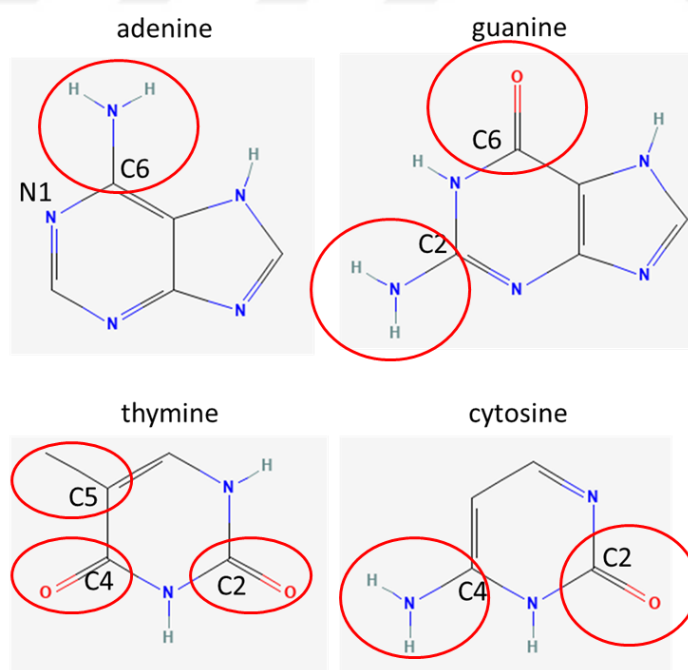


Figure 3. 18 Structures of DNA nucleobases

While celiprolol, acebutolol, and nebivolol are antihypertensive drugs and have the strongest interactions with their respective guanine and adenine nucleobases, oxaprozin and meloxicam are nonsteroidal anti-inflammatory drugs that have the strongest interactions with guanine nucleobases. The main difference between adenine and guanine is that adenine contains an amino group at C6 and an additional double bond between N1 and C6 in its pyrimidine ring, whereas guanine contains an amino group at C2 and a carbonyl group at C6 in its pyrimidine ring (Pediaa, 2021). Two functional groups are present in guanine that cause strong drug interactions. Cytosine has a keto group at C-2 and an amino group at C-4 in the heterocyclic aromatic ring. The heterocyclic aromatic ring of thymine contains two keto groups at C-2 and C-4 and a methyl group at C-5 (Pediaa, 2021). Knowing these basic nucleobase properties lead us to recognize these strong interactions. However, the strong binding energy of the drug with a nucleobase is not sufficient to understand whether this drug can interact with DNA that forms from nucleobase pairs or not.

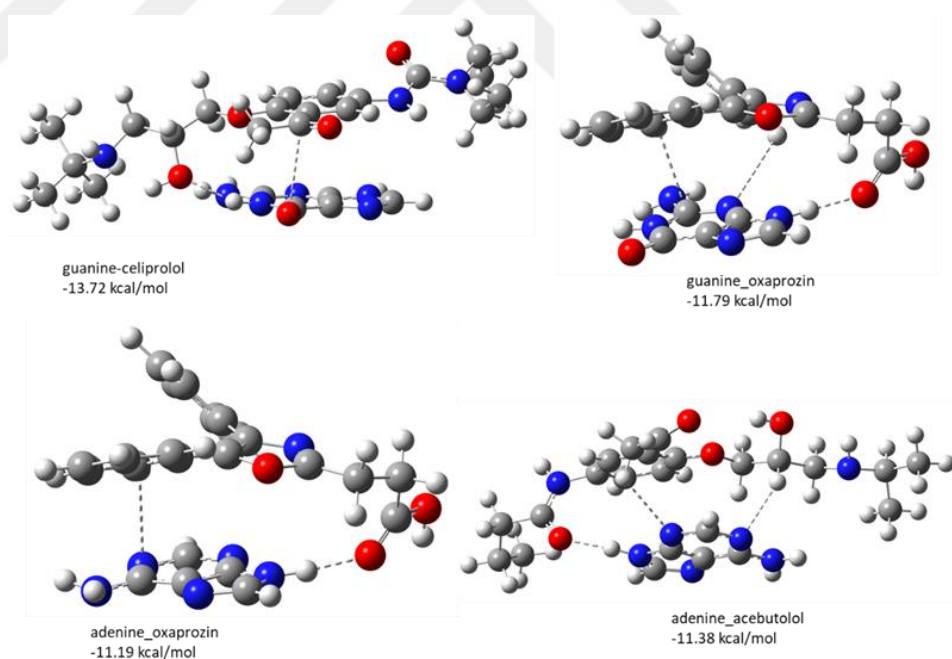


Figure 3. 19 The strongest binding energy of celiprolol, acebutolol, and oxaprozin – nucleobase complexes with the IEF-PCM/ ω B97X-D /6-31+G(d,p) methods

As can be seen from the above Figure 3.19, it is not enough examining only the energy values to understand the drug-nucleobase interaction. Therefore, we also need to study other changes in these complexes, such as the bond length. In the

incoming drug-nucleobase pairs sections, we will see that oxaprozin and mefenamic acid will be the most altered structures of all drug molecules. Therefore, we need to focus more on these drug molecules. Oxaprozin belongs to the 1,3-oxazoles and is a monocarboxylic acid derived from propionic acid. In 2010, S. G. Sagdinc and A. Esmeb investigated the molecular structure, linear and nonlinear optical properties, and electronic properties of oxaprozin as a monomer using Hartree-Fock (HF) and density functional theory (DFT) calculations. The calculated geometrical parameters at HF and DFT at the B3LYP/6-31G(d,p) levels for the OXA monomer were compared with the experimental parameters. Their studies show that there was good agreement between the experimental and calculated geometric parameters. We have used their experimental results to compare our bond lengths. (Sagdinc and Esme, 2010)

Table 3. 17 Experimental and optimized geometrical parameters of oxaprozin [bond length (Å)]

Parameters	Exp ¹	ω B97X-D ²	ω B97X-D ^a	ω B97X-D ^c	ω B97X-D ^g	ω B97X-D ^t
Bond lengths (Å)		6-31+G(d,p)	6-31+G(d,p)	6-31+G(d,p)	6-31+G(d,p)	6-31+G(d,p)
O(2)–C(20)	1.226	1.340	1.332	1.340	1.332	1.372
O(3)–C(20)	1.222	1.214	1.219	1.214	1.219	1.218
C(20)–C(11)	1.448	1.508	1.505	1.508	1.506	1.505
C(11)–C(8)	1.551	1.523	1.522	1.527	1.523	1.522
C(8)–C(7)	1.47	1.490	1.490	1.492	1.490	1.489
C(7)–N(4)	1.334	1.294	1.294	1.295	1.294	1.294
N(4)–C(6)	1.348	1.394	1.394	1.399	1.393	1.395
C(6)–C(5)	1.416	1.364	1.365	1.371	1.366	1.365
C(5)–O(1)	1.357	1.373	1.374	1.384	1.374	1.372
O(1)–C(7)	1.342	1.347	1.345	1.355	1.344	1.345
C(6)–C(10)	1.434	1.473	1.473	1.473	1.472	1.473
C(10)–C(14)	1.423	1.400	1.400	1.403	1.399	1.400
C(5)–C(9)	1.43	1.464	1.463	1.461	1.463	1.463
C(9)–C(13)	1.422	1.401	1.401	1.405	1.402	1.400
C(13)–C(17)	1.406	1.393	1.392	1.392	1.392	1.392
C(17)–C(21)	1.404	1.394	1.394	1.395	1.394	1.393
C(21)–C(16)	1.405	1.395	1.395	1.396	1.394	1.395
C(16)–C(12)	1.41	1.391	1.391	1.392	1.390	1.390
C(12)–C(9)	1.417	1.401	1.401	1.404	1.401	1.401

1 Ref. OXA **2** Computational OXA **a** OXA-adenine **c** OXA-cytosine **g** OXA-guanine

t OXA-tymine

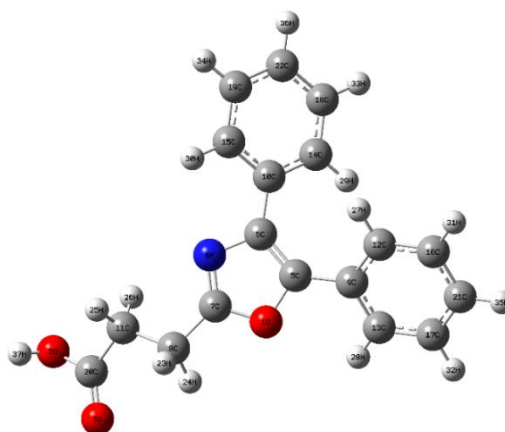


Figure 3. 20 Crystal structure of oxaprozin

The comparison between the experimental and computed values shows that they are well matched and there are almost no change in oxaprozin geometry. The investigation of the X-ray structures revealed that the C–O distance values were in a range from 1.303 to 1.345 Å in other NSAIDs. This value was calculated to be 1.332 - 1.372 Å using the ω B97X-D method for OXA-adenine, OXA-cytosine, OXA-guanine and OXA-thymine, respectively. The ω B97X-D method leads to systematically longer bonds. As you can see from Table 3.17, this allows us to see the change numerically in the interaction of oxaprozin with nucleobases. However, the examining the interaction sites and interaction energies between oxaprozin and nucleobases is not sufficient to understand that drug will disrupt DNA or not. Accordingly, geometry changes in hydrogen bonds of nucleobase pairs should be studied, so drug-nucleobase pair interactions were studied in detail in the next section.

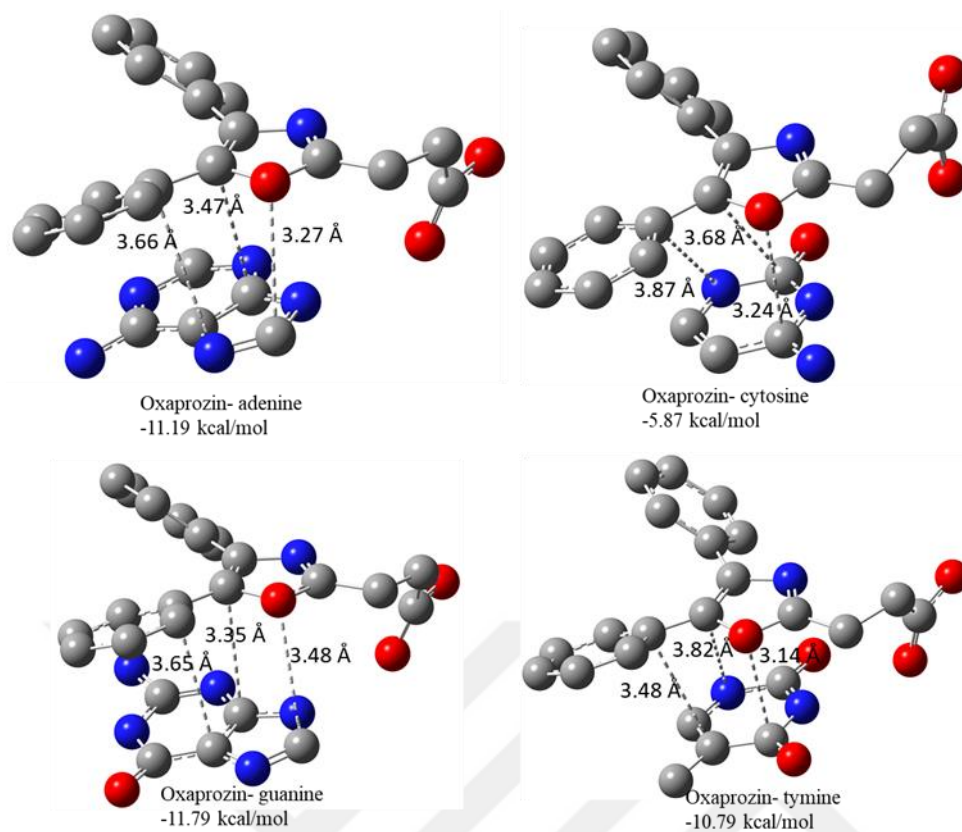


Figure 3. 21 Oxaprozin- nucleobase complexes optimized with the IEF-PCM/ ω B97X-D /6-31+G(d,p) method

Figure 3.21 shows the highly interacting oxaprozin nucleobase structures obtained by the IEF-PCM / ω B97X-D /6-31+G(d,p) method. As expected, oxaprozin shows the largest interaction with the nucleobase guanine and its interaction order with nucleobases is the following: $G > A > T > C$. This sequence resulted from the properties of nucleobases.

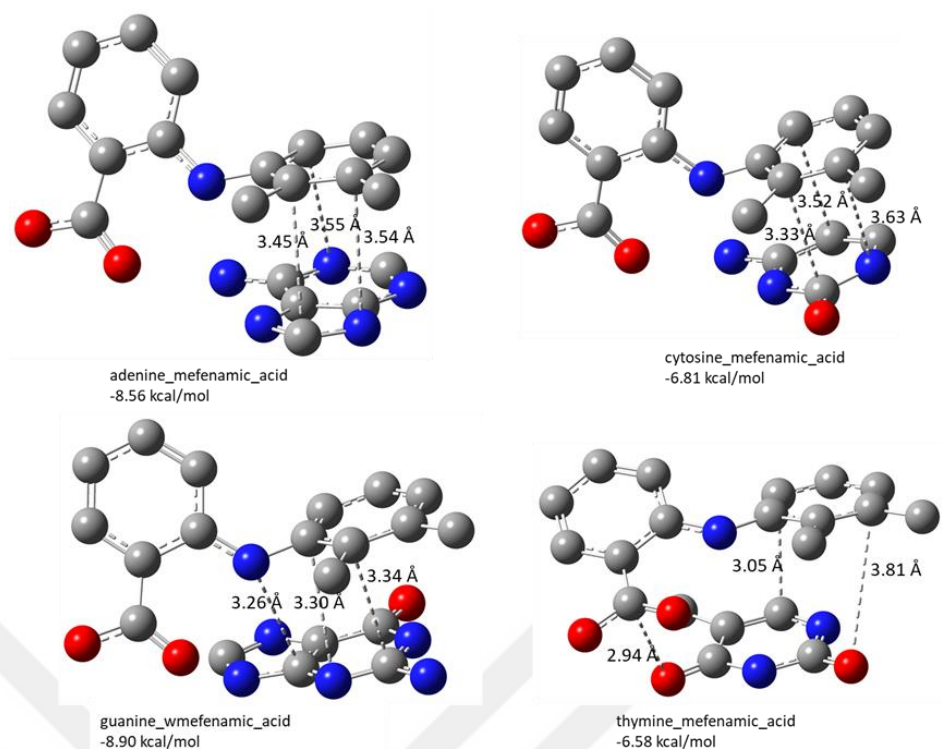


Figure 3. 22 Mefenamic acid- nucleobase complexes optimized with the IEF-PCM/ ω B97X-D /6-31+G(d,p) method

Mefenamic acid is a member of the NSAIDs. Its systematic name, dimethyl phenylamino benzoic acid. Figure 3.22 shows the highly interacting-drug-nucleobase structures obtained with the ω B97X-D method. While mefenamic acid-nucleobase complexes do not have a significant change on geometry, it has the largest interaction with guanine. Depend on energy of most active nucleobases can be ranked as $G > A > C > T$. Besides, thymine interacts with both benzene sites of the drug, while adenine, cytosine, and guanine interact on only a single benzene ring of the drug. In the next section, mefenamic acid gives the most interaction with the A-T nucleobase pair. It gives us an idea that drugs can be interacted with which part of the DNA.

3.5.2 Drug-Nucleobase-Pair Binding Energies

In this part, we investigated the effects of each drug on nucleobase-pairs to determine all drugs-nucleobase pair complexes. All drug-nucleobase pair complexes were optimized at the IEF-PCM/ ω B97X-D/6-31+G(d,p) levels of theory. The binding energies were calculated using eqn 3.5.

$$\Delta E_{bind} = E_0(\text{complex}) - (E_0(\text{drug}) + E_0(\text{base pair})) \quad (3.5)$$



Table 3. 18 ΔE_{bind} (kcal/mol) values of IEF-PCM/ ω B97X-D /6-31+G(d,p) optimized drug–nucleobase pair complexes

Drug Name	Orientation	ΔE_0	E_0 +BSSE	Drug Name	Orientation	ΔE_0	E_0 +BSSE
celiprolol	1_at	-16.52	-13.85	metoprolol	1_at	-13.90	-11.62
	2_at	-7.00	-6.06		2_at	-4.86	-3.88
	3_at	-6.56	-5.56		3_at	-4.87	-4.18
	1_gc	-20.88	-17.59		1_gc	-13.52	-11.01
	2_gc	-7.65	-6.69		2_gc	-11.43	-9.17
	3_gc	-6.50	-5.64		3_gc	-4.90	-4.19
atenolol	1_at	-14.96	-12.65	nabumetone	1_at	-14.56	-12.29
	2_at	-4.32	-3.50		2_at	-7.31	-6.03
	3_at	-10.29	-8.47		3_at	-12.36	-10.36
	1_gc	-14.84	-12.40		1_gc	-7.50	-6.30
	2_gc	-4.10	-3.29		2_gc	-8.63	-7.26
	3_gc	-5.56	-4.81		3_gc	-12.36	-10.13
betaxolol	1_at	-14.95	-12.65	esmolol	1_at	-13.44	-11.36
	2_at	-5.18	-4.33		2_at	-7.00	-5.90
	3_at	-4.21	-3.47		3_at	-11.23	-9.30
	1_gc	-13.85	-11.42		1_gc	-13.51	-11.17
	2_gc	-13.64	-11.30		2_gc	-13.55	-11.48
	3_gc	-5.27	-4.47		3_gc	-4.54	-3.63
bromfenac	1_at	-18.57	-11.69	nebivolol	1_at	*	*
	2_at	-18.56	-11.35		2_at	*	*
	3_at	-14.90	-9.07		3_at	*	*
	1_gc	-18.62	-12.39		1_gc	*	*
	2_gc	-15.76	-9.59		2_gc	1.04	-3.96
	3_gc	-14.05	-8.36		3_gc	*	*
celecoxib	1_at	-13.97	-10.62	nepafenac	1_at	-11.74	-9.44
	2_at	-15.33	-12.25		2_at	-10.36	-8.50
	3_at	-16.20	-13.02		3_at	-8.38	-7.06
	1_gc	-14.84	-11.58		1_gc	-12.93	-10.52
	2_gc	-11.69	-9.43		2_gc	-10.83	-8.53
	3_gc	-15.11	-12.01		3_gc	-12.69	-10.46
mefenamic acid	1_at	-15.23	-12.57	oxaprozin	1_at	-20.68	-17.70
	2_at	-10.38	-8.74		2_at	-17.71	-15.04
	3_at	-12.34	-10.62		3_at	-12.01	-9.40
	1_gc	-14.26	-11.61		1_gc	-15.99	-13.46
	2_gc	-11.22	-9.47		2_gc	-17.75	-14.95
	3_gc	-10.43	-8.36		3_gc	-11.92	-9.25
meloxicam	1_at	-13.22	-9.91	indometacin	1_at	-16.17	-13.39
	2_at	-10.58	-8.26		2_at	-9.64	-7.93
	3_at	-10.70	-8.38		3_at	-16.06	-13.30
	1_gc	-12.92	-9.58		1_gc	*	*
	2_gc	-9.39	-6.95		2_gc	-15.99	-12.88
	3_gc	-11.48	-8.79		3_gc	-15.11	-12.16
diflunisal	1_at	-14.68	-12.08	naproxen	1_at	-10.93	-8.49
	2_at	-11.55	-9.06		2_at	-12.82	*
	3_at	-9.11	-7.61		3_at	-12.84	-10.57
	1_gc	-13.61	-10.95		1_gc	-11.66	-9.18
	2_gc	-14.00	-11.36		2_gc	-12.10	-9.54
	3_gc	-14.44	*		3_gc	-12.78	-10.21
acebutolol	1_at	-17.33	-14.39	bisoprolol	1_at	*	*
	2_at	-7.97	*		2_at	-11.65	-9.79
	3_at	-4.28	-3.62		3_at	-5.55	-4.85
	1_gc	-17.53	-14.44		1_gc	-13.59	-11.32
	2_gc	-3.75	-3.22		2_gc	-12.79	-10.79
	3_gc	-4.46	-3.82		3_gc	-5.61	-4.86

* Calculations did not converge to a stable geometry.

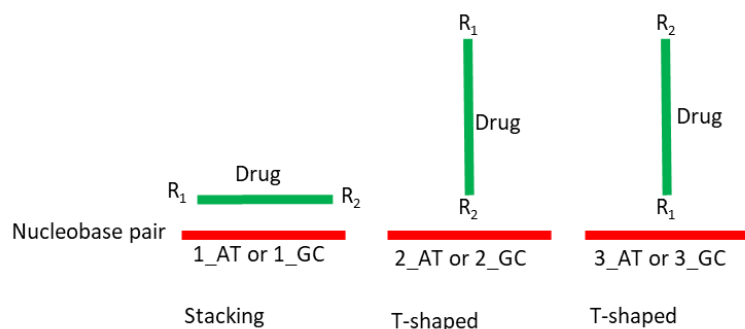


Figure 3. 23 Symbolic of parallel stacking and T-stacking interactions between all drugs and nucleobase pairs using

We examined parallel stacking and T-stacking interactions between all drugs and nucleobase pairs to identify interaction sites on the potential energy surfaces. Table 3.18 shows the binding energy values of the drug-nucleobase pair complexes optimized with the ω B97X-D/6-31+G(d,p) method. Basis set superposition error (BSSE) corrections obtained from the optimized structures were added to compensate the overestimation in binding energies. As in single nucleobase cases, BSSE did not cause much difference in the energies for all the drug structures studied compared to the magnitude of interaction energies.

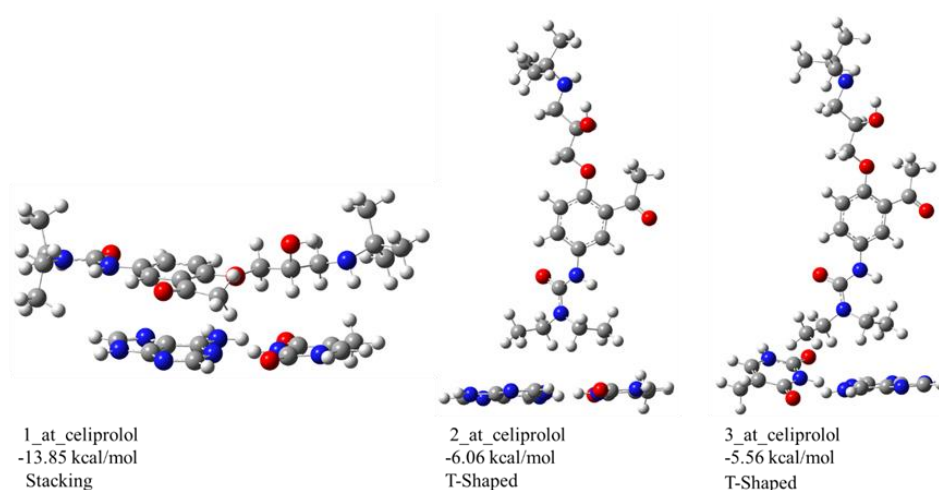


Figure 3. 24 Orientations of celiprolol–nucleobase pair (A-T) complexes optimized by ω B97X-D /6-31+G(d,p)

Figure 3.24 shows the orientation of the drug molecule of celiprolol with the highest interaction with the A-T nucleobase pair.

The drug molecules were placed to interact with the nucleobase pairs in the orientations shown in Figure 3.24. Our calculations show that the strongest interaction for any drug-nucleobase combination is a parallel stacking interaction. It is well known that parallel stacking interactions are highly dependent on dispersion π - π interactions (Azizi and Zoha, 2018) (Rutledge et al., 2009). Moreover, we find that the bond strengths for parallel stacking and T-stacking interactions are very similar in magnitude for nucleobases. On the other hand, the T-stacking interactions are slightly weaker than the parallel stacking interactions. Our drug-nucleobase interaction shows that the π - π interactions play an essential role in determining drug orientations. Parallel stacking interactions are highly dependent on dispersion interactions, so aromatic-aromatic interactions play a somewhat more significant role (Azizi and Zoha, 2018)(Rutledge et al., 2009).

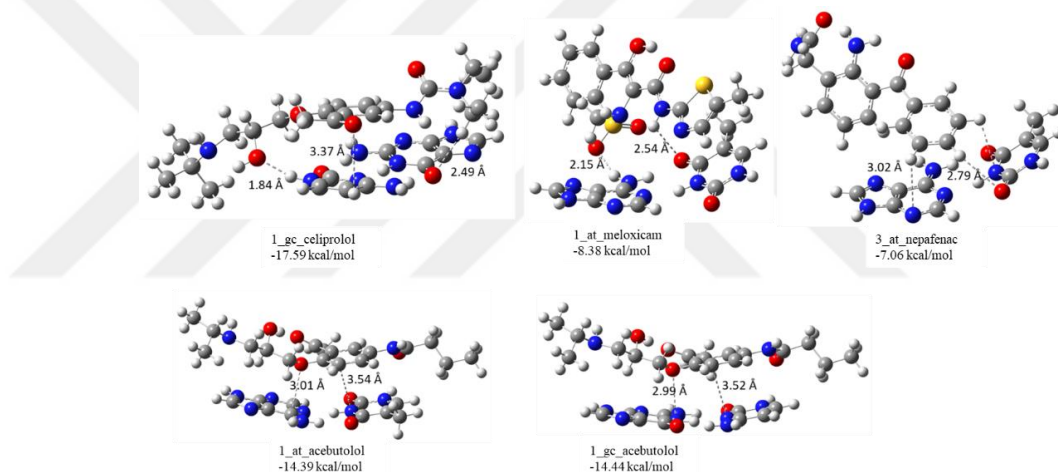


Figure 3. 25 Orientations of drug molecules- nucleobase pairs ((A-T), (G-C)) complexes optimized by ω B97X-D /6-31+G(d,p)

When the interactions of the drug molecules with the nucleobase pairs were examined, no significant changes were observed in the initial and final geometries shown in Figure 3.25. When we examine all complexes based on their energies only, no change is observed in the geometries of those with high interaction energies or those with low interaction energies. Nonetheless, the mefenamic acid and oxaprozin molecules cause large changes in their nucleobase pair complex geometries. Therefore, only binding energies not enough to understand drug-nucleobase interactions.

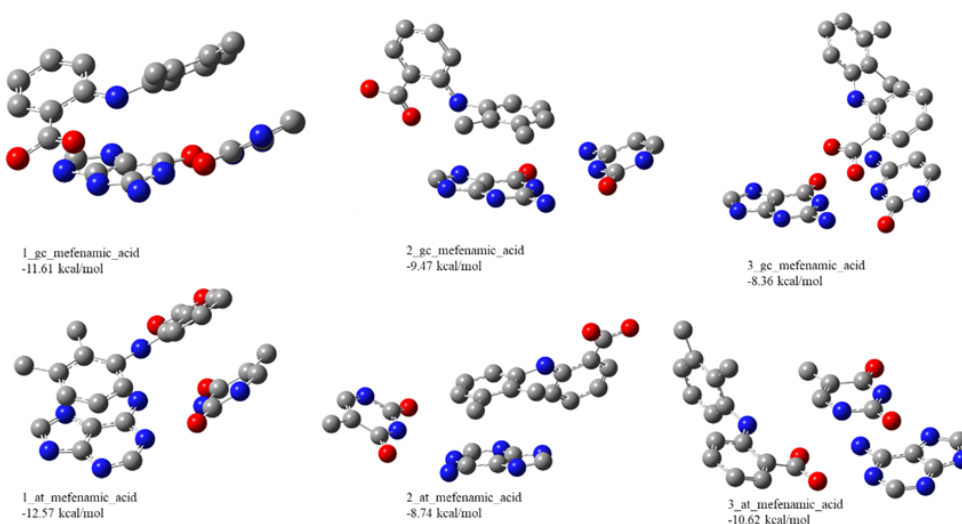
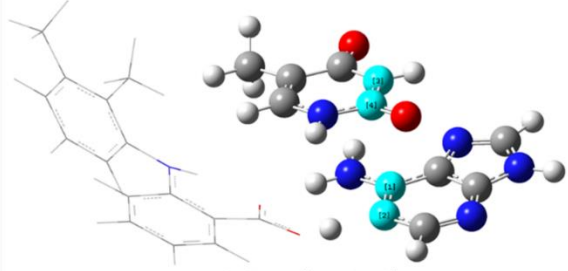


Figure 3. 26 Orientations of mefenamic acid–nucleobase pairs ((A-T), (G-C)) complexes optimized by ω B97X-D /6-31+G(d,p)

The parallel stacking interaction is the largest one for mefenamic acid – nucleobase pairs, GC and AT (see Figure 3.26). The interaction surfaces of the complexes studied shows that dispersion interactions are significant, and electrostatic interactions also play an important role (Rutledge et al., 2009). When mefenamic acid is examined in detail, it is seen that it has the power to separate nucleobase pairs. In addition, H-bonding interactions in nucleic acids play a critical role in the double-helical structure of DNA and parallel stacking interactions that facilitate molecular recognition through replication processes and protein synthesis (Sagdinc and Esme, 2010), which is also crucial in drug-receptor interactions. Typical H-bond strength ranges from 1 to 10 kcal/mol. Moreover, it is known that the strength of a hydrogen bond strongly depends on the geometry. In 2010, Kaila, V. R. I. and their collaborators studied the strength of hydrogen bonds (Fliegl, Lehtonen, Sundholm, and Kaila, 2011). Their research shows that the individual hydrogen bond energies can also be estimated from differences between the base pair initial state energy and base pair final state energy.

Table 3. 19 Estimated Drug effect on H-bonding energy in kcal/mol for adenine–thymine (A–T) and guanine–cytosine (C–G) DNA base pairs with mefenamic acid

Nucleobase pairs	Drug effect on H-bonding
1_at_mefenamic_acid	1.09
2_at_mefenamic_acid	1.06
3_at_mefenamic_acid	5.98
1_gc_mefenamic_acid	1.17
3_gc_mefenamic_acid	1.26



$$\Delta E_{estimated} = E_0(\text{base pair final state energy (without drug)}) - (E_0(\text{base pair initial state energy})) \quad (3.6)$$

To calculate the estimated H-bond strength, the single point energy of the nucleobase pair alone was first calculated by removing the drug from the final geometry of the drug-nucleobase pair complex. The estimated H-bond strength was calculated by taking the difference between the energy of the base pair in the final state (without drug) and the energy of the optimized geometry of the nucleobase pair. The most striking result was obtained when the drug completely breaks nucleobase pair in 3_at_mefenamic acid.

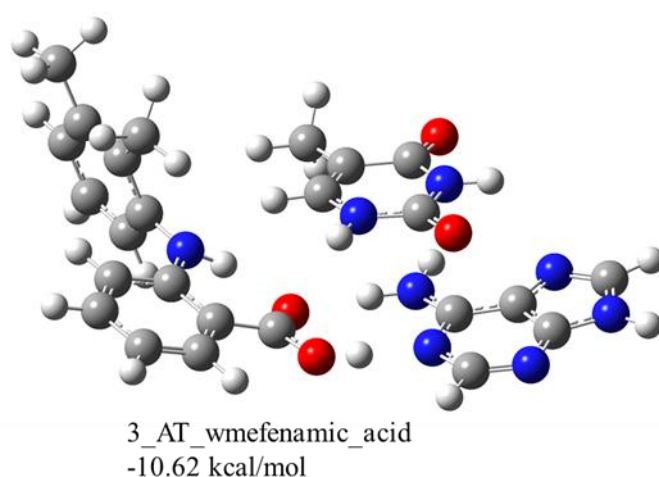


Figure 3. 27 Orientations of 3_at_mefenamic acid (nucleobase pairs ((A-T)) complexes optimized by ω B97X-D /6-31+G(d,p)

As can be seen in the Figure 3.27, the nucleobase pairs are in different axial positions. The thymine nucleobase is located on adenine. This interaction can damage or disrupt the DNA. Therefore, we need to focus more on this drug molecule.

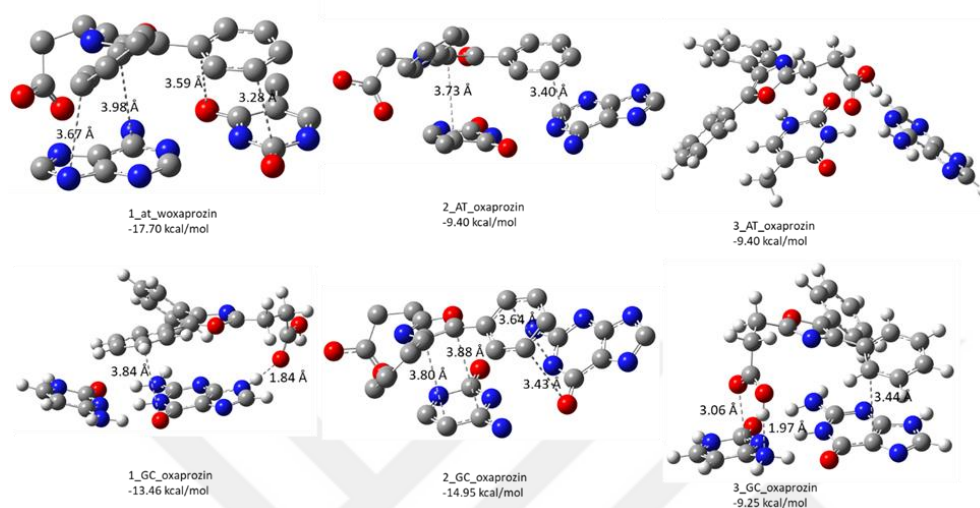
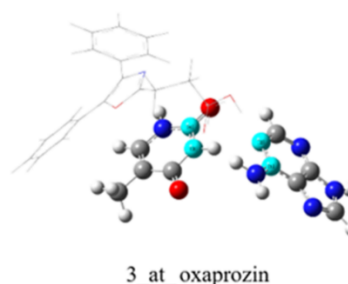


Figure 3. 28 Orientations of oxaprozin –nucleobase pairs ((A-T), (G-C)) complexes optimized by ω B97X-D /6-31+G(d,p)

Oxaprozin is a nonsteroidal anti-inflammatory drug that has the largest interaction with the alignment of the nucleobase pair 1_at (adenine-thymine) –17.70 kcal/mol. However, when the geometries are examined, it is found that the 3_at_oxaprozin orientation is the most dramatic result. Thus, the energy of the interaction alone does not lead us to the correct result.

Table 3. 20 Estimated Drug effect on H-bonding energy in kcal/mol for adenine–thymine (A–T) and guanine-cytosine (C–G) DNA base pairs with oxaprozin

Nucleobase	Drug effect on H-bonding
1_at_oxaprozin	0.51
2_at_oxaprozin	0.73
3_at_oxaprozin	6.30
1_gc_oxaprozin	0.82
2_gc_oxaprozin	0.72
3_gc_oxaprozin	1.52



As shown in Table 3.20, the most dramatic result was obtained by changing the orientation of 3_at_ oxaprozin. As known, the dissociation energy of hydrogen

bonds is an endothermic reaction. The estimated H-bond dissociation energy of mefenamic acid and oxaprozin molecules is about 5.98 kcal/mol and 6.30 kcal/mol, respectively. Examination of the above results indicated that oxaprozin and mefenamic acid probably break the H-bonds between the nucleobase pairs. Therefore, these two drugs need to be further investigated. The common feature of both drugs was their interaction with the adenine-thymine base pair. As can be seen from the geometries of the drug-base pair complexes, there are rather huge changes in nucleobase pairs in terms of both distance and angle. We need to study both the distance and angular changes in the resultant geometries of the nucleobase pairs to predict that the drugs available to us might have side effects on DNA. It is not sufficient to study the changes in the drug-nucleobase pair complex based only on energy, as is the case with mefenamic acid and oxaprozin.

3.5.3 H-Bonds Distance of Drug–Nucleobase Pair

Table 3. 21 H-Bonds Distances (Ångstrom) of the ω B97X-D optimized drug– A-T base pair complexes

Types	Angstroms		Difference In O-H	Difference In N-H
	O-H	N-H		
A-T	1.9075	1.8187		
1_AT_atenolol	1.8979	1.8094	-0.0096	-0.0093
2_AT_atenolol	1.9047	1.8236	-0.0028	0.0049
3_AT_atenolol	1.9222	1.8201	0.0147	0.0014
1_AT_betaxolol	1.905	1.81	-0.0025	-0.0087
2_AT_betaxolol				
3_AT_betaxolol	1.9001	1.8277	-0.0074	0.009
1_AT_bromfenac	1.9796	1.8348	0.0721	0.0161
2_AT_bromfenac	1.90891	1.8575	0.00141	0.0388
3_AT_bromfenac	1.944	1.8875	0.0365	0.0688
1_AT_celecoxib	1.9312	1.78341	0.0237	-0.03529
2_AT_celecoxib	1.90405	1.82331	-0.00345	0.00461
3_AT_celecoxib	1.90209	1.89301	-0.00541	0.07431
1_AT-celiprolol	1.92311	1.78206	0.01561	-0.03664
2_AT-celiprolol	1.8999	1.79946	-0.0076	-0.01924
3_AT-celiprolol	1.91064	1.83111	0.00314	0.01241
1_AT_esmolol	1.91877	1.78796	0.01127	-0.03074
2_AT_esmolol	1.92453	1.86519	0.01703	0.04649
3_AT_esmolol	1.9052	1.81636	-0.0023	-0.00234
1_AT_indometacin	1.87002	1.82763	-0.03748	0.00893
2_AT_indometacin	1.93046	1.82294	0.02296	0.00424
3_AT_indometacin	1.92814	1.77919	0.02064	-0.03951
1_AT_meloxicam	1.90464	1.85813	-0.00286	0.03943
2_AT_meloxicam	1.92935	1.80172	0.02185	-0.01698
3_AT_meloxicam	1.90611	1.85931	-0.00139	0.04061
1_AT_metoprolol	1.89733	1.80803	-0.01017	-0.01067
2_AT_metoprolol	1.91353	1.79974	0.00603	-0.01896
3_AT_metoprolol	1.9074	1.82556	-0.0001	0.00686
1_AT_nabumetone	1.89224	1.79886	-0.01526	-0.01984
2_AT_nabumetone	1.91093	1.81297	0.00343	-0.00573
3_AT_nabumetone	1.91371	1.81059	0.00621	-0.00811
1_AT_nepafenac	1.92098	1.8195	0.01348	0.0008
2_AT_nepafenac	1.90992	1.87233	0.00242	0.05363
3_AT_nepafenac	1.90599	1.81283	-0.00151	-0.00587

Table 3.21 shows that the drugs affect the bond lengths of the hydrogen bonds between the A-T base pair.

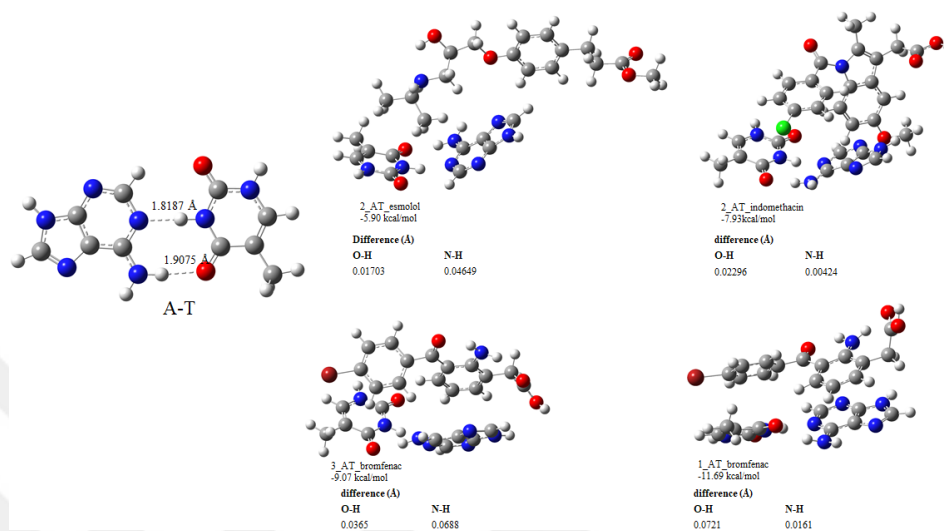


Figure 3. 29 Intermolecular interaction of drug–nucleobase pair complexes

When the results of the optimized geometries of the drug-nucleobase pairs are examined (see Table 3.21 and Figure 3.29), it can be seen that there was no significant change in the lengths of the hydrogen bonds except for oxaprozin and mefenamic acid.

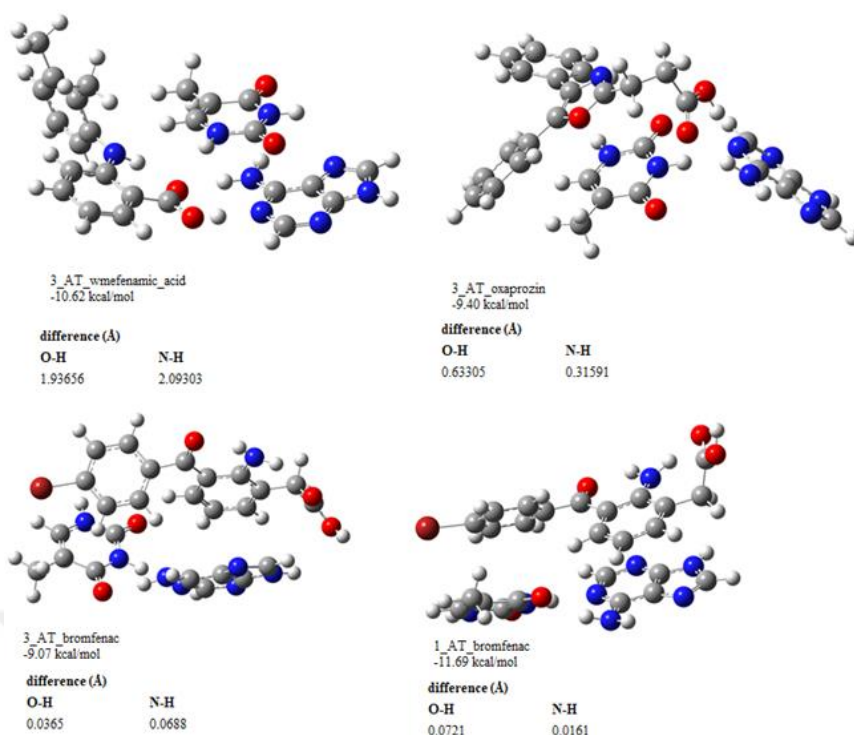


Figure 3.30 Intermolecular interaction of drug–nucleobase pair complexes

As seen from figure 3.30, mefenamic acid causes complete rupture of H-bonds between adenine and thymine nucleobases. Thymine is located above adenine, indicating full breakage.

Table 3. 22 Distances between H-bonds (Ångstrom) of ωB97X-D optimized drug– G-C basepair complexes

Types	Angstroms			Difference In O-H	Difference In N-H	Difference In O-H
	O-H	N-H	O-H			
G-C	1.8237	1.8861	1.8431			
1_GC_atenolol	1.8145	1.8563	1.8425	-0.0092	-0.0298	-0.0006
2_GC_atenolol	1.8244	1.8881	1.8444	0.0007	0.002	0.0013
3_GC_atenolol	1.8386	1.8823	1.8563	0.0149	-0.0038	0.0132
1_GC_betaxolol	1.85551	1.88973	1.83337	0.03181	0.00363	-0.00973
2_GC_betaxolol	1.9196	1.8097	1.882	0.0959	-0.0764	0.0389
3_GC_betaxolol	1.8487	1.884	1.8255	0.025	-0.0021	-0.0176
1_GC_bromfenac	1.88141	1.84491	1.8982	0.05771	-0.04119	0.0551
2_GC_bromfenac	1.83212	1.89702	1.84879	0.00842	0.01092	0.00569
3_GC_bromfenac	1.86594	1.80845	1.88099	0.04224	-0.07765	0.03789
1_GC_celecoxib	1.85908	1.82731	1.86608	0.03538	-0.05879	0.02298
2_GC_celecoxib	1.85707	1.86798	1.84465	0.03337	-0.01812	0.00155
3_GC_celecoxib	1.82465	1.85765	1.86426	0.00095	-0.02845	0.02116
1_GC_celiprolol	1.81349	1.86709	1.82792	-0.01021	-0.01901	-0.01518
2_GC_celiprolol	1.81891	1.87349	1.83496	-0.00479	-0.01261	-0.00814
3_GC_celiprolol	1.8176	1.89051	1.85537	-0.0061	0.00441	0.01227
1_GC_esmolol	1.81473	1.86638	1.82708	-0.00897	-0.01972	-0.01602
2_GC_esmolol	1.81861	1.86851	1.84244	-0.00509	-0.01759	-0.00066
3_GC_esmolol	1.80571	1.88084	1.83976	-0.01799	-0.00526	-0.00334
1_GC_indometacin						
2_GC_indometacin	1.83648	1.8602	1.81057	0.01278	-0.0259	-0.03253
3_GC_indometacin	1.84494	1.87606	1.8468	0.02124	-0.01004	0.0037
1_GC_meloxicam	1.85948	1.85704	1.87518	0.03578	-0.02906	0.03208
2_GC_meloxicam	1.82401	1.88143	1.85154	0.00031	-0.00467	0.00844
3_GC_meloxicam	1.85764	1.8676	1.82442	0.03394	-0.0185	-0.01868
1_GC_metoprolol	1.82619	1.86718	1.84165	0.00249	-0.01892	-0.00145
2_GC_metoprolol	1.81766	1.84868	1.84165	-0.00604	-0.03742	-0.00145
3_GC_metoprolol	1.81816	1.89406	1.84003	-0.00554	0.00796	-0.00307
1_GC_nabumetone	1.82359	1.87969	1.83408	-0.00011	-0.00641	-0.00902
2_GC_nabumetone	1.85301	1.86968	1.8092	0.02931	-0.01642	-0.0339
3_GC_nabumetone	1.8644	1.87536	1.84379	0.0407	-0.01074	0.00069

Table 3.22 shows that the drugs do not considerably affect the bond lengths of the hydrogen bonds between the G-C base pair.

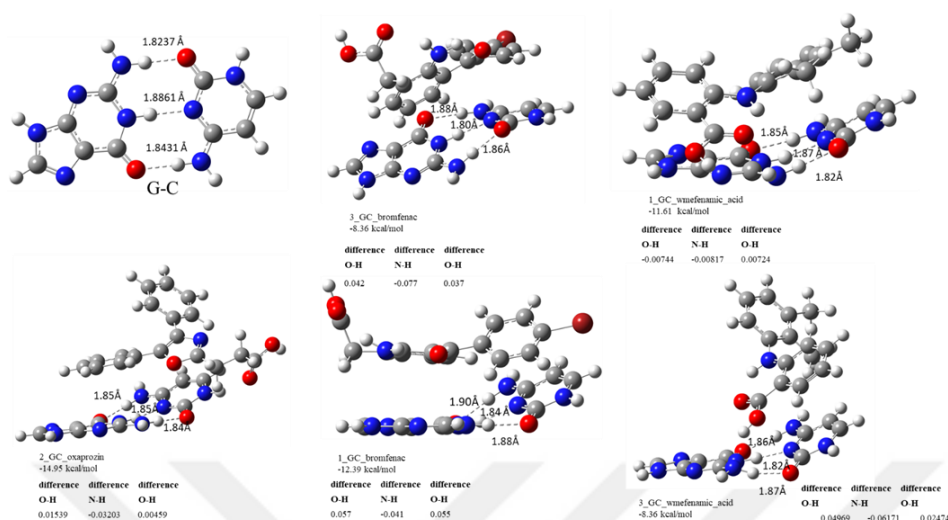


Figure 3. 31 Intermolecular interaction of drug–nucleobase pair complexes

In 2019, N. Balamurugan and colleagues studied the molecular electrostatic potential (MEP) map of mefenamic acid (Charanya, Sampathkrishnan, and Balamurugan, 2019). Their MEP study shows that the oxygen atom of the carbonyl group (C=O) is the most favorable site for electrophilic attack. In our study, mefenamic acid attacks the nucleobase pairs from the same site, and it cannot create a significant change in hydrogen bonding for the guanine-cytosine base pair unlike the adenine-thymine base pair with other drugs.

3.5.4 H-Bonds Angles of Drug–Nucleobase Pair

Table 3. 23 Differences in angles of H-bonds (degrees) of ω B97X-D optimized drug-A-T base pair complexes

Types	Angstroms		Difference In N-H-O	Difference In N-H-N
	N-H-O	N-H-N		
A-T	175.30	179.77		
1_AT_atenolol	174.78	179.30	-0.52	-0.46
2_AT_atenolol	175.35	179.10	0.05	-0.67
3_AT_atenolol	173.99	179.80	-1.32	0.04
1_AT_betaxolol	175.30	179.46	0.00	-0.31
2_AT_betaxolol	170.40	168.56	-4.90	-11.20
3_AT_betaxolol	173.58	172.37	-1.73	-7.40
1_AT_bromfenac	164.34	174.94	-10.96	-4.83
2_AT_bromfenac	174.40	178.12	-0.90	-1.65
3_AT_bromfenac	167.32	171.26	-7.98	-8.50
1_AT_celecoxib	174.77	173.99	-0.53	-5.78
2_AT_celecoxib	175.00	171.22	-0.30	-8.54
3_AT_celecoxib	169.98	167.53	-5.32	-12.23
1_AT-celiprolol	174.41	179.06	-0.89	-0.70
2_AT-celiprolol	174.75	177.11	-0.55	-2.66
3_AT-celiprolol	173.44	173.76	-1.86	-6.01
1_AT_esmolol	174.88	179.59	-0.42	-0.18
2_AT_esmolol	168.16	165.38	-7.14	-14.39
3_AT_esmolol	174.61	175.04	-0.69	-4.72
1_AT_indometacin	174.86	176.96	-0.44	-2.81
2_AT_indometacin	171.93	167.18	-3.37	-12.58
3_AT_indometacin	174.60	179.01	-0.70	-0.76
1_AT_meloxicam	172.37	174.64	-2.94	-5.13
2_AT_meloxicam	174.03	175.44	-1.27	-4.33
3_AT_meloxicam	164.57	165.82	-10.74	-13.95
1_AT_metoprolol	174.29	176.78	-1.01	-2.98
2_AT_metoprolol	174.58	176.19	-0.72	-3.58
3_AT_metoprolol	172.85	171.43	-2.45	-8.34

1_AT_nabumetone	175.38	177.13	0.08	-2.64
2_AT_nabumetone	174.21	175.59	-1.09	-4.17
3_AT_nabumetone	170.70	172.51	-4.60	-7.26
1_AT_nepafenac	169.38	171.91	-5.92	-7.86
2_AT_nepafenac	173.00	176.66	-2.30	-3.11
3_AT_nepafenac	172.39	171.00	-2.91	-8.77
1_AT_oxaprozin	177.50	178.81	2.20	-0.96
2_AT_oxaprozin	173.67	171.83	-1.63	-7.94
3_AT_oxaprozin	109.78	164.09	-65.52	-15.68
1_AT_wmefenamic_acid	172.94	176.48	-2.37	-3.28
2_AT_wmefenamic_acid	170.36	173.24	-4.94	-6.53
3_AT_wmefenamic_acid	60.87	67.43	114.43	112.34

Table 3.23 shows that the drugs affect the angles of the hydrogen bonds between the A-T base pair.

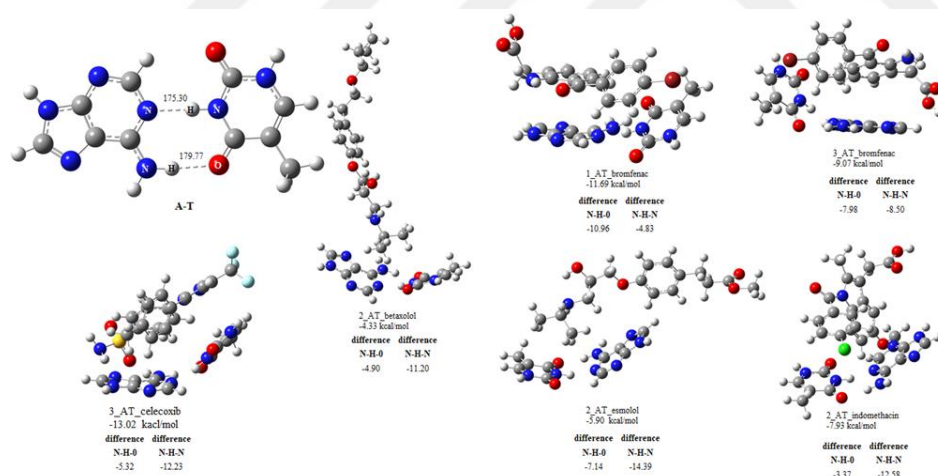


Figure 3. 32 H-bonds angles changes of drug– A-T nucleobase pair complexes

When the results of the optimized geometries of the drug-A-T nucleobase pairs are examined (see Table 3.23 and Figure 3.32), it is seen that there was no significant change in the angles of the hydrogen bonds except for oxaprozin and mefenamic acid.

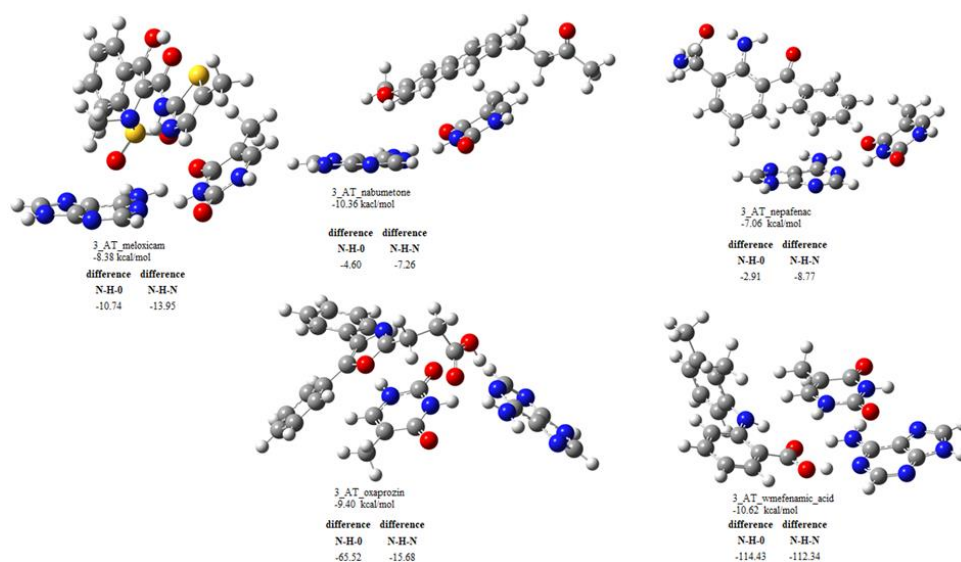


Figure 3.33 H-bonds angles changes of drug– A-T nucleobase pair complexes

However, Figure 3.33 shows that the drug molecules of mefenamic acid and oxaprozin cause significant differences in angles of the H-bonds in the A-T nucleobase pair.

Table 3. 24 Differences Between H-bonds Angles (degrees) of ω B97X-D Optimized Drug–G-T Basepair Complexes

Types	Degrees			Difference in N-H-O	Difference in N-H-N	Difference in O-H-N
	N-H-O	N-H-N	O-H-N			
G-C	179.14	178.67	178.70			
1_GC_atenolol	177.76	177.38	177.53	-1.38	-1.29	-1.17
2_GC_atenolol	178.95	176.99	178.67	-0.19	-1.67	-0.03
3_GC_atenolol	177.28	176.53	176.62	-1.86	-2.13	-2.08
1_GC_betaxolol	174.23	176.17	177.47	-4.91	-2.50	-1.23
2_GC_betaxolol	175.88	172.52	175.12	-3.26	-6.14	-3.57
3_GC_betaxolol	176.39	172.84	177.18	-2.75	-5.82	-1.52
1_GC_bromfenac	177.82	167.84	175.77	-1.32	-10.82	-2.92
2_GC_bromfenac	178.88	171.62	178.17	-0.26	-7.05	-0.53
3_GC_bromfenac	176.19	176.16	174.82	-2.95	-2.51	-3.88
1_GC_celecoxib	176.24	177.92	173.97	-2.90	-0.74	-4.72
2_GC_celecoxib	176.54	177.46	175.91	-2.60	-1.21	-2.79
3_GC_celecoxib	179.26	173.38	174.26	0.12	-5.29	-4.44
1_GC_celiprolol	177.42	176.90	177.74	-1.72	-1.77	-0.96
2_GC_celiprolol	178.73	178.41	178.34	-0.40	-0.26	-0.36
3_GC_celiprolol	176.37	169.96	172.00	-2.77	-8.70	-6.70
1_GC_esmolol	178.25	178.00	178.37	-0.89	-0.66	-0.33
2_GC_esmolol	177.94	169.91	177.75	-1.20	-8.76	-0.95
3_GC_esmolol	177.86	174.29	176.33	-1.28	-4.38	-2.37
1_GC_indometacin						
2_GC_indometacin	178.47	176.93	173.67	-0.67	-1.74	-5.03
3_GC_indometacin	176.26	178.71	174.18	-2.88	0.04	-4.52
1_GC_meloxicam	177.62	178.29	178.04	-1.52	-0.37	-0.65
2_GC_meloxicam	176.51	171.93	171.74	-2.63	-6.73	-6.96
3_GC_meloxicam	172.35	176.44	176.75	-6.79	-2.22	-1.95
1_GC_metoprolol	178.92	178.40	176.68	-0.22	-0.26	-2.02
2_GC_metoprolol	178.56	178.32	1.88	-0.58	-0.35	-176.82
3_GC_metoprolol	178.08	173.85	176.54	-1.06	-4.81	-2.15
1_GC_nabumetone	178.83	178.81	178.73	-0.31	0.14	0.03
2_GC_nabumetone	175.24	176.39	177.89	-3.90	-2.27	-0.81
3_GC_nabumetone	171.34	175.60	176.89	-7.80	-3.07	-1.81
1_GC_nepafenac	176.08	177.42	175.63	-3.06	-1.24	-3.07
2_GC_nepafenac	178.19	179.20	177.31	-0.94	0.53	-1.39
3_GC_nepafenac	177.26	172.50	174.54	-1.88	-6.17	-4.16

Table 3.24 shows that the drugs affect the angles of the hydrogen bonds between the G-C base pair.

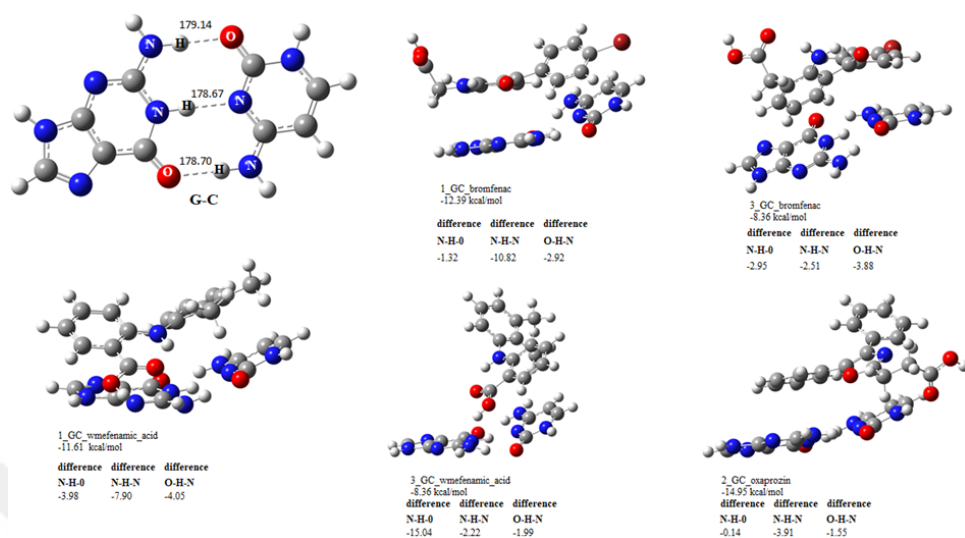


Figure 3. 34 H-bonds angles changes of drug– G-C nucleobase pair complexes

As seen from Figure 3.34, there is not any significant change in the angles of the hydrogen bonds of the drug-G-C nucleobase pairs with the largest interactions, except for mefenamic acid and oxaprozin. On the other hand, Figure 3.34 also shows that the hydrogen atom of the carboxylic acid group on mefenamic acid interacts with one of the nitrogen atoms of guanine, and the benzene moiety of mefenamic acid interacts with cytosine through π - π interactions. These interactions cause the nucleobase pair to bend. This may lead a rupture on G-C nucleobase pairs of DNA interacting with mefenamic acid.

3.6 Computational Investigation Drug-Nucleobase and Drug-Nucleobase Pair Interaction with ω B97M-V And ω B97X-V Methods

3.6.1 Binding Energies of Drugs to Nucleobases

In this part, we have studied the effects of each drug on a single nucleobase to determine the drug-nucleobase interaction. This is because π - π - parallel stacking interactions play a fundamental role in chemical and biological recognition (Meyer et al., 2003). Therefore, the single-point electronic energies of all drug-nucleobase complexes at the ω B97M-V/def2-TZVPP and ω B97X-V/def2-

TZVPP levels of DFT were calculated, and the binding energies were calculated employing the eqn 3.7.

$$\Delta E_e = E_e(\text{complex}) - (E_e(\text{drug}) + E_e(\text{base})) \quad (3.7)$$

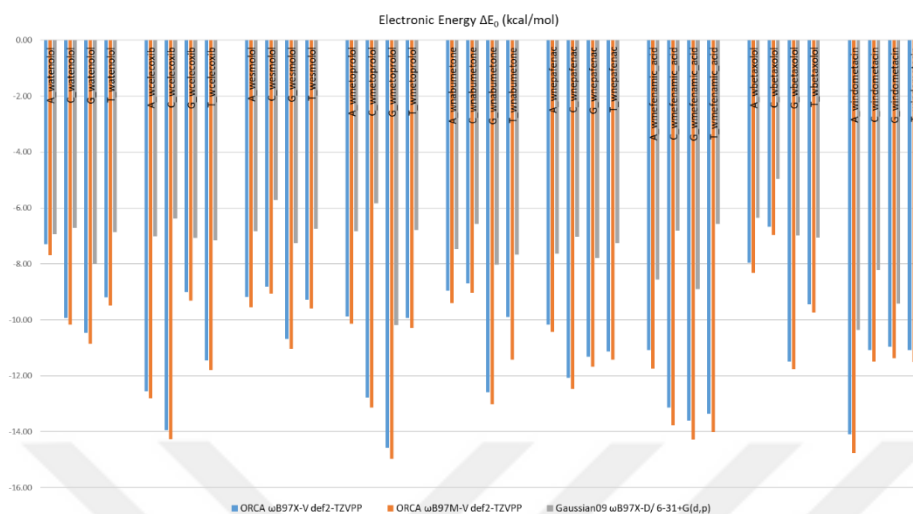


Figure 3.35 Electronic energy values in kcal/mol of drug-nucleobase calculated with ω B97M-V/def2-TZVPP, ω B97X-V/def2-TZVPP and ω B97X-D/6-31+G(d,p)

First, geometry optimization of drug and nucleobase was performed with the ω B97X-D method employing the effect of implicit water solvation on complex formation was studied using the Polarized Continuum Model (IEF-PCM), followed by single point calculations with ω B97M-V and ω B97X-V. When non-covalent drug-nucleobase interactions are studied, the ω B97M-V results give a relatively good estimate of this weak interaction in the ω B97 families, as shown in the Figure 3.35. In 2017, the study by N. Mardirossian and M. Head-Gordon showed that ω B97M-V in ω B97 families gives better results for the non-covalent interactions. Mefenamic acid is a non-steroidal anti-inflammatory drug that has the largest interaction with the guanine nucleobase -14.29 kcal/mol with ω B97M-V/def2-TZVPP.

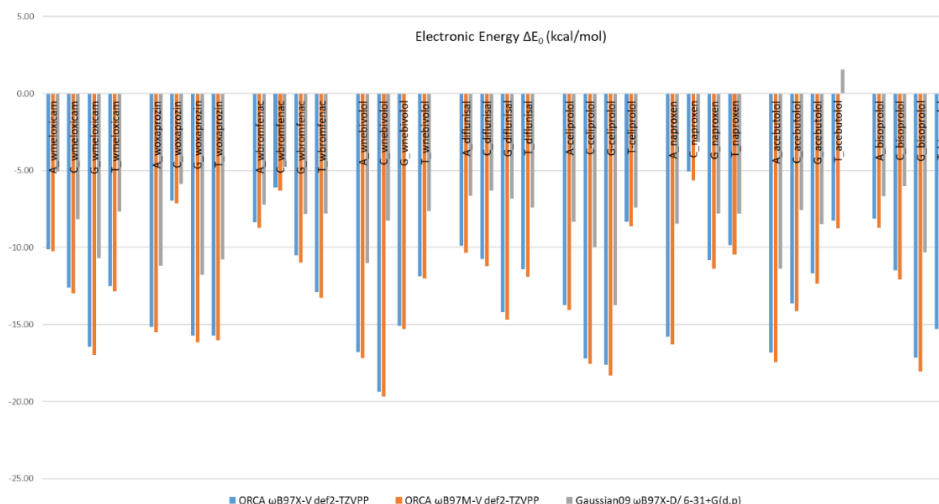


Figure 3. 36 Electronic energy values in kcal/mol of drug-nucleobase calculated with ω B97X-V, ω B97M-V, and ω B97X-D

As shown in Figure 3.36, the cytosine-nebivolol and guanine-celiprolol interactions have the largest interactions and energy values of -19.66 kcal/mol and -18.30 kcal/mol with ω B97M-V/def2-TZVPP, respectively. Oxaprozin gave the largest interaction with guanine with an energy value of -16.17 using the ω B97M-V method. As we know from the previous results, it is likely that oxaprozin and mefenamic acid break the H-bond between the nucleobase pairs. It is not enough to study the changes of the complex of drug and nucleobase pair only based on the energy, as is the case with mefenamic acid and oxaprozin. The common feature of both drugs was their interaction with the guanine nucleobase, and both drugs are non-steroidal anti-inflammatory drugs.

3.6.2 Binding Energies of Drugs – Nucleobase pairs

In this part, we examined the effects of each drug on nucleobase pairs to determine all drug-nucleobase pair complexes. All drug-nucleobase pair complexes were optimized at the IEF-PCM/ ω B97X-D/6-31+G(d,p) level of theory. Binding energies were calculated using eqn 3.8.

$$\Delta E_e = E_e(\text{complex}) - (E_e(\text{drug}) + E_e(\text{basepair})) \quad (3.8)$$

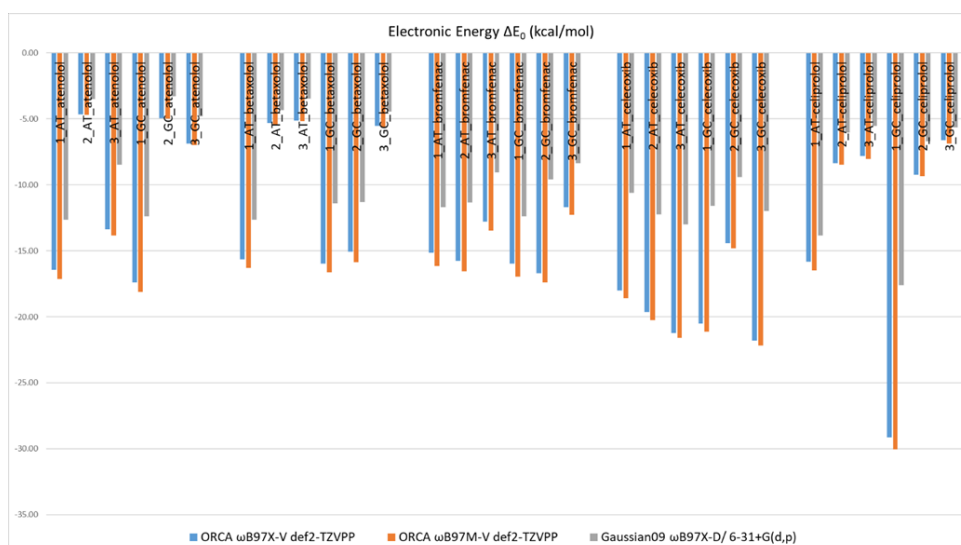


Figure 3.37 Electronic energy values in kcal/mol of drug-nucleobase pairs calculated with ω B97X-V, ω B97M-V, and ω B97X-D

As in the previous section, geometry optimization of drug-nucleobase pairs was performed with the ω B97X-D in the effect of implicit water solvation on complex formation, followed by single-point calculations with ω B97M-V and ω B97X-D. When non-covalent drug-nucleobase interactions are examined, the ω B97M-V results give a relatively good estimate of this weak interaction in the ω B97 families, as shown in Figure 3.37. As can be seen from Figure 3.37, celiprolol showed the largest interaction with the guanine-cytosine nucleobase pair and energy values of -30.05 kcal/mol by the ω B97M-V/def2-TZVPP method. As we know, no significant changes were observed in the initial and final geometries when studying the interactions of drug molecules with nucleobase pairs. However, the drug molecules mefenamic acid and oxaprozin show large changes in their complex geometries.

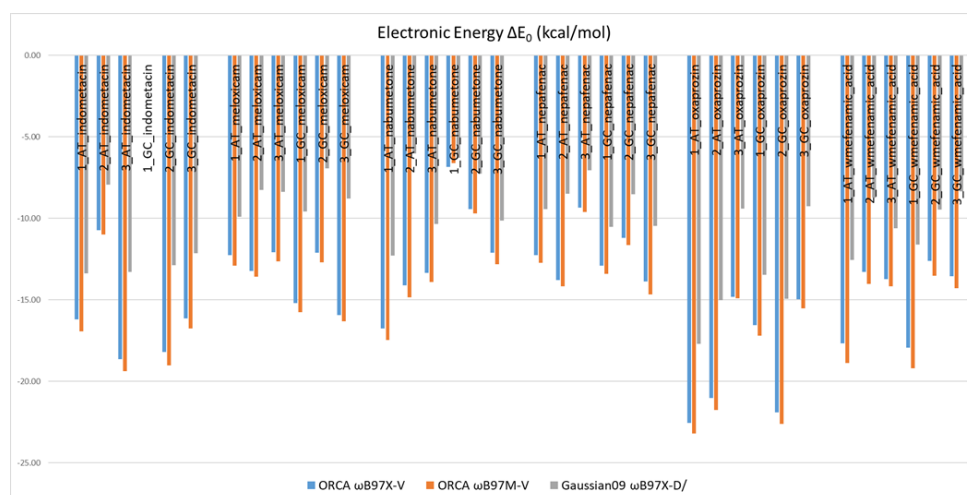


Figure 3.38 Electronic energy values in kcal/mol of drug-nucleobase pair are calculated with ω B97X-V, ω B97M-V, and ω B97X-D

Oxaprozin is a nonsteroidal anti-inflammatory drug that has the largest interaction with the alignment of the nucleobase pair 1_at (adenine-thymine) -23.19 kcal/mol. However, when the geometries are examined, it is found that the 3_at_oxaprozin orientation is the most striking result. Thus, the energy of the interaction alone does not lead us to the correct result, and the same is true for mefenamic acid. For this reason, two structures are important to us. The first is 3_ AT _oxaprozin, and the other is 3_ AT _mefenamic acid. The estimated dissociation energy of H-bonding of oxaprozin and mefenamic acid molecules is about -14.92 kcal/mol and -14.17 kcal/mol, respectively. The common feature of both drugs was their interaction with the adenine-thymine base pair. Moreover, functional has the best performance for interaction energy, followed by ω B97M-V/def2-TZVPP and ω B97X-V/def2-TZVPP. ω B97M-V and ω B97X-V gave improved energy for a single point calculation.

3.6.3 Molecular Electrostatic Potentials (Meps) Of Drug-Nucleobase Pair

Natural Bond Orbital Analysis (NBO) predicts hyperconjugation by estimating the energy of a molecule with the same shape in the absence of electronic delocalization. It's also utilized to figure out what's going on inside the molecule, like intramolecular rehybridization and electron density delocalization. Hyperconjugation is a stabilizing effect that occurs when one occupied orbital overlaps with an adjacent electron-deficient orbital when the orbitals are appropriately oriented. NBO analysis can quantify this non-covalent bonding interaction (anti-bonding interaction), which is expressed as the second-order perturbation interaction energy, $E^{(2)}$. The stronger the giving propensity from electron donors to electron acceptors of the entire system, the larger the $E^{(2)}$ value, the more intense is the interaction between electron donors and electron acceptors.(Shankar Rao, Prasad, Udaya Sri, and Veeraiah, 2016). To study the various second-order interactions between the filled orbital of one subsystem and the empty orbital of another subsystem, the DFT / ω B97X-D method is used. Our goal is to predict the interaction area in the drug with the nucleobase or nucleobase pair. Since each donor NBO (i) and acceptor NBO (j) is involved, the donor-acceptor stabilization energy $E^{(2)}$ associated with the delocalization $i \rightarrow j$ is estimated as(Saji et al., 2020)

$$E^{(2)} = q_i \frac{F_{(i,j)}^2}{\varepsilon_{(j)} - \varepsilon_{(i)}} \quad (7)$$

q_i is the donor orbital occupation (2 for closed shell, 1 for open shell), ε_i , ε_j are diagonal elements (orbital energies), and $F(i,j)$ is the off-diagonal element of the NBO-Fock matrix.

Molecular electrostatic potentials (MEPs), on the other hand, provide extensive information on chemical reactivity or biological activity research. The attack of an electrophilic or nucleophilic agent as the primary event of a chemical reaction is determined by the spatial distribution and values of the electrostatic potential. Furthermore, the binding of a substrate to the active site of a receptor is primarily dependent on the three-dimensional distribution of the electrostatic

potential (Gasteiger, Li, Rudolph, Sadowski, and Zupan, 1994). The electrostatic potential is thus precisely defined for each location in space around a molecule, with negative values indicating positive charge attraction and positive values suggesting positive charge repulsive interactions. As a result, individual molecules' MEPs play a crucial part in their interaction. MEPs can reveal molecular recognition mechanisms like drug-receptor interactions. On the other hand, the MEP ignores the arriving species' polarization of the molecule (Levine, 2014).

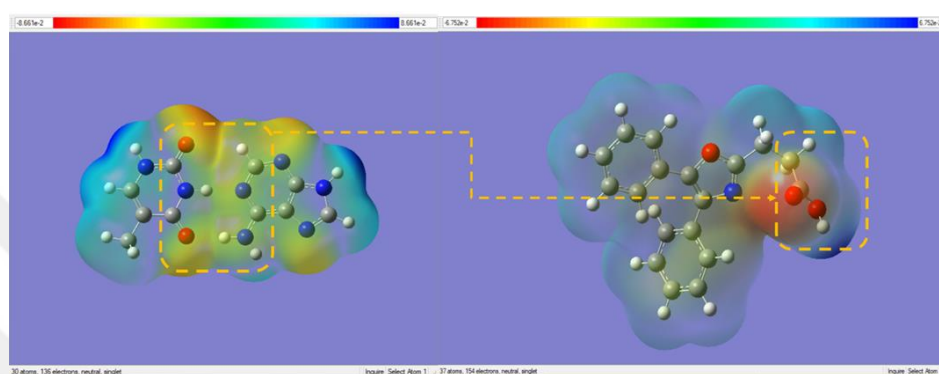


Figure 3. 39 Molecular electrostatic potential maps of A-T nucleobase pair and drug oxaprozin

As shown in Figure 3.39, the adenine-thymine base pair has predominantly negative charges distributed on the surface of the hydrogen bond. In contrast, the drug molecule of oxaprozin has a predominantly positive charge above the hydroxyl group (shown with a region confined by orange-colored dashed lines). Therefore, we can predict that drug-nucleobase interaction might occur in these regions.

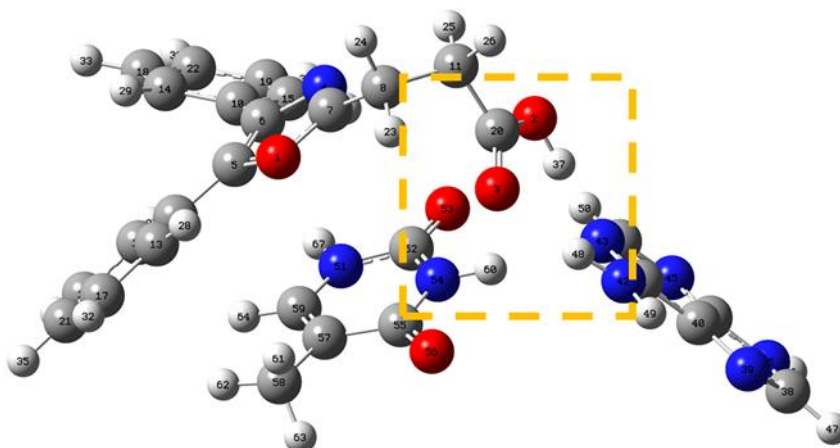


Figure 3. 40 Interaction of Oxaprozin and A-T nucleobase pair result

After optimizing the geometry of oxaprozin-A-T nucleobase pair complex, we found that our prediction devised from the molecular electrostatic potentials is consistent with the optimized geometry of the complex (see Figure 3.40).

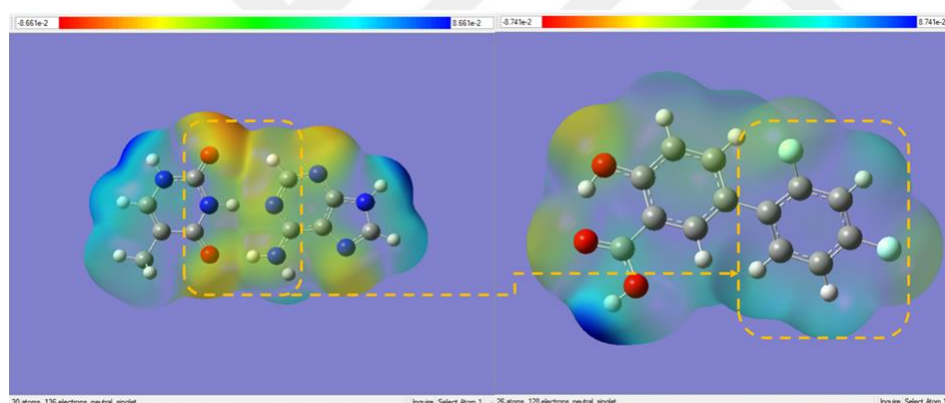


Figure 3. 41 Molecular electrostatic potential maps of A-T nucleobase pair and diflunisal drug molecule

In the region confined by orange-colored dashed lines, it is seen that there are more positive charges on the drug. On the other hand, there are more negative charges on the adenine-thymine nucleobase pair whose part of the region is highlighted in orange border. It is more likely that the drug molecule diflunisal and the nucleobase pair interact on these surfaces.

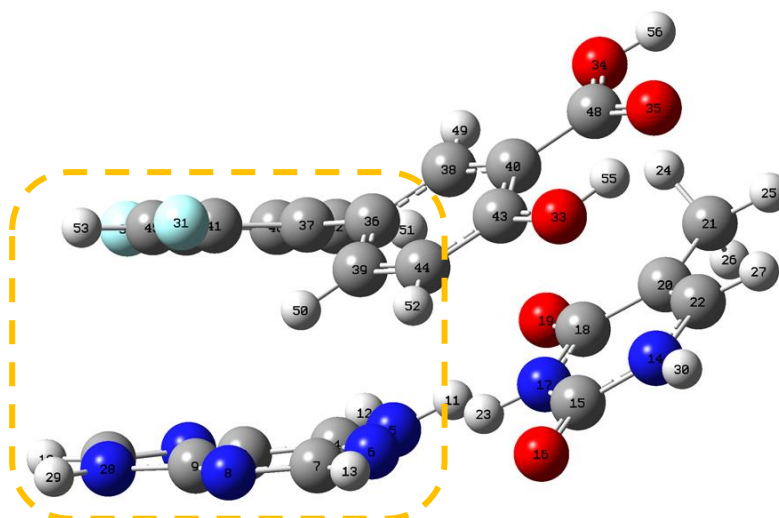


Figure 3. 42 Interaction of diflunisal and A-T nucleobase pair result

As you can see from Figure 3.42, the resultant complex geometry almost perfectly fits our prediction devised from the molecular electrostatic potentials.

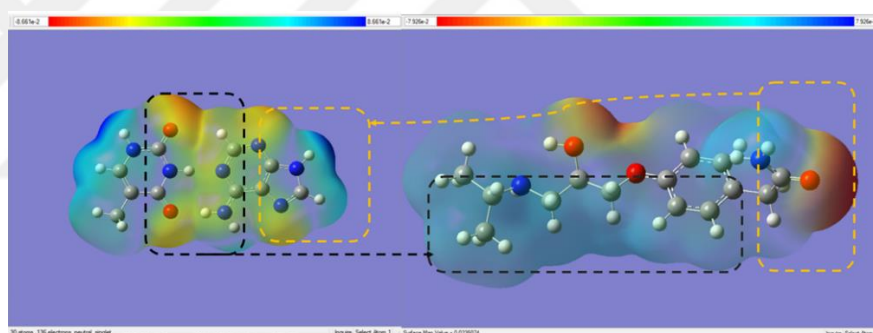


Figure 3. 43 Molecular electrostatic potential maps of A-T nucleobase pair and drug
atenolol

The negative charges are the region marked in black on the adenine-thymine nucleobase pair. The most positive part of the drug surface marked in black will preferentially attack the drug molecule. Similarly, negatively charged part of the drug will attack positively charged part of the nucleobase pair shown with orange color.

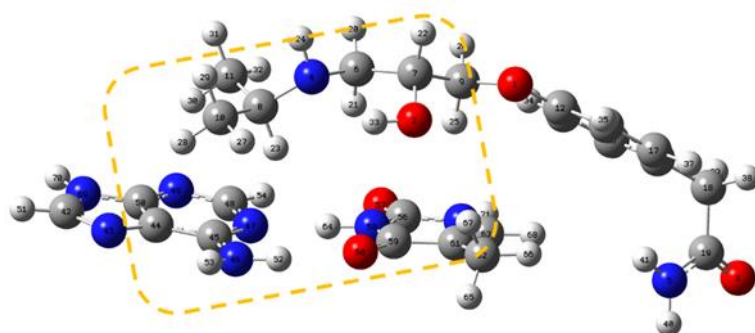


Figure 3. 44 Interaction of atenolol and A-T nucleobase pair result

When we compare the possible prediction with the resultant geometry of the complex, we find that the interaction regions are correct. These graphical maps of MEPs provide insights into molecular recognition processes.

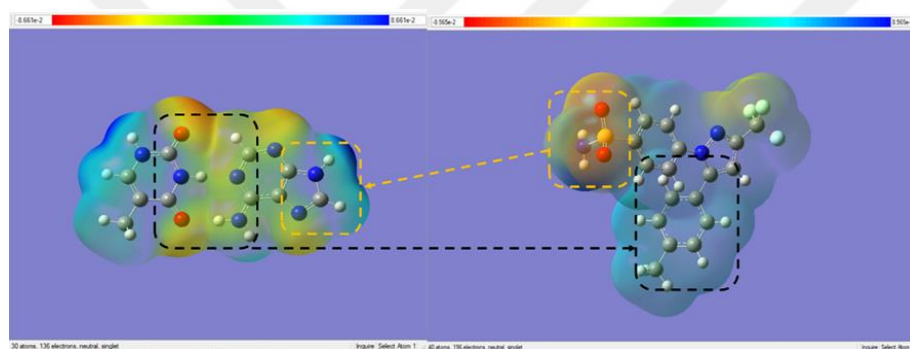


Figure 3. 45 Molecular electrostatic potential maps of A-T nucleobase pair and drug celecoxib

An electrophilic moiety on the drug and the most negative part of the A-T nucleobase pair are marked with black dashed lines. The H-bonding region of the nucleobase pair will attack to the blue part. In addition, other possible attacks are indicated with orange dashed lines. We assume that the interaction occurs from these regions.

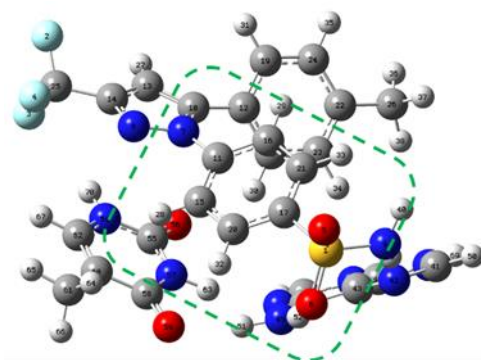


Figure 3. 46 Molecular electrostatic potential maps of A-T nucleobase pair and drug celecoxib

When we compare our result with the optimized geometry of the complex, which is seen in Figure 3.46, we see that our prediction from MEP is correct.

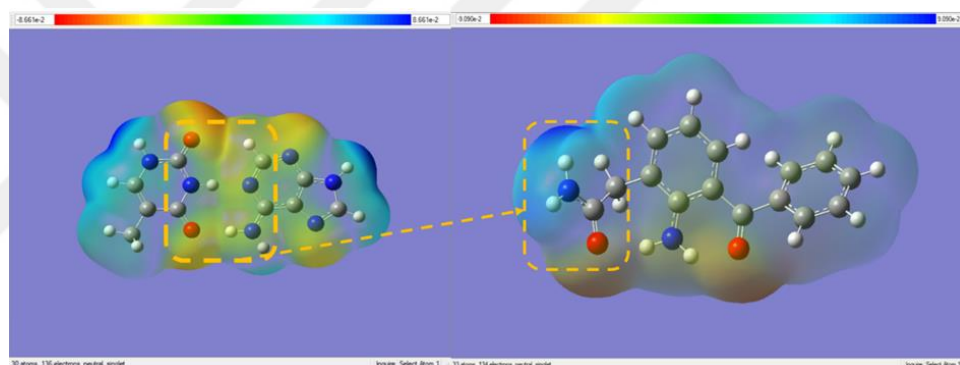


Figure 3. 47 Molecular electrostatic potential maps of A-T nucleobase pair and drug nepafenac

The MEP of nepafenac (the drug molecule) and the A-T nucleobase pair are shown in Figure 3.47. The orange highlighted region of the drug molecule nepafenac has a strong positive charge. As we know, the A-T nucleobase pair has negative charges on hydrogen bonds. Our prediction is that the interactions will occur from these regions, based on the MEP graphical maps.

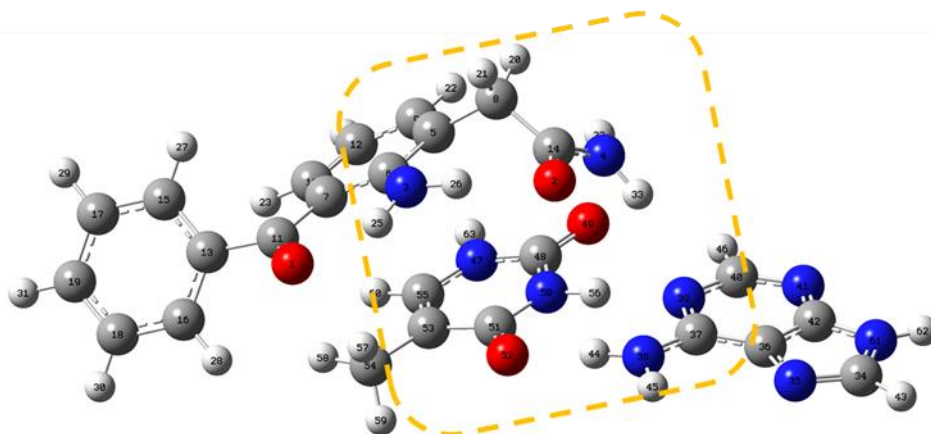


Figure 3. 48 Molecular electrostatic potential maps of A-T nucleobase pair and drug nepafenac

When we checked our result in Figure 3.48 with the above predictions, we found that the interactions occurred in the expected direction. Based on our data and the literature, the MEP analysis provides insight into explaining the donor-acceptor relationship and sheds light on possible interactions.

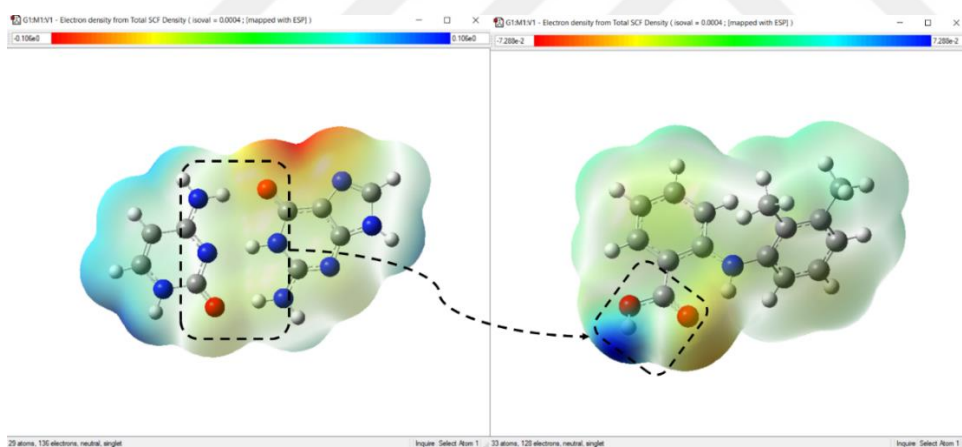


Figure 3. 49 Molecular electrostatic potential maps of G-C nucleobase pair and drug mefenamic acid

The MEP of the drug molecule, mefenamic acid, and the G-C nucleobase pair are shown in Figure 3.49. The black highlighted region of the drug molecule mefenamic acid has a strong positive charge. As we know, the G-C nucleobase pair has negative charges on hydrogen bonds. Our prediction is that the interactions will occur from these regions, based on the MEP graphical maps.

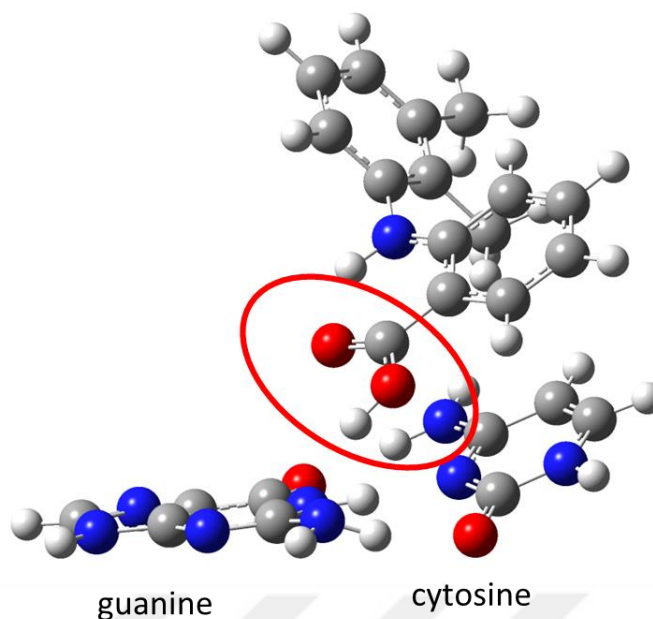


Figure 3. 50 Molecular electrostatic potential maps of G-C nucleobase pair and drug mefenamic acid

When we checked the resultant drug-nucleobase pair complex, we found that the interactions occurred in the expected direction, seen from Figure 3.50. Based on our findings and the literature data, the MEP analysis provides insight into explaining the donor-acceptor relationship and sheds light on possible interactions. In 2019, N. Balamurugan and colleagues studied the MEP of mefenamic acid (Charanya et al., 2019). Their study shows that the MEP plot for understanding the negative region (the region confined with red color in Figure 3.50) was associated with the oxygen atom of the carbonyl group ($C=O$), which is the most favorable site for electrophilic attack. In our study, mefenamic acid attacks the nucleobase pairs from this site. These quantities do not have a well-defined meaning, but these values are usually calculated as the area and volume enclosed by the van der Waals surface or by an isodensity surface with a given ρ -value (Levine, 2014). A very different aspect of predictive utility is the demonstrated ability of NBO descriptors to serve as effective second-order parameters in statistical correlations of experimental data. Such a correlation can be used to explain known experimental trends in the properties of a species or to predictively match the species to the desired target property (Weinhold, Landis, and Glendening, 2016).

3.7 Introduction to QM/MM Part

It has been previously determined by the DFT calculations that mefenamic acid (MEF) and oxaprozin (OXA) are the drugs that can affect the geometric structures of nucleobase pairs. In this section, we present QM/MM simulation of a model DNA and these two drug molecules. Our aim is to observe drug -DNA chain interaction in the presence of water molecules. Therefore, the CP2K program is used for QM/MM MD simulations.

3.7.1 What is CP2K?

Atomistic simulations of solid state, liquid, molecular, periodic, material, crystal, and biological systems can be carried out using the quantum chemistry and solid-state physics software package known as CP2K. For many modeling techniques, such as DFT using the combined Gaussian and plane waves approaches GPW and GAPW, CP2K offers a general framework. DFTB, LDA, GGA, MP2, RPA, semi-empirical techniques (AM1, PM3, PM6, RM1, MNDO,...), and traditional force fields (AMBER, CHARMM,...) are some of the supported theory levels. With the help of the NEB or dimer approach, CP2K can simulate molecular dynamics, metadynamics, Monte Carlo, Ehrenfest dynamics, vibrational analysis, core level spectroscopy, energy minimization, and transition state optimization (Bioexcel, 2021).

3.7.2 Which Levels of Theory Used?

Hybrid quantum mechanics/molecular mechanics (QM/MM) simulations have become a popular tool for investigating large systems, such as enzymes or drug-protein and drug-DNA interactions. In QM/MM methods, the region of the system in which the process takes place is treated at an appropriate level of quantum chemical theory, while the remainder is described by a molecular mechanics force field.

3.7.3 How Did We Define the QM, MM Regions and The Border?

QM region means applying the QM method to atoms in this region. It is very important here to define the boundary between the QM region and the MM region.

The C=C and C-C bonds should not be preferred in border regions as much as possible, because setting them as boundaries can cause problems in charge balance. C-N bonds are more suitable for boundaries. We applied QM/MM MD simulations by taking these into consideration. In our simulations, atoms that are not covalently bonded to each other (drug atoms) were chosen as boundaries.

In the following subtitle, the simulation results of mefenamic acid - DNA interaction was presented. As mentioned in the introduction, mefenamic acid is the drug molecule with a lot of side effects, and it is the drug molecule that causes the most striking changes in geometries when it interacts with nucleobase pairs.

3.8 MD Simulation Of Mefenamic Acid-DNA Interaction with QM/MM Method

Molecular docking simulation was performed by Haddock(Van Zundert et al., 2016),(Honorato et al., 2021). Molecular docking determines the ligand's (drug) optimal binding position with relation to the target or receptor. The idea behind molecular docking is to calculate the binding affinity of various ligand postures and then rank them based on that binding affinity. The ligand's optimal posture corresponds to the position with the lowest binding energy. To produce a ligand-receptor complex, the optimal posture of the ligand is coupled with the receptor structure (Misra and Yadav, 2022). Here, mefenamic acid was docking with DNA in the effect of implicit water solvation. After simulation, the Haddock program ranked the MEF binding affinities, we took the best docking result and used it for the further part of the study. Particularly, the best docking result was subjected to the MD equilibration employing the QM/MM method of the Gromacs+CP2K interface. Then, the result of the MD equilibration was subjected to the production employing only CP2K program.

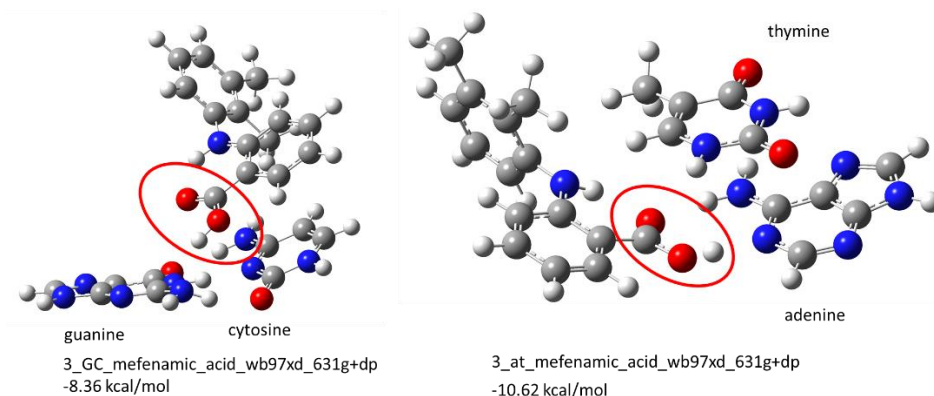


Figure 3. 51 Orientations of mefenamic acid–nucleobase pairs ((A-T), (G-C)) complexes optimized by ω B97X-D /6-31+G(d,p)

From our previous DFT- ω B97X-D results, we know that due to the electron-withdrawing effect of its carboxyl group, mefenamic acid interacts with nucleobase pairs through this part, as seen in Figure 3.51. The ω B97X-D method predicts that while mefenamic acid completely breaks the A-T base pair, it only bends G-C base pair. In the MD simulations our expectation is that mefenamic acid interacts with A-T base pair in a way that it breaks the H-bonds. The docking result of the MEF with DNA (shown in Figure 3.53) indicates that the drug interacts with G-C base pair contrary to our expectation. Most probably the result of the DFT method may not be fully correct for A-T base pair interaction because of the interaction of single base pair rather than the real DNA structure having lots of these base pairs.

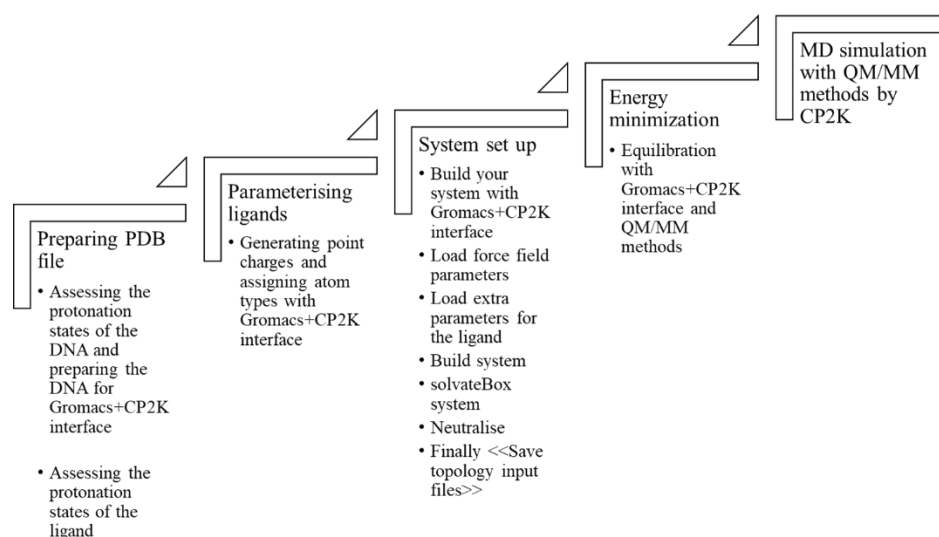


Figure 3. 52 MD simulation steps of mefenamic acid- DNA interactions with Gromacs+CP2K interface of QM/MM method

Figure 3.51 shows that the QM/MM process steps which applied.

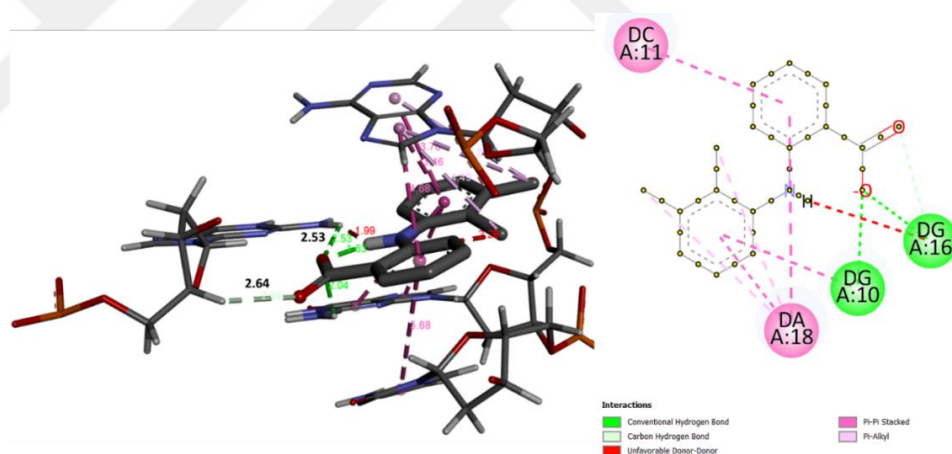


Figure 3. 53 Haddock docking result of the MEF with DNA

This is also our initial structure for the MD simulation. Discovery Studio was used to examine the drug-DNA interaction sides. Our expectation from our previous quantum modeling studies is that there is an interaction between the carboxyl regions of the drug and the DNA bases.

3.8.1 The Process of Setting Up a Simulation System Containing DNA in Complex with Mefenamic Acid

First, our docking result was separated into DNA and drug of mefenamic acid with the grep command and the topology of the system was created with Gromacs. The system topology was prepared in two steps:

- Preparing the DNA topology with pdb2gmx
- Preparation of the ligand topology with external tools (bio2byte).

The force field amber14sb_OL15.ff_corrected-Na-cation-params was chosen for the 24 residue of B-DNA (name of 1BNA from PDB). The force field gaff2 was applied to the ligand with the server tool bio2byte (Bio2byte,2022). Complex was solvated with TIP3P water (~10400 solvent molecules), and 23 Na⁺ ions were added to neutralize the system. The result obtained here was added to the system topology, and then energy minimization was performed. (The number of ions to be added for neutralization is automatically calculated by Gromacs).

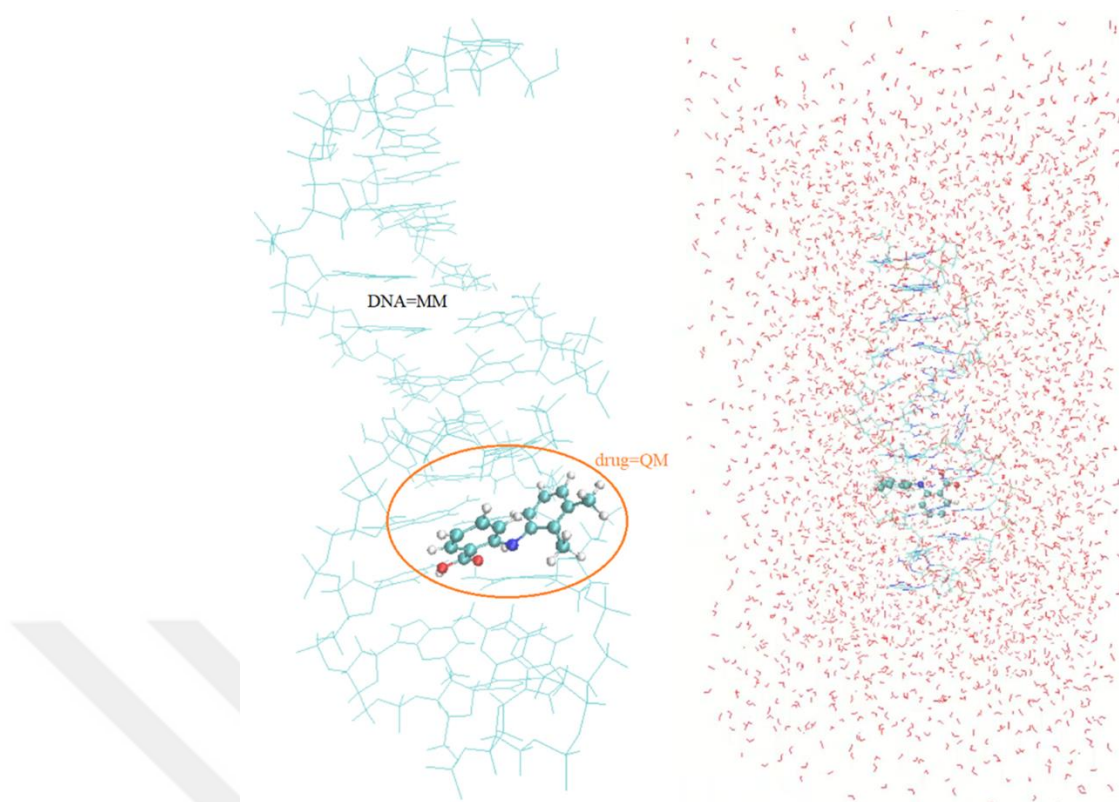


Figure 3. 54 Total system of MEF-DNA

The Figure 3.54 show that all system and it has nearly 10400 water molecules, 24 DNA residues and 33 atoms of drug molecules. Our all system has 11237 atoms. The chemically active region is that drug molecule of mefenamic acid for our own system. We used molecular mechanics for whole region. The QM mechanics region contains 33 atoms in total. The maximum atom number of the quantum mechanical region is 100. when we wanted to add the nucleobase region, then our atomic number approached 100 and this means twice more time and costly. Therefore, we continued our calculations by selecting only the drug region for QM mechanics.

Moreover, the goal of energy minimization is to ensure that there are no steric overlaps or mismatched geometry in the system before the dynamics begin. In this section, only drug was selected for QM region and PBE/DZVP-MOLOPT-GTH method was used in all calculations. When using the DFT method, a larger QM region leads to longer calculations times and many errors due to the boundary regions. Therefore, QM method was used for only drug molecule.

3.8.2. Energy Minimization (EM)

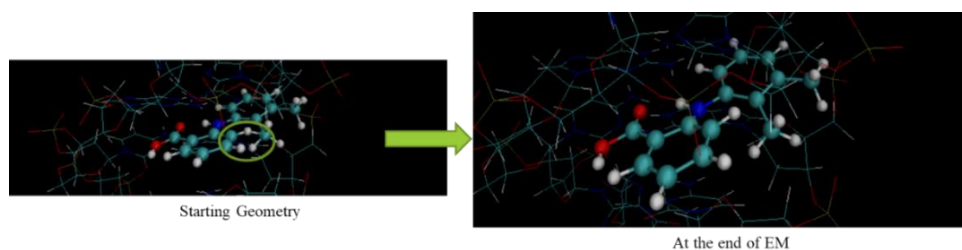


Figure 3. 55 Energy minimization of geometry result

As a result of the docking procedure, the hydrogen atoms in the ligand (ligand= drug of mefenamic acid), are deleted. As with any dynamic simulation preparation, the deleted hydrogen atoms are added back in the preparation of the ligand topology. As a result of this re-addition, it can be clearly seen in the Figure 3.55 that the hydrogens overlap. However, the structure was reached to optimized geometry by energy minimizing using the method QM /MM with Gromaccs+CP2K interface. QM was applied only to the mefenamic acid. EM was performed with a time step of 1 fs in total 50000 steps. Xmgrace software was used for graph plotting (Turner, 2005).

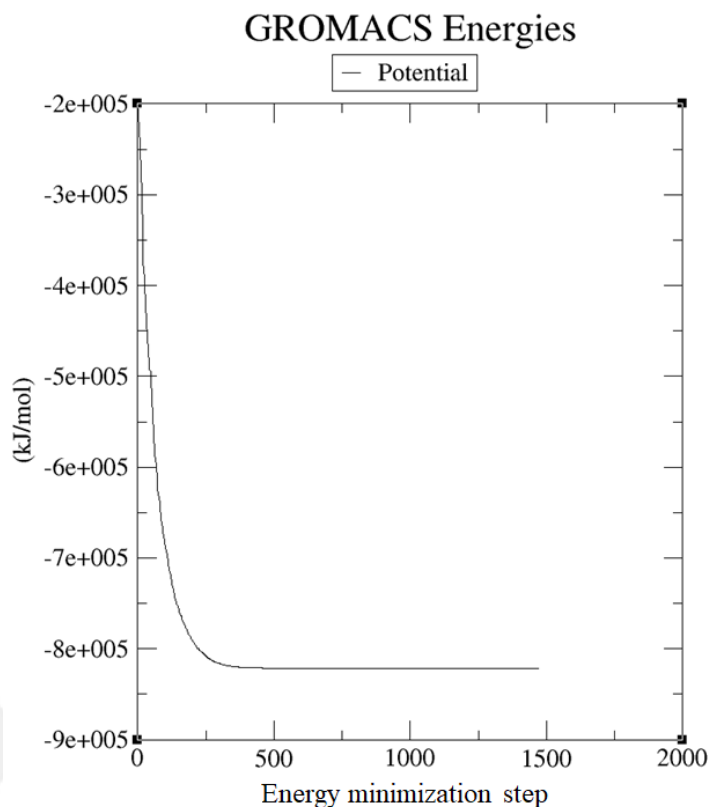


Figure 3. 56 Potential energy from EM

It has been converged as it has reached energy minimization in 1477 steps. Potential energy of -8.21 kJ/mol (-1.96 kcal/mol) as shown in the Figure 3.56. The EM process ensured a reasonable geometry and solvent orientation for the starting structure. After that, real dynamics started.

3.8.3 NVT Equilibration

A solution and ions must be equilibrated around the drug-DNA complex. Unrestrained dynamics may lead to collapse at this point. Solvents are usually optimized within themselves and not necessarily with the solute. The solute (DNA) must be brought to the temperature if we want to simulate and properly aligned. Once we have reached the correct temperature (based on the kinetic energies), we apply pressure until the system reaches the proper density. Equilibration normally occurs in two phases. Phase one is carried out under NVT conditions (constant number of atoms, volume and temperature).

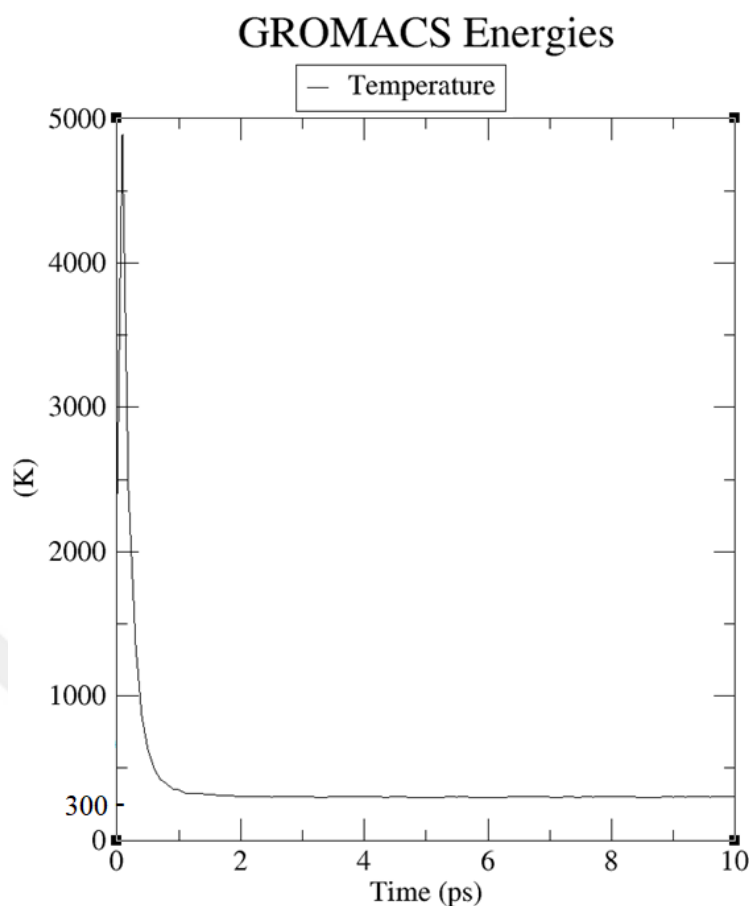


Figure 3. 57 Temperature from NVT equilibrium

The term "isothermal-isochoric" is also used to describe this ensemble. The duration of such a procedure is determined by the content of the system, but in NVT the temperature of the system should reach a plateau at the appropriate value.

A dynamic simulation with a time step of 2 fs in total 50000 steps (100ps) can be completed in 5 minutes using the only MM method with Gromacs. However, a dynamic simulation with a time step of 2 fs in total 5000 steps (10ps) can be completed in 3 days using the QM /MM method with Gromacs+CP2K interface in HPC. Figure 3.57 shows the QM /MM molecular dynamics simulation reaching the optimal temperature after 3 days. Running on 1 node with total 16 cores, 32 processing units for both. Berendsen thermostat was used for NVT ensemble equilibration at a constant temperature of 300K for 10 ps. The temperature of the ensemble reaches the target value (300 K) and remains constant for the rest of the equilibration as shown in Figure 3.57.

3.8.4 NPT Equilibration

The system temperature was stabilized in the previous phase of NVT equilibration. Before data collection can begin, the pressure of the system (and thus the density) must be stabilized. For the pressure calibration, an NPT ensemble is used in which the pressure, temperature and number of particles are constant. An ensemble similar to the experimental conditions is also called an "isothermal-isobaric" ensemble.

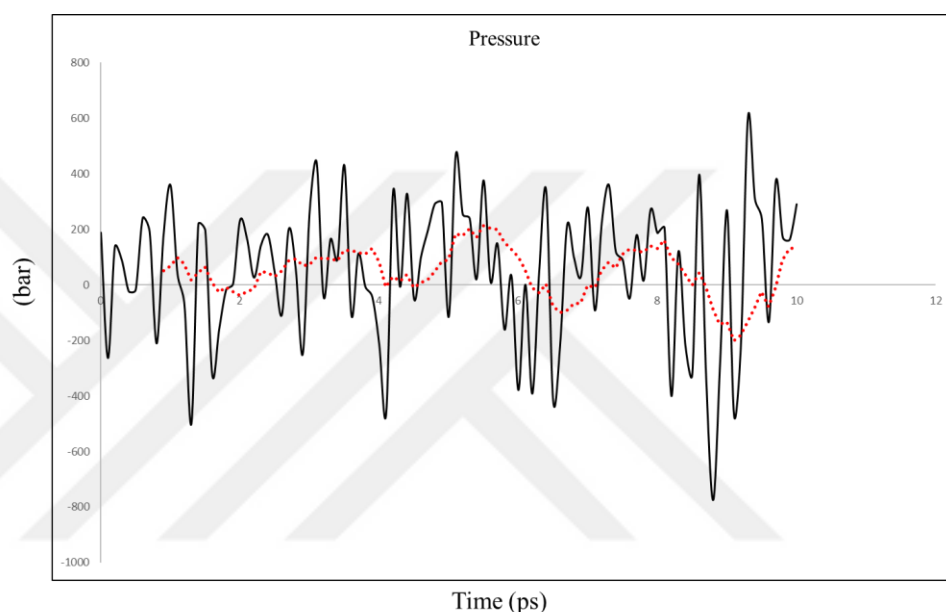


Figure 3. 58 Pressure from NPT equilibrium

The processing time mentioned in the previous section also applies here. A dynamic simulation with a time step of 2 fs in total 50000 steps (100ps) can be completed in 5-6 minutes using the only MM method with Gromacs. However, a dynamic simulation with a time step of 2 fs in total 5000 steps (10ps) can be completed in 7 days using the QM /MM method with Gromacs+CP2K interface in HPC. Running on 1 node with total 16 cores, 32 processing units for both. The NPT ensemble was equilibrated in 5000 steps with integration at 1atm constant pressure.

During the 10-ps equilibration phase, the pressure value widely fluctuates, but this is not surprising as can be seen from Figure 3.58. The red line in the plot represents the running average of the data. During equilibration, the average pressure is 51.82 ± 248.61 bar. Since the reference pressure was set to 1 bar, is the

result acceptable? As evidenced by the large root-mean-square deviation (248.61 bar), pressure fluctuates widely during an MD simulation, which means statistically speaking one can't distinguish any difference between the average (51.82 ± 248.61 bar) and the target/reference value (1 bar). The Berendsen barostat was used for simulation.

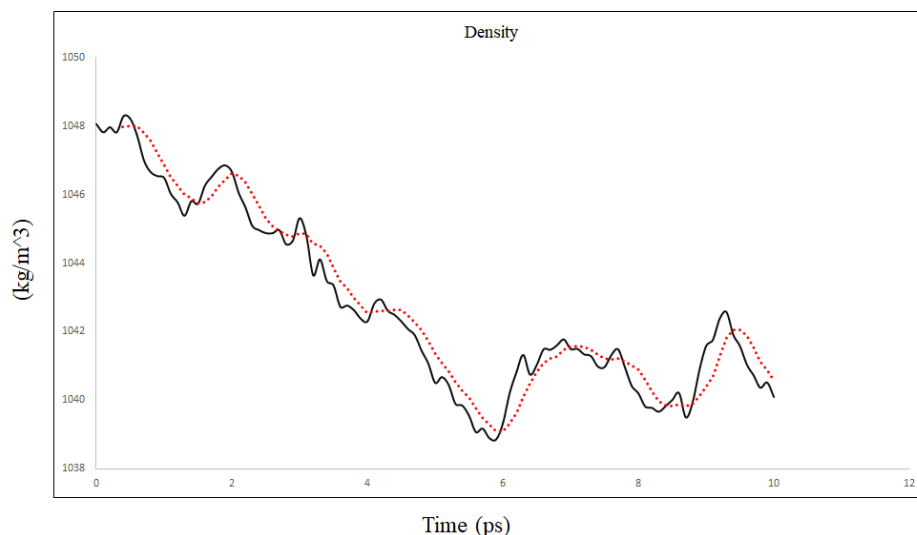


Figure 3. 59 Density from NPT equilibrium

The running average of density is also plotted in red in figure 3.59. The average value over the duration of 10 ps is $1042.77 - 1040 \text{ kg m}^{-3}$. The density values are almost constant over time, indicating that the system is now pressure and density-equilibrated.

3.8.5 Production

Last part of the simulation, production, with a time step of 2 fs in total 25000 steps (50ps) can be completed in 7 day using the QM/MM method with only CP2K program in HPC. Running on 2 node with total 40 processing units. When using the Gromacs+CP2K interface to perform this final simulation, a timeout error occurred (on HPC systems, a job's maximum run time is 15 days), therefore the production part was completed only using the CP2K software.

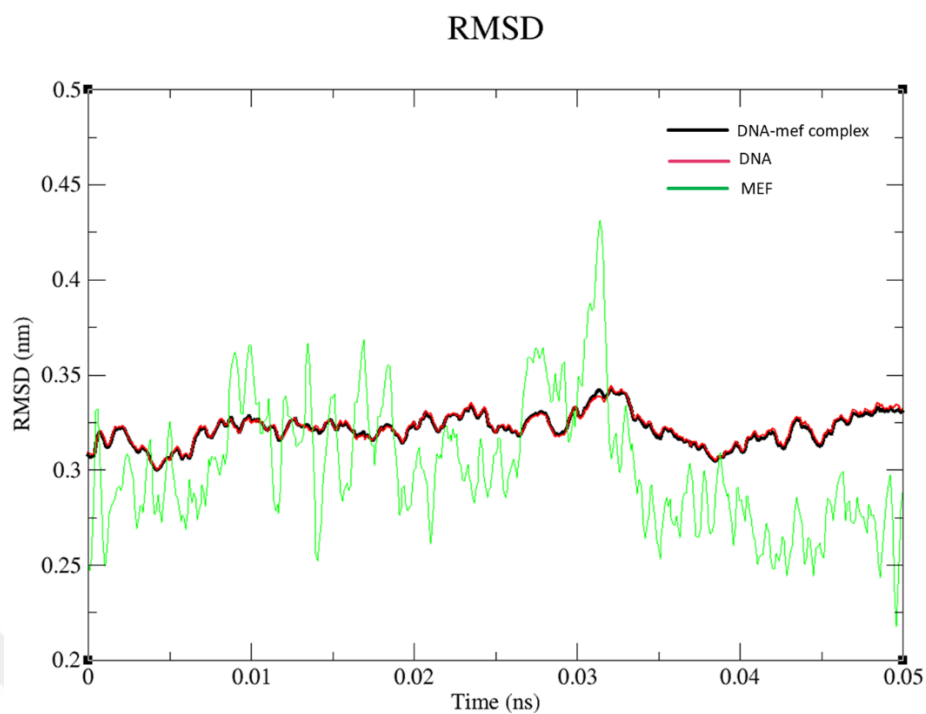


Figure 3. 60 Graphs of the RMSD relative to the structure of crystal DNA

RMSD values are calculated relative to the crystal structure of DNA. By doing so, the overall rotation and translation of the DNA is removed via fitting and the RMSD reported is how much the MEF position has varied relative to the DNA. Moreover, the DNA remains intact during the simulation time. The DNA-MEF complex appears to be a stable complex.

Radius of gyration (total and around axes)

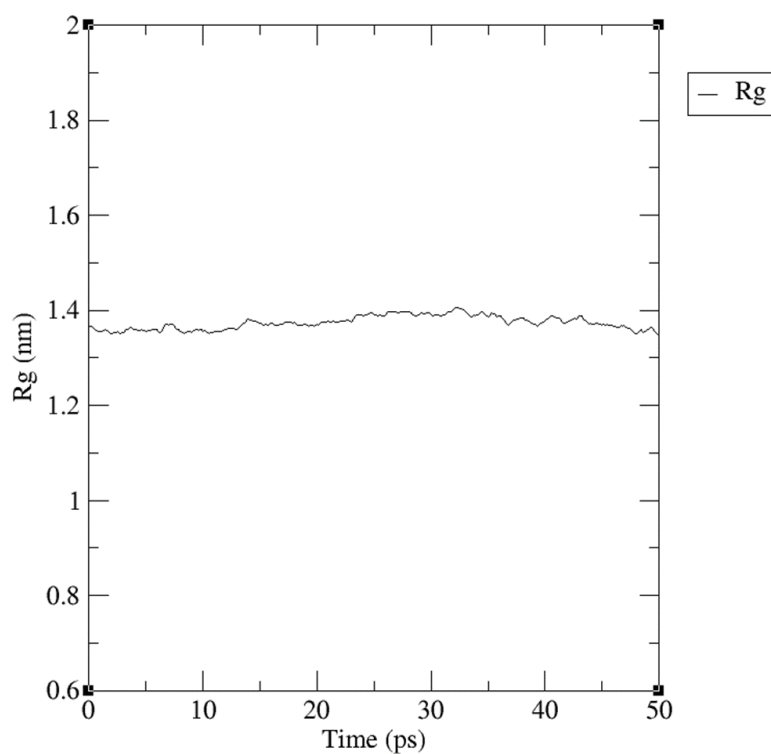


Figure 3. 61 Graphs of the radius of gyration to the structure present in the minimized, equilibrated system

DNA's radius of gyration is a measure of its compactness. If a DNA is stably folded, its R_g will likely remain relatively constant. DNA unfolding causes its R_g to fluctuate. The Figure 3.61 show that R_g has nearly remained relatively constant.

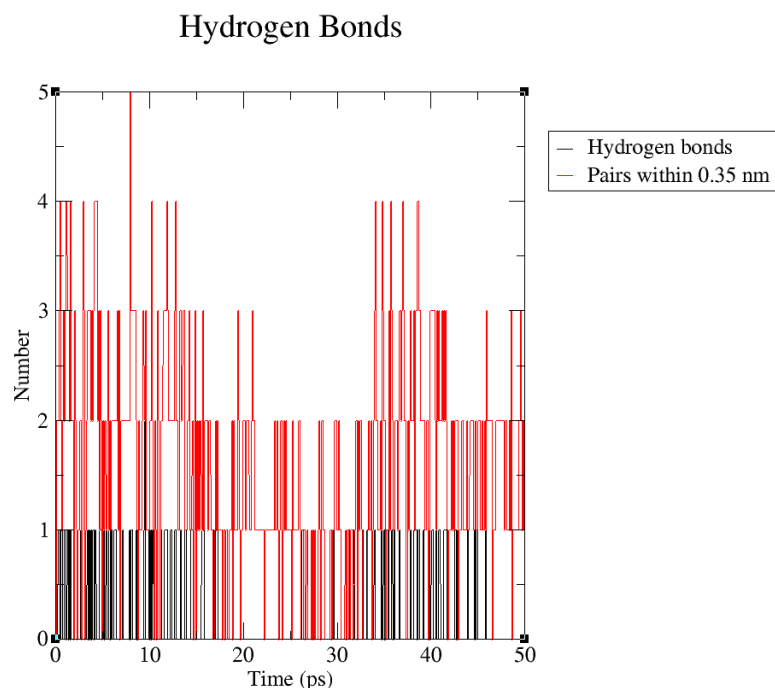


Figure 3. 62 Total hydrogen bonds numbers of MEF-DNA

For the ligand-target structure to be stable, the hydrogen bond is crucial. A complex becomes more stable when there are more hydrogen bonds between the ligand and the target structure. Number of hydrogen bonds formed between target and ligand determines bonding strength. As shown in Figure 3.62, the hydrogen bonds formed between the ligand and the DNA was plotted the simulation time. The graph shows the hydrogen bonds formed between DNA and MEF. Looking at the graph, there are five possible hydrogen bonds between DNA and MEF. In order to see more precise results, the MD simulation must be examined visually.

Figure 3.63 shows the interactions that occur as a result of the production simulation. Interactions have occurred between the cytosine nucleobases of the DNA chain and the drug. The length of an average $\text{-O}\cdots\text{H-}$ bond is around $1.80 \text{ \AA}$ - $1.90 \text{ \AA}$. In our QM /MM simulation result, the distance between the $\text{-O}\cdots\text{H-}$ interactions is almost $1.70 \text{ \AA}$. This means that an H-bond can be formed between the drug and the DNA nucleobase cytosine.

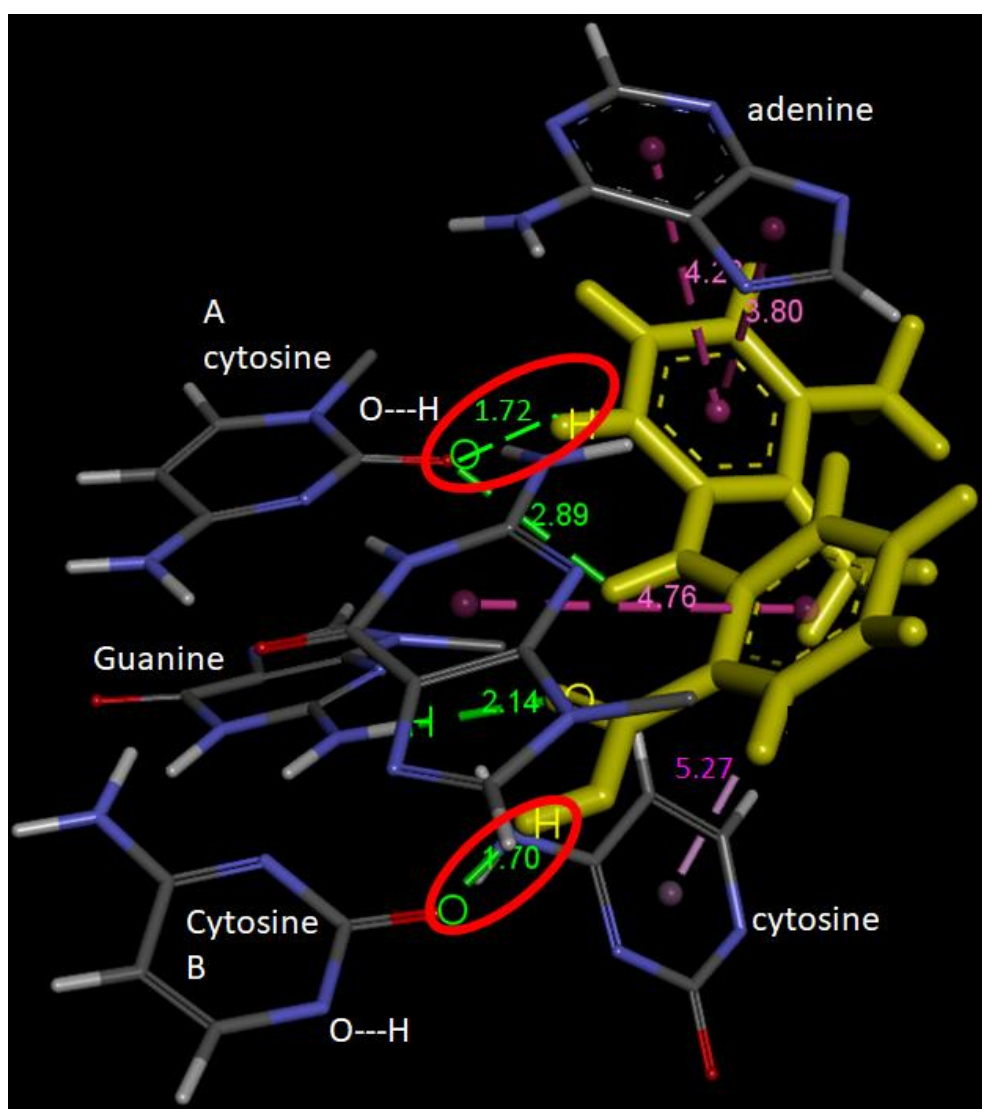


Figure 3. 63 MEF-DNA interactions at the end of production simulation with QM/MM method (Discovery studio was used for drug-DNA interactions) (distance unit is ångstrom)

As shown in Figure 3.63, the yellow molecule is the drug molecule of mefenamic acid. The drug molecule of mefenamic acid interacts mainly with the nucleobase cytosine. During the simulation, the drug molecule interacts with A and B nucleobases belonging to DNA. More changes are observed in the hydrogen bonds within the B nucleobase pair. Moreover, mefenamic acid showed remarkable affinity with cytosine at these interaction sites in DNA with QM/MM method. Figure 3.63 shows that B side cytosine and mefenamic acid has 1.70 ångstrom distance. It means that a new tight hydrogen bond is formed since it is shorter than the normal $\text{O}\cdots\text{H}$ - bond distance (1.80 ångstrom). All these means that mefenamic

acid leads to broken or damaged nucleobase pair. Especially B side. To better understand the interaction taking place here and to compare it with our earlier quantum calculations, the DFT calculations was also applied to the drug molecule with highly interacting cytosine-guanine base pairs, and the interaction energy was determined.

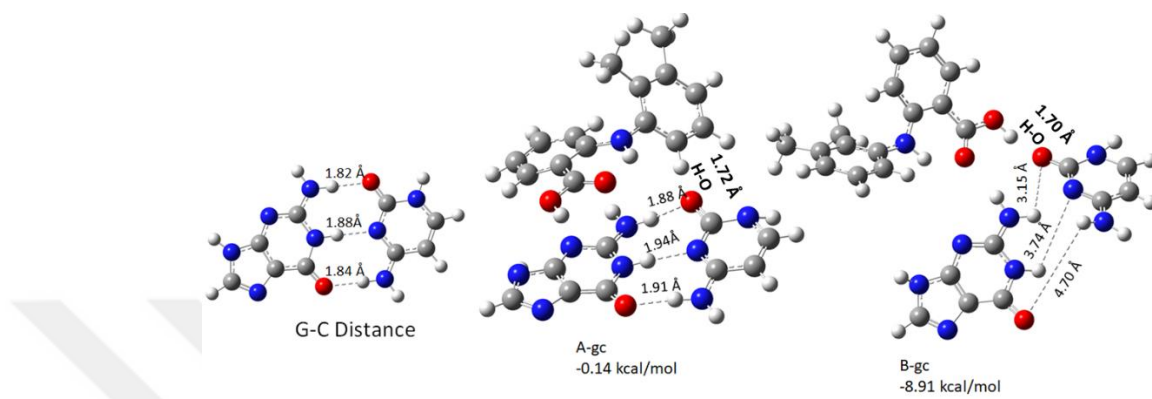


Figure 3. 64 Finding the interaction energy of the highly interacting regions on the DNA chain with ω B97X-D /6-31+G(d,p)

Table 3. 25 Comparison of interaction values in large system with previous quantum model study

Earlier Orientations drug - base pair	E_0 +BSSE	E_0 +BSSE	Last DNA-mef interaction side
1gc_wmefenamic_acid	-11.61	-0.14	A-gc
2gc_wmefenamic_acid	-9.47	-8.91	B-gc
3gc_wmefenamic_acid	-8.36		

As we know that mefenamic acid has led to the most changed structures among all drug molecules in our previous section. It also produced remarkable results in its interaction with DNA. The areas where the mefenamic acid drug molecule interacts in the DNA are seen separately in Figure 3.64. Mefenamic acid showed remarkable affinity with cytosine at these interaction sites. The average distance of the $-O \cdots H$ -bond between them is 1.70 Å. This means that mefenamic acid can bond with DNA, thus causing a change in DNA. Therefore, to better understand these interaction sites, the interaction energy between the cytosine-guanine nucleobase pair and the drug was calculated with ω B97X-D /6-31+G(d,p). In addition, the results obtained were compared with previous quantum models in table 3.25. It is clearly seen from Figure 3.64 that the geometry of the G-C nucleobase pair in the B site is quite

considerably changed. The distances between them are larger than 3 Å. This shows that they have almost twice the distance than normal.

When we examine the magnitude of the interaction energies, it is seen that it is almost the same as our previous model quantum calculations. This increases the reliability of quantum models for non-bonded interactions. Our previous comparisons determined that the ω B97X-D method works better for non-bonded interactions than others. In this part, we were able to check the accuracy of this result, and we saw that this result is still valid. The interaction energy in the B site in DNA is close to the previous quantum model energies.

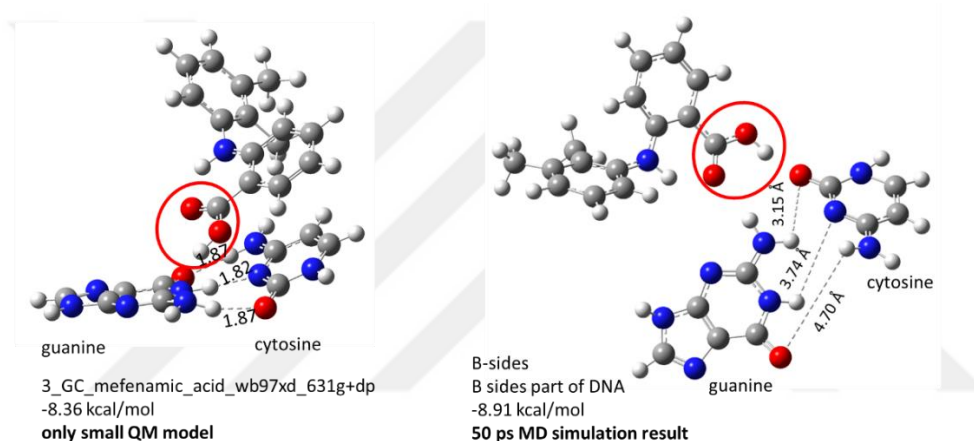


Figure 3. 65 Comparing of QM model result and MD simulation with QM/MM result

Figure 3.65 show the results of QM model and MD simulation with QM/MM method. Both results showed that carboxylic acid group of mefenamic acid is willing to interact with the nucleobase pair.

All these results indicate that the QM model system (DFT) has the ability to show interacting atoms. Therefore, the model QM systems can be considered as trustable guides before applying QM/MM methods to large systems.

3.8.6 Total 100 ps of Production Part

In this part, same system examined in 100 ps. Because the maximum duration in QM/MM simulations are generally 100 ps. However, any critical different interactions were not observed after 50 ps. Therefore, we decided that the 50 ps simulation time is enough for our system.

3.8.7 Experimental Researchs About MEF

In 2001, Weis et al. investigated the role of store-operated calcium entry in human colon cancer cells, which require calcium for proliferation. In colorectal cancer, nonsteroidal anti-inflammatory drugs (NSAIDs) decrease proliferation and angiogenesis. Their results show that inhibition of calcium entry by MEF has an anti-proliferative effect on human cancer cells depending on the dose (Weiss et al., 2001). (NSAIDs) have known that anti-proliferative properties against a variety of cancers. (Morgan, 1996), (Peleg, Lubin, Cotsonis, Clark, and Wilcox, 1996), (Vainio, Miller, and Bianchini, 1997). In 2004, Woo et. al investigated the proliferation inhibitory effect of MEF in liver cancer cells and whether MEF inhibits cell growth via apoptosis. MEF could be a viable drug for chemoprevention of liver cancer, according to the findings (Ho Woo, Han, and Jung, 2004). Based on their DNA histograms taken with a flow cytometer, MEF alone cured the majority of apoptotic cells. MEF also suppresses the proliferation of liver cancer cells in a concentration- and time-dependent manner, according to their research. These experimental studies show that the MEF has a stopping effect on cancer cells, depending on the duration of use and dosage. As known that anti-cancer drugs react with DNA and cellular proteins. Although theoretical studies show that MEF interacts with DNA nucleobases is supported by the experimental studies described here, it cannot be shown as conclusive evidence. The fact that this theoretically observed interaction can also be observed in experimental studies increases the reliability of quantum studies and QM/MM studies.

3.9 MEF Interaction with A-T Nucleobase Pair in DNA with MD Simulation of QM/MM Method

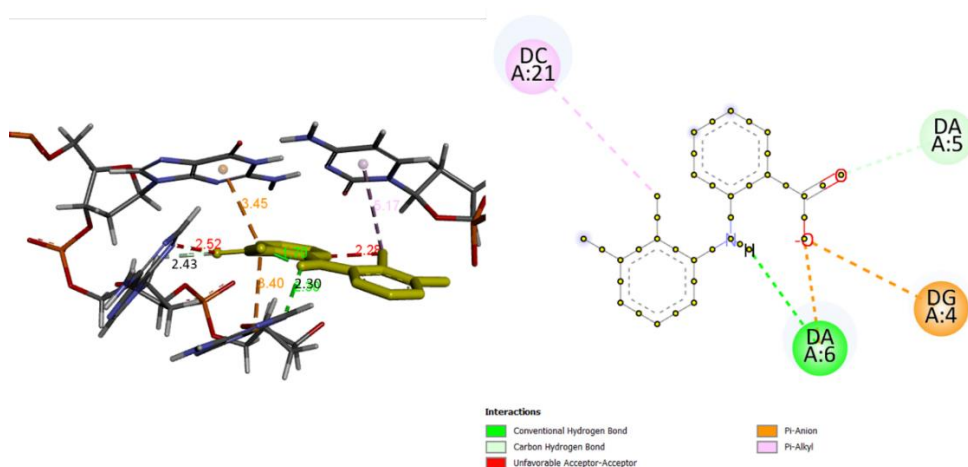


Figure 3. 66 Haddock docking result

Haddock gave 36 docking results, and it also sorts by energy and recommends the best docking result. We have used that one before section. But in this section, we examined possible A-T nucleobase pair region from docking results. Figure 3.66 shows the result of docking. This is also our initial structure for the MD simulation. Discovery Studio was used to examine the drug-DNA interaction sides.

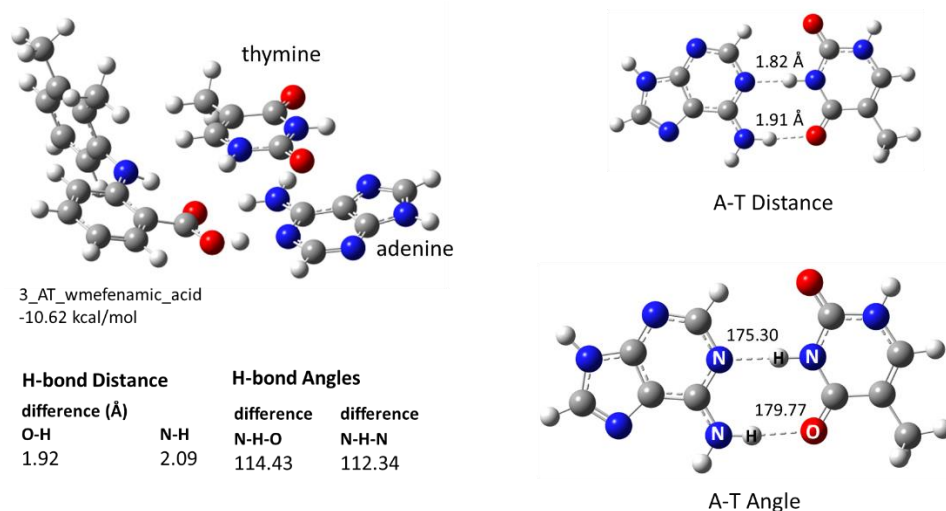


Figure 3. 67 QM model of MEF interaction with A-T nucleobase pair and their H-bonds distance and angles changes

The Figure 3.67 also showed that our expectation from our previous quantum modeling studies is that there is an interaction between the carboxyl region of the drug and the DNA bases.

3.9.1 The Process of Setting up A Simulation System of MEF Interaction with A-T Nucleobase Pair In DNA

Energy minimization was started after all the preparing steps in the previous section were applied in the same way.

3.9.2 Energy Minimization (EM)

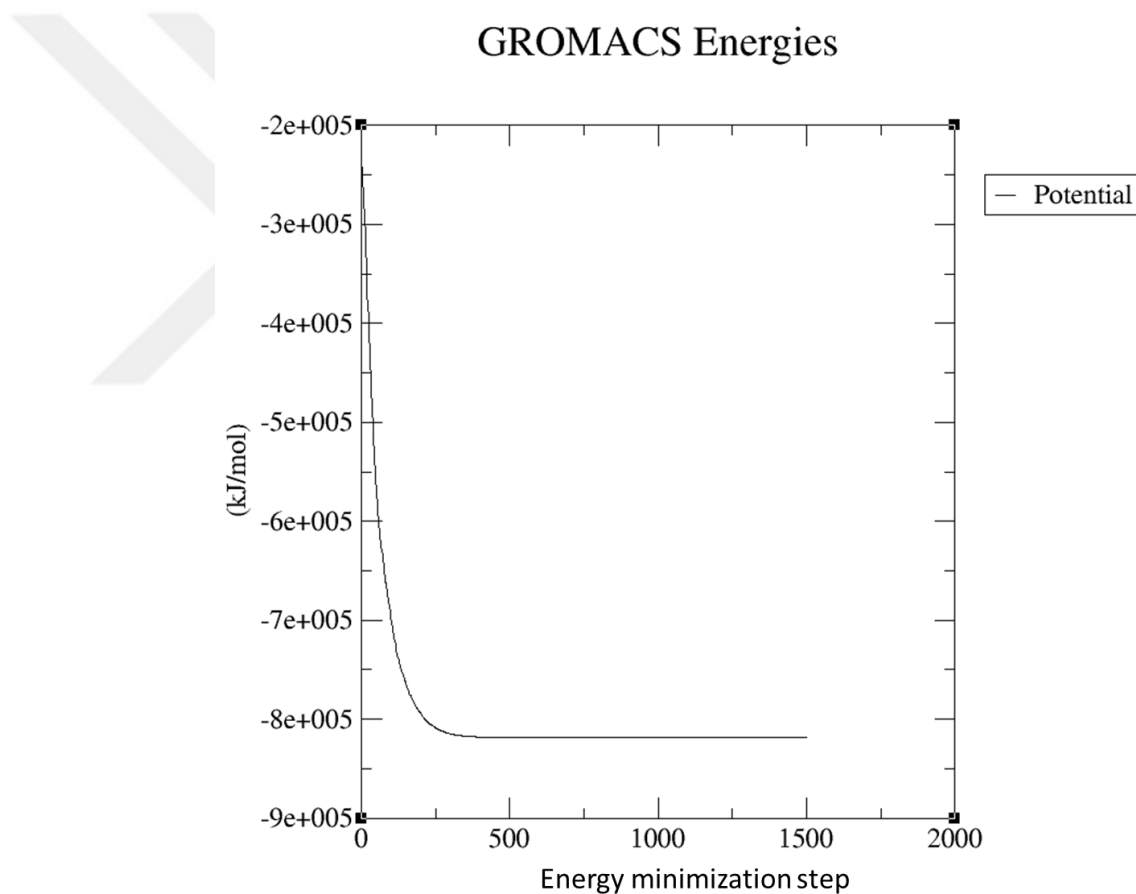


Figure 3. 68 Potential energy from EM

It has been converged as it has reached energy minimization in 1499 steps. Potential energy of -8.19 kJ/mol (-1.95 kcal/mol) as shown in the figure 3.68. The EM process ensured a reasonable geometry and solvent orientation for the starting structure. After that, real dynamics started.

3.9.3 NVT Equilibration

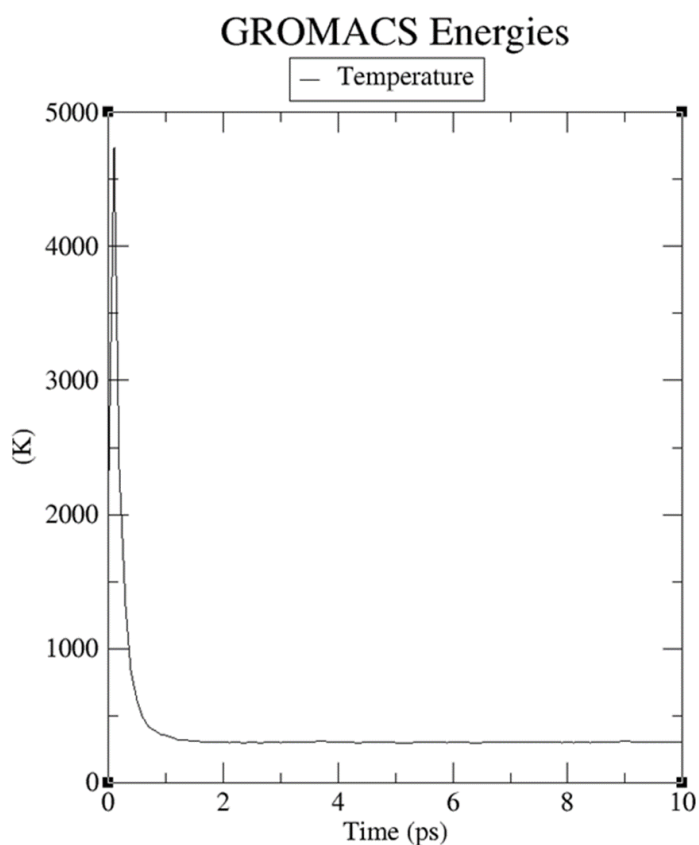


Figure 3. 69 Temperature from NVT equilibrium

A dynamic simulation with with a time step of 2 fs in total 5000 steps (10ps) can be completed in 3 days using the QM /MM method with Gromacs+CP2K interface in HPC. Graph 2 shows the QM /MM molecular dynamics simulation reaching the optimal temperature after 5 day. Running on 1 node with total 16 cores, 32 processing units for both. The temperature of the reaches the 299.17 K (target value of 300 K) and remains constant for the rest of the equilibration as shown in Figure 3.69.

3.9.4 NPT Equilibration

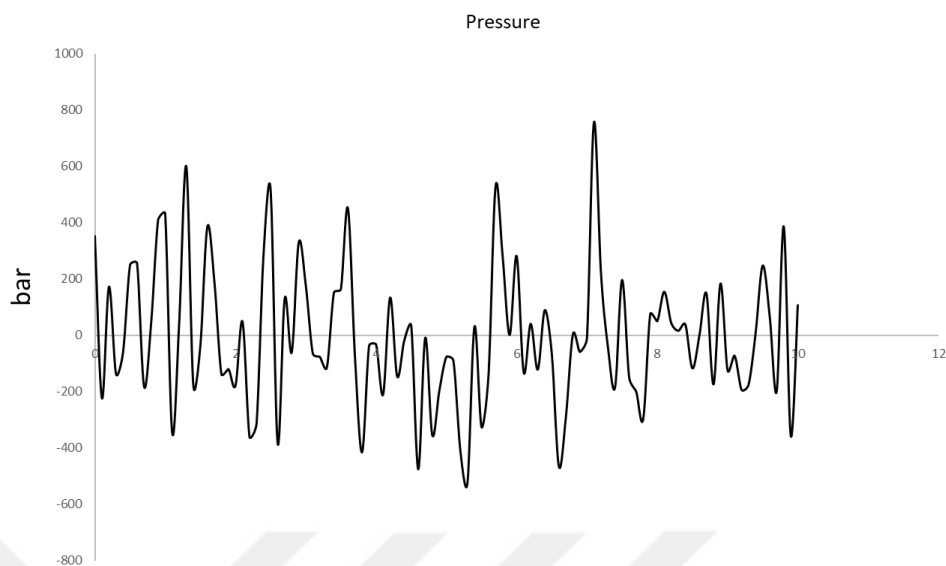


Figure 3. 70 Pressure from NPT equilibrium

A dynamic simulation with a time step of 2 fs in total 5000 steps (10ps) can be completed in 7 days using the QM /MM method with Gromacs+CP2K interface in HPC. Running on 1 node with total 16 cores. During the 10 ps equilibration phase, the pressure value widely fluctuates, but this is not surprising as can be seen from Figure 3.70. The Berendsen barostat was used for simulation.

3.9.5 Production

In this part, same system examined in 50 ps with QM/MM MD simulation.

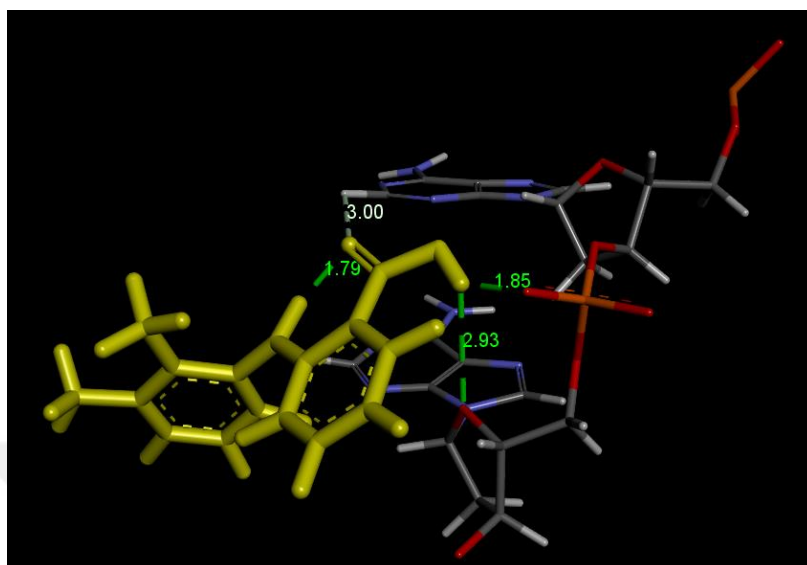


Figure 3. 71 MEF-DNA interactions at the end of production simulation with QM/MM method (Discovery studio was used for drug-DNA interactions) (distance unit is ångstrom)

The only interaction seen as a result of the simulation is between the carboxyl group of the MEF drug molecule and phosphate. The common part of both MEF-DNA simulations is that the MEF drug interacts with DNA through the carboxyl group.

When these interactions were examined with Discovery studio, it was seen that the distance between the two molecules was 1.85 angstroms. Therefore, mefenamic acid has willing to make bond with DNA. Basically, the drug tends to bind with DNA in both mefenamic acid analysis.

3.10 DNA-Oxaprozin Drug Simulation with PM3/MM Methods (QM/MM)

In this part of the study, we utilized the PM3 method, a semiempirical quantum chemical method, to define the QM region and the Amber forcefield to define the MM subsystem. Figure 3.72 shows the steps of QM/MM MD simulation preparation. Molecular docking simulation of OXA- DNA is performed by Haddock, and we used the Haddock application's lowest energy structure for MD simulations.

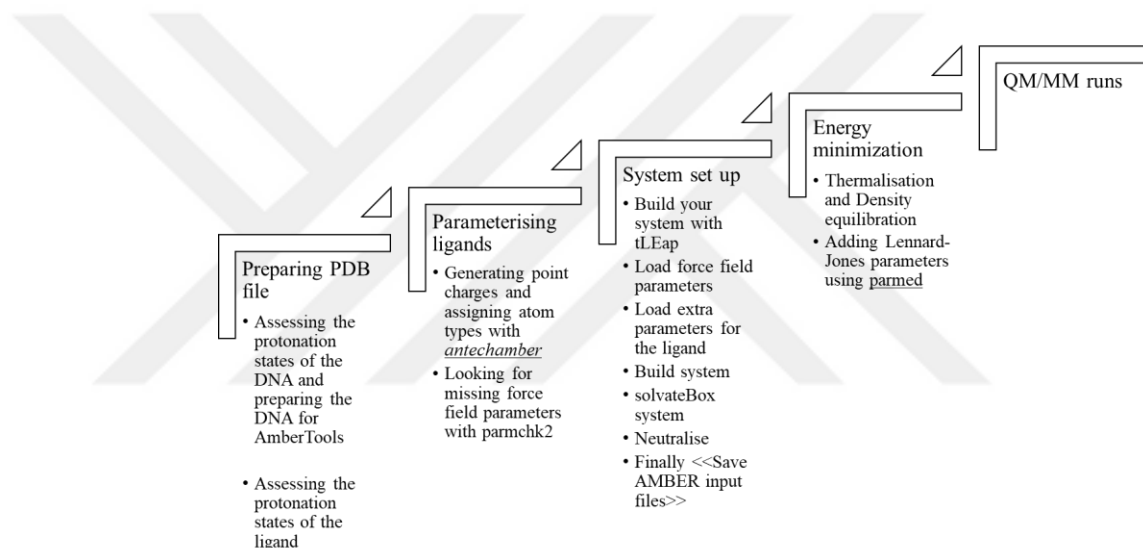


Figure 3. 72 Schematic of MD simulation steps of oxaprozin–DNA interactions with QM/MM method using the Amber and CP2K

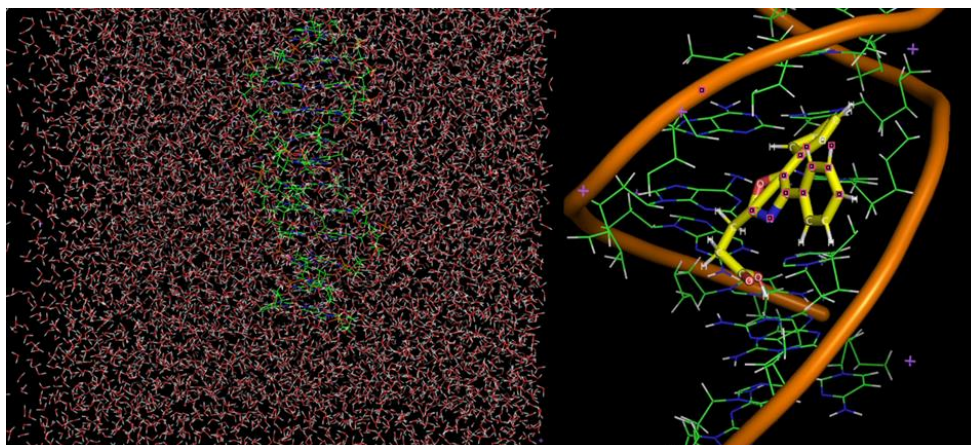


Figure 3. 73 DNA -oxaprozin drug simulation with PM3/MM methods (QM/MM)

Figure 3.73 shows docking result. The whole system of OXA is shown on the left side of Figure 3.73. There are almost 7000 water molecules in the system. 24 residues of B-DNA are used in this simulation. On the right side of Figure 3.73 is shown only DNA and OXA structures. 60 ps is applied in this simulation. In the OXA-DNA interaction, the small method does not cause a change in the geometry of the DNA, however, the interaction sites are in the expected direction. Therefore DNA-OXA QM/MM MD simulation is performed with the DFT method in section 3.11. In this section CP2K is used for QM/MM MD simulation.

3.11 MD Simulation Of OXA-DNA Interaction with QM/MM Method

In this section, molecular docking simulation was performed by Haddock. As in previous parts, we used the Haddock application's lowest energy structure for MD simulations.

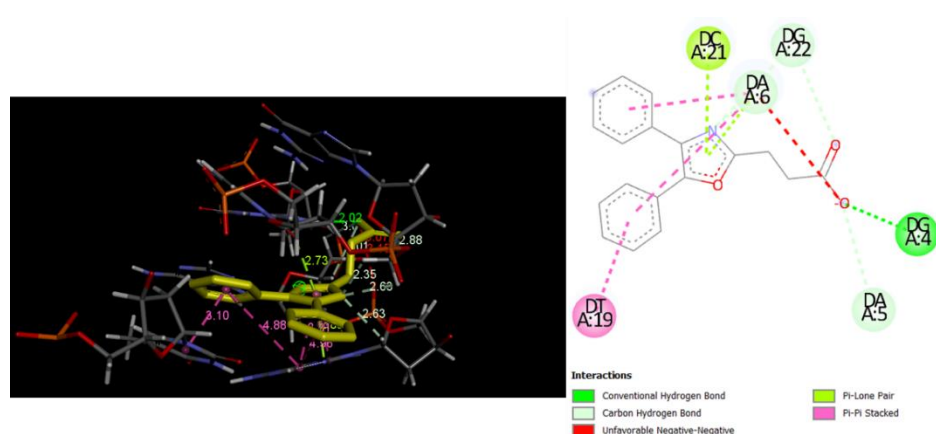


Figure 3. 74 Haddock docking result of OXA interaction with DNA

Figure 3.74 shows the result of docking. This is also our initial structure for the MD simulation. Discovery Studio was used to examine the drug-DNA interaction sides.

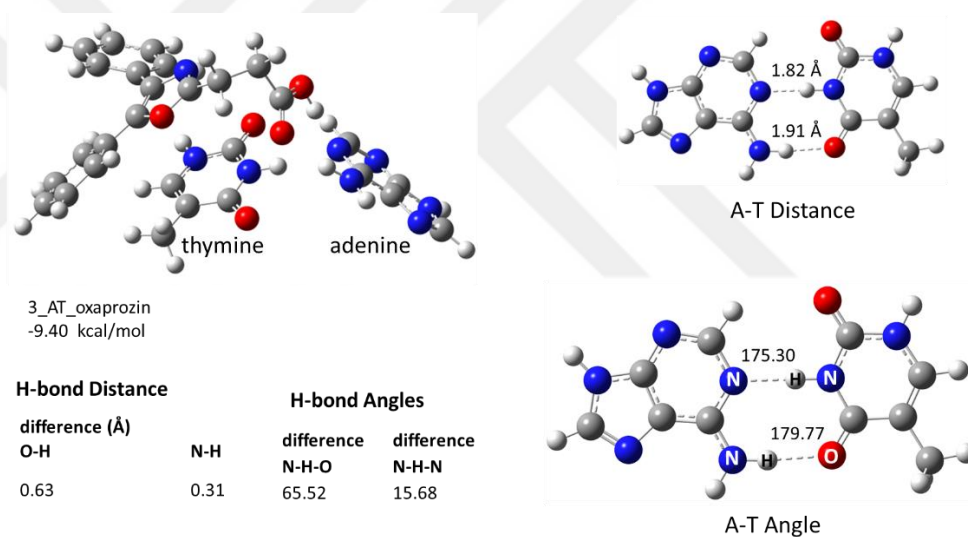


Figure 3. 75 Figure QM model of OXA interaction with A-T nucleobase pair and their H-bonds distance and angles changes

As seen from Figure 3.75 that our expectation from our previous quantum modeling studies is that there is an interaction between the carboxyl and oxazole groups of the drug and the DNA bases.

3.11.1 The Process of Setting Up A Simulation System Of OXA Interaction with DNA

Energy minimization was started after all the preparing steps in the previous section were applied in the same way.

3.11.2 Energy Minimization (EM)

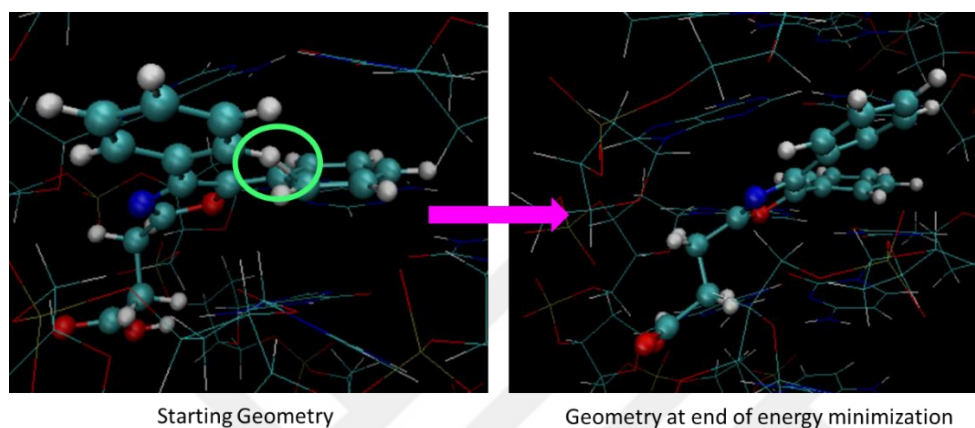


Figure 3. 76 Energy minimization of geometry result

As a result of the docking procedure, the hydrogen atoms in the OXA are deleted. As with any dynamic simulation preparation, the deleted hydrogen atoms are added back in the preparation of the OXA topology. As a result of this re-addition, it can be clearly seen in the Figure 3.76 that the hydrogens overlap. However, the structure was reached to optimized geometry by energy minimizing using the method QM /MM with Gromaccs+CP2K interface. QM was applied only to the oxaprozin of drug molecule. EM was performed with a time step of 1 fs in total 50000 steps.

GROMACS Energies

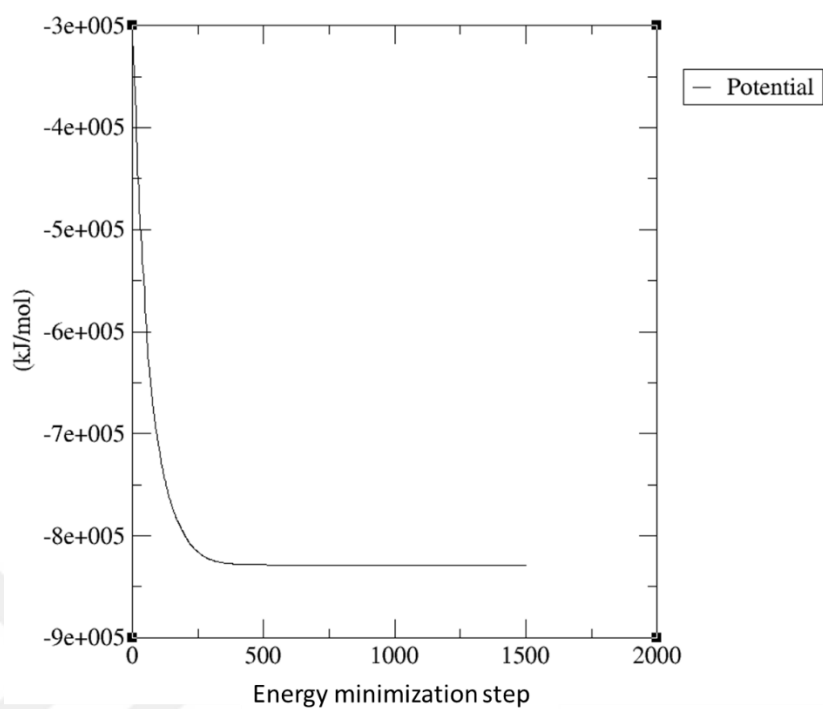


Figure 3.77 Potential energy from EM

It has been converged as it has reached energy minimization in 1506 steps. Potential energy of -8.28 kJ/mol (-1.98 kcal/mol) as shown in the Figure 3.77. The EM process ensured a reasonable geometry and solvent orientation for the starting structure. After that, real dynamics started.

3.11.3 NVT Equilibration

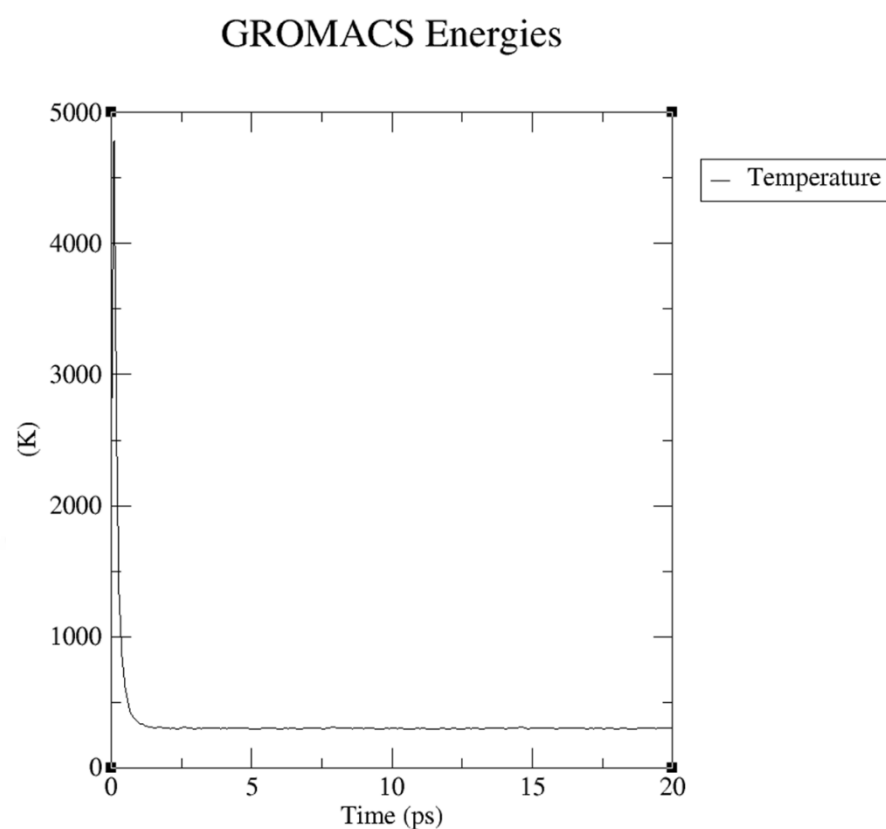


Figure 3. 78 Temperature from NVT equilibrium

A dynamic simulation with a time step of 2 fs in total 10000 steps (20ps) can be completed in 7 days using the QM /MM method with Gromacs+CP2K interface in HPC. Running on 1 node with total 8 cores, 16 processing units. The temperature of the reaches the target value (300 K) and remains constant for the rest of the equilibration as shown in Figure 3.78.

3.11.4 NPT Equilibration

GROMACS Energies

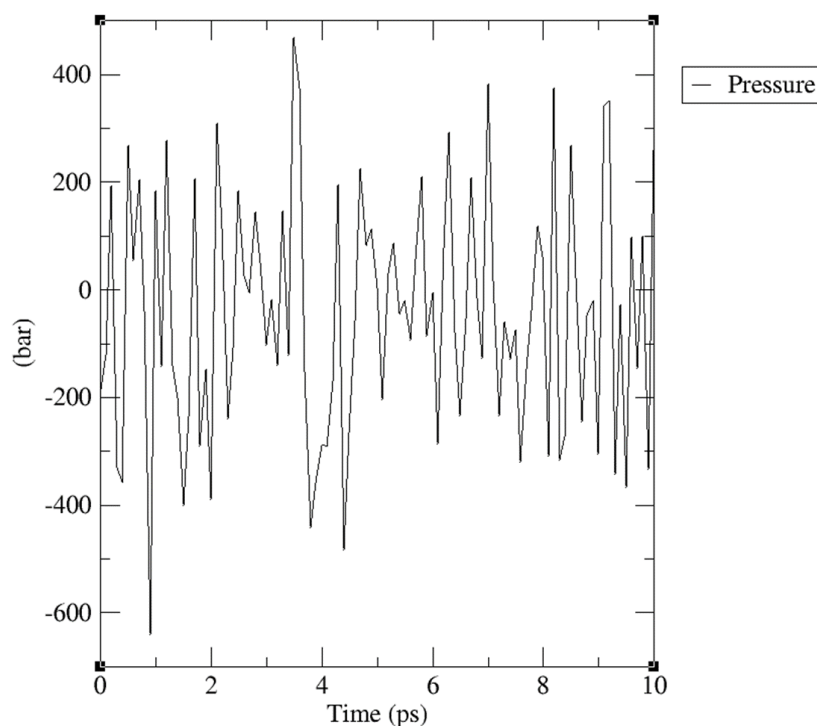


Figure 3. 79 Pressure from NPT equilibrium

A dynamic simulation with a time step of 2 fs in total 5000 steps (10ps) can be completed in 7 days using the QM/MM method with Gromacs+CP2K interface in HPC. Running on 1 node with total 16 cores, 32 processing units for both. During the 10 ps equilibration phase, the pressure value widely fluctuates, but this is not surprising as can be seen from Figure 3.79. Pressure fluctuates widely during an MD simulation and the target/reference value (1 bar). The Berendsen barostat was used for simulation.

3.11.5 Production

In this part, OXA-DNA examined in 50 ps with QM/MM MD simulation.

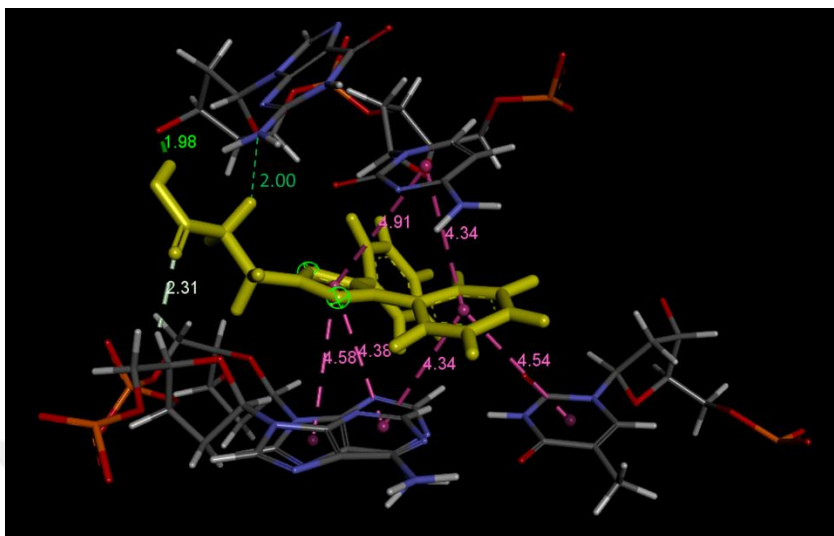


Figure 3. 80 OXA-DNA interactions at the end of production simulation with QM/MM method (Discovery studio was used for drug-DNA interactions) (distance unit is ångstrom)

When these interactions were examined with Discovery studio, it was seen that the distance between the two molecules was 2 angstroms. Therefore, it is not conclusive to define whether there is a bond between the OXA and DNA in there.

4. CONCLUSION

In this study, we calculated drug-nucleobase and drug-nucleobase pair interactions with various DFT methods, and we quantitatively determined the strengths of some selected drugs and DNA-base interactions that may improve our understandings about the side effects of the drugs. The all results mentioned above are summarized in detail below.

➤ To begin with, the dispersion correction models are expected to significantly improve the original B3LYP results. Therefore, we expected that B3LYP-D2 would give a relatively better result than B3LYP. However, the B3LYP-D2 method did not improve the results. The ω B97X-D hybrid functional is dispersion and long-range corrected, so it is considerably better for non-covalent interactions compared to the B3LYP and B3LYP-D2 methods. We can conclude that the ω B97X-D functional performs significantly better than B3LYP and B3LYP-D2. Our study constitutes another indicative of the fact that ω B97X-D works well in dispersion bond systems. The ω B97M-V and ω B97X-V methods are other two DFT methods that describes weakly interacting systems best. However, the gradients of these two methods have not been coded in ORCA, yet. Hence, the ω B97X-D method, which produces energies closer to the CCSD(T) method, is the best choice that can perform geometry optimization.

➤ The most negative binding energy among the drug-single base complexes is between celiprolol and guanine. However, evaluation of interaction energy alone is insufficient to understand drug-nucleobase interaction. Because drug nucleobase interactions were also affected by the distance between the charge donor and the charge acceptor and the electron transitions between purine and pyrimidine bases. They have low symmetry and many non-bonded electrons organized into π -systems((Starikov, 2002),(Charanya et al., 2019)). They have low symmetry and many unbound electrons organized in π -systems. These interactions are much weaker but can have a significant impact on the properties of a molecule. When the effects of drug molecules on hydrogen bond were examined in terms of angle and distance, independent of energy, the DFT calculations point out that mefenamic acid and oxaprozin are the two drugs interacting effectively with both A-T and G-

C base pairs. Consequently, although non-covalent interactions are weak, it can be said that they can significant effect on the properties of a molecule.

➤ We also analyzed their molecular electrostatic potential (MEP) maps to understand the exact interaction sites of the drugs with the nucleobase pairs. We found that the interactions were in the expected direction. Based on our data and the literature, MEP analysis provides insight into explaining the donor-acceptor relationship and sheds light on potential interactions. Among the base pair drug complexes, mefenamic acid has the most prominent attractive interaction between the A-T and G-C base pairs. So, one can conclude that this drug may have possibility to damage to DNA.

➤ Mefenamic acid showed remarkable affinity with cytosine of G-C base pair at the interaction sites (carboxyl group) in DNA with QM/MM method. The average distance of the -O...H- bond between mefenamic acid and cytosine is 1.70 Å. The geometry of the G-C nucleobase pair that the drug affects is quite considerable changed and the distances between the guanine and cytosine nucleobases become larger than 3 Å. This shows that the hydrogen bonds between the G-C base pair is completely broken due to the drug. As can be understood from these interactions, mefenamic acid most probably has the highest possible side effect among all drugs.

➤ To better understand these interaction sites, the interaction energy between the cytosine-guanine nucleobase pair and the drug was calculated with DFT method of ω B97X-D /6-31+G(d,p). When we examine the magnitude of the interaction energy, it is seen that it is almost the same as our previous model quantum calculations. This is an important indication of the reliability of quantum models for non-bonded interactions.

➤ Our previous model quantum calculations (DFT) show that mefenamic acid gives interaction with nucleobase pair through the carboxylic group. Our last QM/MM calculations produce a very similar result. Hence, the DFT methods employed here for non-bonded interactions can be said to be rather reliable, also.

➤ There are experimental evidence supporting that mefenamic acid interacts with DNA nucleobases. This increases the reliability of DFT and QM/MM studies. However, none of these findings are the certain evidence that this drug will damage DNA.



ACKNOWLEDGMENT

There are many cheerful moments in a person's life. In my life, one of these moments is meeting my supervisor Prof. Dr. Armağan KINAL, and Prof. Toomas TAMM. I would like to thank Prof. KINAL and Prof. TAMM because of their sensibility, criticisms, and friendship. I am very glad to meet and work with them. Also, I would like to thank Prof. TAMM and Hellis TAMM because this couple always supported me while I had been in Tallinn, and I always felt at home.

I owe thanks to Dr. Frank Beierlein from the past because he introduced me to MD simulations and put in my mind to QM/MM approach.

I also thank to Prof. Dr. Nursel ACAR SELÇUKİ and Prof. Dr. Ali ÇAĞIR. They made me think differently and freely during my thesis period.

Also, I would like to special thanks to Dr. Dmitry MOROZOV who installing and compiled the Gromacs+ CP2K interface program for me. Therefore, I completed the final stage of this thesis. I wouldn't have been able to complete this without his help.

I would like to special thanks to Ezgi KARACA's group member of Büşra SAVAŞ for her help about to Gromacs.

I'm very grateful to my family for showing patience and tolerance during my working times. Especially, my mum of Nuran Nişli ÖZTÜRK, my father of Ahmet Zeki ÖZTÜRK, my aunt of Nurhan Nişli ÇAĞINER, my uncle of Suat ÇAĞINER and my all family without their love and support, none of this would be possible.

I give thanks to my very dear friend, İmren BAYIL who supports me in every sense and I have become unable to live my happiest and most unhappy moments without her during our friendship lasting for years. I also thank to my special friends Sibel YOYLER and Şevket ERSOY because they never give up supporting me.

ACKNOWLEDGMENT (Continued)

Finally, I thank to TÜBİTAK because they provided me to 2211-C National Ph.D. Scholarship Program in the Priority Fields in Science and Technology (Application number:1649B032001953), and also TÜBİTAK 2214-A - International Research Fellowship Program for Ph.D. Students (Application number: 1059B142000402). Also, this work was supported by 100/2000 YÖK (Council of Higher Education in Turkey) Ph.D. Scholarship- Priority Field Bioinformatics (13.02.2018-13.02.2022). Besides, this thesis was supported by European Regional Development Fund's Doctoral Studies and Internationalization Program DoRa Plus ((fund is provided by the European Regional Development Fund and Estonian government) by Tallin Technology of University, Tallinn/Estonia.

Eventually, most of the calculations reported in this thesis were performed at High Performance and Grid Computing Center (TRUBA resources), ULAKBİM in TÜBİTAK, and High-Performance Computing Center, Tallinn University of Technology from Estonian Scientific Computing Infrastructure.

18 /08/2022

Işıl Öztürk

REFERENCES

- Abbas, H.**, 2017, First-principles study of interaction of serine with nucleobases of DNA and RNA, *Journal of Biological Physics*, 43(1):105–111p.
- Accessdata**, https://www.accessdata.fda.gov/drugsatfda_docs/label/2008/015034s040lbl.pdf (Date of access: 18.03.2022)
- Allinger, N.**, 1977, Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms, *Journal of the American Chemical Society*, 99, pp. 8127–8134p.
- Armakovic, S., Armakovic, S. J., Šetrajcic, J. P., and Šetrajcic, I. J.**, 2012, Active components of frequently used β -blockers from the aspect of computational study, *Journal of Molecular Modeling*, 18(9):4491–4501p.
- Azam, F., Alabdullah, N. H., Ehmedat, H. M., Abulifa, A. R., Taban, I., and Upadhyayula, S.**, 2018, NSAIDs as potential treatment option for preventing amyloid β toxicity in Alzheimer's disease: an investigation by docking, molecular dynamics, and DFT studies. *Journal of Biomolecular Structure and Dynamics*, 36(8):2099–2117p.
- Azizi, H., and Zoha, T.**, 2018, An in-depth study on noncovalent stacking interactions between DNA bases and aromatic drug fragments using DFT method and AIM analysis: conformers, binding energies, and charge transfer, *Applied Biological Chemistry*, 61(2):209–226p.
- Baik, M. H., Friesner, R. A., and Lippard, S. J.**, 2003, Theoretical Study of Cisplatin Binding to Purine Bases: Why Does Cisplatin Prefer Guanine over Adenine, *Journal of the American Chemical Society*, 125(46):14082–14092p.
- Bioexcel**, https://docs.bioexcel.eu/2020_06_09_online_ambertools4cp2k/09qmm/index.html (Date of access: 10.06.2021)
- Bio2byte**, <https://www.bio2byte.be/acpype/> (Date of access: 15.02.2022)
- Boys, S. F. and Bernardi, F.**, 1970, *Mol. Phys.*, 19, 553.
- Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M.**, 1983, CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem.*, 4, 187p.
- Buenker, R. J., Peyerimhoff, S. D.**, 1969, *Chemical Physics Letters*, 3, 37p.

REFERENCES (Continued)

- Burns, L. A., Vázquez-Mayagoitia, Á., Sumpter, B. G., and Sherrill, C. D.,** 2011, Density-functional approaches to noncovalent interactions: A comparison of dispersion corrections (DFT-D), exchange-hole dipole moment (XDM) theory, and specialized functionals, *Journal of Chemical Physics*, 134(8).
- Chai, J. Da, and Head-Gordon, M.** 2008, Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Physical Chemistry Chemical Physics*, 10(44):6615–6620p
- Charanya, C., Sampathkrishnan, S., and Balamurugan, N.,** 2019, Quantum Chemical Computations, Molecular Docking, Vibrational Spectroscopic Analysis, Non-Linear Optical Properties and DFT Calculation of 2-[(2,3-Dimethylphenyl)Amino]Benzoic Acid. *Polycyclic Aromatic Compounds*, 0(0):1–22p.
- Chen, H. Y., Young, P. Y., and Hsu, S. C. N.,** 2009, Theoretical evidence of barrier-free proton transfer in 7-azaindole-water cluster anions, *Journal of Chemical Physics*, 130(16)p.
- Chemistry Libretexts,** <https://chem.libretexts.org/@go/page/13478> (Date of access: 13.12.2020)
- Cmt,** http://cmt.dur.ac.uk/sjc/thesis_dlc/node75.html (Date of access: 02.04.2022)
- Copeland, K. L., Pellock, S. J., Cox, J. R., Cafiero, M. L., and Tschumper, G. S.,** 2013, Examination of tyrosine/adenine stacking interactions in protein complexes. *Journal of Physical Chemistry B*, 117(45):14001–14008p.
- Cossi, M., Rega, N., Scalmani, G. and Barone, V.,** 2003, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, *J. Comput Chem*, 24(6), 669p.
- CP2K,** <https://www.cp2k.org/download> (Date of access: 01.06.2021)
- Drugbank,** <https://www.drugbank.ca/categories/DBCAT002681> (Date of access: 30.11.2019)
- Fliegl, H., Lehtonen, O., Sundholm, D., and Kaila, V. R. I.,** 2011, Hydrogen-bond strengths by magnetically induced currents. *Physical Chemistry Chemical Physics*, 13(2):434–437p.

REFERENCES (Continued)

- FORESMAN, J. B. and FRISCH, Æ.**, 1996, Exploring Chemistry with Electronic Structure Methods, 2nd Edition, Gaussian, Inc., Pittsburgh, PA.
- Frisch, E., Hratchian, H.P., Dennington II, R. D., Keith, T. A., Millam, J., Nielsen A. B., Holder, A. J. and Hiscocks J.**, 2009, GaussView 5 Reference, Gaussian, Inc., Wallingford CT.
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, O., Foresman, J. B., Ortiz, J. V., Cioslowski, J. and Fox, D. J.**, 2010, Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford CT.
- Gangadharana, R. P., and Krishnanb, S. S.**, 2014, Natural Bond Orbital (NBO) population analysis of 1-azanaphthalene-8-ol, *Acta Physica Polonica A*, 125(1):18–22p.
- Gasteiger, J., Li, X., Rudolph, C., Sadowski, J., and Zupan, J.**, 1994, Representation of Molecular Electrostatic Potentials by Topological Feature Maps, *Journal of the American Chemical Society*, 116(11):4608–4620p.
- Gromacs**, <https://manual.gromacs.org/documentation/2022/download.html> (Date of access: 01.02.2022)
- Healthline**, <https://www.healthline.com/health/mefenamic-acid/oral-capsule#:~:text=This%20is%20the%20most%20serious,heart%20failure%2C%20or%20blood%20clot> (Date of access: 24.04.2020)
- Hehre, W. J., Radom, L. P., Schleyer, R., Pople, J. A.**, 1986, Ab Initio Molecular Orbital Theory, Wiley Interscience, New York.

REFERENCES (Continued)

- Hiratani, H., Mizutani, Y. and Alvarez-Lorenzo, C.,** 2005, Controlling drug release from imprinted hydrogels by modifying the characteristics of the imprinted cavities, *Macromol. Biosci.*, 5, 728p.
- Ho Woo, D., Han, I. S., and Jung, G.,** 2004, Mefenamic acid-induced apoptosis in human liver cancer cell-lines through caspase-3 pathway, *Life Sciences*, 75(20):2439–2449p.
- Honorato, R. V., Koukos, P. I., Jiménez-García, B., Tsaregorodtsev, A., Verlato, M., Giachetti, A., ... Bonvin, A. M. J. J.,** 2021, Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem, *Frontiers in Molecular Biosciences*, 8(July):1–7p.
- Intechopen,** <https://www.intechopen.com/online-first/anticancer-drugs-deoxyribonucleic-acid-DNA-interactions> (Date of access: 30.11.2019)
- Jensen, F.,** 2007, *Introduction to Computational chemistry*, John Wiley and Sons Ltd, Second Edition, pp.123.
- Kennard, O.,** 1993, DNA-drug interactions. *Pure and Applied Chemistry*, 65(6):1213–1222p.
- Klamt, A. and Schüümann, G.,** 1993, COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and its Gradient, *J. Chem. Soc. Perkin Transactions*, 2, 799p.
- Levine, I. N.,** 2014, *Quantum Chemistry* Pearson advanced chemistry series.
- Longdom,** <https://www.longdom.org/pharmacovigilance.html> (Date of access: 30.11.2019)
- Mardirossian, N., and Head-Gordon, M.,** 2014, ωb97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Physical Chemistry Chemical Physics*, 16(21):9904–9924p.
- Mardirossian, N., and Head-Gordon, M.,** 2017, Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals, *Molecular Physics*, 115(19):2315–2372p

REFERENCES (Continued)

- Mennucci, B., Tomasi, J., Cammi, R., Cheeseman, J.R., Frisch, M.J., Devlin, F.J., Gabriel, S. and Stephens, P.J.,** 2002, Polarizable continuum model (PCM) calculations of solvent effects on optical rotations of chiral molecules, *J. Phys. Chem. A.*, 106, 6102p.
- Miertuš, S., Scrocco, E. and Tomasi, J.,** 1981, Electrostatic Interaction of a Solute with a Continuum. A Direct Utilizaion of AB Initio Molecular Potentials for the Prevision of Solvent Effects, *Chem. Phys.*, 55, 117p.
- Meyer, E. A., Castellano, R. K., and Diederich, F.,** 2003, Interactions with aromatic rings in chemical and biological recognition, *Angewandte Chemie, International Edition* (Vol. 42).
- Misra, M., and Yadav, A. K.,** 2022, Assessment of Available AMBER Force Fields to Model DNA-Ligand Interactions. *Biointerface Research in Applied Chemistry*, 13(2):1–14p.
- Morgan, G.,** 1996, Non-steroidal anti-inflammatory drugs and the chemoprevention of colorectal and oesophageal cancers, *Gut*, 38(5):646–648p.
- Oliveira, F.M., Barbosa, L.C.A., Fernandes, S.A., Lage, M.R., Carneiro, J.W. D.M. and Kabeshov, M.A.,** 2016, Evaluation of some density functional methods for the estimation of hydrogen and carbon chemical shifts of phosphoramidates, *Computational and Theoretical Chemistry.*, Elsevier B.V., 1090, 218–224pp.
- Öztürk, I., KINAL, A., TOOMAS, T.,** 2020, Computational study on anti-inflammatory and anti-hypertensive drug molecules interaction with base pairs, *Bulletin of Biotechnology* 1(2);:34–38p.
- Peleg, I. I., Lubin, M. F., Cotsonis, G. A., Clark, W. S., and Wilcox, C. M.,** 1996, Long-term use of nonsteroidal antiinflammatory drugs and other chemopreventors and risk of subsequent colorectal neoplasia. *Digestive Diseases and Sciences*, 41(7), 1319–1326.
- Pediaa,** <https://pediaa.com/difference-between-adenine-and-guanine/> (Date of access: 15.05.2021)
- Pediaa,** <https://pediaa.com/difference-between-cytosine-and-thymine/> (Date of access: 15.05.2021)

REFERENCES (Continued)

- Pople, J.A., Krishnan, R., Schlegel, H.B., and Binkley, J.S.,** 1978, Electron correlation theories and their application to the study of simple reaction potential surfaces, *International Journal of Quantum Chemistry.*, John Wiley and Sons, Inc., 14(5), 545–560pp.
- Pople, J., Head-Gordon, M., and Raghavachari, K.,** 1987, Quadratic configuration interaction. A general technique for determining electron correlation energies, *The Journal of Chemical Physics.*, 87(10), 5968pp.
- PubChem,** <https://pubchem.ncbi.nlm.nih.gov/compound/Mefenamic-acid#section=Chemical-and-Physical-Properties> (Date of access: 24.04.2022).
- Purvis, G.D. and Bartlett, R.J.,** 1982, A full coupled-cluster singles and doubles model: The inclusion of disconnected triples, *The Journal of Chemical Physics.*, American Institute of Physics, 76(4):1910–1918pp.
- Richard M.,** 2004, *Electronic Structure, Basic Theory and Practical Methods*, Cambridge.
- Rutledge, L. R., Durst, H. F., and Wetmore, S. D.,** 2009, Evidence for stabilization of DNA/RNA protein complexes arising from nucleobase amino acid stacking and T-shaped interactions, *Journal of Chemical Theory and Computation*, 5(5), 1400–1410p.
- Sagdinc, S. G., and Esme, A.,** 2010, Theoretical and vibrational studies of 4,5-diphenyl-2-oxazole propionic acid (oxaprozin), *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 75(4):1370–1376p.
- Saji, R. S., Prasana, J. C., Muthu, S., George, J., Kuruvilla, T. K., and Raajaraman, B. R.,** 2020, Spectroscopic and quantum computational study on naproxen sodium, *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 226p.
- Scuseria, G.E., Janssen, C.L. and Schaefer, H.F.,** 1988, An efficient reformulation of the closed-shell coupled cluster single and double excitation (CCSD) equations, *The Journal of Chemical Physics.*, American Institute of Physics, 89(12):7382–7387pp.

REFERENCES (Continued)

- Scuseria, G.E., Schaefer, H.F. and Schaefer Iii, H.F.,** 1989, Is coupled cluster singles and doubles (CCSD) more computationally intensive than quadratic configuration interaction (QCISD), J. Chem. Phys. Additional information on J. Chem. Phys., Journal Homepage., 90 (3700).
- Shankar Rao, Y. B., Prasad, M. V. S., Udaya Sri, N., and Veeraiah, V.,** 2016, Vibrational (FT-IR, FT-Raman) and UV-Visible spectroscopic studies, HOMO-LUMO, NBO, NLO and MEP analysis of Benzyl (imino (1H-pyrazol-1-yl) methyl) carbamate using DFT calculations, Journal of Molecular Structure, 1108, 567–582p.
- Shodor.org,** Overview of Computational Chemistry
<http://www.shodor.org/chemviz/overview/ccbasics.html> (Date of access:10.02.2022).
- Simmons, J.,** 2003, An Introduction to Theoretical Chemistry, Cambridge Press, Cambridge.
- Stankiewicz, A., Giebułtowicz, J., Stefański, M., Sikorska, K., Wroczyński, P., and Nalecz-Jawecki, G.,** 2015, The development of the LC-MS/MS method based on S-9 biotransformation for detection of metabolites of selected β -adrenolytics in surface water, Environmental Toxicology and Pharmacology, 39(2):906–916p.
- Starikov, E. B.,** 2002, Quantum chemistry of nucleic acids: How it could help and when it is necessary, Journal of Photochemistry and Photobiology C: Photochemistry Reviews, 3(2):147–164p.
- Tomasi, J., Mennucci, B. and Cammi, R.,** 2005, Quantum Mechanical Continuum Solvation Models, Chem. Rev., 105, 2999p.
- Toupkanloo, H. A., and Rahmani, Z.,** 2018, An in-depth study on noncovalent stacking interactions between DNA bases and aromatic drug fragments using DFT method and AIM analysis: conformers, binding energies, and charge transfer, Applied Biological Chemistry, 61(2):209–226p.
- Turner, P.J. XMGRACE,** 2005, Version 5.1.19. Center For Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, Ore, USA.

REFERENCES (Continued)

- Vainio, H., Miller, A. B., and Bianchini, F.,** 2000, An international evaluation of the cancer-preventive potential of sunscreens. *International Journal of Cancer*, 88(5):838–842p.
- Van Zundert, G. C. P., Rodrigues, J. P. G. L. M., Trellet, M., Schmitz, C., Kastiris, P. L., Karaca, E., ... Bonvin, A. M. J. J.,** 2016, The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes, *Journal of Molecular Biology*, 428(4):720–725p.
- Wang, J., Cieplak, P. and Kollman, P.A.,** 2000, How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules, *J. Comput. Chem.*, 21, 1049p.
- Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. and Case, D.A.,** 2004, Development and testing of a general amber force field, *J. Comput. Chem.*, 25, 1157p.
- Wang, P., Hu, Y., Zhan, H., and Chen, J.,** 2017, Gas-phase conformational preference of the smallest saccharide (glycolaldehyde) and its hydrated complexes with bridged hydrogen bonding, *RSC Advances*, 7(11): 6242–6250p.
- Wasko, M. J., Pellegrine, K. A., Madura, J. D., and Surratt, C. K.,** 2015, A role for fragment-based drug design in developing novel lead compounds for central nervous system targets, *Frontiers in Neurology*, 6(SEP):1–11p.
- Weinhold, F., Landis, C. R., and Glendening, E. D.,** 2016, What is NBO analysis and how is it useful? *International Reviews in Physical Chemistry*, 35(3):399–440p.
- Weiss, H., Amberger, A., Widschwendter, M., Margreiter, R., Öfner, D., and Dietl, P.,** 2001, Inhibition of store-operated calcium entry contributes to the anti-proliferative effect of non-steroidal anti-inflammatory drugs in human colon cancer cells. *International Journal of Cancer*, 92(6):877–882p.
- Wenthold, P. G., and Squires, R. R.,** 2005, Bond Dissociation Energies of F2 and HF2 A Gas-Phase Experimental and G2, 2002–2005.
- YOUNG, D.,** 2001, A Practical Guide for Applying Techniques to Real World Problems, *Computational Chemistry*, Wiley, Interscience, New York.

CURRICULUM VITAE

Name : Işıl Öztürk

Occupation : Graduated from Ege University, Science Faculty,

Department of Chemistry, Bornova-izmir

EDUCATIONAL BACKGROUND

Bachelor Study, 2009-2013 Chemistry Department, Faculty of Science, Ege University, Izmir-Turkey

MS Study, 01.03.2015-28.08.2015 Computer Chemistry Center of Chemistry and Pharmacy, Friedrich-Alexander- University Erlangen-Nürnberg-Germany (Erasmus +)

2014-2016 Physical Chemistry (English), Graduate School of Natural and Applied Sciences, Ege University, İzmir-Turkey

PhD study, 07.01.2021 – 08.07.2022 Tallinn University of Technology (TalTech) Tallinn/ ESTONIA (Visitor student)

PhD Study, 06.09.2017-18.08.2022 Ege University Izmir/Turkey

POSTER PRESENTATIONS

12-13 October 2018-CPC–XII (12th Chemical Physics Congress),
Computational Study on Carboxymethylation and Hydrazination
Mechanisms of Gluconate, Safranbolu/TURKEY

10-14 September 2017- METU (Middle East Technical University) 29th
National Chemistry Congress Ankara/Turkey A Computational Study on
Carboxymethylation Mechanism of Gluconate in gas and water phase: A
DFT study

19.12.2014 MolChem2014, Interaction of Gluconate Derivatives with
Magnetite (Fe₃O₄) Surface in α -D glucose coated iron oxide nanoparticles:
A DFT study, Istanbul, Turkey

9.03.2015- 11.03.2015 Molecular Modelling Workshop 2015, A
Computational Study on Carboxymethylation Mechanism of Gluconate,
Erlangen-Germany

19-22 October 2014- AIHLS -German-Turkish Year of Research education
and Innovation, Computational Investigation of Interaction of Gluconate
Derivatives with Magnetite (Fe₃O₄) Surface in α -D-glucose Coated Iron
Oxide Nanoparticles Kuşadası / Turkey

ORAL PRESENTATIONS

04-11 June 2021, BioExcel Summer School on Biomolecular Simulations
Computational Study on Anti-Inflammatory and Anti-Hypertensive Drug
Molecules Interaction with Base Pairs

19-20 March 2020, 3rd International Eurasian Conference on Biological and
Chemical Sciences (EurasianBioChem 2020), Computational study on anti-
inflammatory and anti-hypertensive drug molecules interaction with base
pairs, Ankara, Turkey

SUMMER SCHOOL

04-11 June 2021, BioExcel Summer School on Biomolecular Simulations
(Venue: Remote)

25-29 June 2018 Summer School of Feza Gürsey Center - Subject of
Protein structure and interaction in Biophysics Summer School – Boğaziçi
University/ İstanbul TURKEY

19 July- 3 Aug 2019 Summer school of Mustafa Akgül - Free Software
Summer Camp – Abant İzzet Baysal University/Bolu TURKEY

SCHOOLORSHIP

03.02.2022 – 08.07.2022 TÜBİTAK Vice President Science
Fellowships and Grant Programmes Department
REF: 53325897-115.02-121630 (Subject: 2214/A - International
Doctoral Research Fellowship Program) Research Scholarship Program
2020 – 1st term) Application number: 1059B142000402 ((from
TURKEY) (for TALTECH))

15.08.2021- 31.03.2022 The Dora Plus scholarship of the European
Regional Development Fund and Estonian government from Tallin
University of Technology

01.06.2021-07.07.2021 ERASMUS+ Tallinn University of Technology
(TALTECH) Tallinn/ ESTONIA

2020 -2022 2211/C National PhD Scholarship Program in the Priority
Fields in Science and Technology (Application term:2020/2) Application
number:1649B032001953 (TÜBİTAK – TURKEY)

05.02.2018 – 13.02.2022 100/2000 YÖK (Council of Higher Education) PhD
Scholarship: in the field of Bioinformatics (from TURKEY)

01.03.2015-28.08.2015 ERASMUS+ Friedrich-Alexander University (FAU)
Erlangen-Nürnberg / GERMANY

LIST OF PUBLICATIONS

1. **Öztürk, I.**, Kınal, A, Tamm, T. (2020). Computational study on anti-inflammatory and anti-hypertensive drug molecules interaction with base pairs, **Bulletin of Biotechnology**, 1(2), 34–3.
2. **Öztürk, I.**, Şanlıer, Ş., and Kınal, A. J. (2020). Determination of Gluconate Binding Properties on Magnetite Surface and Investigation of Carboxymethylation and Hydrazination Mechanisms of the Gluconated Magnetite Surface: A Computational Study, **Jotcsa**, 7(1), 169–178.