



REPUBLIC OF TURKEY
YEDİTEPE UNIVERSITY
GRADUATE SCHOOL OF HEALTH SCIENCES

MOLECULAR MODELING AND ACTIVITY STUDIES OF
THE COMPOUNDS ACTIVE ON THE *ESCHERICHIA COLI*
AND *STAPHYLOCOCCUS AUREUS* DNA GYRASE B
ATPASE ACTIVE SITE

MASTER OF PHARMACEUTICAL CHEMISTRY

GÜRMEK KAYNAR

ISTANBUL 2012



A THESIS SUBMITTED TO YEDİTEPE UNIVERSITY
THE GRADUATE SCHOOL OF HEALTH SCIENCES
PHARMACEUTICAL CHEMISTRY

MOLECULAR MODELING AND ACTIVITY STUDIES OF
THE COMPOUNDS ACTIVE ON THE *ESCHERICHIA COLI*
AND *STAPHYLOCOCCUS AUREUS* DNA GYRASE B
ATPASE ACTIVE SITE

INSTRUCTORS

[t f 0F q± Dr. Bctmp'DGTM





Prof. Dr. Dilek DEMİR EROL

ISTANBUL 2012

TEZ ONAYI FORMU

Kurum : Yeditepe Üniversitesi Sağlık Bilimleri Enstitüsü
Program : Farmasötik Kimya Yüksek Lisans Programı
Tez Başlığı : Molecular Modeling and Activity Studies of the Compounds Active on the Escherchia coli and Staphylococcus aureus DNA Gyrase B ATPase Active Site
Tez Sahibi : Gürmen KAYNAR
Sınav Tarihi : 24.02.2012

Bu çalışma jürimiz tarafından kapsam ve kalite yönünden Yüksek Lisans Tezi olarak kabul edilmiştir.

	Unvanı, Adı-Soyadı (Kurumu)	İmza
Jüri Başkanı:	Prof. Dr. Dilek DEMİR EROL Yeditepe Üniversitesi Farmasötik Kimya Anabilim Dalı	
Tez danışmanı:	Yrd. Doç. Dr. Barkın BERK Yeditepe Üniversitesi Farmasötik Kimya Anabilim Dalı	
Üye:	Prof. Dr. Hülya AKGÜN Yeditepe Üniversitesi Farmasötik Kimya Anabilim Dalı	
Üye:	Yrd. Doç. Dr. Esra Önen BAYRAM Yeditepe Üniversitesi Farmasötik Kimya Anabilim Dalı	

ONAY

Bu tez Yeditepe Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca yukarıdaki jüri tarafından uygun görülmüş ve Enstitü Yönetim Kurulu'nun 23/03/2012 tarih ve 2012/03-02 sayılı kararı ile onaylanmıştır.

Prof. Dr. Selçuk YILMAZ
Sağlık Bilimleri Enstitüsü Müdürü

ACKNOWLEDGEMENTS

There are a number of people I am deeply grateful without whom this thesis would be completed.

I thank Prof. Dr. Dilek Demir Erol and Assis. Prof. Dr. Barkın Berk, for believing in me and for their patience, guidance and support during my graduate study and the thesis process. They created a candidate scientist from a curious college graduate. I have learned so much from them.

I thank Mahmut Özbek and Erkan Er for participating in my thesis, for their insight, and guidance.

I thank Can Dokak for being a great Sensei and a friend.

I thank Özgür Karban for making our company building a nice place to live.

I thank my dear friend Aygöl balaban for being my best friend, with her these years were the best years of my life.

Finally, I thank my parents for their love, patience, and never ending support and Begüm Pilatin for being the best friend, the best cook, the best therapist, best stylist and being the better half of me.

ABSTRACT

For the design of new ligands, the use of computational methods and algorithms, such as HTVS (High Throughput Virtual Screening), docking-scoring and statistical methods such as ROC curves (Receiver Operating Characteristic curves) are among the popular methods used today. Generally, to use these methods, there must be X-Ray crystallography data or a homology model presenting macromolecule and ligand structures to study interactions.

In alternative antimicrobial clinical treatment, the development of molecules based on DNA gyrase enzyme inhibition is important in cases such as *Escherichia coli* (*E. coli*) and *Staphylococcus aureus* (*S. aureus*) which not only have Gram - / Gram + distinction but also multiple drug resistance.

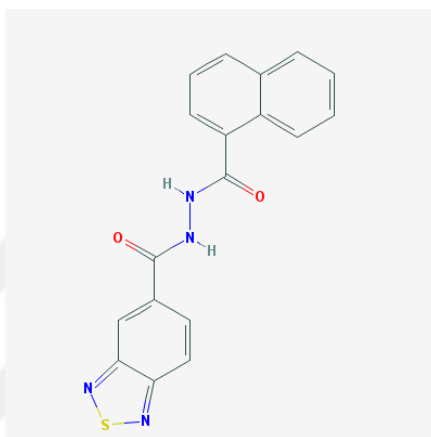
Recently published information regarding *E. coli* and *S. aureus*, crystallographic data of DNA gyrase B (PDB id: 3G7E and 3G7B) give us an opportunity to use the methods of computer usage-based drug research in determining gram +/- selective inhibitor compounds.

When we investigated the amino acid-ligand interactions of both *E. coli* and *S. aureus*, DNA gyrase B active sites with the help of crystallographic data and compared these interactions with the previous literature belong to DNA gyrase B ATPase inhibitors, we determined that some of the water molecules have major impact during these interactions in terms of selectivity.

In this study, first, trial and test sets were prepared by consequent enrichment of the 5000 and 50000 ZINC databases with known *E. coli* and *S. aureus* DNA GyrB ATPase inhibitor molecules. Then, the trial set was evaluated, considering the contribution of water molecules on interactions, the trial set was screened on the active site of the ATPase *E. coli* and *S. aureus* DNA gyrases using crystallographic data and a HTVS method. The analysis of subset docking score led to the identification of novel interaction patterns.

When these interaction patterns were screened over the test set similarly, 20 maximum scored compounds were determined and further tested against novobiocin standard with gel based *E. coli* and *S. aureus*. supercoiling assays, their activity and selectivity.

The highest scoring N'-(1-naphthylcarbonyl)-2,1,3-benzothiadiazole-5-carbohydrazide structure showed a selective inhibition toward *E. coli* and *S. aureus* DNA gyrase B ATPases.



Keywords: HTVS, ROC curves, docking, *Escherichia coli*, *S. aureus*, DNA gyrase B ATPase

ÖZET

Yeni ligandların tasarlanmasında, HTVS (Yüksek çıktılı sanal tarama teknikleri, High Throughput Virtual Screening), docking gibi bilgisayar gerektiren metod ve algoritmalar yanısıra, ROC eğrileri (alıcı işletim karakteristik eğrileri, Receiver Operating Characteristic curves) gibi istatistiksel yöntemlerin kullanılması günümüz popüler yöntemleri arasındadır.

Genelde, bu yöntemlerin kullanılabilmesi için etkileşmeyi incelemeye olanak verecek ligand ile beraber kristallendirilmiş makromolekül yapısının X-Ray kristalografik datası veya homoloji modeli bulunmalıdır.

Antimikrobiale tedavide DNA giraz enziminin inhibe edilmesine dayalı moleküllerin geliştirilmesi özellikle *Escherichia coli* (*E. coli*) ve *S. aureus* (*S. aureus*) gibi bir taraftan Gram-/Gram+ ayrımına sahip, diğer taraftan çoklu ilaç rezistansına sahip mikroorganizmaların alternatif klinik tedavisi açısından önemlidir.

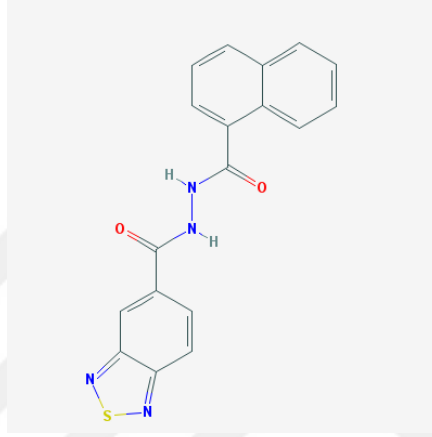
Son yıllarda yayınlanan *E. coli* ve *S. aureus* DNA Giraz B (Pdb id:3G7B ve 3G7E) yöresine ait kristalografik bilgileri, gram+/- selektif inhibitör bileşiklerin belirlenmesinde, bilgisayar kullanıma dayalı ilaç araştırma yöntemlerini kullanma olanağı vermiştir.

E. coli ve *S. aureus* DNA Giraz B kristalografik verileri yardımıyla aktif yöre amino asit-ligand etkileşimlerini inceleyip, bu etkileşimleri DNA Giraz B ATPaz inhibitorlerine ait literatür verileriyle karşılaştırdığımızda, etkileşimde seçicilik açısından bazı su moleküllerinin etkin katkısının olduğu belirledik.

Bu çalışmada, ilk olarak ZINC veri bankasından elde edilen 5000 ve 50000 bileşiğe, bu yöreye etkinlikleri kanıtlanmış bileşiklerin katımıyla, zenginleştirilmiş deneme ve test setleri oluşturulmuştur. Daha sonra su moleküllerinin etkileşmeye olası katkı modelleri gözönüne alınarak *E. coli* ve *S. aureus* DNA giraz B ATPaz kristalografik verileri yardımıyla aktif yöreler HTVS yöntemi kullanılarak deneme

setinde taranmış, elde edilen docking skorları ROC eğrileri yöntemiyle değerlendirilmiş ve etkin etkileşim kalıpları saptanmıştır. Bu etki kalıpları, test seti üzerinde aynı yöntemlerle tarandığında, en yüksek skor alan 20 molekülün mikroorganizmalar arası selektivite ve etkinlik düzeyleri, in-vitro *E. coli* ve *S. aureus* jel tabanlı Supercoiling testleri yardımıyla novobiyosin standardına karşı belirlenmiştir.

Aktivite değerlendirmesinde, en yüksek skoru alan N'-(1-naftilkarbonil)-2,1,3-benzotiyadiazol-5-karbohidrazit yapısı *E. coli* ve *S. aureus* DNA gyrase B ATPaz ları arasında selektif inhibisyonu gerçekleştirmiştir.



Anahtar Kelimeler: HTVS, ROC eğrileri, docking, *Escherichia coli*, *S. aureus*, DNA gyrase B ATPaz

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
ÖZET	vi
LIST OF ABBREVIATIONS AND SYMBOLS	x
LIST OF FIGURES	xi
LIST OF TABLES	1
1. INTRODUCTION	2
2. GENERAL INFORMATION	10
2.1. TOPOISOMERASES	10
2.1.1. TOPOISOMERASES TYPE IA	11
2.1.3. DNA GYRASE	13
2.2. COMPUTATIONAL METHODS	14
2.2.1. STRUCTURE BASED DRUG DESIGN	15
2.2.2. DOCKING AND SCORING	16
2.2.3. HIGH THROUGHPUT VIRTUAL SCREENING-HTVS/VIS	17
2.2.4 RECEIVER OPERATING CHARACTERISTIC CURVES – ROC CURVES	18
2.3. DATABASES	19
2.3.1. RSCB DATABASE	20
2.3.2. ZINC DATABASE	22
2.4. THE METHODS ARE USED FOR DETERMINATION OF ACTIVITY	23
2.4.1. SUPERCOILING ASSAY	23
2.4.2. GEL BASED DRUG INHIBITION ASSAY	23
3. MATERIALS AND METHODS	24
3.1 PREPARATION OF PROTEIN STRUCTURES	25
3.2 PREPARATION OF TRIAL AND TEST SETS	25
3.3 PREPARATION OF GRID FILES	25
3.4 DOCKING AND SCORING	26
3.5 ROC CURVE EVALUATION	27
3.6 BIOLOGICAL EVALUATION	28
4. RESULTS	29
4.1. TRIAL SET DOCKING, SCORING AND ROC CURVES EVALUATION	29
4.2. TEST SET DOCKING, SCORING AND ROC CURVES EVALUATION	38
4.3. BIOLOGICAL ACTIVITY	47

5. DISCUSSION AND CONCLUSION	55
REFERENCES.....	59
VITA	62



LIST OF ABBREVIATIONS AND SYMBOLS

Escherichia coli

E. coli

Staphylococcus aureus

S. aureus

World Health Organization

WHO

Tuberculosis

TB

HTVS

High Throughput Virtual Screening

ROC curves

Receiver Operating curves

DNA

Deoxyribonucleic acid

ATP

Adenosine-5'-triphosphate



LIST OF FIGURES

FIGURE 1: CHEMICAL STRUCTURE OF LIGANDS FROM 3G7E (1A) AND 3G7B (1B).....	3
FIGURE 2: H-BONDS BETWEEN THE <i>E. COLI</i> BINDING SITE AND THE LIGAND.	4
FIGURE 3: THE HYDROPHOBIC INTERACTIONS BETWEEN THE LIGAND AND <i>E. COLI</i> BINDING SITE	5
FIGURE 4: H-BONDS BETWEEN <i>S. AUREUS</i> BINDING SITE AND THE LIGAND	6
FIGURE 5: HYDROPHOBIC INTERACTIONS BETWEEN THE <i>S. AUREUS</i> BINDING SITE AND THE LIGAND	7
FIGURE 6: THE ACTIVE SITE-LIGAND INTERACTIONS IN <i>E. COLI</i> AND <i>S. AUREUS</i>	7
FIGURE 7: <i>E. COLI</i> AND <i>S. AUREUS</i> OVERLAPPED AS PAIRS.....	8
FIGURE 8: DOMAIN STRUCTURE OF TYPE IA TOPOISOMERASES.....	12
FIGURE 9: THE STEPS IN PDB DATA PROCESSING. ELLIPSES REPRESENT ACTIONS AND RECTANGLES DEFINE CONTENT.....	21
FIGURE 10: A. THE ZINC SEARCH TOOL B. THE ZINC DATABASE BROWSER.	23
FIGURE 11: SUPERCOOLING ACTIVITY OF <i>E. COLI</i> AND <i>S. AUREUS</i> HATA! YER İŞARETİ TANIMLANMAMIŞ.	
FIGURE 12: GEL DOCUMENTATION SOFTWARE AND SUBSEQUENT STATISTICAL ANALYSIS.....	24
FIGURE 13: EXPLANATION OF GELS	29
FIGURE 14: 3G7B DOCKED WITHOUT WATERS, SELECTED AND RANKED ACCORDING TO MAXIMUM DOCKING SCORE.....	30
FIGURE 15: 3G7B DOCKED WITHOUT WATERS, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE	30
FIGURE 16: 3G7B DOCKED WITH WATER 235, SELECTED AND RANKED ACCORDING TO MAXIMUM DOCKING SCORE.....	31
FIGURE 17: 3G7B DOCKED WITH WATER 235, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE	32
FIGURE 18: 3G7B DOCKED WITH WATERS 235 AND 263, SELECTED AND RANKED ACCORDING TO MAXIMUM DOCKING SCORE.....	33
FIGURE 19: 3G7B DOCKED WITH WATERS 235 AND 263, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE	33
FIGURE 20: 3G7E DOCKED WITHOUT WATERS, SELECTED AND RANKED ACCORDING TO MAXIMUM DOCKING SCORE.....	34
FIGURE 21: 3G7E DOCKED WITHOUT WATERS, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE	35
FIGURE 22: 3G7E DOCKED WITH WATER 408, SELECTED AND RANKED ACCORDING TO MAXIMUM DOCKING SCORE.....	36
FIGURE 23: 3G7E DOCKED WITH WATER 408, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE	36
FIGURE 24: 3G7E DOCKED WITH WATERS 408 AND 443, SELECTED AND RANKED ACCORDING TO MAXIMUM DOCKING SCORE.....	37
FIGURE 25: 3G7E DOCKED WITH WATERS 408 AND 443, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE	37
FIGURE 26: 3G7B DOCKED WITH WATERS 235 AND 263 IN HTVS ALGORITHM, SELECTED AND RANKED ACCORDING TO MAXIMUM DOCKING SCORE.....	38
FIGURE 27: 3G7B DOCKED WITH WATERS 235 AND 263 IN HTVS ALGORITHM, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE	39
FIGURE 28: 3G7B DOCKED WITH WATERS 235 AND 263 IN SP ALGORITHM AFTER A CUT OFF OF 20000 COMPOUNDS, SELECTED AND RANKED ACCORDING TO MAXIMUM E- MODEL SCORE.....	40
FIGURE 29: 3G7B DOCKED WITH WATERS 235 AND 263 IN XP ALGORITHM AFTER A CUT OFF OF 20000 COMPOUNDS, SELECTED AND RANKED ACCORDING TO MAXIMUM E- MODEL SCORE.....	40
FIGURE 30: 3G7E DOCKED WITHOUT WATERS, SELECTED AND RANKED ACCORDING TO MAXIMUM DOCKING SCORE.....	41

FIGURE 31: 3G7E DOCKED WITHOUT WATERS, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE	42
FIGURE 32: 3G7E DOCKED WITH WATER 408, SELECTED AND RANKED ACCORDING TO MAXIMUM DOCKING SCORE.....	43
FIGURE 33: 3G7E DOCKED WITH WATER 408, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE	43
FIGURE 34: 3G7E DOCKED WITH WATER 408 IN SP ALGORITHM AFTER A CUT OFF OF 20000 COMPOUNDS, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE.....	44
FIGURE 35: 3G7E DOCKED WITH WATER 408 IN XP ALGORITHM AFTER A CUT OFF OF 20000 COMPOUNDS, SELECTED AND RANKED ACCORDING TO MAXIMUM E-MODEL SCORE.....	44
FIGURE 36: 1 MG/20 μ L (W/V) DILUSION <i>E. COLI</i> DNA GYRASE GEL ELECTROPHORESIS RESULTS, SUPER-COILED (R. PLASMID+GYRASE), RELAXED (R. PLASMID), NOVABIOCIN (R. PLASMID+GYRASE+NOVABIOCIN)	48
FIGURE 37: 1 MG/50 μ L (W/V) DILUSION <i>E. COLI</i> DNA GYRASE GEL ELECTROPHORESIS RESULTS, SUPER-COILED (R. PLASMID+GYRASE), RELAXED (R. PLASMID), NOVABIOCIN (R. PLASMID+GYRASE+NOVABIOCIN)	48
FIGURE 38: 1 MG/100 μ L (W/V) DILUSION <i>E. COLI</i> DNA GYRASE GEL ELECTROPHORESIS RESULTS, SUPER-COILED (R. PLASMID+GYRASE), RELAXED (R. PLASMID), NOVABIOCIN (R. PLASMID+GYRASE+NOVABIOCIN)	49
FIGURE 39: 1 MG/20 μ L (W/V) DILUSION <i>S. AUREUS</i> DNA GYRASE GEL ELECTROPHORESIS RESULTS, SUPER-COILED (R. PLASMID+GYRASE), RELAXED (R. PLASMID), NOVABIOCIN (R. PLASMID+GYRASE+NOVABIOCIN)	49
FIGURE 40: 1 MG/50 μ L (W/V) DILUSION <i>S. AUREUS</i> DNA GYRASE GEL ELECTROPHORESIS RESULTS, SUPER-COILED (R. PLASMID+GYRASE), RELAXED (R. PLASMID), NOVABIOCIN (R. PLASMID+GYRASE+NOVABIOCIN)	49
FIGURE 41: 1 MG/100 μ L (W/V) DILUSION <i>S. AUREUS</i> DNA GYRASE GEL ELECTROPHORESIS RESULTS, SUPER-COILED (R. PLASMID+GYRASE), RELAXED (R. PLASMID), NOVABIOCIN (R. PLASMID+GYRASE+NOVABIOCIN)	50

LIST OF TABLES

TABLE 1: CLASSIFICATION OF TOPOISOMERASES	11
TABLE 2: COMBINED AREAS UNDER ROC CURVES OF APPLIED GRID FILES BELONGING TO DOCKING EXPERIMENTS.	45
TABLE 3: COMPOUNDS WHICH RECEIVED HIGHEST E-MODEL SCORE DURING WATERS 235 AND 263 INCLUDED XP DOCKING TO 3G7B AND SELECTED FOR “GEL BASED SUPERCOILING ASSAY”	46
TABLE 4: COMPOUNDS WHICH RECEIVED HIGHEST E-MODEL SCORE DURING WATER 408 INCLUDED XP DOCKING TO 3G7E AND SELECTED FOR “GEL BASED SUPERCOILING ASSAY”	47
TABLE 5: ACTIVITIES ZINC CODES AND MAXIMUM E-MODEL SCORES OF SELECTED COMPOUNDS DURING <i>E. COLI</i> AND <i>S. AUREUS</i> DNA GYRASE SUPERCOILING ASSAYS IN DIFFERENT CONCENTRATIONS.	51



1. INTRODUCTION

Escherichia coli (*E. coli*) and *S. aureus* (*S. aureus*) are two main, very well known representatives of Gram-negative and positive bacteria subclasses. *E. coli* is commonly found in the normal lower intestine flora of warm-blooded organisms (endotherms) where as *S. aureus* is frequently found as a part of the normal skin flora (1).

When usual balances of both floras are disturbed by immune system deficiencies or other factors, both behave in opportunistic manner. Virulent strains of *E. coli* can cause gastroenteritis, urinary tract infections, and neonatal meningitis. In rarer cases, these strains are also responsible for hemolytic-uremic syndrome, peritonitis, mastitis, septicemia and Gram-negative pneumonia (1).

S. aureus can cause a range of illnesses from minor skin infections, such as pimples, impetigo, boils (furuncles), cellulitis folliculitis, carbuncles, scalded skin syndrome, and abscesses, to life-threatening diseases such as pneumonia, meningitis, osteomyelitis, endocarditis, toxic shock syndrome, bacteremia, and sepsis (1).

A good example for this opportunistic behavior is the tuberculosis (TB) case in Human Immunodeficiency Virus (HIV) infected persons (2). According to the World Health Organization (WHO), the increasing values of 8-10 million people each year are infected by TB. Especially pulmonary tuberculosis in which primary factors are *E. coli* and *S. aureus*, causes death in most cases of HIV (3).

In clinical practice, the biggest problem encountered in TB is the multi-drug resistance. Unfortunately, in nearly 40 years, no new drugs except rifampicin and rifabutin have been introduced on the market and the traditional clinical combination treatment performed using these drugs can not prevent infection resistance anymore (4). Before leading to major results, specific precautions treatments for *E. coli* and *S. aureus* infections in the early stages of HIV infections are almost certainly a solution.

In classical antibacterial compound design, when it is intended to prevent DNA replication or transcription at the level cell division, usually DNA gyrase enzyme is targeted for gram-negative bacteria, and Topoisomerase IV enzyme is targeted for gram-positives (5).

DNA gyrase, an enzyme unique to prokaryotes, has been implicated in almost all processes that involve DNA and majorly catalyses the ATP-dependent introduction of closed circular double-helix DNA to negative super-helices. Structurally it is a tetramer and includes an A2B2 motive. A chain is connecting to the DNA and is responsible for opening, breaking and re-unification of the structure, and the energy necessary for this process, is provided in B sub-chain by ATP hydrolysis. Although efficient inhibitors of this protein have been known for more than 20 years, none of them have enjoyed prolonged pharmaceutical success (5).

It is only recently that the X-ray crystal structures and binding dynamics of *E. coli* (Pdb id: 3G7E) and *S. aureus* (Pdb id: 3G7B) DNA Gyrase B with their inhibitors “prop-2-yn-1-yl {[5-(4-piperidin-1-yl-2-pyridin-3-yl-1,3-thiazol-5-yl)-1H-pyrazol-3-yl]methyl} carbamate” (1a) and “methyl ({5-[4-(4-hydroxypiperidin-1-yl)-2-phenyl-1,3-thiazol-5-yl]-1H-pyrazol-3-yl}methyl)carbamate” (1b) were released by PDB data Bank (Figure 1).

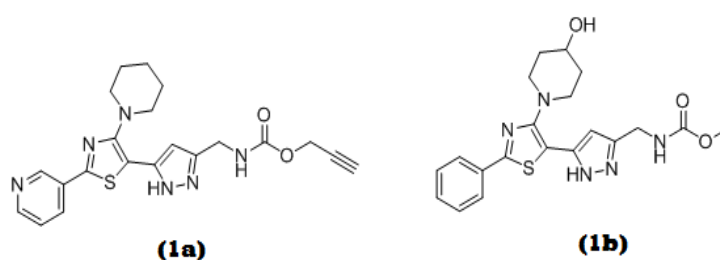


Figure 1: Chemical structure of ligands from 3G7E (1a) and 3G7B (1b)

When 6°A radius ligand surrounding binding region amino acids of these two crystal structured DNA gyrase B ATPase were isolated with above stated ligands (1a) and (1b), underneath interaction patterns were identified;

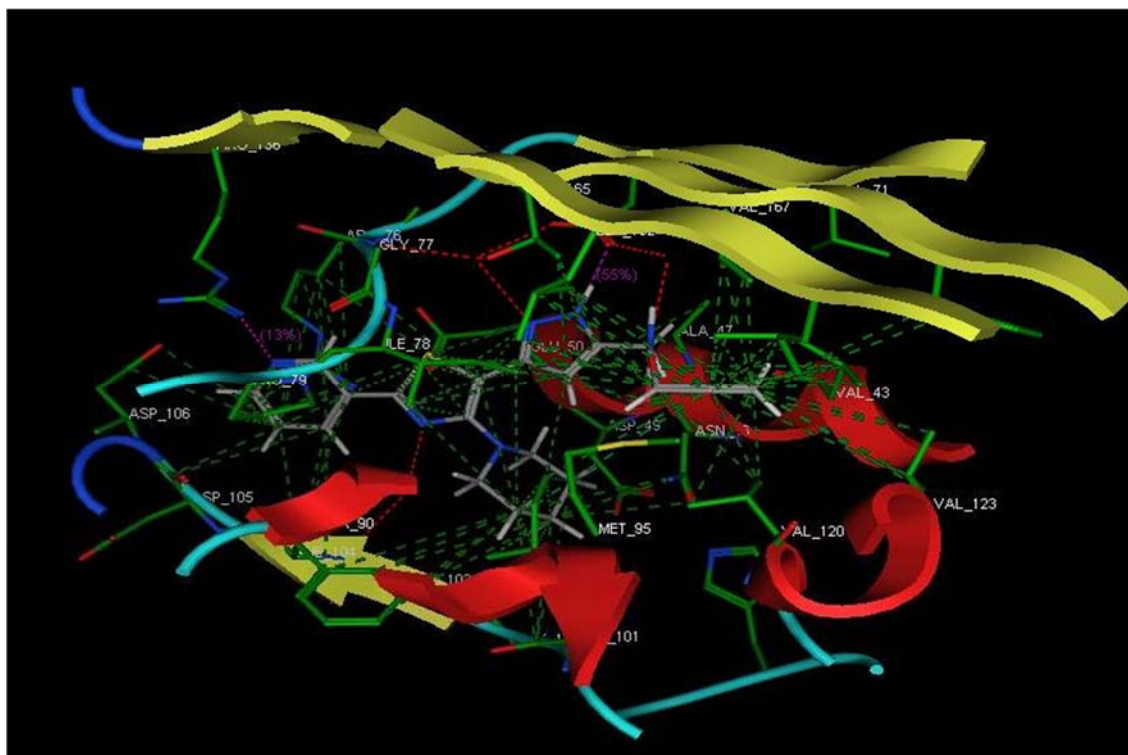


Figure 3: The hydrophobic interactions between the ligand and *E. coli* binding site

Between the compound (1b) and the binding site of *S. aureus* (Pdb id: 3G7B), the water molecules show a dominant activity of setting boundaries to positioning. Only one nitrogen atom of the pyrazol ring system has H bond donor relationship with Asp 81 and carbonyl group of carbamate side chain possibly acts as a hydrogen bond acceptor from Asn 54 (Figure 4).

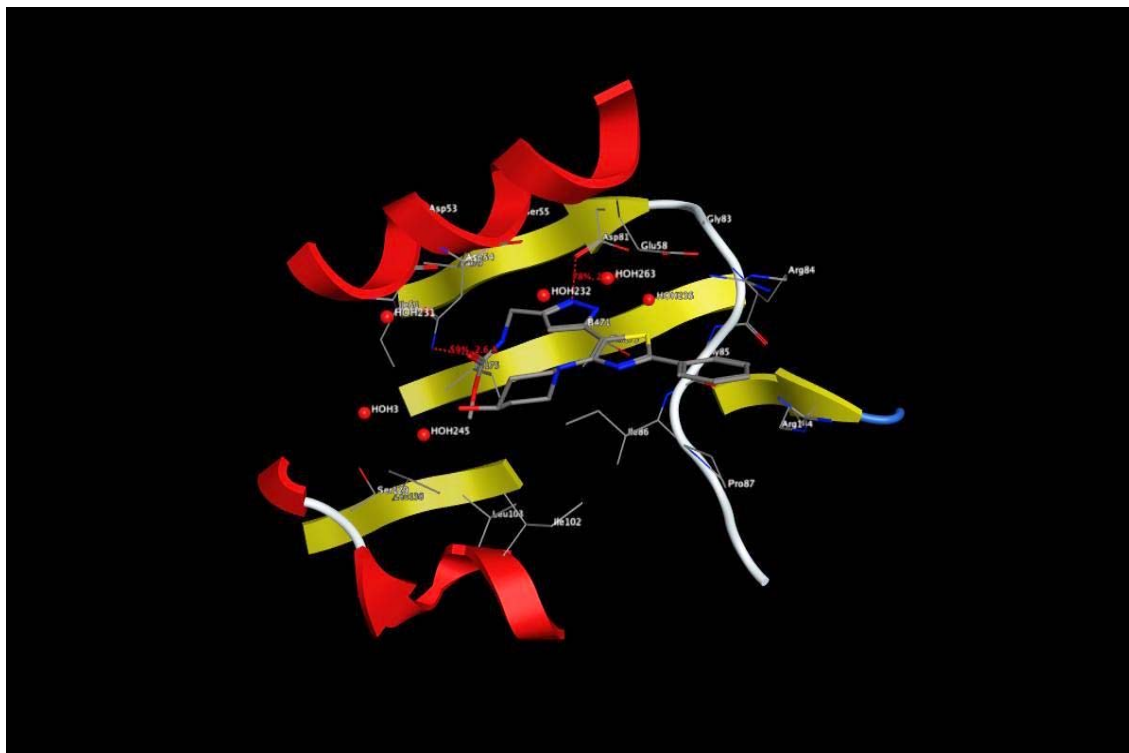


Figure 4: H-bonds between *S. aureus* binding site and the ligand

The phenyl ring system is in arene-cation relationship with Arg 84 where Ile 86 and Gly 85 amino acids create hydrophobic interactions with thiophen ring system. In addition the pyrazol is in the same type of interactions with the amino acids Asn 54, Glu 58 and Ile 86 (Figure 5).

Using BlastP both amino-acid sequences were aligned and scored by matrix adjustment composition method to verify their similarities. The entire identities have 100/206 (49%) and on the basis of Positives 127/206 (62%) similarities (6).

When the amino acids of the both protein sequences has been overlapped as pairs (Figure 7), in the whole chain, the pairwise root-mean-square deviation (RMSD) matrix value is 1.50, and the value belong to backbone is 1.19.



Figure 7: *E. coli* and *S. aureus* overlapped as pairs

These scorings and the interactions above suggest that positioning of some water molecules especially HOH 408, 443 in *E. coli* and HOH 235, 263 in *S. aureus* DNA Gyrase B ATPase site are very important for their bridge and boundary functions that might cause fractional differences among ligand designs for the both DNA Gyrase B's in terms of specific selectivity.

Molecular docking and HTVS are two basic tools which have widespread usage in computer aided drug design.

Docking software consists of two basic elements; the simulation algorithms that produce “exposure-pose” or “pose” which determine how structures such as ligand-protein or protein-protein interact with each other and the scoring algorithms made of certain mathematical functions that rank these poses. The types and algorithms of these two functions differ from software to software.

Usually all docking softwares determine poses that are responsible for the most probable interactions. However, up to now, scoring functions that supposed to reflect the relationship between activity and pose are unable to show high achievement as many different biological parameters, including basic solvation parameters can not be revealed to the mathematical algorithms sufficiently.

In all virtual screening studies, the test sets that consist of compounds selected from databanks are being enriched by using cluster of compounds that experimentally proven to be active on the related target previously. When the ratio of active compounds in the test set is known and when the correct pose is determined, it is possible to measure statistically and numerically the quality of hypothesis that is screening based on by using methods such as ROC curves.

One of the main objectives in this study is to suggest water concerned new inhibitor-active site interaction patterns that can contribute developing selective Escherichia coli and *S. aureus* DNA Gyrase B ATPase inhibitors. Then using this patterns, perform a series of HTVS experiments based on stepwise docking algorithms to find new selective hit molecules. Last test all of our hypothesis by in-vitro screening assays.

During this study, first 36 *invitro* experimented active DNA Gyrase ATPase inhibitor ligands (true positives) were added to the 5000 trial and 50000 test sets of compounds for enrichment purposes and afterward prepared for docking based HTVS.

Then the trial set was subjected to a series of docking processes, by using water molecules excluded or included interaction patterns so called “unrestricted” and “restricted” GRID files. Scoring and further ROC curve evaluation of these files identify which interaction pattern should be considered as “important” or “necessary” for inhibition process.

Later, the selected GRID files were similarly experimented over test set. Finally, selected and combined twenty top scoring compounds were subjected to gel-based supercoiling assays of both microorganisms for inhibition measurement, hypothesis testing and hit finding.

2. GENERAL INFORMATION

In the design of antibacterial compounds, when prevention of DNA (Deoxyribonucleic acid) replication or transcription at the level of cell division is intended, the inhibition of DNA gyrase enzyme often referred to simply as gyrase and /or topoisomerases are among the popular targets.

2.1. TOPOISOMERASES

The DNA topoisomerases are essential for DNA replication, transcription, recombination, as well as for chromosome compaction and segregation. Several families and subfamilies of the two types of DNA topoisomerases (I and II) have been described in the three cellular domains of life (Archaea, Bacteria and Eukarya), as well as in viruses infecting eukaryotes or bacteria.

The main families of DNA topoisomerases, Topo IA, Topo IB, Topo IC (Topo V), Topo IIA and Topo IIB (Topo VI) are not homologous, indicating that they originated independently ([7](#)).

The Classification of Topoisomerases are given below the Table 1.

Topoisomerase ^a	Subfamily Type	Subunit Structure	Size(s)
Eubacterial DNA topoisomerase I (<i>E. coli</i>)	IA	Monomer	865
Eubacterial DNA topoisomerase III (<i>E. coli</i>)	IA	Monomer	653
Yeast DNA topoisomerase III (<i>S. cerevisiae</i>)	IA	Monomer	656
Mammalian DNA topoisomerase III α (human)	IA	Monomer	1001
Mammalian DNA topoisomerase III β (human)	IA	Monomer	862
Eubacterial and archaeal reverse DNA gyrase (<i>Sulfolobus acidocaldarius</i>)	IA	Monomer	1247
Eubacterial reverse gyrase (<i>Methanopyrus kandleri</i>) ^c	IA	Heterodimer	A, 358 B, 1221
Eukaryotic DNA topoisomerase I (human)	IB	Monomer	765
Poxvirus DNA topoisomerase (vaccinia)	IB	Monomer	314
Eubacterial DNA gyrase (<i>E. coli</i>)	IIA	A ₂ B ₂ hetero-tetramer	GyrA-875 GyrB-804
Eubacterial DNA topoisomerase IV (<i>E. coli</i>)	IIA	C ₂ E ₂ hetero-tetramer	ParC-752 ParE-630
Yeast DNA topoisomerase II (<i>S. cerevisiae</i>)	IIA	Homodimer	1428
Mammalian DNA topoisomerase II α (human)	IIA	Homodimer	1531
Mammalian DNA topoisomerase II β (human)	IIA	Homodimer	1626
Archaeal DNA topoisomerase VI (<i>Sulfolobus shibatae</i>)	IIB	A ₂ B ₂ hetero-tetramer	A, 389 B, 530

Table 1: Classification of Topoisomerases

2.1.1. TOPOISOMERASES TYPE IA

The topoisomerases belonging to the type IA subfamily share the following properties ([8](#), [9](#), [10](#));

- (a) They are all monomeric,
- (b) Cleavage of a DNA strand is accompanied by covalent attachment of one of the DNA ends to the enzyme through a 5' phosphodiester bond to the active site tyrosine,
- (c) All require Mg (II) for the DNA relaxation activity.

- (d) Negative supercoils are substrates for the relaxation reaction which do not go to completion and consistent with the last two point interaction,
- (e) All enzymes in this subfamily require an exposed single-stranded region within the substrate DNA,
- (f) When the process has been examined, in addition to the ability to relax negative supercoils, these enzymes can catalyze the knotting, unknotting, and interlinking of single-stranded circles as well as the knotting, unknotting, catenation, and decatenation of gapped or nicked duplex DNA circles.

Some domain structures of type IA topoisomerases aligned with respect to *E. coli* Topo IA are given at Figure 8.

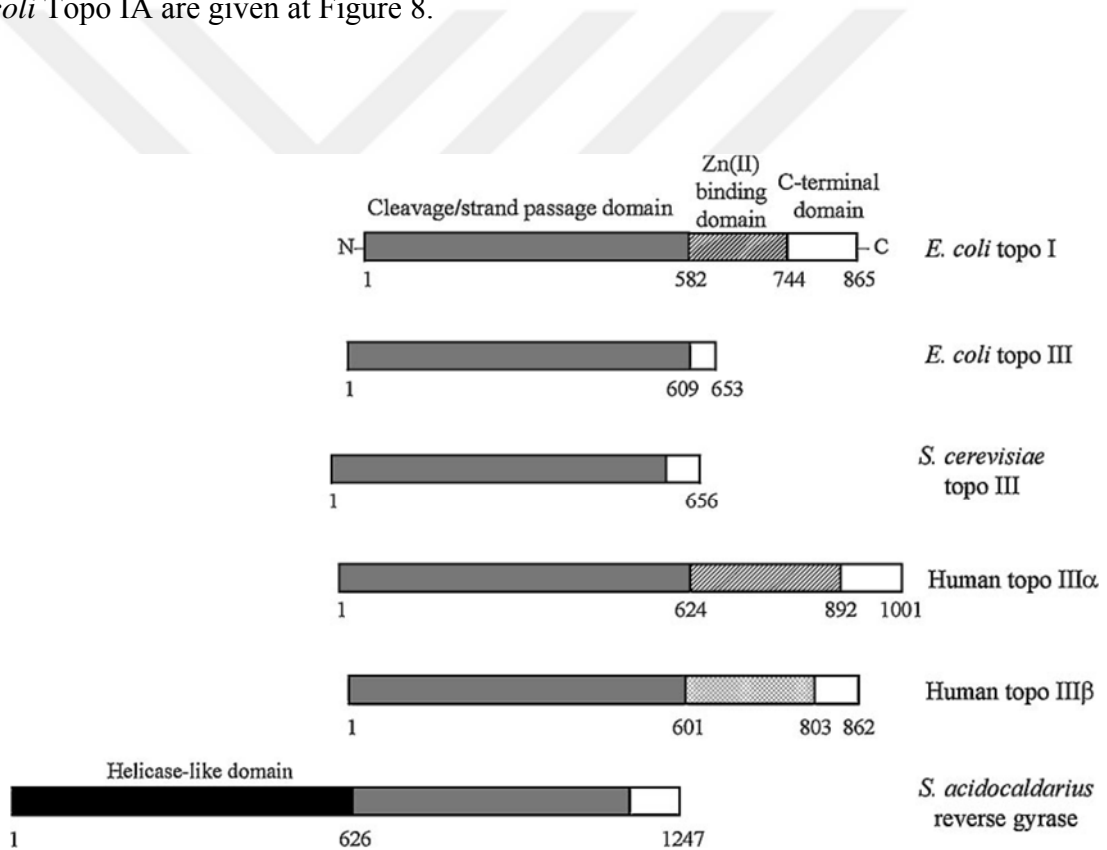


Figure 8: Domain structure of type IA topoisomerases

2.1.2. TOPOISOMERASES TYPE II

The following properties are shared by all type II topoisomerases (8);

- (a) The dimeric enzymes bind duplex DNA and cleave the opposing strands with a fourbase stagger (topoisomerase VI may generate a two-base stagger) (9).
- (b) Cleavage involves covalent attachment of each subunit of the dimer to the 5' end of the DNA through a phosphotyrosine bond.
- (c) A conformational change pulls the two ends of the cleaved duplex DNA apart to create an opening in what is referred to as the gated or G-segment DNA. A second region of duplex DNA from either the same molecule (relaxation, knotting or unknotting) or a different molecule (catenation or decatenation), referred to as the transported or T-segment, is passed through the open DNA gate. This feature of the reaction explains why the linking number is changed in steps of two when the supercoiling of a circular DNA is changed
- (d) The reactions require Mg(II), and ATP (Adenosine-5'-triphosphate) hydrolysis is required for enzyme turnover and rapid kinetics, although one cycle of relaxation or decatenation/catenation can occur in the presence of the nonhydrolyzable analog of ATP, ADPNP (10, 11, 12).
- (e) The crystal structures of several members, including the structurally distinct topoisomerase VI, reveal that the active site tyrosines are situated in a helix-turn-helix (HTH) motif found within a domain that strongly resembles the DNA binding region of the *E. coli* catabolite activator protein (CAP). In addition, acidic residues within a Rossmann fold on the opposing protomer appear to collaborate with the HTH region of the CAP-like domain to assemble the active site for catalysis and may be involved in metal ion binding in some cases (9, 13).

2.1.3. DNA GYRASE

Two topoisomerase-producing genes are for DNA gyrase, a type II DNA topoisomerase, and the other for topoisomerase I, excluding the topoisomerase IV gene, another type II DNA topoisomerase involved in the DNA replication process. In such cases, DNA gyrase is the only type II enzyme responsible for the negative supercoiling of DNA by relaxing and decatenating positively supercoiled DNA (10).

DNA gyrase, which is found in all bacteria, is an adenosine triphosphate (ATP)-dependant hydrolytic enzyme, a proven target for antibacterial chemotherapy ([11](#)).

DNA gyrase comprises two subunits, GyrA and GyrB, forming functional heterodimer A₂B₂. The GyrA subunit is responsible for DNA breakage and reunion. This catalytic process involves large conformational changes in the enzyme, which are triggered by the binding and hydrolysis of ATP on the GyrB subunit ([12](#))

2.2. COMPUTATIONAL METHODS

The process of drug design created a new branch named 'computer-aided drug design' besides the conventional approaches in 1990s with the beginning of internet and 2000s due to computer-software industry pushing the limits of technology and speed.

This branch walking arm in arm with the traditional techniques reduced costs and workload of the pharmaceutical industry not only by launching new molecules to the market faster but also brought up the 'rational drug design' perspective to new centuries agenda.

Computer-aided drug design consists of a series of methodological simulation combinations that use molecular modeling techniques such as molecular mechanics and dynamics with cheminformatics, bioinformatics and related branches.

The whole idea lying beneath these simulations is that; in physiological environment, pharmacological response is related to the receptor/enzyme based interaction of the three-dimensional chemical structure with the macromolecule.

Basically, these methodologies can be divided into two approaches depending on the available data; 'ligand-based' and 'structure-based' drug design.

In ligand-based drug design, three-dimensional crystal structure of the macromolecule could not be obtained. Thus, researchers try to establish the correlations

between structure-efficacy by using structural and / or physicochemical properties of molecules which precise biological activities are known. Pharmacophore modeling, QSAR and 3D-QSAR can be represented as the most used methods in this field.

In structure-based drug design, three-dimensional structure of the macromolecule is obtained by crystallographic techniques or homology modeling. The interaction pattern in physiological environment between protein structured macromolecule and ligand is determined by simulation techniques such as docking, de-novo design and pharmacophore modeling.

Whatever technique is used, the resulting molecules which are thought to acquire pharmacological response should be subjected to biological activity tests.

Biological activity tests usually are based on chemometric, fluorometric, radiometric and/or colorimetric measurement. Response versus concentration evaluation is obtained by result of the interaction between the molecule and the isolated receptor and/or enzyme system. Optimized and automated in high-speed is known as 'high throughput screening techniques'[\(13\)](#)[\(14\)](#)

2.2.1. STRUCTURE BASED DRUG DESIGN

Structure-based drug design (or direct drug design) relies on knowledge of the three dimensional structure of the biological target obtained through methods such as X-Ray crystallography or NMR spectroscopy [\(15\)](#)

In case structure of a target is not available, it may be possible to create a homology model or in otherwords comparative model of the target based on the experimental structure of a related protein. Homology modeling relies on the identification, alignment and reconstructing of an atomic-resolution model of the target protein (query sequence) from its amino acid sequence and an experimental three-dimensional structure of a related homologous protein (template sequence). It has been shown that protein structures are more conserved than protein sequences amongst

homologues although sequences falling below a 20% sequence identity can have very different structure([16](#)).

By using the structure of the biological target, candidate drugs that are predicted to bind with high affinity and selectivity to the target may be designed using interactive graphics, automated computational and the intuition of a medicinal chemist.

Current methods for structure-based drug design can be divided roughly into two categories. The first category is about “finding” ligands for a given receptor, which is usually referred as database searching or virtual screening (VS, HTVS). In this case, a large number of potential ligand molecules are screened to find those fitting the binding pocket of the receptor by using simulations such as docking-scoring. The second category is about “building” ligands, which is usually referred as *de-novo* design. In that case, ligand molecules are built up within the constraints of the binding pocket by assembling small pieces that can be either individual atoms or molecular fragments in a stepwise manner ([17](#), [18](#), [19](#))

2.2.2. DOCKING AND SCORING

Molecular recognition, in other words, 'docking'; can be described as the simulation of the preferred conformations of a small molecule or peptide structure throughout the binding pocket of the target macromolecule in order to create a stable complex. Strength of the preferred relative conformations or in other words the bond strength between the two molecules is measured by scoring functions ([17](#)).

Docking software, consists of two basic elements;

- 1- Simulation algorithms that make up 'pose' (exposure), verifying the interaction of ligand-protein or protein-protein structures.
- 2- Certain mathematical algorithms which score these poses.

The main objective of employing a combination of these two functions, is to minimize the total free energy of the the system during the optimized interactions of ligand and protein structured macromolecule ([18](#)).

Usually, the algorithmic part of the software that creates the pose, matches the surfaces of protein and ligand structures or measures ligand-protein interaction energies and rotational movements around single bonds. In addition, it determines in very high probability the poses responsible for the interactions ([19](#), [20](#), [21](#), [22](#), [23](#), [24](#), [25](#)).

On the other hand, scoring functions are supposed to illustrate the strength of non-covalent interaction (binding-affinity) between two docked molecules. These functions based on some mathematical formulas derived from abbreviated molecular dynamic equations where many different biological parameters, including basic solvation parameters cannot be reflected sufficiently ([26](#), [27](#)).

2.2.3. HIGH THROUGHPUT VIRTUAL SCREENING-HTVS/VS

In computer aided drug design, high throughput virtual screening techniques - virtual screening (HTVS/VS) have fairly common use. Detailed information about this topic is illustrated and summarized in many valuable books ([17](#), [18](#), [19](#), [20](#), [21](#), [22](#)).

Simply, HTVS; can be summarized as screening of chemical structures containing libraries with the help of *in-silico* techniques in order to identify compounds may be connected to a drug target protein receptor or enzyme structure ([19](#)).

In the case of ligand-based drug design, the hypothesis is built on pharmacophore obtained by matching VS application and evaluated by comparing the effective molecules and/or the common properties constituting the biological activity (H-bond donor or acceptor groups, hydrophobic elements, etc.) with conformational and structural compliance of chemical libraries ([20](#)). If the number of ligands are not enough for pharmacophore modeling, the chemical similarity screening (chemical similarity searching) also can be possible ([21](#)).

In the structure-based design, docking and scoring are the basic applications to simulate the interactions between the structures in the chemical libraries and the crystallographic data and / or homology model by (22), (23).

Two studies come forward in the relevant literature research related to VS of DNA Gyrase A and B active regions.

The first one is obtained from the NCI database. In this study 140,000 compounds with molecular weight less than 500 were virtually screened based on docking for their activity on both DNA Gyrase A and B active regions. With in all, 10 compounds having highest score were further tested in-vitro by '*E. coli* DNA Gyrase supercoiling assay' where highest enzymatic activity was determined in 2 compounds that are structurally different from novobiocin and cyclothialidine, the most non selective inhibitors (24).

In an other study, the MDL- ACD and MDL-SCD databases fragments were docked to GyrB active site of theADPNP Gyr-DNA crystal structure (PDB 1EI1) and 2-aminobenzimidazole and indolin-2-one ring systems were shown to establish similar interactions just like ADPNP. Further, 10 structures having these ring systems were tested in terms of activity and optimized (25).

2.2.4 RECEIVER OPERATING CHARACTERISTIC CURVES – ROC CURVES

The main purpose of using the "receiver operating characteristic" curve method is to evaluate the ability of a given test to discriminate between two populations. This helps decision-making in fields which a wrong judgment may have serious consequences including clinical diagnosis and economic strategies. When virtual screening is used to speed-up the drug discovery process in pharmaceutical research, taking the right decision upon selecting or discarding a molecule prior to in vitro evaluation has vital importance. Characterizing both the ability of a virtual screening workflow to select active molecules and the ability to discard inactive ones, the ROC curve approach is a usefull decision tool. (26)

As a case study, the first virtual screening work flow focused on metabotropic glutamate receptor subtype 4 (mGlu4R) agonists was reported. Six compounds out of 38 selected and tested *in vitro* were shown to have agonist activity on this target of therapeutic interest (27)

The visual analysis of the ROC curve was useful to have an overall picture of the classification and to select the right threshold that marks the boundary between active and inactive classes (28)

The method can be applied successfully for GPCR's as well as crystal structures and homology model (29, 30)

2.3. DATABASES

Multiple scenarios are feasible for selecting compounds for a screening. The scenario of choice can vary significantly from one project to another and is dictated by the level of knowledge regarding the target structures and/or bioactive ligands (29).

When the three dimensional (30) structure and location of binding pocket(s) of a target are known, docking and structure based virtual screening have proven to be valuable methods for successful identification of novel bioactive molecules (31). On the contrary, when the 3D structure of the target is unknown, pharmacophore and virtual screening approaches play a predominant role in filtering compound collection(32). Finally, when no information is available regarding the 3D structure of the target or bioactive ligands, successful identification of hit compounds will rely on enrichment of compound collections and determination of the best possible molecular starting points (30).

In order to perform screening, large searchable collections of compounds are needed. Fortunately, multiple sources of both virtual compounds (developed on the basis of Markush reaction (33), or synthetic feasibility rules (34), and real compounds exist, in both private corporate or public (Zinc database , NCI diversity) collections, as

well as numerous commercial libraries that can be either focused or random in content (35).

Strategies for the compilation and organization of these compound collections are company dependent and of course, project dependant, considering the target type and the screening approaches (virtual and/or high-throughput screening (HTS) (30).

In addition to compound databases, ligand knowledge bases such as World drug index (WDI), MDLDrug Data Report (MDDR), WOMBAT, AurSCOPE1 (AureusPharma, Paris, France), ChemBioBase1 (Jubilant OrganosysLtd., New Delhi, India), GVKBIO (Hyderabad,India) are of high interest and widely used in the chemoinformatic approach to drug discovery (36).

Direct searches in journals and patents *via* available search engines also provide valuable sources of compound information.

For determination of 3D structures of our interest proteins, the Protein Data Bank (PDB) (www.pdb.org) archive, a single worldwide repository of information about the 3D structures of large biological molecules, including proteins and nucleic acids were used (2).

In our screening study ZINC; a free database of 13 million ready-to-dock commercially available and purchasable 3D formatted compounds for virtual screening, provided by the Shoichet Laboratory in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF) were used (35).

2.3.1. RSCB DATABASE

The Protein Data Bank first established at Brookhaven National Laboratories (BNL) (37) in 1971 as an archive for biological macromolecular crystal structures. In the 1980s the number of deposited structures began to increase dramatically due to the improved technology for all aspects of the crystallographic process, nuclear magnetic resonance (NMR) methods. By the early 1990s guidelines published by the

International Union of Crystallography (IUCr) requiring data deposition for all structures and the management of the PDB became the responsibility of the Research Collaboratory for Structural Bioinformatics (RCSB). Up to now, 74601 structures were deposited and served for the users.

The key systematic data flow for the Protein Data Bank Data consists of data deposition, annotation and validation. These steps are part of the fully documented and integrated data processing system shown in Figure 9.

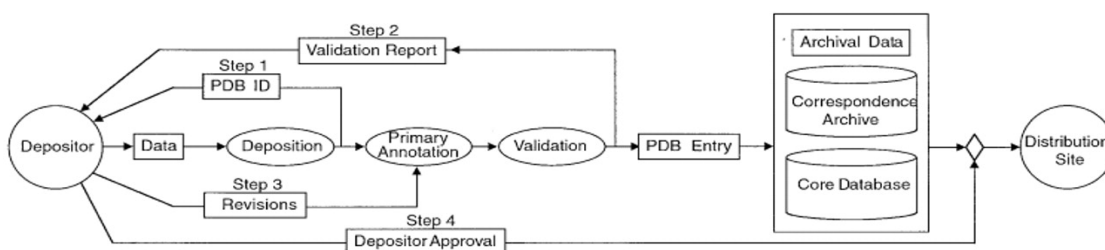


Figure 9: The steps in PDB data processing. Ellipses represent actions and rectangles define content.

Current status information, comprised of a list of authors, title and release category, is stored for each entry in the core database and is made accessible for query via the net interface.

- PDB Data

The primary information stored in the PDB archive consists of coordinate files for biological molecules. These files list the atoms in each protein, and their 3D location in space. These files are available in several formats (PDB, mmCIF, XML). A typical PDB formatted file includes a large "header" section of text that summarizes the protein, citation information, and the details of the structure solution, followed by the sequence and a long list of the atoms and their coordinates. The archive also contains the experimental observations that are used to determine these atomic coordinates.

- Visualizing Structures

PDB files can be viewed directly using a text editor, visualization and/or simulation programs used in computed aided drug design. Also online tools, such as the ones on the RCSB PDB website, allow one to search and explore the information under the PDB header, including information on experimental methods and the chemistry and biology of the protein. These programs often include analysis tools that allow you to measure distances and bond angles and identify structural features.

- **Coordinate Files**

In a typical entry, one will find a diverse mixture of biological molecules, small molecules, ions, and water, often, with the names and chain IDs to help sort these out. In structures determined from crystallography, atoms are annotated with temperature factors that describe their vibration and occupancies that show if they are seen in several conformations. NMR structures often include several different models of the molecule [\(2\)](#).

2.3.2. ZINC DATABASE

ZINC is a downloadable database of 3D compounds built from the catalogs of major vendors, provided by the Shoichet Laboratory in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF).

A Web server has been established to distribute the ZINC database, allowing investigators to search, browse, subset, and download some or all of the molecules in SMILES, mol2, SDF, and DOCK flexibase formats.

Users may search ZINC based on several criteria (Figure 10). One can either limit molecular properties such as net charge and molecular weight, individual ZINC database registration codes, the unique serial number assigned to each substance in ZINC, choosing a text file of codes or SMILES strings to upload or by simply drawing a molecular structure/substructure using the Java Molecular Editor (JME) (Figure 10A). Molecules matching any of these criterias specified will be found in the browser page (Figure 10B) with vendor may also be specified.

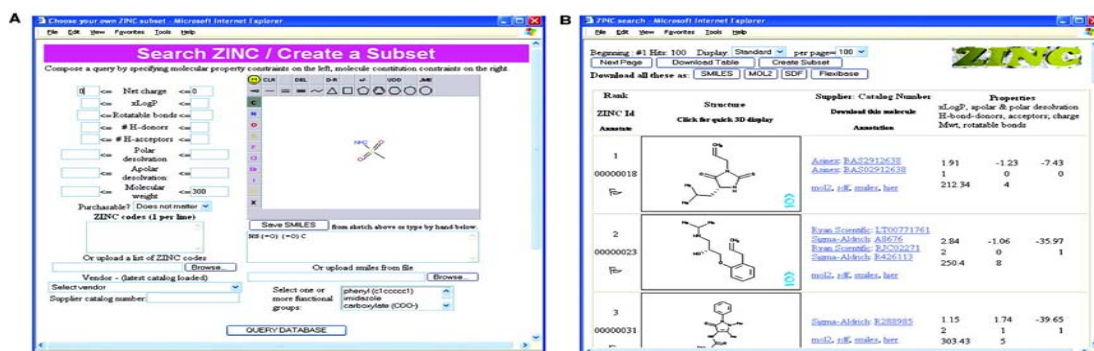


Figure 10: A. The ZINC search tool B. the ZINC database browser.

The ZINC database provides 3D molecules in several formats compatible with most docking programs. The Web-based interface is fast and supports moderately complex queries (35).

2.4. THE METHODS ARE USED FOR DETERMINATION OF ACTIVITY

2.4.1. SUPERCOILING ASSAY

DNA topoisomerases such as bacterial topoisomerase II (gyrase) convert relaxed circular DNA into supercoiled DNA. The DNA Topoisomerase II (Gyrase) Assay Kit is based on the principle that the supercoiled DNA and relaxed DNA yield different fluorescent intensity when interact with fluorescence dye H19. The relaxed DNA suppresses the fluorescent intensity much more than the supercoiled DNA. Therefore, when the relaxed DNA is converted into its supercoiled form, the fluorescent signal increases. The change of fluorescence intensity is used to measure the supercoiling reaction.

2.4.2. GEL BASED DRUG INHIBITION ASSAY

The IC_{50} for inhibition of gyrase supercoiling is visually assessed as the concentration of compound which leads to a 50% reduction of the supercoiled band and

the appearance of a spread of slower migrating topoisomers above. This is then verified using gel documentation software and subsequent statistical analysis. (Figure 11)

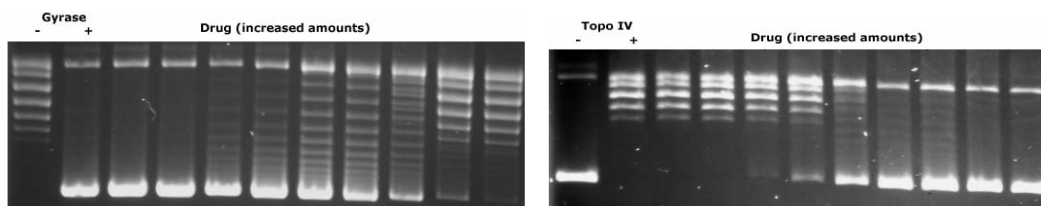


Figure 11: Gel documentation software and subsequent statistical analysis

The IC_{50} for inhibition of relaxation is obtained in a similar way. 50% inhibition is determined visually as being the compound concentration at which the relaxed band is reduced by 50% and a supercoiled band of topoisomers becomes apparent. Gel documentation software and statistical analysis is then used to confirm the result

3. MATERIALS AND METHODS

PDB files belong to “Crystal structure of *E. coli* Gyrase B co-complexed with inhibitor” and “*S. aureus* Gyrase B co-complex with inhibitor” (PDB ID; 3G7E and 3G7B) were downloaded from RCSB Protein Data Bank (2). Preparation of protein structures, trial and test sets, GRID files, docking and scoring were performed by algorithms belong to modules of Maestro (Schrodinger Inc, USA). Special “ROC Curves SVL” ve “Ligand-receptor contacts (visualization+scoring) SVL” belong to MOE (Chemical Computing Group Inc., Canada) software were used during preparation of ROC curves and interaction graphics. Computations performed by special designed Dell Precision T7500 work stations of Yeditepe University, Faculty of Pharmacy.

After screening, compounds presumed to be active were synthesized and purchased from Molport Chemicals (Letonia) with minimum 99,5 % purity and used after checking their LC-MSMS spectral and elemental analysis results. Biological activities of compounds were assessed by *E. coli* and *S. aureus* Gyrase Supercoiling

Assay Kits (Inspiralis Inc. UK- Gyrase Supercoiling Assay kits-K0003 and SAS4002) using BIO-RAD (CA, USA) gel electrophoresis and imaging systems against novobiocin (CAS; 1476-53-5, AppliChem, Germany) as a reference.

3.1 PREPARATION OF PROTEIN STRUCTURES

Subsequently, 3G7E and 3G7B PDB files were downloaded and subjected to Protein preparation wizard workflow of Maestro, for hydrogen insertion and rotamer adjustment, and H-bond optimization using OPLS 2005 as the energy parameters.

3.2 PREPARATION OF TRIAL AND TEST SETS

This process, begins with reduction of 1442716 compounds in the content belong to “clean-leads-subset (# 11)” (LogP values < 3.5, molecular weight < 350 and the number of rotational bonds ≤ 7) of the Zinc data base to 50000 randomly selected compounds as Test set. Trial set was evaluated from Test Set with a second random selection of 5000 from this group.

Further these two sets in SDF format were transferred to the Lig Prep function of Maestro to be prepared for their tautomers and ionized forms at various pH levels (pH= 7 ± 2). Later duplicates were eliminated by an automatic script. Activity proven 36 ligands were prepared in the same manner and added for enrichment purposes to both sets before docking and scoring.

3.3 PREPARATION OF GRID FILES.

All GRID files belong to both microorganisms DNA Gyrase ATPase active site were prepared using Maestro’s Glide- Receptor Grid Generation tool. Receptor binding pockets defined by picking 6 Å surrounding of both ligands existing in each PDB file. During preparation, the original ligands were excluded and a scaling factor of 1.0 and a partial charge cutoff of 0.25 were used as parameters for the Van der Walls radius scaling factor.

For *S. aureus* Gyrase B active site (PDB ID; 3G7B), three different GRID files were prepared;

- Without water molecules (without restrictions/ places the compounds considering the original ligand as centroid)
- with HOH 235 (with restrictions to make bond either with HOH 235 or define the position of ligand considering the position of HOH 235)
- with HOH 235 and 263 (with restrictions to make bond either with HOH 235 and /or HOH 263 or define the position of ligand considering the position of HOH 235 and 263)

Similarly for *E. coli* Gyrase B (PDB ID; 3G7E), three different GRID files were prepared;

- Without water molecules (without restrictions/ places the compounds considering the original ligand as centroid)
- with HOH 408 (with restrictions to make bond either with HOH 408 or define the position of ligand considering the position of HOH 408)
- with HOH 408 and 443 (with restrictions to make bond either with HOH 408 and /or HOH 443 or define the position of ligand considering the position of HOH 408 and 443)

3.4 DOCKING AND SCORING

All docking experiments were performed by using Maestro's Glide-docking tool. The basic settings for HTVS, SP and XP algorithms were set as;

- treating receptor as rigid and ligands as flexible,
- dock without using core pattern comparison algorithm
- use constrains from GRID files if needed,
- write 1 000 000 poses per docking run and perform top 5 poses a post-docking minimization.

In trial sets for both microorganisms DNA Gyrase ATPase active site, HTVS algorithm were preformed by using each GRID file stated.

With the enriched test set of *S. aureus* Gyrase B active site (PDB ID; 3G7B), restricted GRID file with HOH 235 and 263 (with restrictions to make bond either with HOH 235 and /or HOH 263 or define the position of ligand considering the position of HOH 235 and 263) was used. 50000 prepared compounds were included with their isomers and tautomers during docking with HTVS algorithm whereas a cut-off of 20000 were experimented for SP and re scored in place by XP algorithms for computational and time concerns.

For *E. coli* Gyrase B (PDB ID; 3G7E) active site, enriched test set was docked by using GRID file without water molecules (without restrictions/ place the compounds considering original ligand as centroid) and GRID HOH 408 and 443 (with restrictions to make bond either with HOH 408 and /or HOH 443 or define the position of ligand considering the position of HOH 408 and 443) in HTVS mode for detailed evaluation of trial set ROC curves. Additional dockings were accomplished with a cut-off of 20000 in SP and scored in place by XP modes for computational and time concerns. All dockings were ranked according to their docking score and e-model score for further evaluation of ROC curves.

3.5 ROC CURVE EVALUATION

All ranked results of docking experiments belong to each microorganism were than transferred to the MOE software (Chemical Computing Group Inc., Canada) as SDF files and compiled as databases. Later databases ranked according to their docking and e-model scores as giving actives as 1 and non-predicted as 0 and processed with “ROC Curves SVL” using appropriate thresholds. Percentage evaluations were made by using multiplication of area under curve with 100.

After assessment of the curves and poses, total twenty compounds, where nine of them which received highest e-model score during waters 235 and 263 included XP docking to 3G7B and eleven compounds which received highest e-model score during water 408 included XP docking to 3G7E selected for gel based inhibition assay against standard novobiocin.

3.6 BIOLOGICAL EVALUATION

All twenty compounds were tested with the “Gel Based inhibition Assay” at 1mg/20, 50, 100 microliter concentrations both in *Staphylococcus aureus* (for 3G7B) and *Escherichia coli* (for 3G7E) with novobiocin standard for comparison purposes.

DNA gyrase supercoiling assays were performed with a Gyrase Supercoiling Assay Kit (Inspiralis) according to the manufacturer’s instructions and analyzed by monitoring the conversion of relaxed pBR322 plasmid to its supercoiled form using DNA gel electrophoresis. Essentially, 1 U of either *E. coli* or *S. aureus* DNA gyrase was first diluted in 5×gyrase buffer and incubated in an assay buffer (35 mM Tris HCl (pH 7.5), 24 mM KCl, 4 mM MgCl₂, 2 mM DTT, 1.8 mM spermidine, 1 mM ATP, 6.5% (w/v) glycerol, and 0.1 mg/mL BSA), with 0.5 µg of pBR322 plasmid and purchase twenty compound dilutions at 37 °C for 30 min. Reactions were stopped with the addition of stop dye (40% sucrose, 100 mM Tris HCl (pH 7.5), 1 mM EDTA, and 0.5 mg/mL bromophenol blue) and loaded onto TAE agarose gel (1%). Gels were visualized using a gel documentation system (Bio-Rad ChemiDoc). Since high levels of DMSO are known to affect DNA gyrase activity, titration was used to determine the minimum amount of DMSO to be used in the assays, and 5% DMSO (with negligible or no effect on the gyrase) was chosen to dilute the compounds (38).

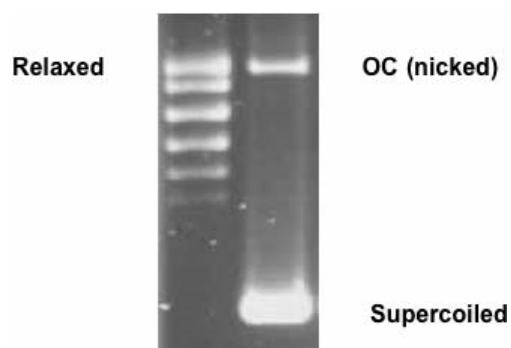


Figure 122: Standard view expected for Gyrase Supercoiling Assay (Relaxed; without gyrase, OC; Open coiled, Supercoiled; with the addition of active gyrase)

4. RESULTS

4.1. TRIAL SET DOCKING, SCORING AND ROC CURVES EVALUATION

The trial sets which consist of randomly selected 5000 compounds enriched with 36 active ligands were docked with Glide-docking HTVS mode by prepared as explained GRID files of *E. coli* and *S. aureus*.

Meanwhile during the docking processes the Novobiocin structure was bound to its original position with an RMSD range of 0.83-0.94 Å and showed similar interactions with its original crystallographic data in both cases.

The ROC curves belong to HTVS docking of *S. aureus* Gyrase B (PDB ID: 3G7B) trial set without water molecules, with HOH 235, with HOH 235 and HOH 263 are given in Figures 13-18

When the poses obtained in the absence of water molecules (without restrictions/ place the compounds considering original ligand as centroid) are selected for maximum docking and e-model scores and are moved to ROC curves, the consequent area under curves are 0,6268 and 0,9388. According to the results, the score for e-model separates the true positives from false ones 93,8 percent successfully.

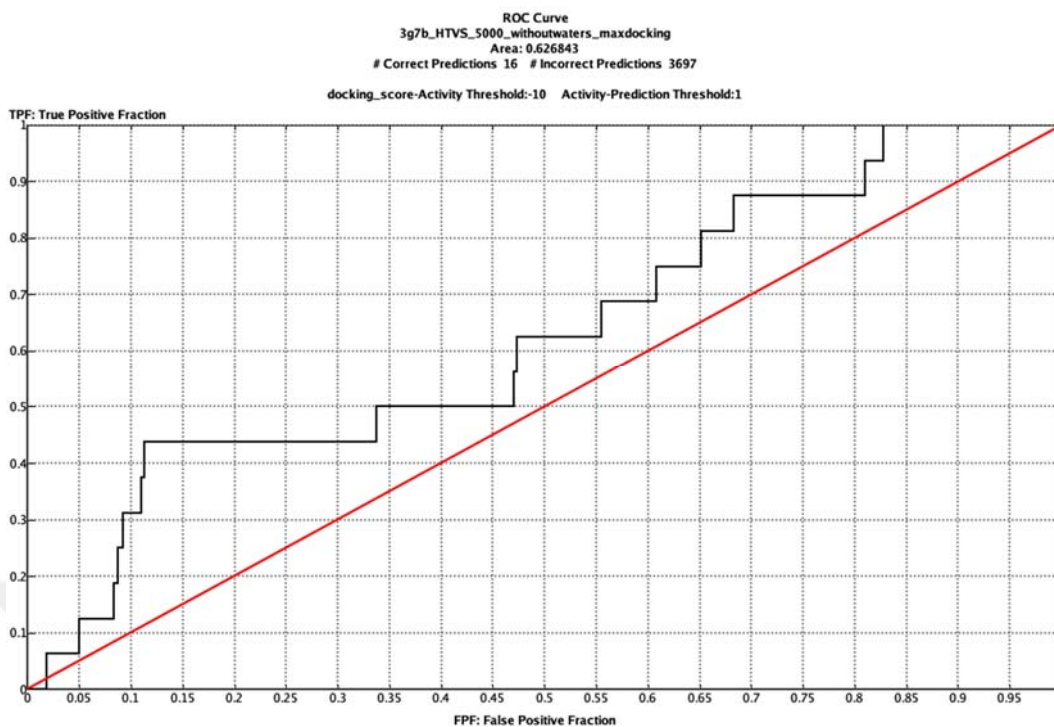


Figure 13: 3G7B docked without waters, selected and ranked according to maximum docking score

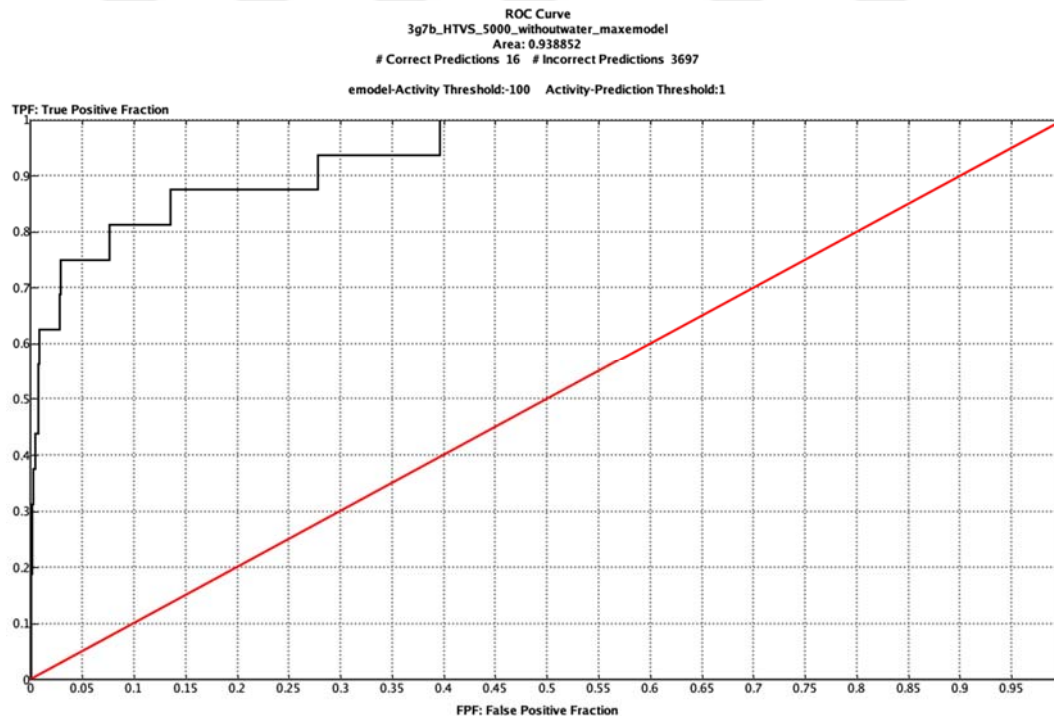


Figure 13: 3G7B docked without waters, selected and ranked according to maximum e-model score

Concerning poses with HOH 235 comprising GRID file (with restrictions to make bond either with HOH 235 or define the position of ligand considering the position of HOH 235), the areas for docking and e-model scores were found to be 0,6628 and 0,9265. According to the results, the process for e-model separates the true positives 92,6 % where docking score separates 66,28 %.

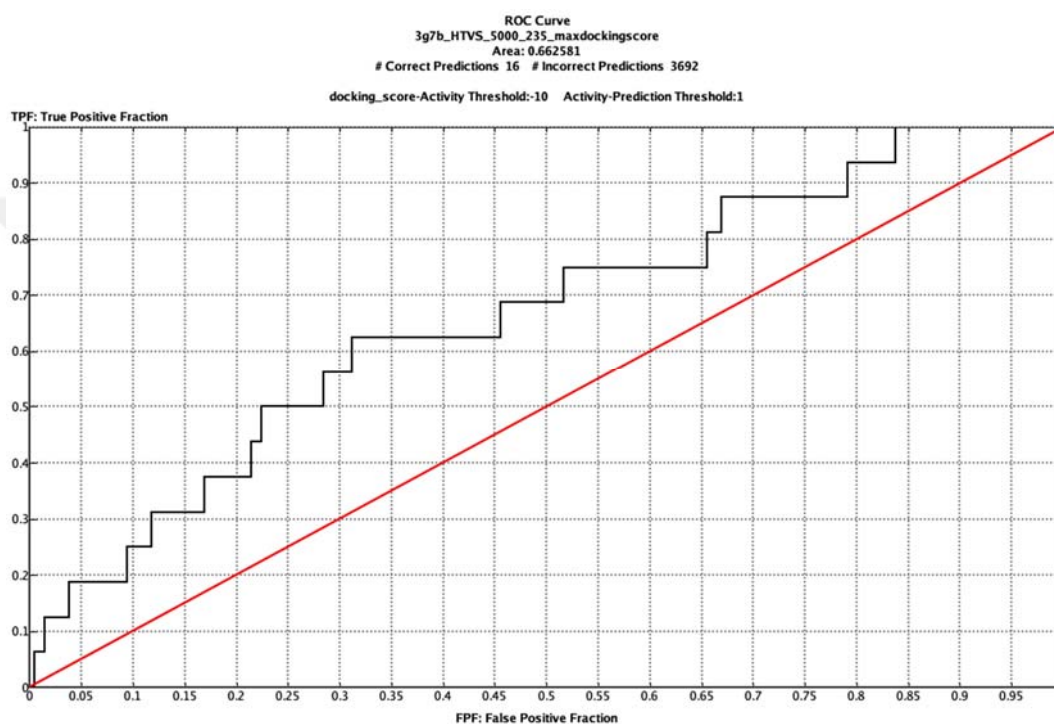


Figure 14: 3G7B docked with water 235, selected and ranked according to maximum docking score

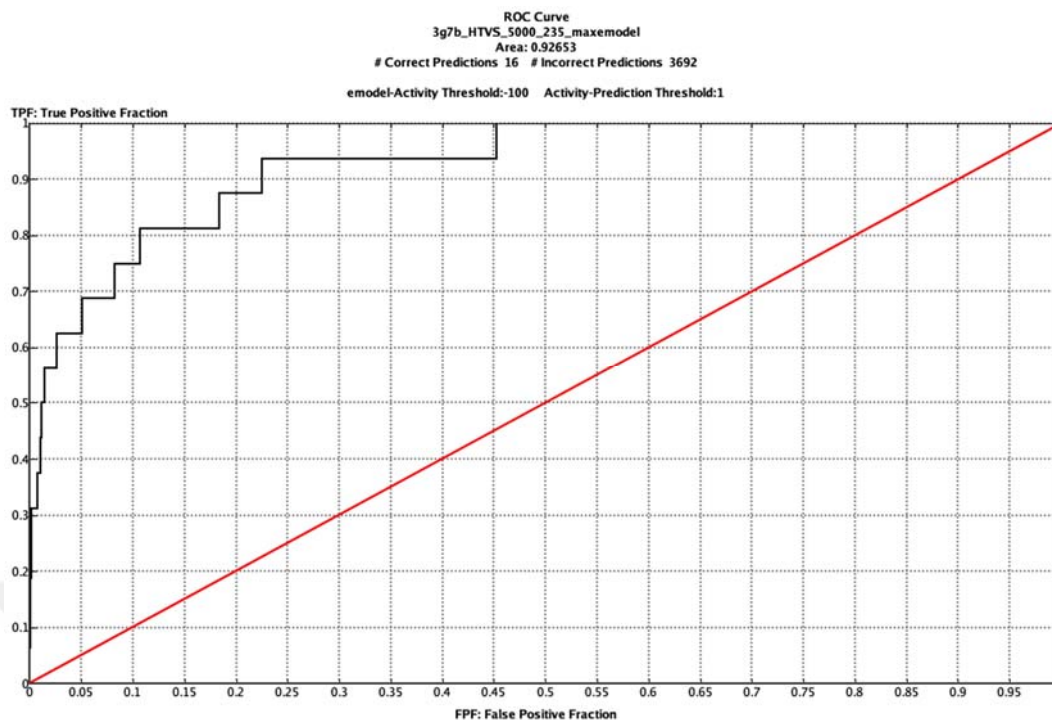


Figure 16: 3G7B docked with water 235, selected and ranked according to maximum e-model score

With HOH 235 and 263 (with restrictions to make bond either with HOH 235 and /or HOH 263 or define the position of ligand considering the position of HOH 235 and 263) GRID files, when the poses are selected for maximum docking and e-model scores and moved to ROC curves, they found out to be 0,6580 and 0,9520 respectively. According to the results, the process for e-model scores separate the true positives 95,2 percent successfully.

These finding oblige us to continue with HOH 235 and 263 GRID file for further test set experiments.

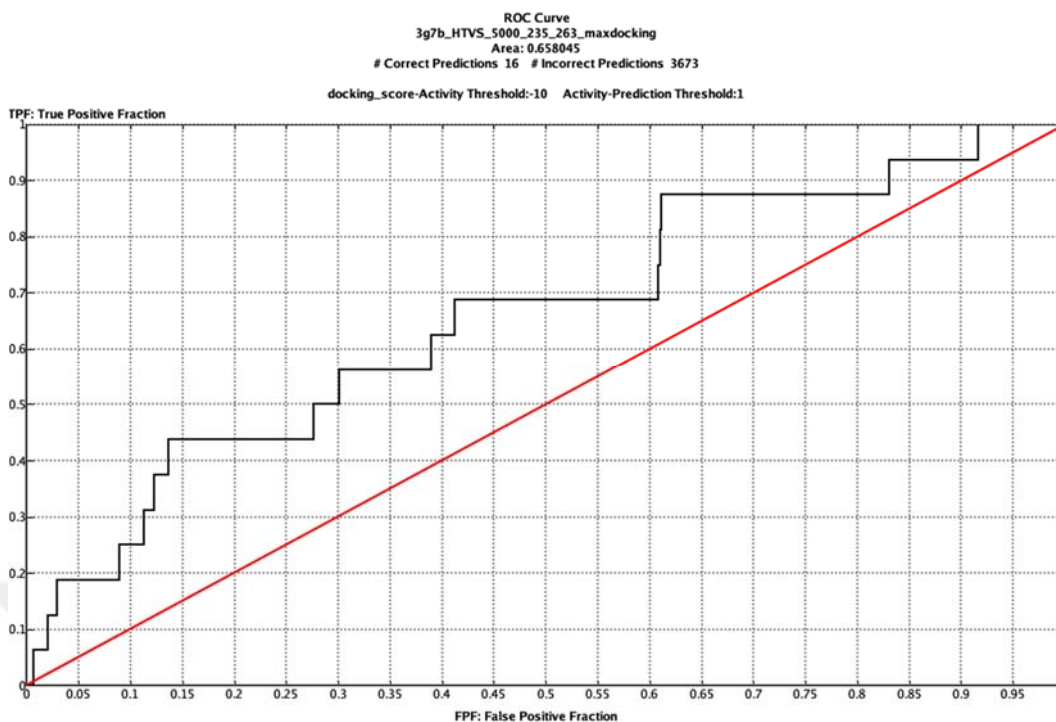


Figure 15: 3G7B docked with waters 235 and 263, selected and ranked according to maximum docking score

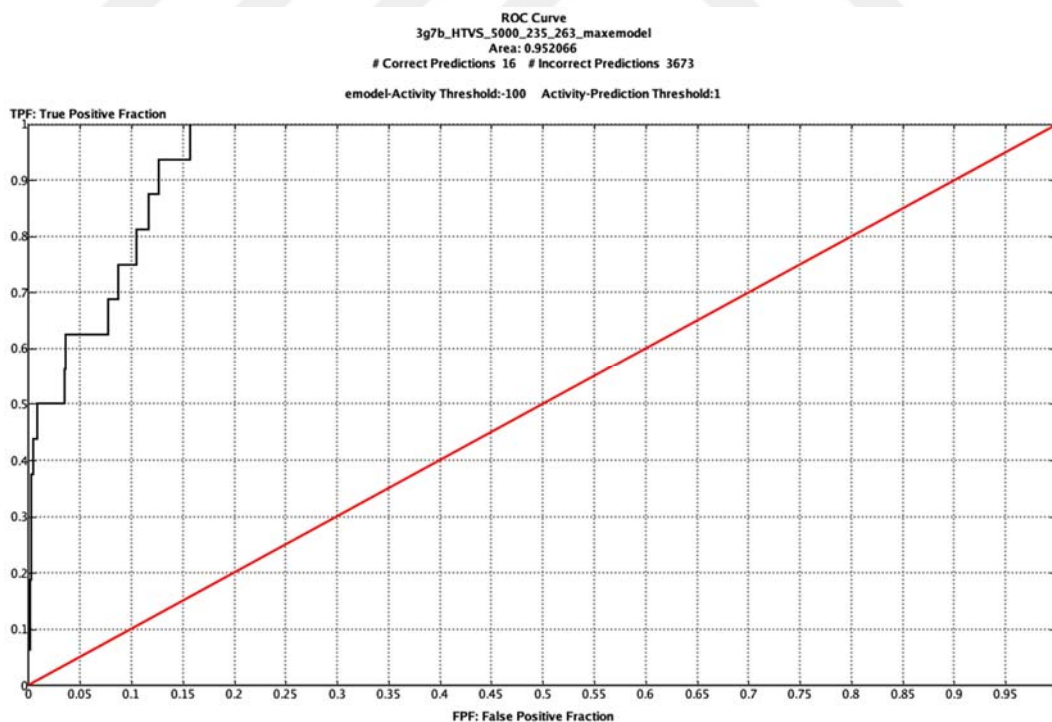


Figure 18: 3G7B docked with waters 235 and 263, selected and ranked according to maximum e-model score

The ROC curves belong to HTVS docking of *E. coli* Gyrase B (PDB ID: 3G7E) trial set without waters, with HOH 408, with HOH 408 and HOH 443 are given in Figures 19-24

With the without waters (without restrictions/ place the compounds considering original ligand as centroid) GRID file, when the poses are selected, ranked for maximum docking and e-model scores, then moved to ROC curve, the area under curves found 0.49 and 0.64. This means that, the process for e-model separates the true positives from false positive 64.85 percent ratio.

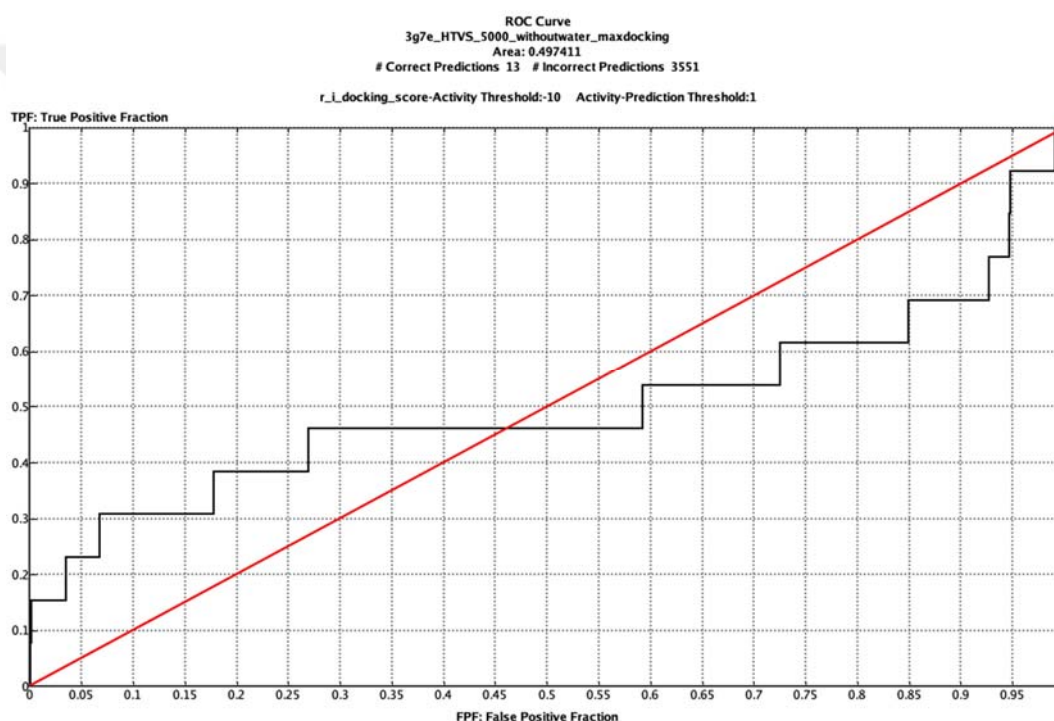


Figure 19: 3G7E docked without waters, selected and ranked according to maximum docking score

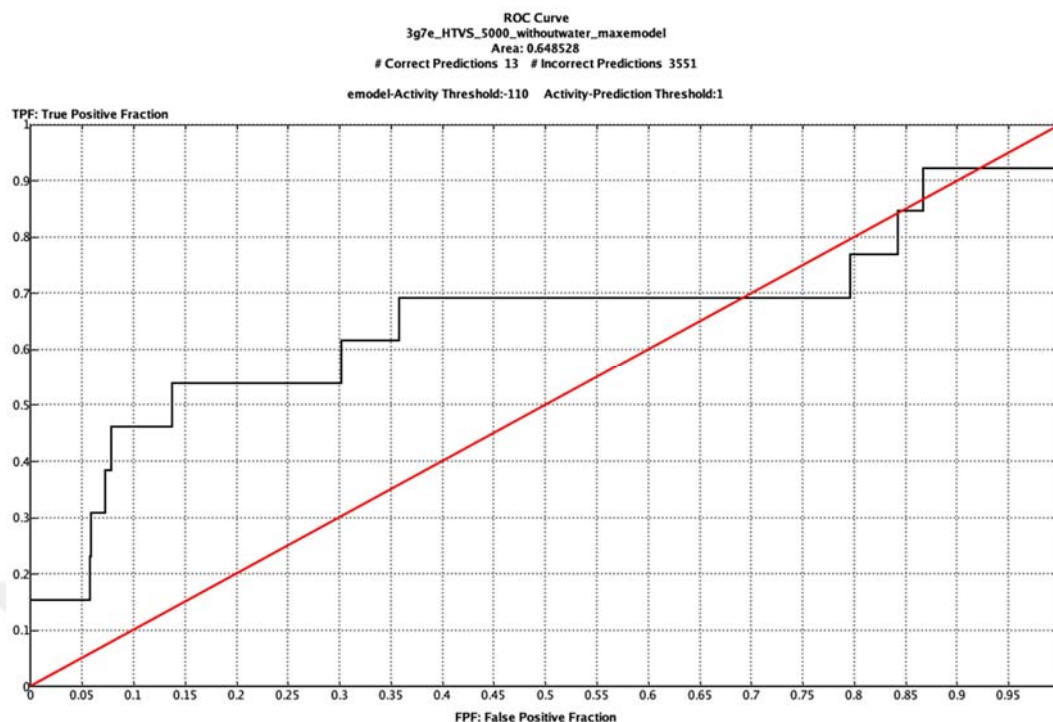


Figure 160: 3G7E docked without waters, selected and ranked according to maximum e-model score

HTVS dockings with HOH 408 (with restrictions to make bond either with HOH 408 or define the position of ligand considering the position of HOH 408) or with HOH 408 and 443 (with restrictions to make bond either with HOH 408 and /or HOH 443 or define the position of ligand considering the position of HOH 408 and 443) gave similar results.

During evaluation with HOH 408 the areas beneath curves were 0,62 and 0,59. According to the result, the process for docking score separated the true positives 62,57 percent ratio. On the other hand ROC curves with HOH 408 and 443 GRID files, areas for maximum docking and e-model scores were 0,39 and 0,57. These results suggested the process for e-model separates the true positives 57,02 percent effectively.

Conflicting results from without water and with HOH 408 GRID files lead us to perform a replicate experiment with test set.

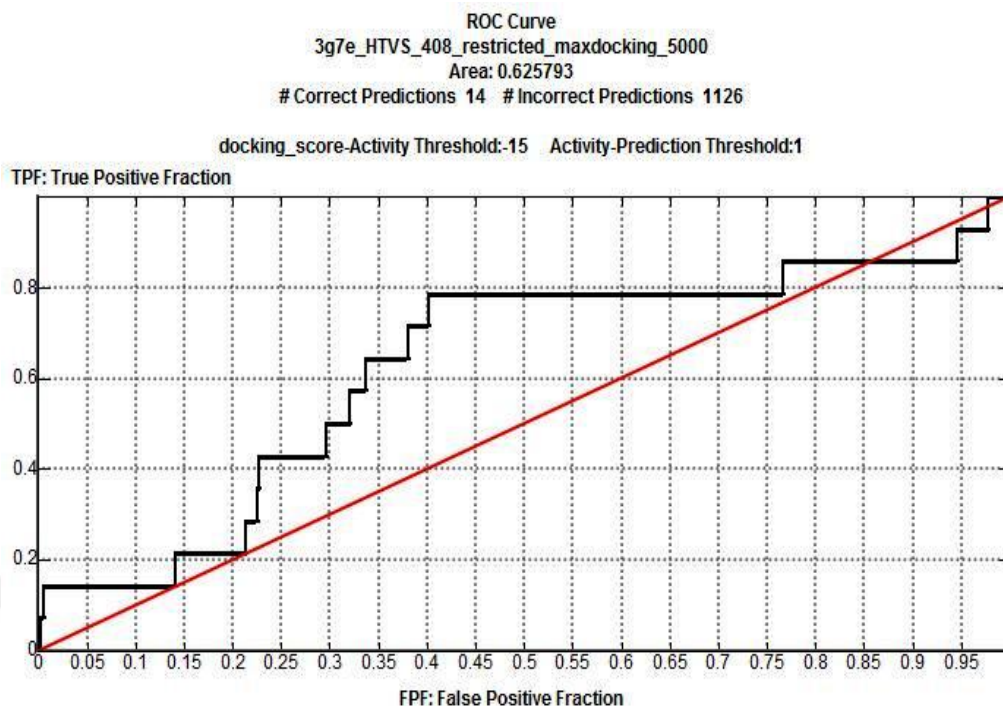


Figure 171: 3G7E docked with water 408, selected and ranked according to maximum docking score

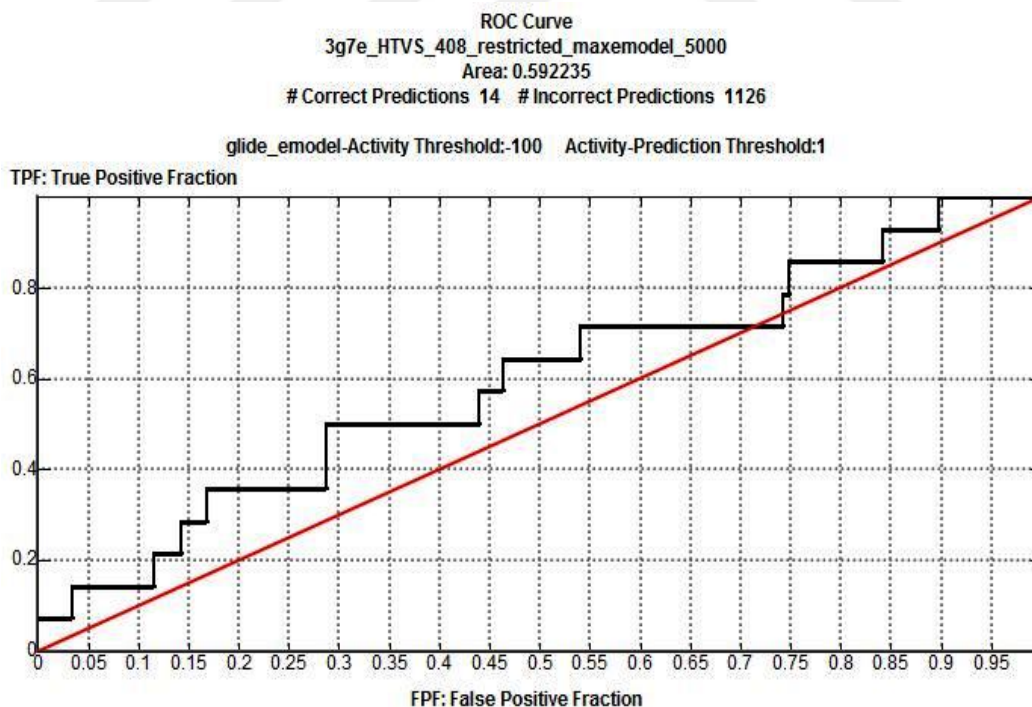


Figure 22: 3G7E docked with water 408, selected and ranked according to maximum e-model score

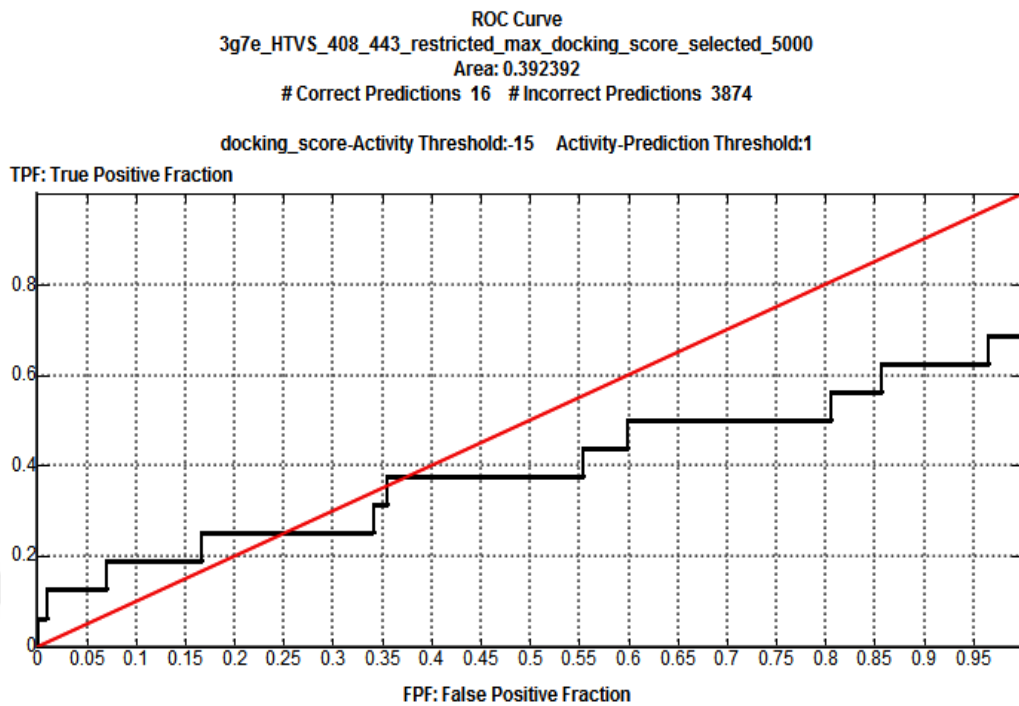


Figure 23: 3G7E docked with waters 408 and 443, selected and ranked according to maximum docking score

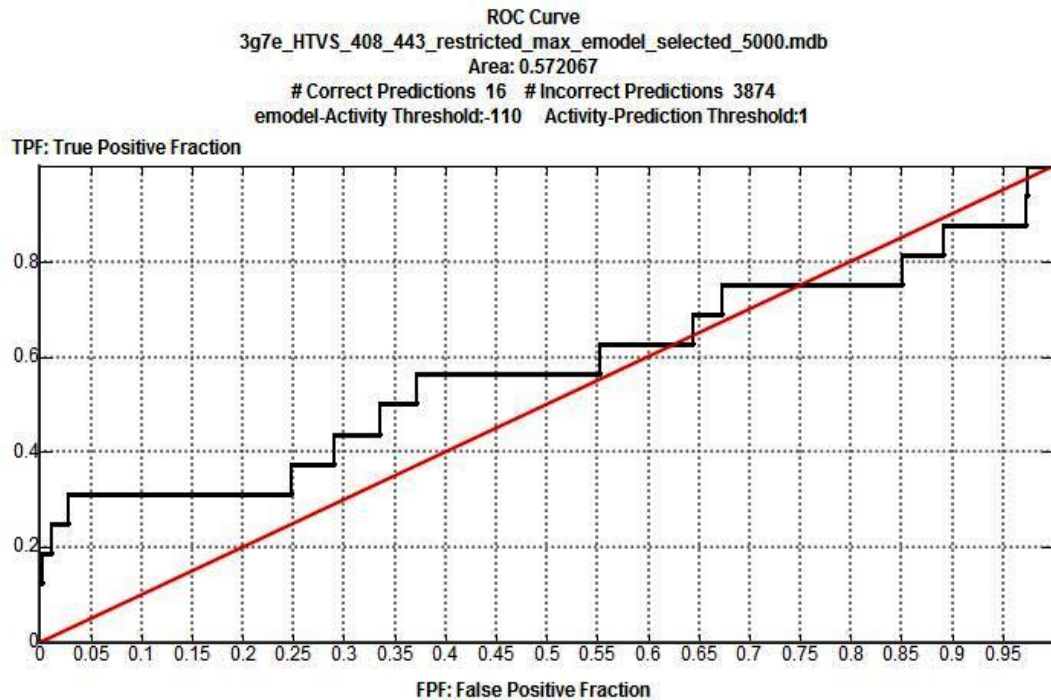


Figure 24: 3G7E docked with waters 408 and 443, selected and ranked according to maximum e-model score

4.2. TEST SET DOCKING, SCORING AND ROC CURVES EVALUATION

In Test set evaluations for *S. aureus* Gyrase B (PDB ID: 3G7B), GRID file with waters 235 and 263 (with restrictions to make bond either with HOH 235 and /or HOH 263 or define the position of ligand considering the position of HOH 235 and 263) was used. For decision to choose the type of scoring function for further experiments, ROC curves were produced both with rankings according to HTVS maximum docking and e-model scores. (Figures 25 and 33)

The areas under ROC curves for docking and e-model scoring functions were 0.58 and 0.84 therefore SP and XP evaluations were made only based on e-model scoring functions with a cut-off of maximum scoring 20000 compounds.

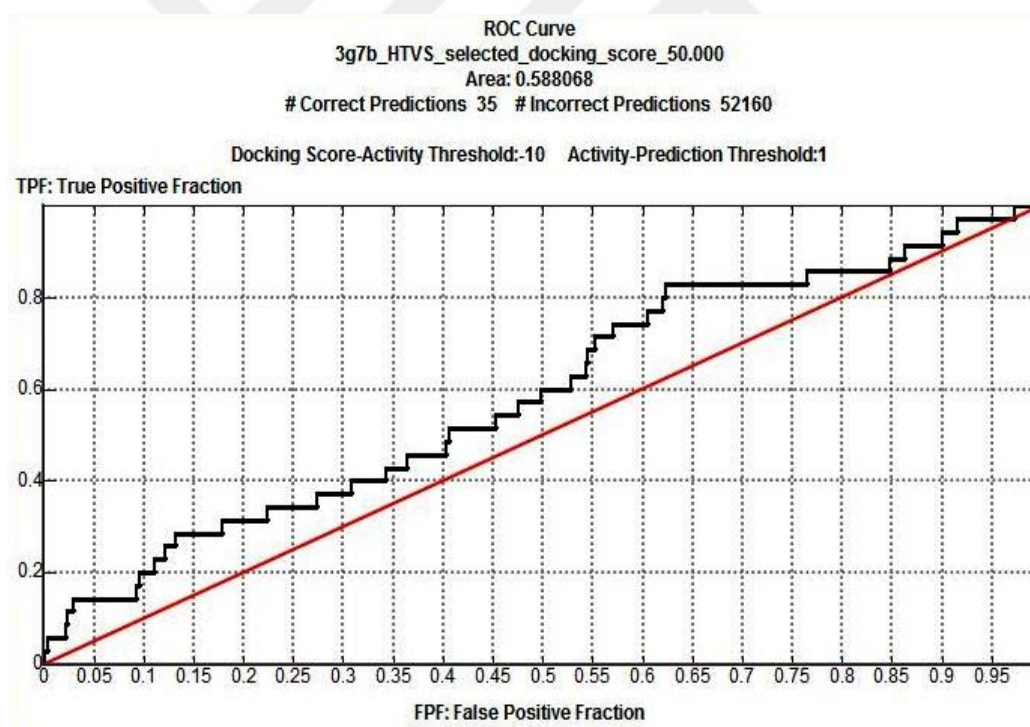


Figure 25: 3G7B docked with waters 235 and 263 in HTVS algorithm, selected and ranked according to maximum docking score

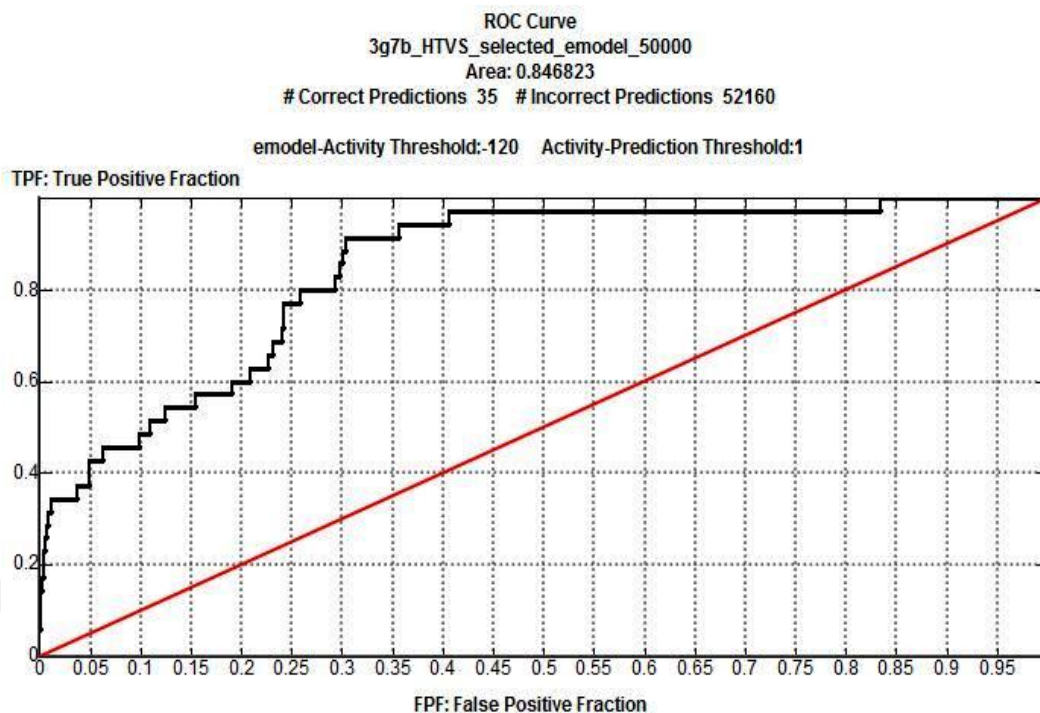


Figure 26: 3G7B docked with waters 235 and 263 in HTVS algorithm, selected and ranked according to maximum e-model score

In SP docking of test set for *S. aureus* Gyrase B (PDB ID: 3G7B) when the poses are selected for e-model scores and moved to ROC curve, the process for e-model scores separate the true positives from false ones with 78,05 percent ratio (Figure 27).

When the compounds rescored with the same GRID file with XP algorithm, the area under curve drop down to 77,73 percent (Figure 28).

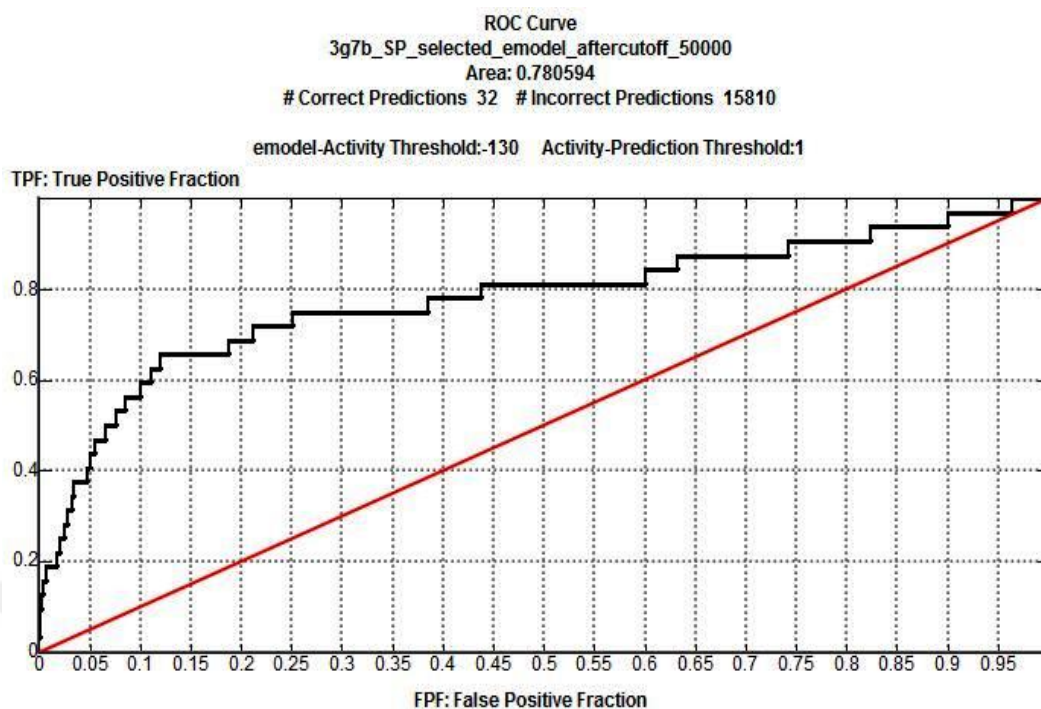


Figure 27: 3G7B docked with waters 235 and 263 in SP algorithm after a cut off of 20000 compounds, selected and ranked according to maximum e-model score

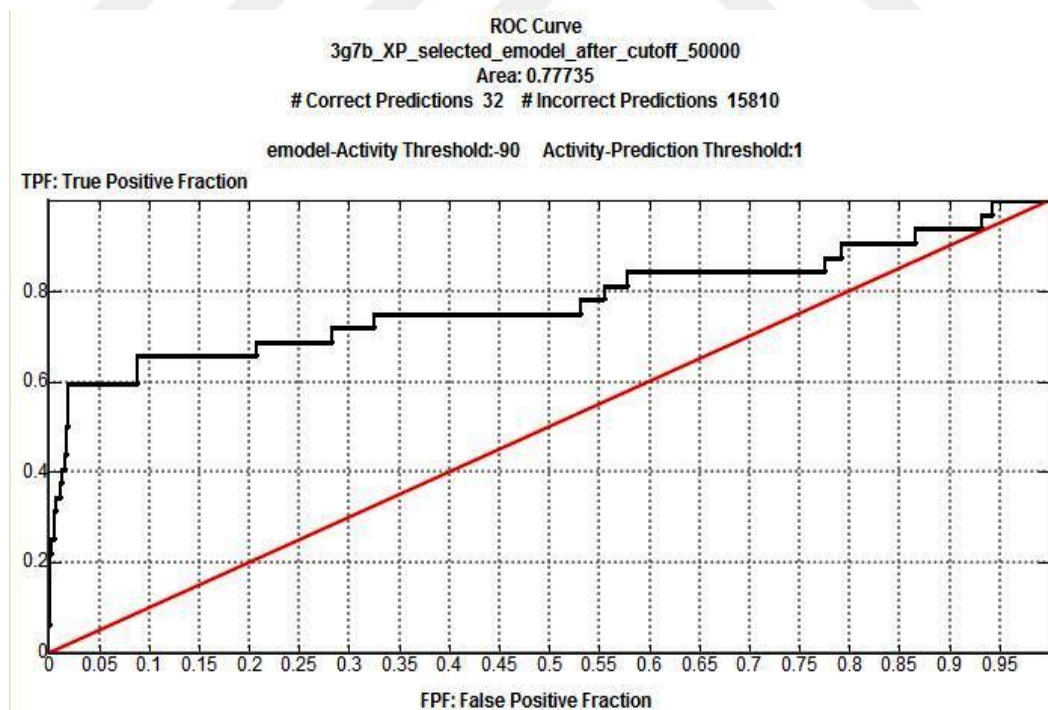


Figure 188: 3G7B docked with waters 235 and 263 in XP algorithm after a cut off of 20000 compounds, selected and ranked according to maximum e-model score

As contradictory results were produced from the trial set belong to *E. coli* Gyrase B (PDB ID: 3G7E), we run HTVS experiment of test set both with without waters (without restrictions/ place the compounds considering original ligand as centroid) and with HOH 408 (with restrictions to make bond either with HOH 408 or define the position of ligand considering the position of HOH 408) GRID files to see if the results were affected from the random selected compounds of the trial set. The results were then evaluated for both scoring algorithms for judgement.

The percent ratio of true positives from false positives using area under curves without waters and with HOH 408 for docking scores found 44,89% and 53,27% respectively. Same percents for e-model scores with without waters and HOH 408 for docking scores initiated 55,23 and 64.07 % consequently.

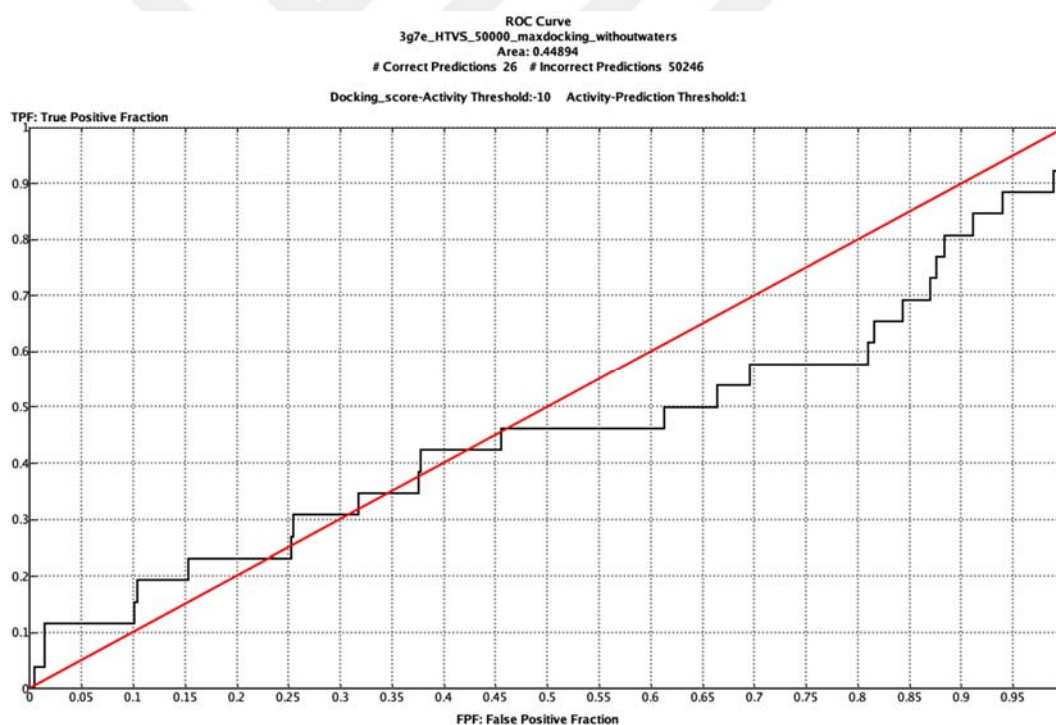


Figure 2919: 3G7E docked without waters, selected and ranked according to maximum docking score

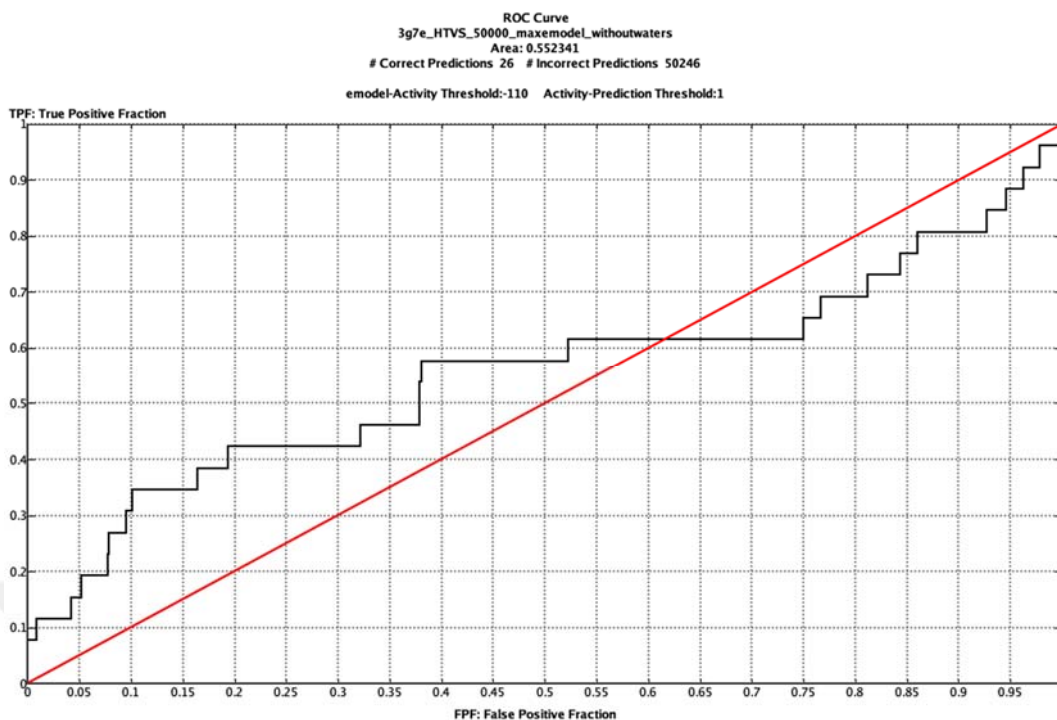


Figure 200: 3G7E docked without waters, selected and ranked according to maximum e-model score

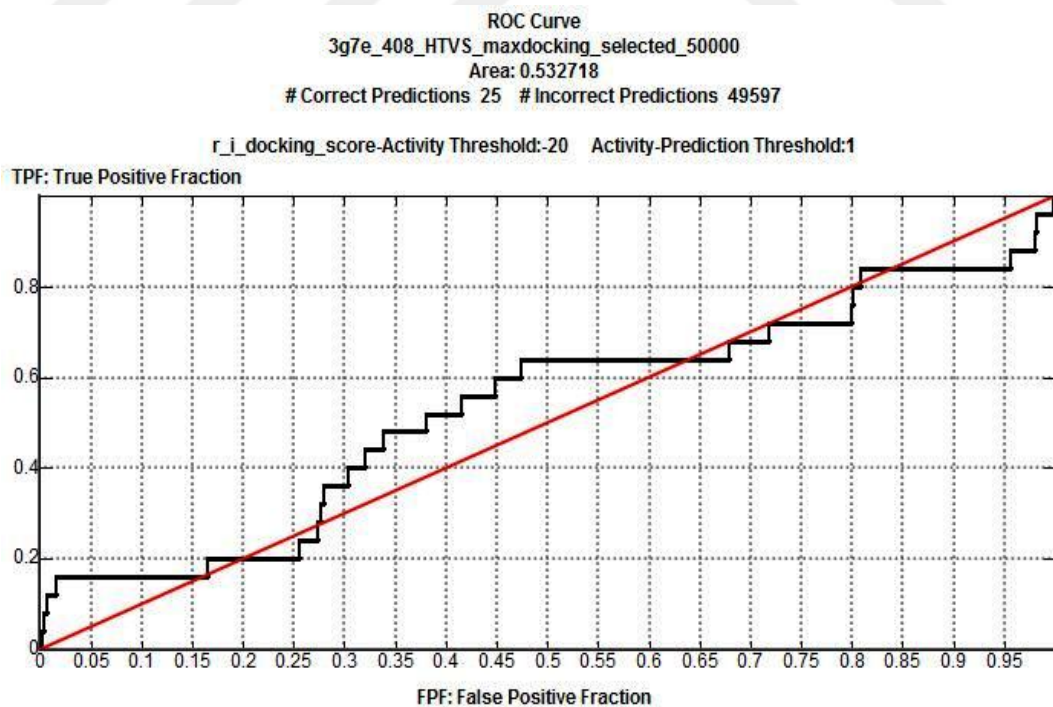


Figure 211: 3G7E docked with water 408, selected and ranked according to maximum docking score

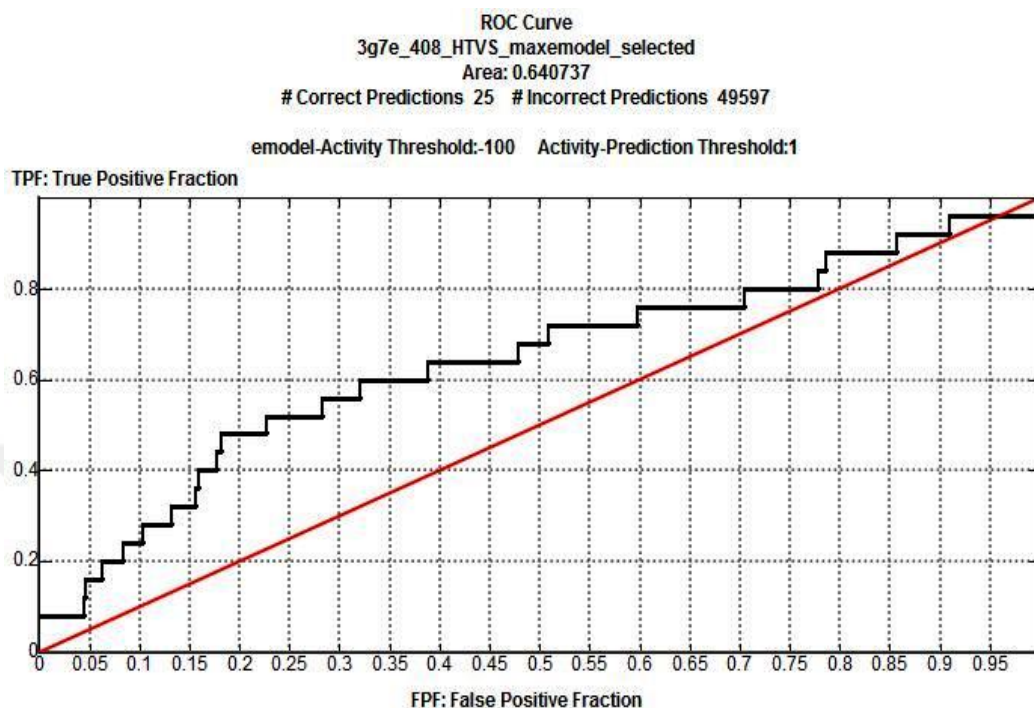


Figure 32: 3G7E docked with water 408, selected and ranked according to maximum e-model score

The results belong to HOH 408 GRID file scoring with e-model were consistent with the previous experiments and algorithm logic, further SP and re-scoring XP experiments performed with the same conditions after a cut-off of 20000 compounds for time and computational concerns.

During evaluation with HOH 408 the areas under curves with SP and XP algorithms, ranking according to e-model scores were 0,68 and 0,63. According to rescoring XP results, the process for e-model score separated the true positives with 63,19 percent ratio.

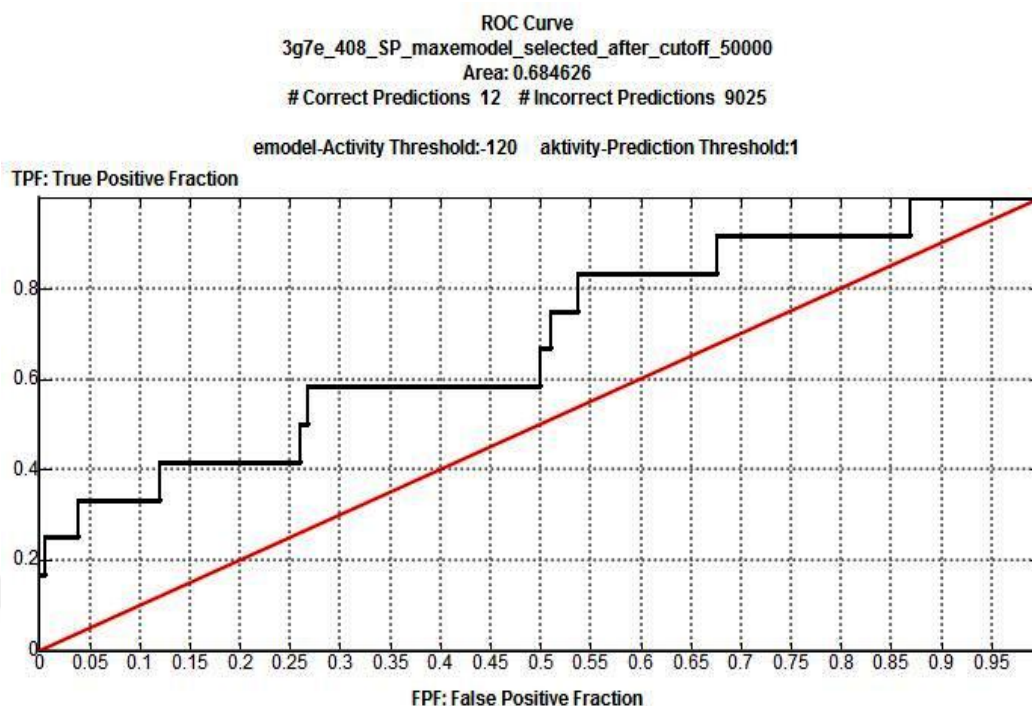


Figure 33: 3G7E docked with water 408 in SP algorithm after a cut off of 20000 compounds, selected and ranked according to maximum e-model score

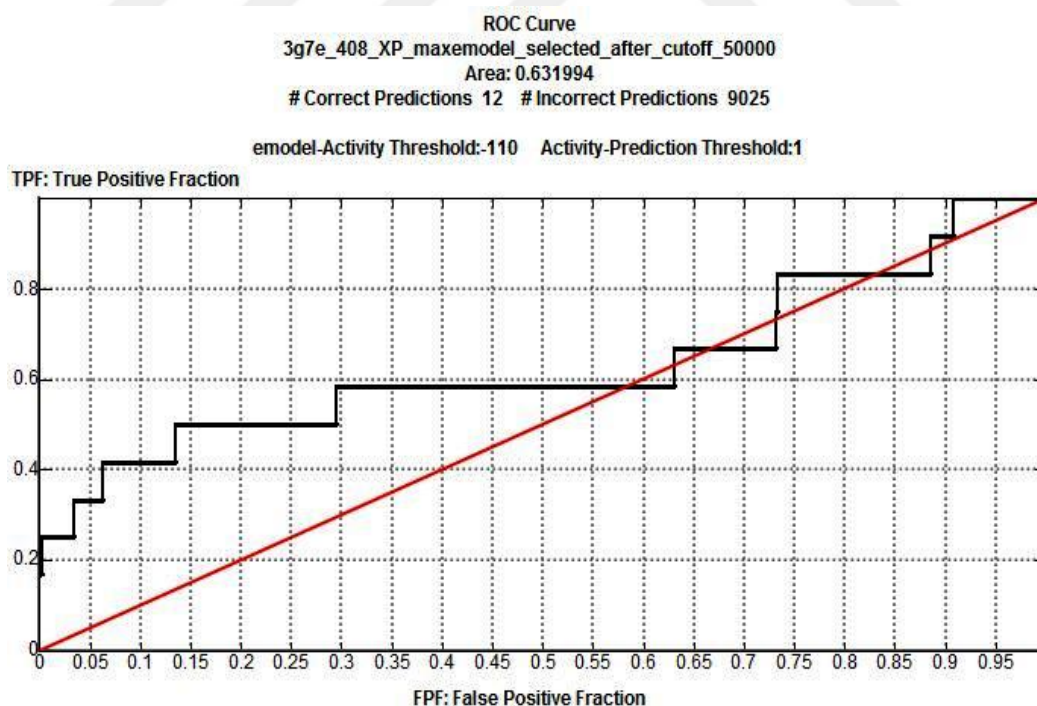


Figure 34: 3G7E docked with water 408 in XP algorithm after a cut off of 20000 compounds, selected and ranked according to maximum e-model score

Combined ROC area under curve tables elucidated from of all docking with stated GRID files are given in table 2

3G7B Trial Set		
	Max Docking Score	Max E Model Score
HTVS without water	0,6268	0,9388
HTVS with water 235	0,6625	0,9265
HTVS with waters 235 and 263	0,658	0,952
3G7E Trial Set		
	Max Docking Score	Max E Model Score
HTVS without water	0,4974	0,6585
HTVS with water 408	0,6257	0,5922
HTVS with waters 408 and 443	0,3923	0,572
3G7B Test Set with waters 235 and 263		
	Max Docking Score	Max E Model Score
HTVS	0,588	0,8468
SP		0,7805
XP		0,7773
3G7E Test Set		
	Max Docking Score	Max E Model Score
HTVS without water	0,4489	0,5523
HTVS with water 408	0,5327	0,6407
SP with water 408		0,6846
XP with water 408		0,6319

Table 2: Combined areas under ROC curves of applied GRID files belonging to docking experiments.

Highest scoring, purchased compounds belonging to *S. aureus* and *E. coli* test sets, selected with XP algorithms using with HOH 235 and 263 and with HOH 408 GRID files for further supercoiling assays are listed in decending orders in tables 3 and 4.

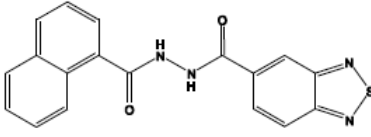
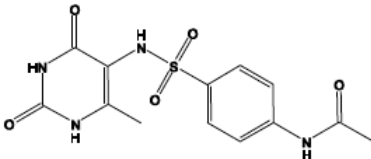
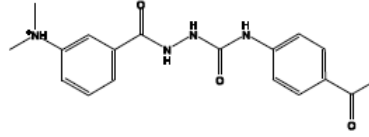
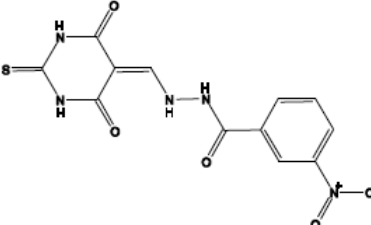
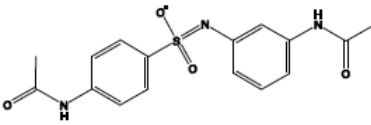
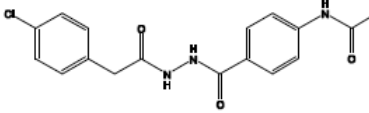
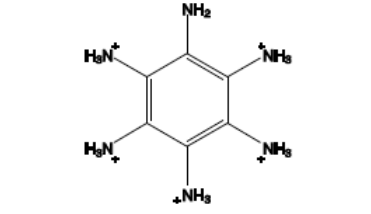
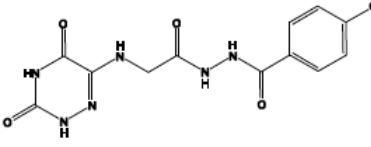
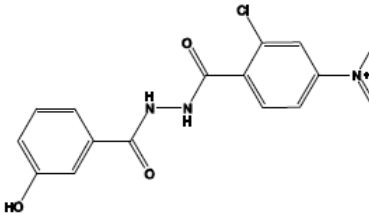
<p>1 ZINC00155744</p> 	<p>2 ZINC00368254</p> 	<p>3 ZINC00610552</p> 
<p>4 ZINC03067000</p> 	<p>5 ZINC00029706</p> 	<p>6 ZINC00441071</p> 
<p>7 ZINC05286123</p> 	<p>8 ZINC00055814</p> 	<p>9 ZINC00054801</p> 

Table 3: Compounds which received highest e-model score during waters 235 and 263 included XP docking to 3G7B and selected for “Gel Based Supercoiling Assay”

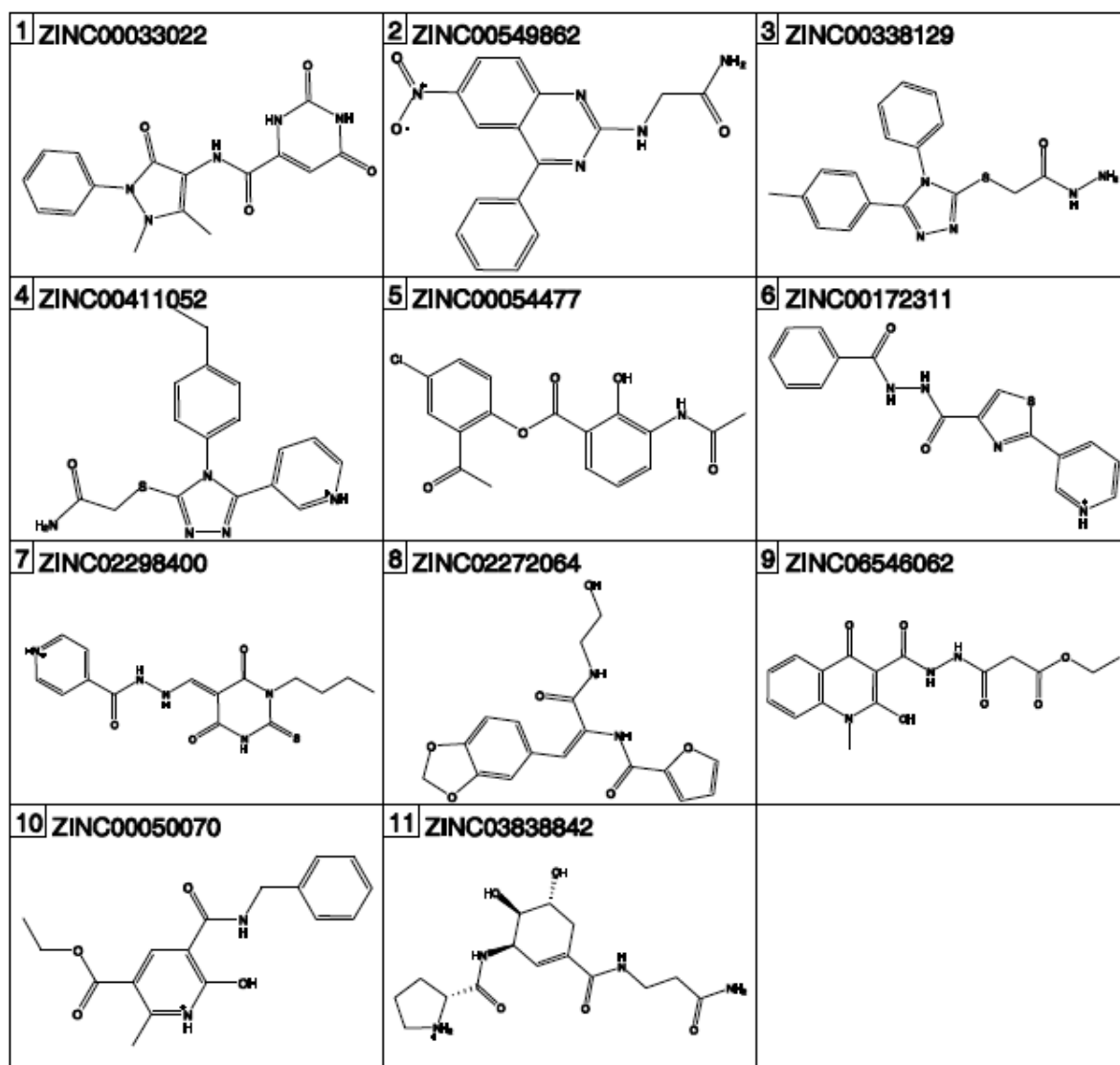


Table 4: Compounds which received highest e-model score during water 408 included XP docking to 3G7E and selected for “Gel Based Supercoiling Assay”

4.3. BIOLOGICAL ACTIVITY

After the screening, 11 compounds from *E. coli* test set, 9 compounds from *S. aureus* total 20 having maximum e-model scored custom made compounds, tested against standard novabiocin by using *E. coli* and *S. aureus* Gyrase Supercoiling Assay Kits (Inspiralis). All compounds applied in a range of 1 mg/20 μ L, 1 mg/50 μ L ve 1 mg/100 μ L (w/v) dilutions to gel electrophoresis according to instructions of

manufacturer. Post run staining was completed by using ethidium bromide solution. All results were combined from figure 35 to 40. First 9 compounds (*S. aureus* 1-9) belong to *S. aureus* test set where remaining (*E. coli* 1-11) belong to *E. coli*. All numbering and structures are coherent with the tables 3-5 and also harmonious with descending order maximum e-model scores.

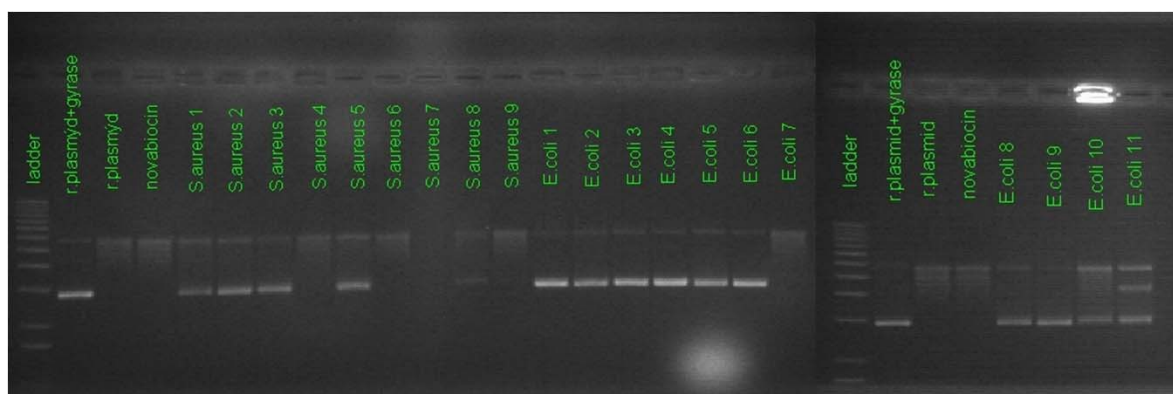


Figure 22: 1 mg/20 μ L (w/v) dilution *E. coli* DNA Gyrase gel electrophoresis results, super-coiled (r. plasmid+gyrase), relaxed (r. plasmid), novabiocin (r. plasmid+gyrase+novabiocin)

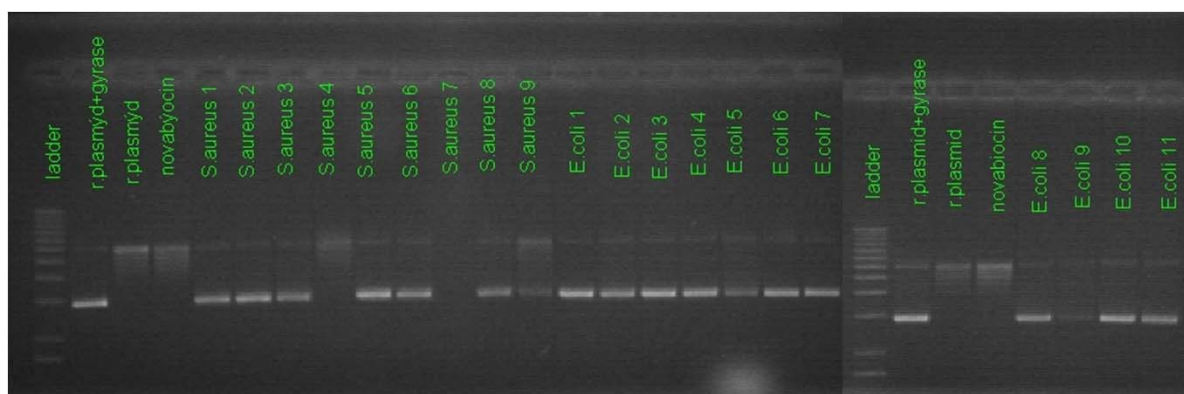


Figure 23: 1 mg/50 μ L (w/v) dilution *E. coli* DNA Gyrase gel electrophoresis results, super-coiled (r. plasmid+gyrase), relaxed (r. plasmid), novabiocin (r. plasmid+gyrase+novabiocin)

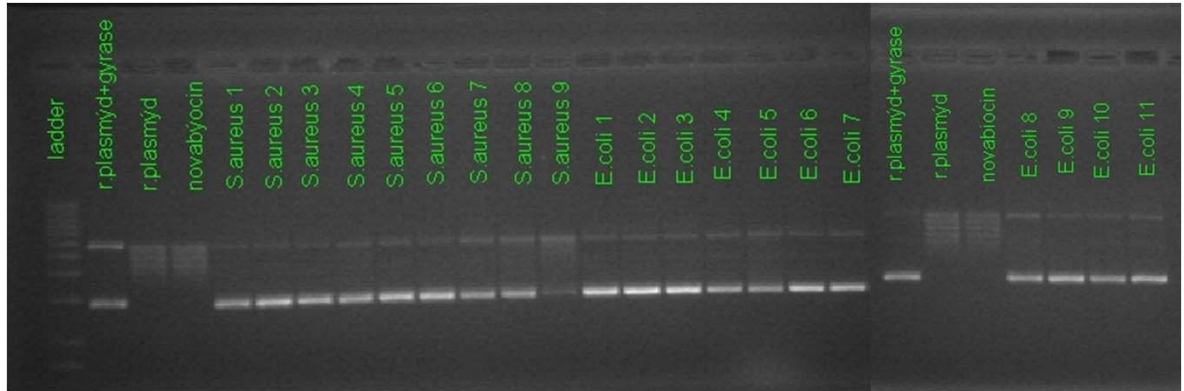


Figure 24: 1 mg/100 μ L (w/v) dilution *E. coli* DNA Gyrase gel electrophoresis results, super-coiled (r. plasmid+gyrase), relaxed (r. plasmid), novabycin (r. plasmid+gyrase+novabycin)

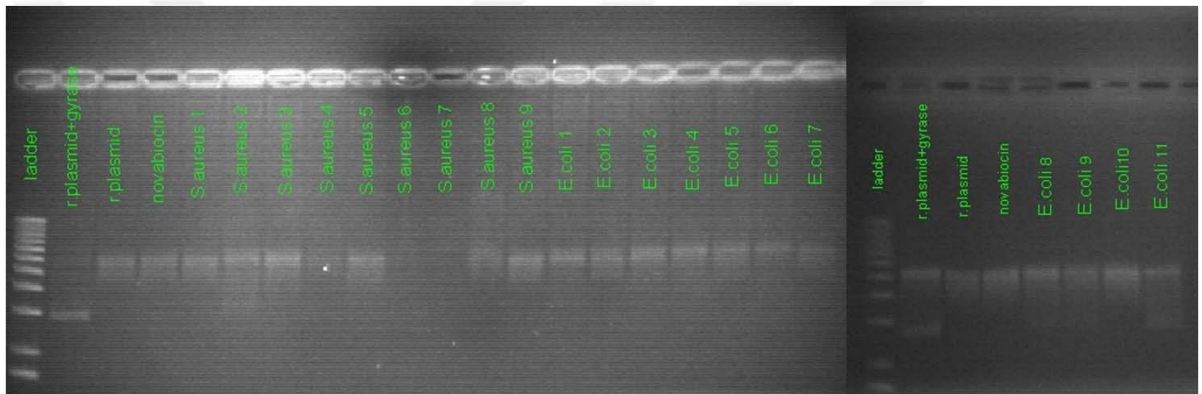


Figure 258: 1 mg/20 μ L (w/v) dilution *S. aureus* DNA Gyrase gel electrophoresis results, super-coiled (r. plasmid+gyrase), relaxed (r. plasmid), novabycin (r. plasmid+gyrase+novabycin)

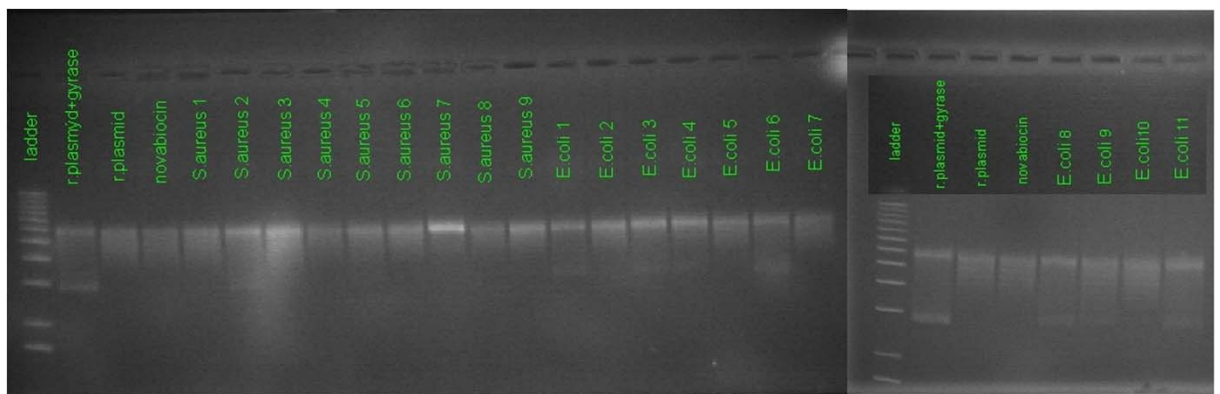


Figure 26: 1 mg/50 μ L (w/v) dilution *S. aureus* DNA Gyrase gel electrophoresis results, super-coiled (r. plasmid+gyrase), relaxed (r. plasmid), novabycin (r. plasmid+gyrase+novabycin)

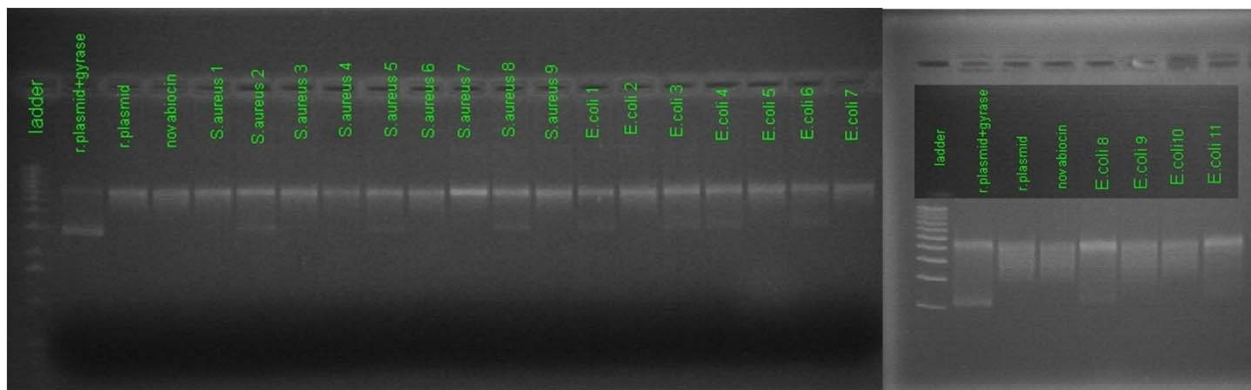


Figure 40: 1 mg/100 μ L (w/v) dilution *S. aureus* DNA Gyrase gel electrophoresis results, super-coiled (r. plasmid+gyrase), relaxed (r. plasmid), novabiocin (r. plasmid+gyrase+novabiocin)

Table 5 shows a detailed cross chart of activities for compounds in *E. coli* and *S. aureus* DNA Gyrase supercoiling assays.

Compound No	Zinc Code	Max e-model Score	<i>E. coli</i> Supercoiling Assay			<i>S. aureus</i> Supercoiling Assay		
			1/20	1/50	1/100	1/20	1/50	1/100
<i>S. aureus</i> 1	ZINC00155744	-80675377	-	-	-	+	+	+
2	ZINC00368254	-80016497	-	-	-	-	-	-
3	ZINC00610552	-79495529	-	-	-	-	-	-
4	ZINC03067000	-78063271	+	+	-	+	+	+
5	ZINC00029706	-77099167	-	-	-	+	+	-
6	ZINC00441071	-76471031	+	-	-	+	+	+
7	ZINC05286123	-75941658	+	+	-	+	+	+
8	ZINC00055814	-75807533	+	-	-	+	+	-
9	ZINC00054801	-75307159	+	-	-	+	+	+
<i>E. coli</i> 1	ZINC00033022	-96186478	-	-	-	-	-	-
2	ZINC00549862	-95777328	-	-	-	+	+	+
3	ZINC00338129	-94271538	-	-	-	-	-	-
4	ZINC00411052	-91462143	-	-	-	-	-	-
5	ZINC00054477	-90465988	-	-	-	+	+	+
6	ZINC00172311	-90200745	-	-	-	+	-	-
7	ZINC02298400	-88609749	+	-	-	+	+	+
8	ZINC02272064	-88418015	-	-	-	-	-	-
9	ZINC06546062	-88114059	-	-	-	+	-	-
10	ZINC00050070	-87938751	-	-	-	+	+	+
11	ZINC03838842	-87907578	-	-	-	-	-	-

Table 5: Activities zinc codes and maximum e-model scores of selected compounds during *E. coli* and *S. aureus* DNA Gyrase supercoiling assays in different concentrations.

5. DISCUSSION AND CONCLUSION

Enriched trial and test set of compounds randomly selected from 1.442.716 containing ZINC data base “clean-leads-subset (# 11)” were prepared, docked and their scores were evaluated by ROC curves for selective *E. coli*/*S. aureus* DNA Gyrase B site activities.

Ligand-macromolecule interactions of both *E. coli* (Pdb ID: 3G7E) and *S. aureus* (Pdb ID: 3G7B) DNA Gyrase B with their inhibitors were inspected visually and evaluated by “Ligand-receptor contacts (visualization+scoring) SVL” of the MOE software and apparently noticed that the water molecules, HOH 235, HOH 263 in *S. aureus* Gyrase B (PDB ID: 3G7B) and HOH 408 and 443 in *E. coli* (Pdb ID: 3G7E) have significant bridging and binding roles.

Henceforth, we decided to produce unrestricted (without water) and restricted GRID files which can dock the ligands in a manner making bond either with these waters or define the position of ligand considering the position of the waters. Moreover, we docked trial set with HTVS of Maestro’s Glide docking algorithm not only with unrestricted ones but also restricted GRID’s and evaluate the results with ROC curves of two different scoring functions to make a judgement whether these water molecules play a role or not during interaction.

In all of the HTVS experiments run according to these files, e-model scores separate true positive from false ones more successfully when compared to docking scores except *E. coli*’s with HOH 408 GRID file where docking score had an area under curve of 0,6257.

Besides, *E. coli*’s unrestricted file receives better scores than with HOH 408 and with HOH 408 and 443 files in general. This is obviously interesting as e-model score algorithm designed such in a manner for reflecting the solvation parameters perfectly to docking score by calculating the positions of the incidental waters.

Also there is an inconsistency of having these scores as pyrazol ring system belong to original ligand “prop-2-yn-1-yl{[5-(4-piperidin-1-yl-2-pyridin-3-yl-1,3-thiazol-5-yl)-1H-pyrazol-3-yl]methyl} carbamate” (1) synchronized with Ile 78, Gly 77, Thr 165, Asp 73 and HOH 408.

For fear that the random picking of compounds in trial set might be the reason of this unpredicted situation, we repeated the HTVS dockings of the test set with previously higher scoring unrestricted and with HOH 408 GRID files. When the number of subjects increased from 5000 to 50000 the inconsistency disappeared and e-model score belong to with HOH 408 GRID file received an area under curve of 0,6407. In conclusion we decided to make further SP and XP simulations for *E. coli* by using this file.

In trial set HTVS experiments belong to *S. aureus* (PDB ID: 3G7B), with HOH 235 and 263 received the highest area under curve result of 0,952 in e-model score ROC curve evaluation which means that ligands that bind to these site have to make a H-bond with HOH 235 and must bind a space restricted between these waters. All results belonging to this structure were very consistent with a an average 30% difference between the docking score and e-model score which means the ROC curves separate the true positives in very high frequency if the e-model scores were taken as a base. Restricted GRID file with HOH 235 and 263 was used in further evaluations for *S. aureus* DNA gyrase ATPase site.

In the period of SP and evaluate score in place by XP mode docking, a test set of 50000 compounds were assessed for at least 10 poses in both active sites. Rescoring in place was preferred to reduce computational load and time.

With the chosen files, gained from XP e-model descending score listing, ROC area under curve results were 0.6319 and 0.7773 for *E. coli* and *S. aureus* active sites respectively.

During these processes, although ROC area under curves decreased from HTVS to XP algorithms gradually, combination of top scoring twenty compounds in the content of each set were selected for in-vitro *E. coli* and *S. aureus* DNA Gyrase gel based supercoiling assay against general standard novabiocin.

The compounds and standard were prepared and tested between 1/20-1/100 mg/ μ l concentrations.

Eleven compounds coming from the test set of *E. coli* except compound 7 in 1/20mg/ μ l dilution showed no activity over *E. coli* DNA gyrase supercoiling whereas compound 2,5,7 and 10 showed promising inhibiting activity in *S. aureus* DNA gyrase supercoiling assay.

In contrast, except for compound 2 and 3 of *S. aureus* test set all the compounds in all concentrations showed remarkable activity in *S. aureus* DNA gyrase supercoiling assay where compound 4 and 7 showed a moderate one in *E. coli* DNA gyrase supercoiling.

Besides, highest scoring compound 1 and 5 from *S. aureus* test set, *E. coli* test set compounds 2, 4, and 10 selectively inhibited *S. aureus* DNA Gyrase supercoiling without effecting *E. coli*.

Clearly, *invitro* results are interesting showing that the ROC curve evaluation sincerely efficient and consistant for either separating actives from inactives or deciding the type of interaction. In addition, the evaluation results of at least 90% or more have to be processed for further testing.

On the other hand, it is inevitable that water molecules not also have an important impact on positioning but have contribution to selectivity of ligands over inhibiting DNA Gyrase ATPase binding site.

During the process of this study, finding these selective ligands not only will direct us for extensive reviewing of binding modes with other waters but also force us to plan synthesizing more efective derivatives using other facilating computational tools such as molecular interaction field (MIF) for our future studies.



REFERENCES

1. Ryan KJ, Ray CG, Sherris JC. Sherris medical microbiology. 5th ed. New York: McGraw Hill Medical; 2010. xiv, 1026 p. p.
2. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. *Nucleic Acids Research*. 2000;28(1):235-42.
3. Kayser FH. Medical microbiology. Stuttgart ; New York, NY: Georg Thieme Verlag; 2005. xxvi, 698 p. p.
4. Tomioka H, Namba K. [Development of antituberculous drugs: current status and future prospects]. *Kekkaku* : [Tuberculosis]. 2006;81(12):753-74. Epub 2007/01/24.
5. Lewis RJ, Tsai FT, Wigley DB. Molecular mechanisms of drug inhibition of DNA gyrase. *BioEssays* : news and reviews in molecular, cellular and developmental biology. 1996;18(8):661-71. Epub 1996/08/01.
6. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schaffer AA, et al. Protein database searches using compositionally adjusted substitution matrices. *The FEBS journal*. 2005;272(20):5101-9. Epub 2005/10/13.
7. Forterre P, Gribaldo S, Gadelle D, Serre M-C. Origin and evolution of DNA topoisomerases. *Biochimie*. 2007;89(4):427-46.
8. Wang JC. DNA topoisomerases. *Annual review of biochemistry*. 1996;65:635-92. Epub 1996/01/01.
9. Nichols MD, DeAngelis K, Keck JL, Berger JM. Structure and function of an archaeal topoisomerase VI subunit with homology to the meiotic recombination factor Spo11. *The EMBO journal*. 1999;18(21):6177-88. Epub 1999/11/02.
10. Champoux JJ. DNA topoisomerases: structure, function, and mechanism. *Annual review of biochemistry*. 2001;70:369-413. Epub 2001/06/08.
11. Ferrero L, Cameron B, Manse B, Lagneaux D, Crouzet J, Famechon A, et al. Cloning and primary structure of *S. aureus* DNA topoisomerase IV: a primary target of fluoroquinolones. *Molecular Microbiology*. 1994;13(4):641-53.
12. Maxwell A. DNA gyrase as a drug target. *Trends in Microbiology*. 1997;5(3):102-9.
13. Agresti JJ, Antipov E, Abate AR, Ahn K, Rowat AC, Baret JC, et al. Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proceedings of the National Academy of Sciences of the United States of America*. 2010;107(9):4004-9. Epub 2010/02/10.
14. Hann MM, Oprea TI. Pursuing the leadlikeness concept in pharmaceutical research. *Current opinion in chemical biology*. 2004;8(3):255-63. Epub 2004/06/09.
15. Jhoti H, Leach AR. Structure-based drug discovery. Dordrecht: Springer; 2007. xii, 249 p. p.
16. Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. *The EMBO journal*. 1986;5(4):823-6. Epub 1986/04/01.
17. Lengauer T, Rarey M. Computational methods for biomolecular docking. *Current opinion in structural biology*. 1996;6(3):402-6. Epub 1996/06/01.
18. Alvarez J, Shoichet B. Virtual screening in drug discovery. Boca Raton: Taylor & Francis; 2005. 470 p. p.

19. Rester U. From virtuality to reality - Virtual screening in lead discovery and lead optimization: a medicinal chemistry perspective. *Current opinion in drug discovery & development*. 2008;11(4):559-68. Epub 2008/07/05.
20. McInnes C. Virtual screening strategies in drug discovery. *Current opinion in chemical biology*. 2007;11(5):494-502. Epub 2007/10/16.
21. Willett P, Barnard JM, Downs GM. Chemical similarity searching. *Journal of Chemical Information & Computer Sciences*. 1998;38(6):983-96.
22. Kroemer RT. Structure-based drug design: docking and scoring. *Current protein & peptide science*. 2007;8(4):312-28. Epub 2007/08/19.
23. Cavasotto CN, Orry AJ. Ligand docking and structure-based virtual screening in drug discovery. *Current topics in medicinal chemistry*. 2007;7(10):1006-14. Epub 2007/05/19.
24. Ostrov DA, Hernandez Prada JA, Corsino PE, Finton KA, Le N, Rowe TC. Discovery of novel DNA gyrase inhibitors by high-throughput virtual screening. *Antimicrobial agents and chemotherapy*. 2007;51(10):3688-98. Epub 2007/08/08.
25. Oblak M, Grdadolnik SG, Kotnik M, Jerala R, Filipic M, Solmajer T. In silico fragment-based discovery of indolin-2-one analogues as potent DNA gyrase inhibitors. *Bioorganic & medicinal chemistry letters*. 2005;15(23):5207-10. Epub 2005/10/06.
26. Roca J, Wang JC. DNA transport by a type II DNA topoisomerase: evidence in favor of a two-gate mechanism. *Cell*. 1994;77(4):609-16. Epub 1994/05/20.
27. Triballeau N, Acher F, Brabet I, Pin JP, Bertrand HO. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *Journal of medicinal chemistry*. 2005;48(7):2534-47. Epub 2005/04/02.
28. Rizzi A, Fioni A. Virtual screening using PLS discriminant analysis and ROC curve approach: an application study on PDE4 inhibitors. *Journal of chemical information and modeling*. 2008;48(8):1686-92. Epub 2008/08/02.
29. Langer T, Hoffmann R, Bryant S, Lesur B. Hit finding: towards 'smarter' approaches. *Current opinion in pharmacology*. 2009;9(5):589-93. Epub 2009/07/07.
30. Jacoby E, Schuffenhauer A, Popov M, Azzaoui K, Havill B, Schopfer U, et al. Key aspects of the Novartis compound collection enhancement project for the compilation of a comprehensive chemogenomics drug discovery screening collection. *Current topics in medicinal chemistry*. 2005;5(4):397-411. Epub 2005/05/17.
31. Waszkowycz B. Towards improving compound selection in structure-based virtual screening. *Drug discovery today*. 2008;13(5-6):219-26. Epub 2008/03/18.
32. Hristozov DP, Oprea TI, Gasteiger J. Virtual screening applications: a study of ligand-based methods and different structure representations in four different scenarios. *Journal of computer-aided molecular design*. 2007;21(10-11):617-40. Epub 2007/11/17.
33. Barnard JM, Downs GM, von Scholley-Pfab A, Brown RD. Use of Markush structure analysis techniques for descriptor generation and clustering of large combinatorial libraries. *Journal of molecular graphics & modelling*. 2000;18(4-5):452-63. Epub 2001/01/06.
34. Fink T, Reymond JL. Virtual exploration of the chemical universe up to 11 atoms of C, N, O, F: assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physicochemical properties, compound classes, and drug discovery. *Journal of chemical information and modeling*. 2007;47(2):342-53. Epub 2007/01/31.

35. Irwin JJ, Shoichet BK. ZINC--a free database of commercially available compounds for virtual screening. *Journal of chemical information and modeling*. 2005;45(1):177-82. Epub 2005/01/26.
36. Ghose AK, Herbertz T, Salvino JM, Mallamo JP. Knowledge-based chemoinformatic approaches to drug discovery. *Drug discovery today*. 2006;11(23-24):1107-14. Epub 2006/11/30.
37. Bernstein FC, Koetzle TF, Williams GJ, Meyer EF, Jr., Brice MD, Rodgers JR, et al. The Protein Data Bank: a computer-based archival file for macromolecular structures. *Journal of molecular biology*. 1977;112(3):535-42. Epub 1977/05/25.
38. Maxwell A, Burton NP, O'Hagan N. High-throughput assays for DNA gyrase and other topoisomerases. *Nucleic Acids Res*. 2006;34(15):e104.



VITA

Gürmen KAYNAR was born in Istanbul in 03.03.1982. He has graduated from Fenerbahce High school, has graduated from Science and Mathematics faculty of Eskisehir Osmangazi University in 2004.

His working career start with Basel Pharmaceuticals in 2006 as a R&D specialist and his main working topics were about searching the patents of active pharmaceutical ingredients (API), carrying on the licencing procedures following finding the appropriate API supplier, working with R&D to be out of scope of the patents and/or finding new API suppliers, providing the critical analysis (XRD, DSC, NMR) of the API and carrying on the correspondences with the API suppliers from abroad, searching national and international patents and trademarks for R&D activities, developing strategies for the lawsuits of Intellectual Property rights, comparative evaluation of the patent and trademark searches, preparing patent and trademark applications and carrying on the concerning procedures. After the resigning from Basel Pharmaceuticals in 2010, he was started for Deva Pharmaceuticals for same working topics in an Executive position.

From 2011 to present time he was working for World Medicine Pharmaceuticals in a supervisor position and his related topics are preparing patent search and evaluation reports, supporting of new patent applications, evaluation of API, searching new API sources, providing technical support in funding's of government, technical support in lawsuits.