THE REPUBLIC OF TURKEY
BAHÇEŞEHİR UNIVERSITY

IDENTIFICATION OF NOVEL HIT MOLECULES AGAINST B-CELL LEUKEMIA/LYMPHOMA-2 (BCL2)

Master's Thesis

GURBET TUTUMLU

İSTANBUL, 2019

THE REPUBLIC OF TURKEY
BAHÇEŞEHİR UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

IDENTIFICATION OF NOVEL HIT MOLECULES AGAINST B-CELL LEUKEMIA/LYMPHOMA-2 (BCL2)

Master's Thesis

GURBET TUTUMLU

SUPERVISOR: PROF. DR. SERDAR DURDAĞI

İSTANBUL, 2019

THE REPUBLIC OF TURKEY
BAHÇEŞEHİR UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES BIOENGINEERING

Name of the thesis: IDENTIFICATION OF NOVEL HIT MOLECULES AGAINST B-CELL LEUKEMIA/LYMPHOMA-2 (BCL2)

Name/Last Name of the Student: Gurbet TUTUMLU

Date of the Defense of Thesis: 09.01.2019

The thesis has been approved by the Graduate School of Natural and Applied Sciences.

Assist. Prof. Dr. Yücel Batu Salman
Graduate School Director
Signature

I certify that this thesis meets all the requirements as a thesis for the degree of Master Of Sciences.

Assoc. Prof. Dr. Gülay BULUT
Program Coordinator
Signature

This is to certify that we have read this thesis and we find it fully adequate in scope, quality and content, as a thesis for the degree of Master of Sciences.

Examining Comittee Members	Signature
Thesis Supervisor Prof. Dr. Serdar DURDAĞI	----------
Member Assist. Prof. Dr. Timuçin AVŞAR	----------
Member Assist. Prof. Dr. Özge ŞENSOY	----------

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my thesis advisor Prof. Dr. Serdar Durdağı for the regular support of my master study and research, for his motivation, guidance, patience, and immense knowledge. I could not have imagined having a better advisor and mentor for my master thesis study. I would like to thank to all Durdagi Lab members for collaboration especially; Ismail Erol, Busecan Aksoydan, Berna Doğan had imaginable experience to share with me. Busecan was the most helpful person I have ever met. Whenever I need Ismail he was instantly beside me. Ismail was a polymath person that I know in my life was a kind of stimulant to work more when things get awful. Berna always motivatated me. I am also grateful to Işık, Taha, Saima, Ümit, Canberk, Gülru, Salma also supported me everytime. All these people have lots of contributions in this thesis.

I am thankful to my lovely family. They always supported and never lose their faith on me. Especially my mother inspired me to do always better, even in the hardest conditions. I am thankful to my fiance for supporting me in all conditions.

Concisely, all these people reminded me I would never walk alone in different parts of my life. Without their supports, I would not be able to come this point.

Istanbul, 2019 Gurbet TUTUMLU

ABSTRACT

IDENTIFICATION OF NOVEL HIT MOLECULES AGAINST

B-CELL LEUKEMIA/LYMPHOMA-2 (BCL2)

Gurbet TUTUMLU

Bioengineering

Thesis Supervisor: Prof. Dr. Serdar DURDAĞI

January 2019, 51 pages

B-cell leukemia*/lymphoma*-2 (BCL2) family proteins are one of the important oncogenic targets used in anti-cancer drug designing studies. In this study, in order to identify new hits against BCL2, both ligand- and structure-based techniques were applied. Initially 212520 small molecules were retrieved from Specs SC database and binary QSAR based models from MetaCore/MetaDrug platform were used to screen against a defined therapeutic activity, "cancer". Crystal structure of BCL-2 (Protein Databank ID: 4LXD) and NMR structure of BCL-2 (Protein Databank ID: 1YSW,1YSI,2O2F) used for target-driven drug screening methods (i.e., molecular docking). These selected compounds were then used in Glide/high throughput virtual screening (HTVS) at the binding pocket of BCL2. Filtered structures were then used in more sophisticated molecular docking protocols (i.e., Glide/SP, Glide/IFD and GOLD); their docking scores were compared with known inhibitors (i.e., positive controls) and selected molecules were considered for molecular dynamics (MD) simulations. Selected hit molecules which are the molecules that have higher docking scores than positive controls and they construct known crucial interactions were used in MD simulations. Moreover, structural and dynamical properties of ligand-protein complex is investigated for better understanding of hit molecules at the binding pocket of the target.

Keywords: Molecular Docking, *De Novo* Design, Molecular Dynamics (MD) Simulations, Cancer, BCL-2

ÖZET

B-HÜCRELİ LÖSEMİ / LENFOMA-2'YE KARŞI YENİ MOLEKÜLLERİN KEŞFEDİLMESİ

Gurbet TUTUMLU

Biyomühendislik

Tez Danışmanı: Prof. Dr. Serdar DURDAĞI

Ocak 2019, 51 sayfa

B-cell leukemia/lymphoma-2 (BCL-2) proteinleri anti-kanser ilaç tasarım çalışmalarında kullanılan önemli onkojenik hedef yapılardan birisidir. Bu çalışmada, BCL2'ye karşı yeni ligant moleküllerinin belirlenmesi amacıyla, hem ligand bazlı hem de yapı temelli teknikler uygulanmıştır. Başlangıçta 212520 küçük molekül Specs SC veritabanından alındı ve MetaCore / MetaDrug platformunda ikili QSAR tabanlı modeller örneğin "kanser-QSAR modeli" kullanılarak tarama gerçekleştirildi. Seçilen moleküller için BCL-2'nin kristal yapısı (Protein Databank ID: 4LXD) ve BCL-2'nin NMR yapısı (Protein Databank ID: 1YSW, 1YSI, 2O2F) hedefe yönelik ilaç tarama yöntemleri (moleküler yerleştirme) kullanılarak test edildi. Bu seçilmiş bileşikler daha sonra BCL2'nin bağlanma cebinde Glide / yüksek verimli sanal taramada (HTVS) kullanıldı. Glide/HTVS'den seçilen yapılar için daha kompleks moleküler yerleştirme protokolleri (Glide / SP, Glide / IFD ve GOLD) uygulanarak doking skorları bilinen inhibitörlerle (pozitif kontrol) karşılaştırıldı. Pozitif kontrollerden daha yüksek doking skorlarına sahip olan ve bilinen çok önemli ligant-rezidü etkileşimlerini koruyan moleküller için moleküler dinamik (MD) simülasyonlar uygulandı Ayrıca, hedefin bağlanma bölgesinde ligand-protein kompleksinin yapısal ve dinamik özellikleri araştırılmıştır.

Anahtar Kelimeler: Moleküler Kenetlenme, *De Novo* Tasarım, Moleküler Dinamik (MD) Simülasyon, Kanser, BCL-2

İÇİNDEKİLER

TABLES

FIGURES

ABBREVIATIONS

2D	:	Two Dimensional
3D	:	Three Dimensional
A	:	H-bond Acceptor Site
Ala	:	Alanine
Arg	:	Arginine
Asp	:	Asparagine
Asp	:	Aspartic Acid
A1	:	Alpha-1 blocker
BAK	:	Bcl-2-Antagonist/Killer 1
BAX	:	Bcl-2-Associated X Protein
BCL-W	:	BCL2 like 2
BCL-2	:	B-Cell Lymphoma-2
BCL-xL	:	B-Cell Lymphoma-Extra Large
BFL-1	:	BCL-2 Related Protein A1
BH	:	Bcl-2 Homology
BAD	:	Bcl-2 Associated Agonist Of Cell Death
BID	:	BH3 Interacting Domain Death Agonist
BIK	:	Bcl-2 Interacting Killer
BIM	:	Bcl-2 Interacting Mediator Of Cell Death
BOK	:	Bcl-2 Related Ovarian Killer
BP	:	High Blood Pressure
C-terminal	:	Carboxy-Terminal
CADD	:	Computer Aided Drug Design

Cys	:	Cystein
FDA	:	Food and Drug Administration
GBSA	:	Generalized Born SurfaceArea
Gln	:	Glutamine
Glu	:	Glutamic Acid
Gly	:	Glycine
H	:	Hydrophobic Site
H-bond	:	Hydrogen Bond
His	:	Histidine
HRK	:	Activator of Apoptosis Harakiri
HTVS	:	High Throughput Virtual Screenig
IC	:	Intracellular
IFD	:	Induced Fit Docking
Ile	:	Isoleucine
L	:	Ligand
LBDD	:	Ligand Based Drug Design
Leu	:	Leucine
LINCS	:	Linear Contraint Solver
Lys	:	Lysine
MCL-1	:	Myeloid Cell Leukemia 1
MD	:	Molecular Dynamics
Met	:	Methionine
MM	:	Molecular Mechanics
N	:	Negatively Charged

N-terminal	:	Amino-Terminal
NMR	:	Nuclear Magnetic Resonance
NOXA	:	Phorbol-12-myristate-13-acetate-induced protein 1
OMM	:	Outer Mitochondrial Membrane
OPLS	:	Optimized Potential for Liquid Simulations
OPM	:	Orientation in Protein Membranes
P	:	Positively Charged
PBC	:	Periodic Boundary Conditions
PDB	:	Protein Databank
Phe	:	Phenylalanine
P	:	Protein
Pro	:	Proline
PUMA	:	p53 Upregulated Modulator of Apoptosis
QM	:	Quantum Mechanics
QPLD	:	Quantum-Polarized Ligand Docking
QSAR	:	Quantitative Structure Activity Relationship
RMSD	:	Root Mean Square Deviation
RMSF	:	Root Mean Square Fluctuation
SBDD	:	Structure Based Drug Design
SD	:	Steepest Descent
Ser	:	Serine
SP	:	Standard Precision
Thr	:	Threonine
TM	:	Transmembrane

TNF	:	Tumor Necrosis Factor
TNFR	:	Tumor Necrosis Factor Receptor
Trp	:	Tryptophan
Tyr	:	Tyrosine
Val	:	Valine
VMD	:	Visual Molecular Dynamics
XP	:	Extra Precision

SYMBOLS

Acid Dissociation Constant	:	pKa
Activity Constant	:	K_A
Angstrom	:	Å
Aromatic Contributions of Binding Free Energy	:	ΔG_{arom}
Binding Free Energy	:	ΔG
Dissociation Constant	:	K_d
Femtosecond	:	fs
H-bond Contributions of Binding Free Energy	:	$\Delta G_{H\text{-}bond}$
Inhibitory Constant	:	K_i
Ionic Contributions of Binding Free Energy	:	ΔG_{ionic}
Kilodalton	:	kDa
Ligand	:	[L]
Lipophilic Contributions of Binding Free Energy	:	ΔG_{lipo}
Micromolar	:	µM
Microsecond	:	µs
Nanosecond	:	ns
Picosecond	:	ps
Protein	:	[P]
Protein Ligand Complex	:	[PL]
Rotational Binding Free Energy	:	ΔG_{rot}

1. INTRODUCTION

Targeted cancer therapy is one of the strategies used to precisely eliminate malignant tumor cells with minimal effect on neighboring normal cells. It works by targeting upregulated proteins in cancer cells with concomitant low expression in normal cells. Conventional anticancer methods including chemotherapy and radiation aim to kill the actively proliferating cells without differentiating between healthy and tumor cells. These generalized therapies cause number of side effects as they affect the body's normal cells and can also trigger secondary tumor formation when administered alone. However, these classical therapy methods can be used in combination with drugs that can specifically disrupt various cellular processes, such as DNA repair, to enhance their effect. Dysregulation of apoptosis is a characteristic feature of many cancers and drugs that restore the apoptotic function of proteins (i.e., B cell lymphoma-2 (BCL-2)) can be used as anticancer agents (Radha, G. *et al.*, 2017).

The development of a drug against cancer, even though the molecular mechanisms involved in certain pathologies, such as cancer, and causal relationships have been investigated since the mid-1980s, is still now a challenging task (Jagani, H. *et al.*, 2016). As noted by Hanahan and Weinberg, the multistage development of tumors consists of six biological capabilities widely known as cancer symptoms; sustaining proliferative signaling, avoiding growth suppressors, triggering invasion and metastasis, enhancing replicative continuity, inducing angiogenesis and resisting cell death (Hanahan, D. *et al.*, 2000, Hanahan, D. *et al.*, 2011). The ability of cancer cells to escape from programmed cell death remains a critical feature of these six indicators (Mohamad Rosdi, M.N. *et al.*, 2018).

Programmed cell death called apoptosis, is a molecular pathway, which results in cell self-destruction of the end of physiological function, or after damaging the genetic material (Igney, F.H. *et al.*, 2002, Reed, J.C., 2002, Verma, S. *et al.*, 2015). In developed organisms, apoptosis carries out a various function, such as removal of injured or dangerous cells, formation of different organs and structures during fetal growth, and continuation of tissue homeostasis (Hetts, S.W., 1998, Kollek, M. *et al.*, 2016). The extrinsic and the intrinsic pathways are two well-defined basic apoptosis pathways that

use a variety of stimulated signaling elements (Kollek, M. *et al.*, 2016). The extrinsic pathway is activated by outer stimulation of death receptors. Death receptors are members of the tumor necrosis factor (TNF) receptor family which has an intracellular signaling domain, that is able to accumulate and trigger caspase-8 followed by operation of effector caspases including caspase-3, -6 or -7. This accumulation is not accompanied by direct participation of B cell lymphoma-2 (BCL-2) (Eimon, P.M. *et al.*, 2010, Wu, H. *et al.*, 2018, Youle, R.J. *et al.*, 2008). In relation to the role of these organelles, the intrinsic pathway also called mitochondrial pathway, is initiated by variety of cytotoxic damages or growth signals, some of which have genetic instability. A group of proteins named as BCL-2 tightly regulates this process subsequently leads to the activation of caspase-9 (Cory, S. *et al.*, 2003, Eimon, P.M. *et al.*, 2010, Youle, R.J. *et al.*, 2008).

All members of the BCL-2 protein family have retained sequence patterns regarded as the BCL-2 homology (BH) domains and could be divided into 3 main classes. The first class of proteins is made up of the pro-apoptotic activator BH domain 3 (BH3)-only proteins (e.g., BIM, BID, PUMA, BIK, NOXA, HRK, BAD) that contain just BH3. They act as molecular guardians that connect outer spurs to the mitochondrial pathway. The following group contains the proapoptotic effectors which are multidomain proteins (e.g., BAX, BAK) and each of them has 3 BH domains. These proteins disrupt the integrity of mitochondrial outer membrane, leading to the free movement of cytochrome C into the cytoplasm, initiating downstream caspase activity, and ultimately terminating the cell when triggered.

The last class of BCL-2 family is the antiapoptotic proteins (e.g., BCL-XL, BCL-2, MCL-1, BCL-W). All of these members consist of 4 BH domains and keeps cells safe by segregating their proapoptotic peers. The most important point in promoting apoptosis is increased amount of BH3-only proteins or switching off one of its antiapoptotic BCL-2 counterparts (Chung, C., 2018, Dewson, G. *et al.*, 2010, Fesik, S.W., 2005, Mohamad Rosdi, M.N. *et al.*, 2018).

The idea of BH3 mimetics as promising anticancer drugs inspired by the conclusion that plenty of cancers rely on BCL-2 family proteins, and that the interaction between these proteins occurs through specific BH domains (Oltersdorf, T. *et al.*, 2005, Solderquist,

R.S. *et al.*, 2016). A genuine BH3 mimetic is that it mimics the BH3 domain of a proapoptotic BCL2 protein, which inactivates antiapoptotic family members by filling the their BH3 binding pocket. For example, ABT-737 molecule was issued as a potential inhibitor of BCL-2 and BCL-XL which takes up their BH3-binding domain, further triggers apoptosis in diversified cancer types (Solderquist, R.S. *et al.*, 2016, Tse, C. *et al.*, 2008). Ensuing pharmaceutical trials directed to clinical trials with ABT-263 (navitoclax), which increased bioavailability, indicated efficacy in leukemia and a few other neoplasia, but also revealed toxicities with leading dose restrictions of neutropenia and thrombocytopenia (Gandhi, L. *et al.*, 2011, Tse, C. *et al.*, 2008). The thrombocytopenia was connected to the blockage of BCL-XL, as BCL-XL is essential for the survival of platelets (Zhang, H. *et al.*, 2007). More recently, ABT-199 (venetoclax) was designed as a selective BCL-2 inhibitor and it evades the issue of thrombocytopenia (Souers, A.J. *et al.*, 2013).

Tumor growth is inhibited by apoptotic cell death, which acts as an innate barrier. Avoidance of apoptosis is a key hallmark of cancer and comprises a crucial process in resistance to chemotherapeutics. This phenomenon led peculiar approaches in anticancer therapies focusing on apoptosis including suppression of survival factors which their overexpression is detected in numerous malignancies. In the group of survival factors, BCL-2 proteins are one of the families that step forward for drug discovery studies (Billard, C., 2013, Hanahan, D. *et al.*, 2011, Lessene, G. *et al.*, 2008).

1.1 THEORETICAL BACKGROUND

1.1.1 Computer-Based Drug Design

The development of a new drug is long and laborious and costly process (Leelananda, S.P. *et al.*, 2016). In order to find a way to address the problem of financial costs and time, researchers have conducted many studies using computational technics and in many cases they have found innovative therapeutic approaches (Chiba, S. *et al.*, 2017, Yoshino, R. *et al.*, 2017, Yoshino, R. *et al.*, 2015).

Drugs can be described as ligands which act as therapeutic agents by binding to specific proteins. The discovery of a drug has been a very important topic because diseases cause various problems for the society, however, the process requires a lot of time and it is financially costly. According to reports, the market demanded about 10 to 15 years and also more than $ 800 million during the process until it reaches to the market (Paul et al., 2010). Therefore, decreasing both time and costs for developing a drug by discovering innovative techniques is crucial. The competent methods that have been applied to make the drug development process more efficient are the computational approaches. (Mavromoustakos, T., 2011).

The computational methods applied in drug discovery could be categorized into two broad techniques: Structure-Based Drug Discovery (SBDD) and Ligand-Based Drug Discovery (LBDD) (Güner, O.F., 2000, Tropsha, A., 2010, Wakui, N. *et al.*, 2018). While in the SBDD technique the data (i.e., structural information) for the target structure can be retrieved using X-Ray diffraction, NMR or electron microscopy techniques, in the LBDD, mainly there is no available structural information about the target, thus quantitative structure-activity relationships (QSAR) is considered. Molecular docking and molecular dynamics (MD) simulations are effective computational tools for the protein-ligand interactions in SBDD. Atomic resolution dynamics of biomolecules can be determined using MD. MD uses Newton principles of motion and tries to explain the flexibility of proteins and it can be applied to different kind of biomolecules like nucleic acids, proteins, and biomembranes (Wakui, N. *et al.*, 2018) (Arora, K. *et al.*, 2009, Gapsys, V. *et al.*, 2013, Hayes, R.L. *et al.*, 2012, Levine, Z.A. *et al.*, 2014, Sekijima, M. *et al.*, 2003, Sodt, A.J. *et al.*, 2014, Yildirim, A. *et al.*, 2014). Moreover, MD simulations can explore the underlying mechanisms of ligands in complex with proteins (Wakui, N. *et al.*, 2018) (Bártová, I. *et al.*, 2004, Buch, I. *et al.*, 2011, Doruker, P. *et al.*, 2000, Shan, Y. *et al.*, 2011). Using MD, it is possible to determine lead compounds with rapid processes compared to conventional drug design studies (Alonso, H. *et al.*, 2006).

Molecular Mechanics (MM) and Quantum Mechanics (QM) methods can be used to calculate the total energy of an investigated molecular structure. QM calculations have two methods as semi-empirical and non-empirical methods which need knowledge about

electrons of all system and that requires computer technologies. Semi-empirical methods estimate energy using data from experiments and takes into account only valence electrons. Non-empirical methods work on computational approximations like electronic density. The system's potential energy is calculated in molecular modeling by force field method which is also called MM.

1.1.1.1 Structure-Based Drug Design

In structure-based drug design approach, information of target protein structure should be known. There are two options, if the experimental structure data of the protein is revealed, it is possible to retrieve and use its spatial coordinates from the protein data bank (PDB). If there is no available experimental data, then homology modeling can be used to construct the 3D model of the target protein. In this approach, firstly sequence-based search can be utilized to find similar proteins (templates), and then target sequence can be aligned to the template, finally, the amino acid sequence of the target protein is mapped to the known 3D structure of template protein and model structure can be obtained. When the structure of the protein is obtained, then the binding site of the ligands should be identified. (Berman, H.M. *et al.*, 2002). If the crystal structure of the target protein is crystallized with the small molecule at the active site (i.e., holo form), this information can be used to dock other molecules. If the crystal structure is in apo form (without any ligand), then binding site information can be obtained from experimental mutation studies or computational methods can be applied to identify the active site.

Molecular docking is an approach to determine the conformations of bioactive molecules at the binding pocket of the receptor. It calculates non-bonded interactions such as hydrogen bonding, hydrophobic, electrostatic and van der Waals interactions between protein and ligand molecule and predicts free energy of binding ΔG (Gibbs free energy-docking score) which corresponds to the sum of all these non-bonded terms. Low docking score (i.e., high ΔG values) corresponds to the unfavorable binding pose, and low ΔG corresponds to the favorable binding conformation at the binding pocket. Most docking programs, due to computational costs, only takes ligand flexibility into account while neglecting protein flexibility (Kitchen, D.B. *et al.*, 2004).

Various docking algorithms and programs are available like DOCK, AutoDock, AutoDock Vina, FlexX, GOLD, HADDOCK, Glide, etc (Ewing, T.J. *et al.*, 2001, Friesner, R.A. *et al.*, 2004, Jones, G. *et al.*, 1997, Rarey, M. *et al.*, 1996). Each of these programs has its own algorithm to search different conformations of ligands at the binding site and scoring function to rank molecules according to their binding affinities. In this thesis, the Glide docking algorithm was considered which uses a hierarchical filter to predict bioactive ligand conformations at the binding pocket. In Maestro molecular modeling package, Glide docking algorithm with different precisions are available, namely, starting from the most complicated one quantum-polarized ligand docking (QPLD), induced fit docking (IFD), extra precision (XP), standard precision (SP) and high throughput virtual screening (HTVS). Millions of compounds can be screened against a specific target *in silico* using HTVS (Schrödinger, 2010). The main differences between SP and XP algorithms are their scoring functions and penalties like desolvation and charge. XP uses more complicated scoring functions and applies charge and desolvation penalties.

In IFD algorithm, initial docking is performed using SP or XP mode, then generated poses are refined using Prime module of Maestro to provide flexibility to the surrounding residues at the binding pocket. This enables active site residues to adopt more favorable orientations to maximize ligand-protein interactions. Finally, refined protein-ligand complexes are used to re-dock ligands back to the binding pocket and scored.
Although QPLD is the algorithm that provides the one of the most reliable bioactive poses and docking scores, it is the one of the most time-consuming docking protocols compared to the others. This algorithm combines QM calculations which provide high-level accuracy and MM which enables speed.

Docking algorithms and programs provide the static picture of protein-ligand complexes and do not contain physiological conditions such as temperature. Since all biological processes are dynamic and take place in physiological conditions, these conditions are required to correctly represent the nature of dynamical cellular processes. This can be acquired by MD simulations.

MD simulation uses Newton's law of motion to calculate the potential energy of each atom. Potential energy term calculated by taking into account non-bonded interactions, bond angles and lengths, and dihedral angles. MD method enables to calculate atomic positions and give insights about the dynamic nature of protein-ligand complexes within a certain period of time, yielding MD trajectories.

1.1.1.2 Ligand-Based Drug Design

Ligand-based drug design basically relies on the ligand structures and experimental binding data, no prior knowledge of the structure of the protein is needed. The most applied method in ligand-based drug design is the QSAR. (Verma, J. *et al.*, 2010). This method arises from the conclusion that molecules with similar structures will exhibit similar binding properties. In this method, functional groups on each ligand of the ligand set are aligned on a template molecule to obtain a universal pharmacophore model. Further, the generated pharmacophore model can be validated by using another set of ligands with known experimental binding data. Where the 3D structure of the protein is not revealed, this method allows the discovery of new molecules capable of binding on the same target.

Depending on their experimental binding affinity, molecules can be categorized into three classes;

i. Inactive molecules (low binding affinity, ie. millimolar)
ii. Moderate molecules (micromolar to high nanomolar affinity)
iii. Active molecules (high binding affinity, ie. picomolar)

In 3D QSAR method; six different features are used to define functional groups of studied molecules.

A: Hydrogen Bond Acceptor

D: Hydrogen Bond Donor

H: Hydrophobic

P: Positively Charged

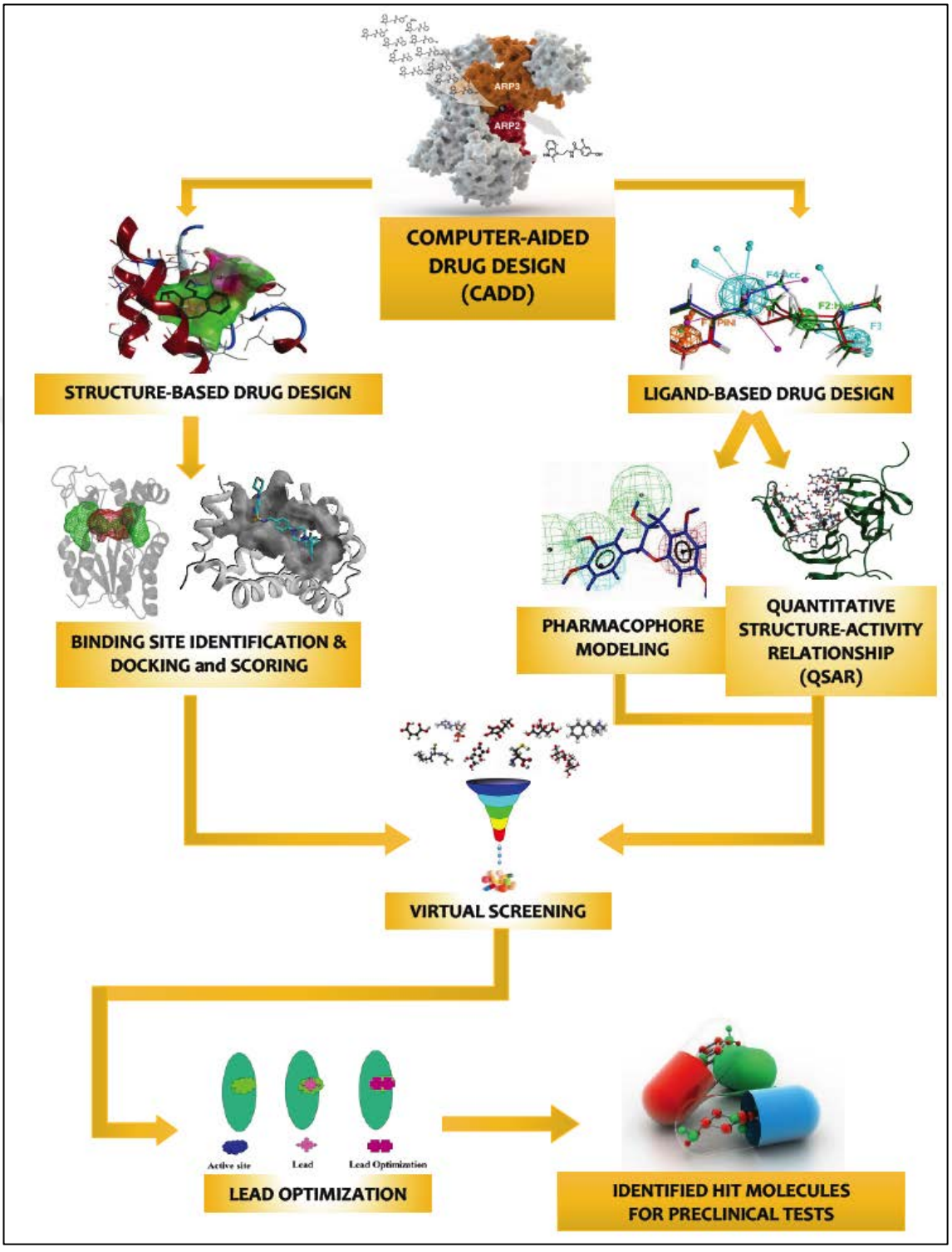
N: Negatively Charged

R: Aromatic Ring

A pharmacophore model is generated by aligning of all molecules regarding their functional groups. Two different sets of molecules used for this aim; (i) a training set and (ii) test set. Generation and validation of the pharmacophore model are done using these two sets of molecules. Validation of the generated pharmacophore model is achieved by statistical parameters and external test set (which consist of molecules that have experimental binding data, true unknowns).

A virtual screening scheme which is applied within the scope of this thesis can be seen in Scheme-1.1

Figure 1.1: Applied virtual screening workFlow of CADD



1.1.2 Cancer

Cancer was an attractive topic among scientists because of its darkness since the 3rd quartile of 20th century. It is one of the most mortal disease killing millions of people in each year. It is predicted by WHO (2012) that the number of people lost their lives because of cancer will increase from 7.6 million (2008) to 13 million (2030) (Bray, F. *et al.*, 2012, Jemal, A. *et al.*, 2011, Mohamad Rosdi, M.N. *et al.*, 2018). Moreover, it has been claimed that the incidence of cancer is particularly high in less developed regions, and according to estimates, 19 million people per year will die because of cancer (Mohamad Rosdi, M.N. *et al.*, 2018).

According to Hannah and Weinberg (Hanahan, D. *et al.*, 2000, Hanahan, D. *et al.*, 2011), cancer has some several major characteristics. The signals, insensitivity to inhibition of growth, metastasis, and tissue running, not having a limit of replication potential, angiogenesis, evasion of apoptosis, instability in the genome are some of them. However, the most crucial hallmark of the cancer is thought as that cancer cells can evade apoptosis (Mohamad Rosdi, M.N. *et al.*, 2018). That evading, or in other words failing to undergo apoptosis of cancer cells make them more advantageous over normal cells. While targeting cancer therapy, the proteins that have an important role in cancer progression are targeted to lose their function. BCL2, an anti-apoptotic protein, is a competent protein for this, because it is highly upregulated in many cancers by targeting the cell for therapy. Although BCL2 has been in the interest of researchers for over 30 years, only a few have been identified as having clinically significant effect (Radha, G. *et al.*, 2017).

Cancer is yet a troubling issue for physicians to cure although the underlying molecular mechanisms and etiology was found out during the last three decades. If new aims which are responsible for tumors could be detected, novel methods/approaches and anticancer molecules can be investigated by scientists in this field. The well-known molecular pathway components can be useful for researchers to investigate innovative methods. Studies are trying to elucidate the development process of molecules which are aiming the cancerogenic components but not disturbing normal cells (Hetts, S.W., 1998, Jagani, H. *et al.*, 2016).

There are plenty of chemotherapeutic agents which can induce apoptosis after cell damage of drugs and they are being used to treat cancer (Brown, J.M. *et al.*, 2005, Fesik, S.W., 2005, Jagani, H. *et al.*, 2016). Moreover, if the apoptotic pathway has some defects, it makes cancer cells more resistant to drugs. There are many studies claiming that BCL-2 protein can develop resistance to drugs in specific cancer forms (Davis, J.M. *et al.*, 2003, Jagani, H. *et al.*, 2016, Sartorius, U.A. *et al.*, 2002, Schmitt, C.A. *et al.*, 2000).

A common type of leukemia in is chronic lymphocytic leukemia (CLL) and the rate of people have it in the United States (US) is 4.1/100.000. According to NCL (2016), around 15.000 people are diagnosed with CLL every year and 4.500 patients die. For the diagnosis, the median age is 71 and 70% of patients is over 65 and also half of the patients are between 65-84. The more people will be die because of CLL and it seems that it will increase in the coming decade because of the increase in mean age of the population (Hallek, M., 2015, Itchaki, G. *et al.*, 2016)

1.1.3 Targeting BCL-2 For Cancer Therapy

BCL-2 has thought as the one of the most efficient structures thanks to its functions in both tumor development process and metastasis. Inhibition of BCL-2 protein can be considered as an innovative treatment approach for cancer, as it will induce apoptosis in cancer-causing cells.(Jagani, H. *et al.*, 2016, Wang, S. *et al.*, 2003)

Furthermore, the BCL-2 gene has an effect on the process of developing resistance to chemotherapeutic agents by cancer cells. (Leber et al. 2010). So far, a few molecules were identified as the inhibitor of BCL-2 proteins (Jagani, H. *et al.*, 2016, Kang, M.H. *et al.*, 2009).

The lack of balance between proapoptotic and anti-apoptotic proteins causes the accumulation of cells. Thus, cancer therapies based on apoptosis induction cannot work efficiently enough. Cancer therapy targets the BCL-2 gene family due to the fact that it is an important point in the survival of cells (Petros, A.M. *et al.*, 2004, Ruefli-Brasse, A. *et al.*, 2017).

1.1.4 Mechanism of Apoptosis

In a regulated process, a primary and basic mechanism that eliminates old, damaged and unnecessary cells is called apoptosis; the cell death in a genetically programmed manner. That process of stopping signals of apoptosis is the essential characteristic of cancer. It is also very important for chemoresistance, maintenance of tumors and also oncogenesis (Hanahan, D. *et al.*, 2000, Souers, A.J. *et al.*, 2013). Therefore, a promising method to induce apoptosis of tumor cells and some of the anticancer agents which are BCL-2 inhibitors is to aim BCL-2 by binding to its BH3 binding pocket (Chen, G. *et al.*, 2017).

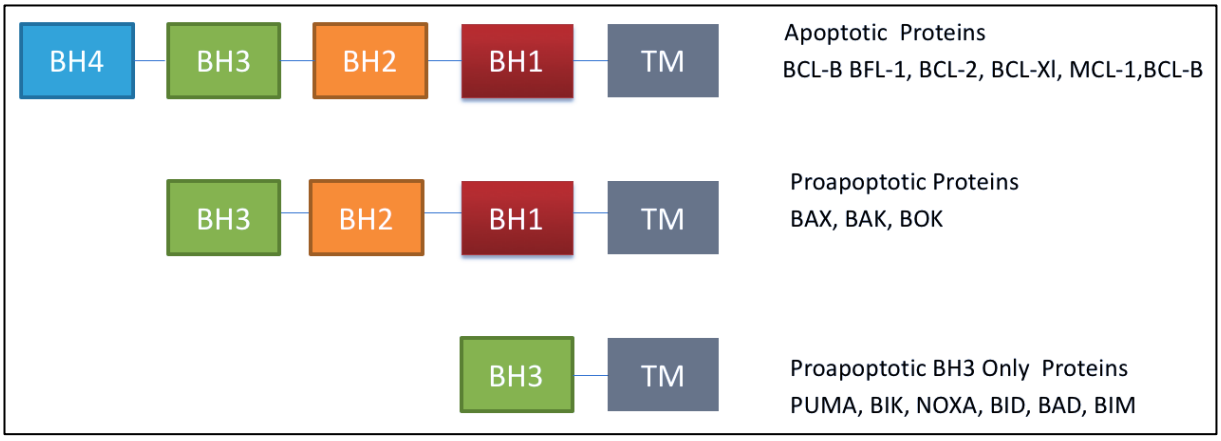
In the essence, though developmental process, apoptosis is responsible for the programmed deaths of cells in order to provide homeostasis(Acoca, S. *et al.*, 2011, Elmore, S., 2007, Krammer, P.H. *et al.*, 1994, Meier, P. *et al.*, 2000). There are two main pathways of apoptosis: the intrinsic and the extrinsic mode. The intrinsic pathway is triggered by intracellular signals in which the mitochondria play an important role. Cytotoxic stimuli, such as growth factor withdrawal, developmental signals, and anticancer drugs, activate the intrinsic pathway. However, the extrinsic pathway includes external signals that bind to cell death receptors on the cell surface leading to the formation of the death induced signaling complex (DISC)(Delbridge, A.R. *et al.*, 2016, Elmore, S., 2007, Radha, G. *et al.*, 2017). For example, the interaction of FAS-FASL on T cells and natural killer cells result in activation of caspases leading to morphological changes such as chromatin condensation, nuclear fragmentation, membrane blebbing, apoptotic body formation, and cell death (Okada, H. *et al.*, 2004, Radha, G. *et al.*, 2017).

2. LITERATURE REVIEW

2.1 APOPTOSIS AND BCL-2

Apoptosis is a term used to describe suicides of cells after completing their missions or serious damage in the genetic structure (Verma, S. *et al.*, 2015). It has a crucial role in the healthy development of the organism, for homeostasis and for the elimination of damaged cells. Therefore, a break in apoptotic activity could lead to serious problems in the health of an organism (Petros, A.M. *et al.*, 2004). It is claimed that it has an evolutionary root managed by BCL-2 protein family (Igney, F.H. *et al.*, 2002, Reed, J.C., 2002). The BCL-2 family proteins are classified into antiapoptotic and proapoptotic proteins. While the antiapoptotic family contains 4 BH domains (BH-1 to BH-4), proapoptotic family proteins are divided into BH123 and BH3 only subfamily. BH123 subfamily do not include BH4 domain and it interacts with antiapoptotic proteins through their BH3 domain in the hydrophobic pocket. However, the BH3 only subfamily proteins exhibit sequence homology within the BH3 motif and they do not have other domains (Pinto, M. *et al.*, 2004, Verma, S. *et al.*, 2015).

Figure 2.1: Representation of BCL-2 family proteins.

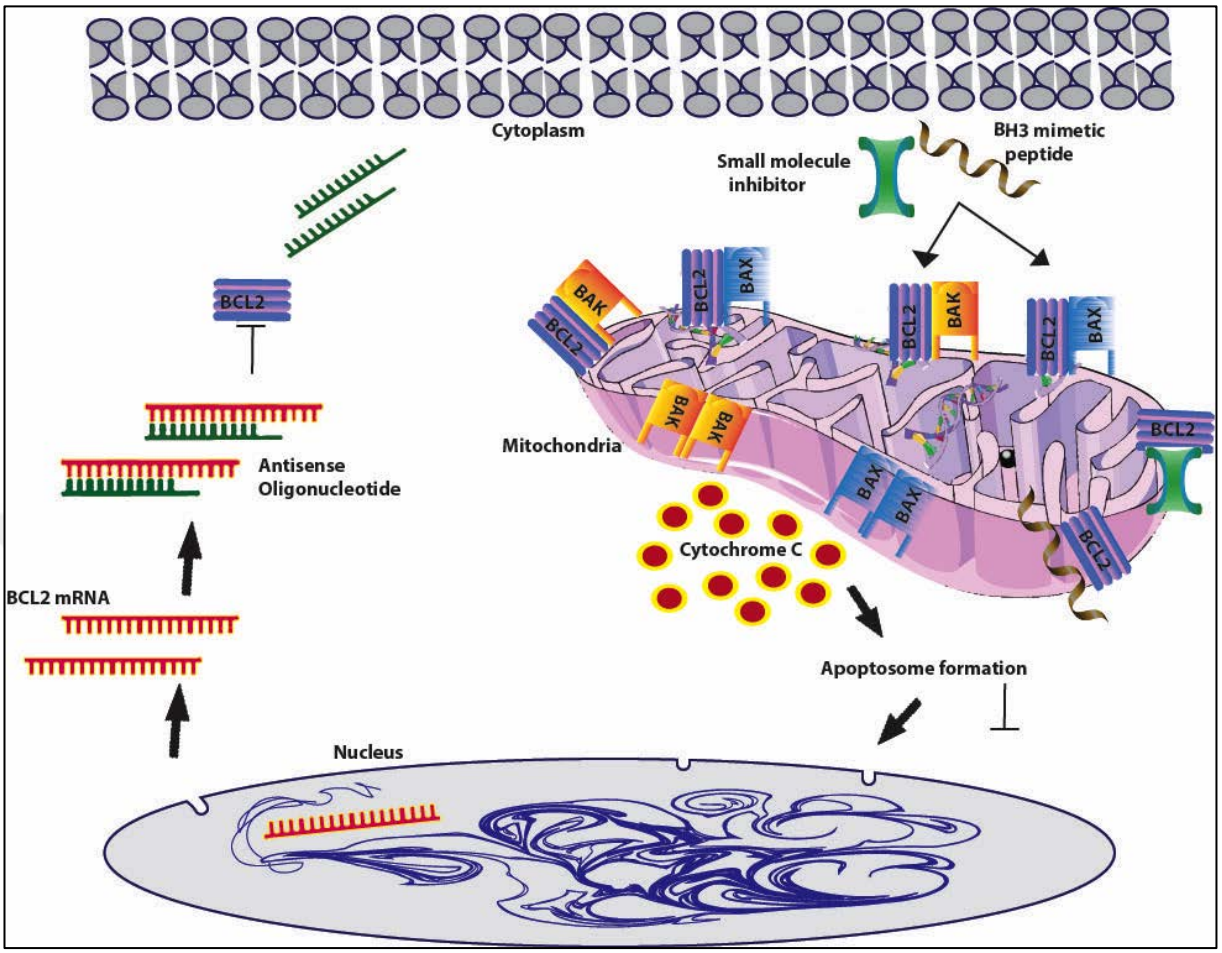


The antiapoptotic BCL2 family protein consists of BCL2 (26 kDa) and the related proteins such as BCL-xL (26 kDA), MCL-1 (37 kDa), BCL-W (20 kDa), BCL-B (22 kDa), and BCL2A1 (20 kDa). In the absence of apoptotic stimulation, the antiapoptotic BCL2 protein binds to BAK and BAX (proapoptotic BH123 proteins) on the outer membrane of the mitochondria, thereby inhibiting their activation. Together with the

apoptotic stimuli, the BH3-only proteins can bind to BCL-2 for blocking its interactions with BH123 subfamily proteins, or it can also directly bind to BH123 proteins in order to activate them (Radha, G. *et al.*, 2017, Rossé, T. *et al.*, 1998). When activated, the effector proteins BAK and BAX oligomerize to form pores in the outer mitochondrial membrane and thus enable soluble protein cytochrome C and other proapoptotic proteins such as Smac/DIABLO to diffuse from the mitochondrial intermembrane space and to proceed with their pro-apoptotic activities (Chipuk, J.E. *et al.*, 2010). This leads to proteolytic caspase activation followed by DNA fragmentation and eventually cell death (Radha, G. *et al.*, 2017, Rossé, T. *et al.*, 1998). Therefore, upregulation of BCL-2 protein prevents the cell from being exposed to apoptosis, whereas inhibition of BCL-2-BH123 interaction may induce apoptosis (Radha, G. *et al.*, 2017, Rossé, T. *et al.*, 1998). There are three main strategies for targeting the BCL-2 proteins in cancer therapy (Figure 2.2). (i) the use of the antisense oligonucleotide so that it can bind to complementary BCL-2 mRNA and thereby block the expression at the protein level; (ii) usage of peptides against the BCL-2 protein; (iii) identification of small molecule inhibitors that can bind to the hydrophobic groove of BCL-2 proteins to obstruct their interactions with proapoptotic partners (Ebrahim, A.S. *et al.*, 2016, Radha, G. *et al.*, 2017, Thomas, S. *et al.*, 2013, Verdine, G.L. *et al.*, 2007, Yip, K. *et al.*, 2008).

Schematic showing various strategies to target BCL-2. In one of the strategies, antisense oligonucleotide binds to BCL-2 mRNA leading downregulation of expression of *BCL-2* gene. In the second strategy, BH3 mimetic peptides are used. In the third strategy, small molecule inhibitors are utilized. In last two cases, association of proapoptotic BAK or BAX proteins with antiapoptotic protein, BCL-2 is disrupted upon administration of peptides or small molecule inhibitors, leading to BAK/BAX homodimerization, cytochrome C release, caspase activation, DNA fragmentation and eventually cell death.

Figure 2.2: BCL-2 targeting strategies.



The first gene encoding for a BCL-2 family member was identified in human follicular lymphoma about 25 years ago. In this type of cancer, a translocation between chromosomes 14 and 18 occurs. Such rearrangement brings the BCL-2 gene under the transcriptional control of the immunoglobulin enhancer element, causing potent upregulation of the BCL-2 gene transcription. A few years later using mouse models, it was demonstrated that an overexpressed BCL-2 gene blocks apoptosis of B-cells and extends the survival of certain hematopoietic cell lines upon growth factor deprivation. Taken together, these studies have established BCL-2 as an anti-apoptotic protein with proto-oncogenic potential. Since then, the family has grown significantly, and to date, based on the sequence similarity to BCL-2, over twenty family members have been identified in humans. BCL-2 family homologs have been identified in virtually all metazoans (Blaineau, S.V. *et al.*, 2009, Hardwick, J.M. *et al.*, 2013, Hockenbery, D. *et al.*, 1990, Tsujimoto, Y. *et al.*, 1985).

In 2001, Enyedy et al. proposed a novel class of nonpeptide small molecule acting at the BH3 site of BCL-2. First, they used the NMR structure of BCL-XL to model BCL-2 three dimensional structure and then the National Cancer Institute molecule database docked to a homology model of BCL-2 to identify inhibitors. They selected 35 possible binders and evaluated these inhibitors in the myeloid leukemia cell line, HL-60 (Enyedy, I.J. *et al.*, 2001).

Kitada et al. used 50 different polyphenol molecules and applied molecular docking simulation to unravel their binding conformation and calculate binding energies, and successfully identified gossypol and purpurogallin as inhibitors of BH3 binding to BCL-XL. Also, they identified GLU96, PHE97, ARG100, TYR101, VAL126, LEU130, ARG139, TYR195 and ALA200 as important residues in Gossypol binding, among these residues, GLU96, and ARG139 found to make hydrogen bond (Kitada, S. *et al.*, 2003).

In 2005, Yin et al. used the docking method to identify the inhibition effect of terephthalamide derivatives at the BCL-XL/Bak BH3 complex. They found that terephthalamide derivative compound 26 with an IC_{50} of 35 μM in HEK293 cells. In combination with molecular docking and NMR calculations, ALA89, GLY94, TYR101, ARG102, LEU108, LEU112, ALA119, ARG139, and ALA142 identified as crucial residues in compound 26 binding to BCL-XL (Yin, H. *et al.*, 2005).

Another structure-based molecular design approach used to identify new BCL-2 inhibitors, starting from gossypol a natural product from cotton seeds and discovered compound 5 (TW-37) with a Ki of 290 nM to BCL-2. In molecular docking studies, residues ARG146 and ASN143 found to make H-bonds with the compound 5, which have an important role in BIM BH3 peptide binding interface (Wang, G. *et al.*, 2006).

Xing et al. first evaluated a series of reported inhibitors' (antimycin A3, chelerythrine, gossypol, HA12-1, and BHI-1) selectivity for BCL-W, BCL-2, and BCL-XL, but all studied molecules were not shown any selectivity. Then they synthesized a series of BHI-

1 based analogs and found that by derivatization of BHI-1, binding affinities and profiles could be varied (Xing, C. *et al.*, 2007).

Another structure-based approach applied to propose new flavonoid based compounds to mimic BIM BH3 against BCL-2, and by using molecular docking in combination with in vitro cell line assays Tang et al. identified compound 4 as potent. Binding pose evaluation of compound 4 revealed that ASN143 and TYR202 have the critical role in the inhibition of BCL-2 (Tang, G. *et al.*, 2007).

Lama et al. investigated protein-protein interactions of BCL-XL in complex with BIM, BAK, and BAD BH3 using MD simulations. Both of BIM BH3 and BAD BH3 showed strong hydrophilic and hydrophobic interactions with BCL-XL, but BAK BH3 failed to show same strong interactions. Authors speculated that due to BAK BH3 lacked ARG residue at the interaction interface, (this important positively charged residue changed with ALA) BAK BH3 showed weak binding affinity compared to other tested BH3s (Lama, D. *et al.*, 2008).

Bernardo et al. synthesized analogs of two natural compounds, chelerythrine, and sanguinarine, respectively, and evaluated derivatives of two compounds which act at two different sites. They applied molecular docking and identified sanguinarine derivative 35 and 42, that interact GLU129 and ASP133 at BH1 domain, and VAL135, TYR173, and GLU184 at BH groove, respectively with potent IC_{50} values (Bernardo, P.H. *et al.*, 2008).

In 2008, Tang et al. designed and synthesized a series of acylpyrogallols against anti-apoptotic BCL-2. They combined their in vitro effort with molecular docking simulation to shed light into binding conformations of synthesized compounds. Their most compound 9, which showed nM affinity, interacted via H-bonds between TYR108, ASN143, ARG146 and TYR202 at the binding pocket of BCL-2 (Tang, G. *et al.*, 2008).

Rodriguez et al. conducted a molecular docking study of series of benzoylurea derivative compounds to probe important residues at the binding pocket of BCL-XL. The most potent compound, 14 found to interact via π-π interaction with PHE146, and from its

carboxyl group, compound 14 interact with two positively charged ARG132 and ARG139 residues. The authors also compared compound 14 with other inactive compounds to reveal important pharmacophore groups to have high binding affinity against BCL-XL (Rodriguez, J.M. *et al.*, 2009).

In 2010, Lama et al. investigated 6 different length BH3 peptides (BAK 16-mer, BAD 25-mer, BAD 16-mer wt, BAD 16-mer mutant, BIM 33-mer, and BIM 16-mer) binding to the anti-apoptotic BCL-XL protein using MD simulations. They proposed that, length and sequence of BH3 peptide are critical points, which could facilitate peptidomimetic ligand design studies (Lama, D. *et al.*, 2011).

Another pharmacophore-based rational design of BH3 mimetic inhibitor design study came out in 2011, where the authors generated their pharmacophore model using MD trajectories of BCL-XL – BH3 (HRK, BAX, BAK, and BID) complex. Pharmacophore model was composed of two H-bond acceptor and two hydrophobic groups. Obtained model used to screen ~500.000 compounds from 4 different chemical libraries. Three molecules, UBQF14, UBQF17 and UBQF7, had 7, 8, and >25 μM affinites in fluorescence polarization assays, respectively. These 3 compounds were subjected to MD simulations and further free energy calculation methods such as MM-PBSA and MM-GBSA to reveal time dependent behaviors and interacting partners (hot spot) in terms of conformational analysis. In free energy decomposition, authors suggested PHE97, ARG100, TYR101, PHE105, LEU130, ALA142, and TYR195 in UBQF14 case, PHE97, ARG100, TYR101, LEU130, ALA142, and PHE146 in UBQF17 case as crucial aminoacids that acting as hot spots (Pinto, M. *et al.*, 2011).

Acoca et al. investigated obatoclax and ABT-737 binding conformations using molecular docking and dynamical demeanour using MD simulations on BCL-2 and BCL-XL. Also they conducted post MD free energy calculation to predict binding free energies and critical aminoacids. Furthermore, identified crucial aminoacids mutated computationally to alanine to detect their contributions (Acoca, S. *et al.*, 2011).

Another structural investigation of BCL-2 was conducted to predict the structure of flexible loop domain (FLD) which contain critical aminoacids such as SER87, THR74, SER70, THR69, THR56, and ASP34 using MD simulations and principal component analysis (PCA) by Raghav and coworkers. This study attempted to reveal unstructured loop domain of BCL-2 (namely, FLD), which plays important role in binding of other partners to BCL-2 (Raghav, P.K. *et al.*, 2012).

High temperature MD simulation was applied to study stability and thermal unfolding of BAX protein of BCL-2 family at 300, 400, and 500 K. It was found that, increased temperature caused the BAX peptide to lose its secondary structure, interestingly conserved its globular structure via hydrogen bonds, salt bridges and hydrophobic interactions (Rosas-Trigueros, J.L. *et al.*, 2011).

Li et al. designed, synthesized and biologically evaluated a set of rhodanine-based acylsulfonamide compounds against BCL-2. Synthesized molecules were subjected to molecular docking to evaluate their binding modes and important residues. All compounds were found to be tight binders in terms of docking scores in simulations and the most potent molecules, compound 10 and 11 had 20 and 25 nM Ki values (Li, H.-q. *et al.*, 2012).

In 2013, Maity et al. investigated BCL-XL dynamics both in implicit water and membrane environment. They performed 10 different simulations (also repetition simulations) in total ~1.6 μs in different mediums. Authors found that, BAK BH3 decreased binding energy in membrane, compared to the water environment (Maity, A. *et al.*, 2013).

Azam et al. used GOLD and Autodock molecular docking programs to investigate 104 structurally diverse compounds set on BCL-XL, and identified a new binding site covering TYR120, GLU121, VAL161, SER164, ARG165, ALA168, TRP169, and THR172 residues. The authors speculated that this unique binding site would open new possibilities to discovering novel compounds (Azam, S.S. *et al.*, 2014).

Kamath et al. synthesized indole-coumarin compound bearing halogen (namely, N'-((2-(2-oxo-2H-chromen-3-yl)-1H-indol-3-yl)methylene)benzohy-drazides) and evaulated these compounds computationally using molecular docking at BCL-2 and BCL-XL targets. TYR108 and GLN118 found to be important for the binding of most potent compounds; 5g and 5h, in BCL-2 structure. In BCL-XL protein, in the case of 5g, a h-bond between molecule and ARG139 was found. In the case of 5h, molecule interacted with BCL-XL mostly via hydrophobic interactions (Kamath, P.R. *et al.*, 2016).

Recently, Ziedan et al. conducted virtual screening study to discover novel oxadiazole derivatives against BCL-2 inhibition at the BH3 binding site. Also, they assayed in human breast cancer cell line (MDA-MB-231) and most potent compound 16j had 4.27 μM IC_{50} value. Binding mode of compound 16j was evaluated by means of molecular docking simulations and ARG143 found to be important via π-cation interaction, also binding pocket consisted of PHE101, GLY142, GLU149, ALA146, PHE109, VAL153, PHE150, VAL130, LEU134, and TYR105 (Ziedan, N.I. *et al.*, 2017).

Flavonoids have been reported to possess potent cytotoxic activity against a number of cancer cell lines. lonchocarpin (34) (flavonoids) having antitumor activity isolate from traditional herbal medicine Pongamia pinnata (L.) Pierre and to show the implicit chemical mechanism. Gang Chen *et al*, 2017 reported new compound lonchocarpin having anticancer activity. This study showed that the cytotoxic activities of lonchocarpin in 10 lung cancer cell lines and it exhibited 97.5% activity at a dose of 100 μM in the H292 cell line. 3D-QSAR study of 37 flavonoids from P. pinnata was also performed, and the results obtained showed that the hydrophobic interaction could be the crucial factor for the antitumor activity of lonchocarpin. Docking studies revealed that lonchocarpin bound stably to the BH3-binding groove of the BCL-2 protein with hydrophobic interactions with ALA146. Also, lonchocarpin significantly reduced cell proliferation via modulating BAX/Caspase-9/Caspase-3 pathway. These results showed that lonchocarpin is a potentially useful natural agent for cancer treatment (Chen, G. *et al.*, 2017).

In 2001 Alexei Degterev *et al*, identified a series of novel small molecules (BH3Is) that inhibit the binding of the BAK BH3 peptide to BCL-xL. NMR analyses revealed that BH3Is target the BH3-binding pocket of BCL-xL. Reported inhibitors specifically block the BH3-domain-mediated heterodimerization between BCL-2 family members in vitro and in vivo and induce apoptosis. Results of this study indicate that BH3-dependent heterodimerization is the key function of anti-apoptotic BCL-2 family members and is required for the maintenance of cellular homeostasis (Degterev, A. *et al.*, 2001).

3. METHODS

3.1 METACORE/METADRUG ANALYSIS

MetaCore/MetaDrug platform from Clarivate Analytics provides a comprehensive tool to analyze the pharmacodynamic and pharmacokinetic profiles for the screening molecules. MetaCore database contains 9166 drug molecules (small molecules and biologics) and not all of small molecule drugs were used for QSAR models building (each QSAR models were built with different number of training and test set compounds). This tool can predict the therapeutic activity of molecules for the common 25 diseases as well as predictions of 26 different toxicities using binary QSAR disease models. It uses the property of *Tanimoto Prioritization* (TP) to find the similarity between analyzed compounds and compound sets (training and test set molecules) in the QSAR models based on similar fragments found in the structure. These QSAR models were prepared with a diverse set of compounds based on experimental evidence of their activity/function on a certain protein of interest, and then tested with validation sets. The "therapeutic activity prediction" feature predicts the potential therapeutic activity of the compounds. The predicted QSAR values (between 0 and 1) greater than 0.5 indicate the potential therapeutic activity. The model performance in MetaCore was evaluated using Cooper statistics parameters: specificity, sensitivity, accuracy and Matthews Correlation Coefficient (MCC). Sensitivity, specificity and accuracy measure correctly predicted positives, correctly predicted negatives, and closeness to the true values, respectively. The closer the value of these coefficients to 1.00, the better the quality of the model. MCC is a correlation between observed and predicted values and its values ranges between -1.00 to +1.00 (+1.00, perfect prediction and -1.00, inverse prediction). The correlation coefficient (R^2) and root mean squared error (RMSE), must be also checked which may be linked with the accuracy of a model where a higher R^2 and low RMSE indicates higher model accuracy. The QSAR model with the highest specificity, sensitivity, accuracy and the MCC was selected in MetaDrug for each particular activity or toxicity tested. The prediction of a therapeutic activity or toxic effect is calculated based on the ChemTree ability to correlate structural descriptors to that property using the recursive partitioning algorithm. The default ChemTree parameters in MetaCore/MetaDrug that gave the best results were used (i.e., path length 5, the maximum number of segments in a multiway

split 3, p-value threshold Bonnferoni 0.99, p-value multiway split 0.99, and number of random trees 50). The training set used in MetaCore/MetaDrug includes molecules that possess the property (positives) and chemicals that do not have such property (negatives) in approximately equal numbers. The marketed drugs were used if their number was greater than 100 in the disease QSAR models; drug candidates in clinical trials and preclinical compounds with *in vivo* activity have been added to the training set. The drugs that have been annotated to cause a particular toxic effect were used for the prediction of toxic effects (Kanan, T. *et al.*, 2019).

MetaCore/MetaDrug uses developed binary QSAR models for therapeutic activity prediction as well as pharmacokinetic and ADME/toxicity profile predictions for the investigated molecules. For this aim, a small molecule library (Specs-SC) that have 212520 available drug-like compounds as well small molecules extracted from available literature were submitted to MetaCore/MetaDrug platform. The "therapeutic activity prediction" feature predicts the potential therapeutic activity of the chemical. The predicted QSAR values (between 0 and 1) greater than 0.5 indicate the potential therapeutic activity, however in this study, we used a higher threshold value (≥0.80) for the "cancer" therapeutic activity using MetaCore^{TM}/MetaDrug^{TM} Cancer QSAR model with 1.00 being the highest value. 5356 molecules had the cancer therapeutic activity values equal or higher than 0.8. These molecules were then subjected to 26 different common toxicity QSAR models included in MetaCore^{TM}/MetaDrug^{TM} platform and non-toxic 250 compounds were selected for structure-based studies.

3.2 PROTEIN PREPARATION

Crystal and NMR structures of the BCL-2 were retrieved from the Protein Data Bank, (PDB entries: 4LXD (Souers, A.J. *et al.*, 2013) and 1YSW (Oltersdorf, T. *et al.*, 2005), 1YSI (Oltersdorf, T. *et al.*, 2005), 2O2F (Bruncko, M. *et al.*, 2007). The missing atoms of the proteins, such as hydrogen and side chains atoms were added and the water molecules within the catalytic domains were kept using protein preparation wizard of Schrodinger suite (Madhavi Sastry, G. *et al.*, 2013). The systems were assigned at a physiological pH by PROPKA code (Søndergaard, C.R. *et al.*, 2011), and finally in order to relax the proteins, the side chains atoms were minimized with the (OPLS2005)

forcefield method (Banks, J.L. *et al.*, 2005). The binding sites of the BCL-2 were verified based on the co-crystallized ligands and these enzymatic sites together with the water molecules were considered as docking boxes in the simulations.

3.3 LIGAND PREPARATION

All the screening compounds used in this study were downloaded from from Specs SC database (212520 molecules) (Specs SC database). All the downloaded molecules were prepared using LigPrep module of Maestro molecular modeling suite with the OPLS2005 forcefield. The ions, small molecules that were used to help crystallization and water molecules that were not in the vicinity (> 5 Å) of co-crystallized ligand were removed. PROPKA (Søndergaard, C.R. *et al.*, 2011) was used to adjust protonation states of amino acids and Epik (Shelley, J.C. *et al.*, 2007) was used to predict ionization and tautomeric states of the co-crystallized ligand at the physiological pH of 7.4.

3.4 MOLECULAR DOCKING SIMULATIONS

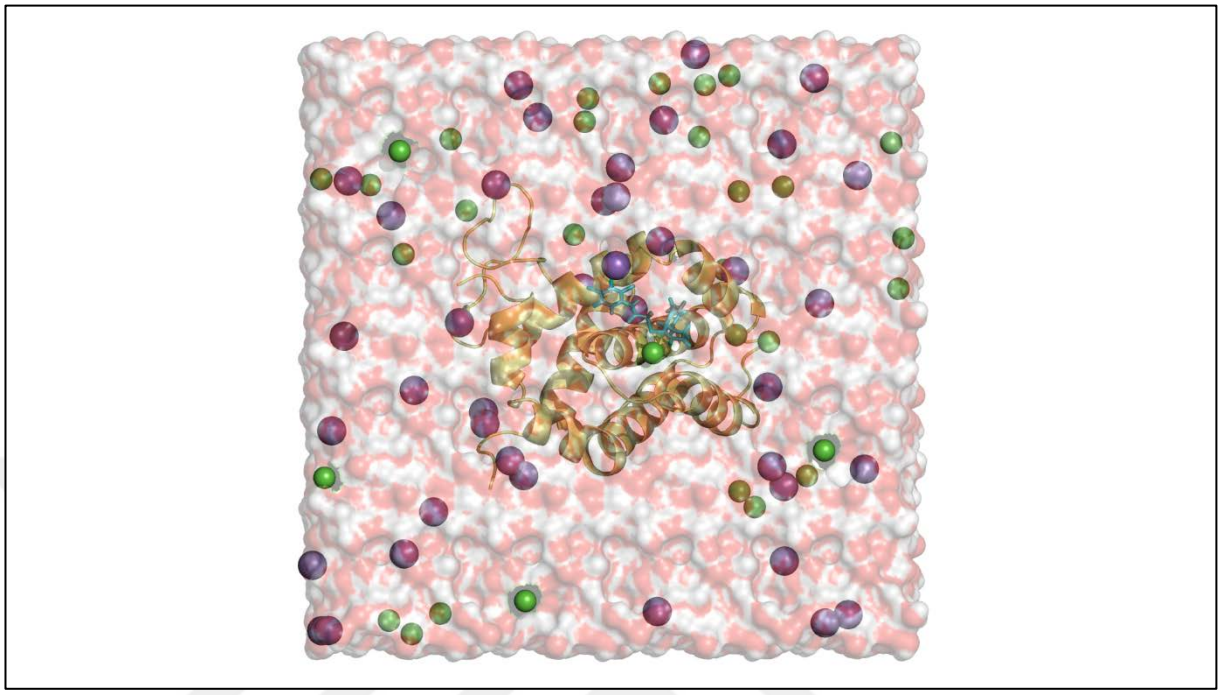
Algorithms used in the docking studies included Glide's standard precision (SP) (Friesner, R.A. *et al.*, 2004, Halgren, T.A. *et al.*, 2004), extra precision (XP) (Friesner, R.A. *et al.*, 2006) and Induced Fit Docking (IFD) (Sherman, W. *et al.*, 2006) modules in Maestro with flexible ligand sampling. A receptor grid box was generated, and certain amino acids were allowed to rotate their side chains to add flexibility to the used target. IFD method comprises 3 consequent stages, including (i) docking of the compounds while the receptor is rigid; (ii) refining the complex amino acids within 5Å of the ligand using Prime module (Jacobson, M.P. *et al.*, 2002, Jacobson, M.P. *et al.*, 2004) of the Schrodinger Suite; (iii) finally redocking of the compounds inside the refined domains. We also used GOLD docking program (Jones, G. *et al.*, 1997) that is based on a genetic algorithm for docking flexible ligands into protein binding sites. Water molecules in the binding pocket were allowed to switch on and off and to rotate around their three principal axes during the docking. The binding pocket was shaped by residues within 10 Å vicinity of co-crystallized ligand. Ligand flexibility were achieved by: (i) flipping ring corners and utilizing ring conformations extracted from the Cambridge Structural Database (CSD); (ii) flipping amide bonds; (iii) flipping pyramidal nitrogen atoms; (iv) flipping

planar trigonal nitrogen atoms while rotating ring-NHR or ring-NR_1R_2 groups; (v) rotating protonated carboxylic acid groups; and (vi) enabling default torsion angle distributions extracted from the CSD. For each ligand 100 docking poses were requested. Default settings were used for population (population size, 100) and genetic operations (number of operations, 100000) steps.

3.5 MOLECULAR DYNAMICS (MD) SIMULATIONS

The top-IFD poses of the selected compounds from the lowest binding energies were submitted to the MD simulations. The complexes were hydrated in the cubic box with explicit TIP3P water molecules that have 10 Å thickness from surfaces of protein. The MD simulations were performed for apo and holo systems of both enzymes by Desmond code (Bowers, K.J. *et al.*, 2006). The particle mesh Ewald method (Darden, T. *et al.*, 1993) was used to calculate the long-range electrostatic interactions. A cut-off radius of 9 Å was used for both van der Waals and Coulombic interactions. Nose-Hoover thermostat (Hoover, W.G., 1985, Nosé, S., 1984) and Martyna-Tobias−Klein protocols (Martyna, G.J. *et al.*, 1994) were used to adjust the temperature and pressure of the systems at 310 K and 1.01325 bar, respectively. The time-step was assigned as 2.0 fs. The default values were used for minimization and equilibration steps and finally 100-ns production run was performed for each simulation. The details of the simulation protocols were described in the previous studies (Durdagi, S. *et al.*, 2016, Salmas, R.E. *et al.*, 2017).

Figure 3.1: Graphical representation of MD simulation box.



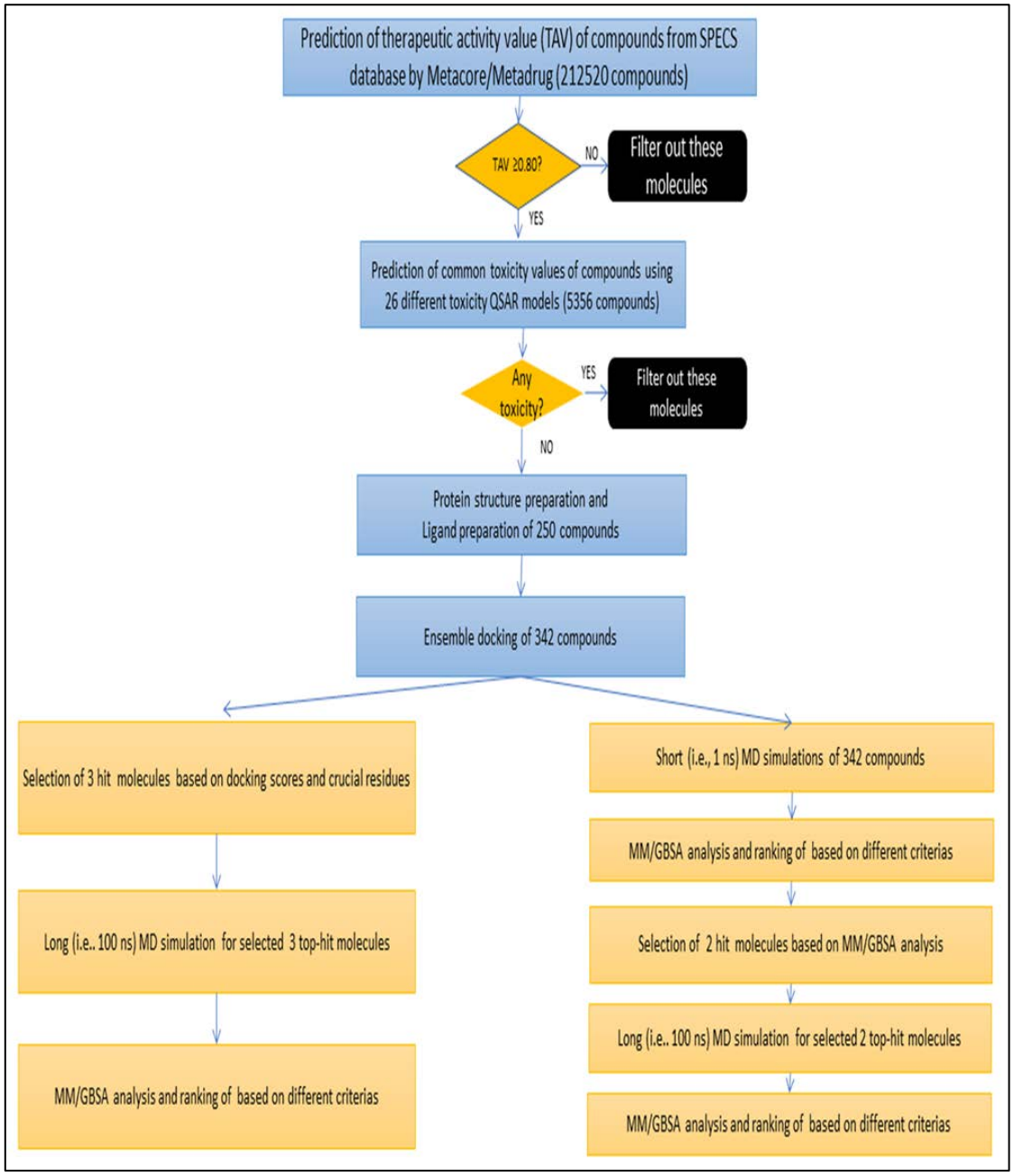
Graphical representation of MD simulation box. Protein is depicted as cartoon in yellow color, ligand is depicted as stick in cyan color, water molecules are depicted as surface representation. Na^+, green; Cl^-, purple. (Figure 3.1)

3.6 MOLECULAR MECHANICS/GENERALIZED BORN SURFACE AREA (MM/GBSA) CALCULATIONS

Schrodinger's Prime module was used to execute molecular mechanics generalized Born surface area (MM/GBSA) binding free energies calculations of selected ligand-protein complexes (Onufriev, A. *et al.*, 2004) 100 trajectory frames were chosen from the last 50-ns (50ns-100ns) MD simulations (i.e., 1 frame/1 ns). For the prediction of free binding energies of complexes, OPLS2005 forcefield and VSGB 2.0 solvation model (Li, J. *et al.*, 2011) were applied.

All applied procedure for virtual screening in this study has been summarized in Scheme 3.2.

Figure 3.2: Applied virtual screening workflow.



4. RESULTS AND DISCUSSION

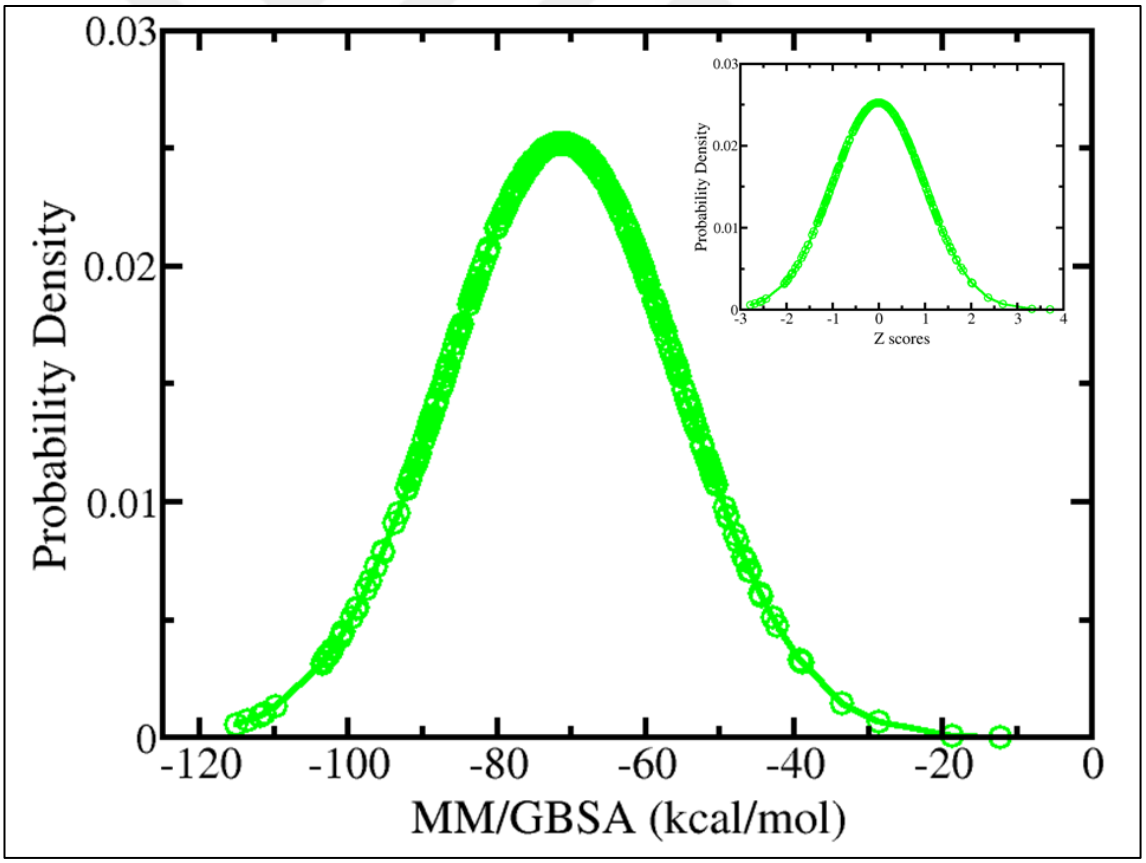
In this study, we virtually screened 212520 compounds that are drug-like small molecules from Specs-SC database. These drug-like molecules are available at the database. In the first step of the screening study, the molecules were filtered based on their therapeutic activity values (TAV) predicted by MetaCore^{TM}/MetaDrug^{TM} binary QSAR models. As we only kept compounds that have predicted activity values of at least 0.80 and higher, 5356 molecules were forwarded for next step. These selected molecules were used in 26-common toxicity QSAR models (i.e., cardiotoxicity, neurotoxicity, nefrotoxicity, hepatotoxicity, liver necrosis, kidney necrosis, etc.) of MetaCore^{TM}/MetaDrug^{TM}. Out of 5356 molecules only 250 molecules showed no toxicities in these all 26 QSAR models. Using LigPrep, the 3D structures of selected hit compounds were generated. Possible stereoisomers of compounds were also considered if they contain chiral centers. Since our database contained drug-like molecules, their chiral centers were mostly well-defined, in case that there were no defined chiral centers for a molecule, we have retained all possible stereoisomers. At the end, there were 342 structures in total, with the possible tautomeric and ionic states of the selected hit compounds for docking studies. These selected molecules as well as known 20 BCL-2 inhibitors from literature, as positive control molecules, were used in docking simulations. After ligand preparation, together with possible tautomeric and protonation states, the number of known BCL-2 molecules number reached to 54. Thus, in total 396 molecules (342 molecules from screening and 54 molecules from known BCL2 inhibitors) were considered in docking. All molecules were docked using Glide/SP and GOLD docking programs. In docking, instead a single docking target protein, 3 BCL2 target proteins obtained by NMR (i.e., PDB IDs: 1YSW, 2O2F, and 1YSI) as well as one target obtained by X-Ray diffraction method (i.e., PDB ID: 4LXD) were considered. Based on docking scores, 3 molecules (compound **43** (Specs ID: AO-081/41887762; compound **58** (Specs ID: AJ-292/12931005); and compound **243** (Specs ID: AN-698/40780701) and a reference molecule for BCL2 (3703601) were used in Glide/IFD docking and molecular Dynamics (MD) simulations.

In the selection of hit molecules, two strategies were considered (i) selection of molecules based on docking scores and application of long (100-ns) MD simulations to these

selected molecules; (ii) applying short (i.e., 1-ns) MD simulations for all molecules from MetaCore/MetaDrug (i.e., for 342 molecules) and selection of molecules based on average MM/GBSA free energy values.

Hence, we have used the top-docking poses for all considered structures and performed short (i.e., 1 ns) MD simulations. The trajectories were written at 10 ps frequency and 100 snapshots/poses were obtained throughout MD simulations. Then, we calculated the binding free energies for all 100 MD trajectories and the average free energies of binding scores were ranked. When we ranked the compounds based on their MM/GBSA energies, two new molecules (compound **258** (Specs ID: AK-968/12163470) and compound **292** (Specs ID: AK-968/11842328) were also identified. (Figure 4.1)

Figure 4.1: The normal distribution curves for MM/GBSA energies.



The inset graphs show the normalized distribution curves and Z-scores.

While Table 4.1 shows the molecules that have high docking scores using Glide docking algorithms (SP and IFD), Table 4.2 shows molecules that have high docking scores obtained by GOLD (ChemScore DG). Although compound #300 was identified in both docking algorithms, it's docking score was not too high especially in GOLD. Based on both docking scores in GOLD and Glide/IFD as well as conserved crucial amino acids, following molecules were considered for further investigations: Compounds **43**, **58** and **243**. Since compounds **258** and **292** were also gave promising MM/GBSA results, we considered these five compounds.

Table 4.1: Selected molecules that show high docking scores in Glide/SP and Glide/IFD. (Values are in kcal/mol)

1YSI (Glide/SP)		1YSW (Glide/SP)		2O2F (Glide/SP)		4LXD (Glide/SP)		4LXD (IFD)		4LXD Known Inhibitors (IFD)	
Molecule ID	Docking Score	Molecule ID	Docking Score	Molecule ID	Docking Score	Molecule ID	Docking Score	Molecule ID	Docking Score	Molecule ID	Docking Score
151	-7.243	37	-7.007	140	-7.951	73	-7.951	243	-13.215	3703601	-16.381
324	-7.138	56	-6.612	276	-7.757	71	-7.901	58	-12.818	3703620	-15.963
296	-7.002	26	-6.313	197	-7.737	197	-7.642	43	-12.556	217354	-15.498
302	-6.922	276	-6.294	118	-7.581	118	-7.508	55	-12.524	3703599	-14.789
35	-6.831	206	-6.267	243	-7.355	243	-7.507	142	-12.457	3703616	-14.317
								258	-10.993		
								292	-10.686		

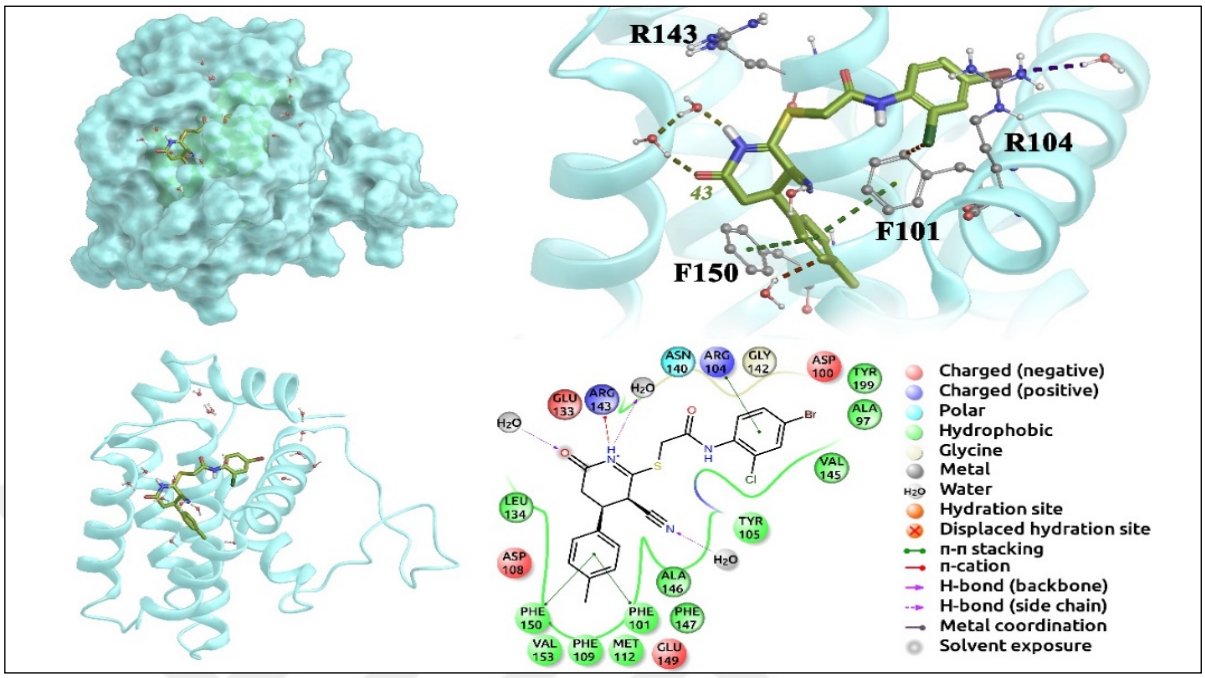
Table 4.2: Selected molecules that show high docking scores in GOLD docking algorithm. (Values are in kcal/mol)

1YSI (GOLD)		1YSW (GOLD)		2O2F (GOLD)		4LXD (GOLD)		4LXD Known Inhibitors(GOLD)	
Molecule ID	Docking Score	Molecule ID	Docking Score	Molecule ID	Docking Score	Molecule ID	Docking Score	Molecule ID	Docking Score
321	-9.493	306	-8.782	300	-6.384	300	-5.932	3703599	-7.336
330	-8.140	300	-7.327	191	-6.153	230	-5.411	2031015	-5.645
306	-6.527	332	-6.636	210	-6.065	320	-5.057	473446	-4.877
298	-6.139	298	-6.600	298	-4.980	299	-4.995	2031025	-4.666
320	-5.373	233	-6.465	192	-4.614	110	-3.955	3703601	-4.143

2D and 3D ligand interaction diagrams

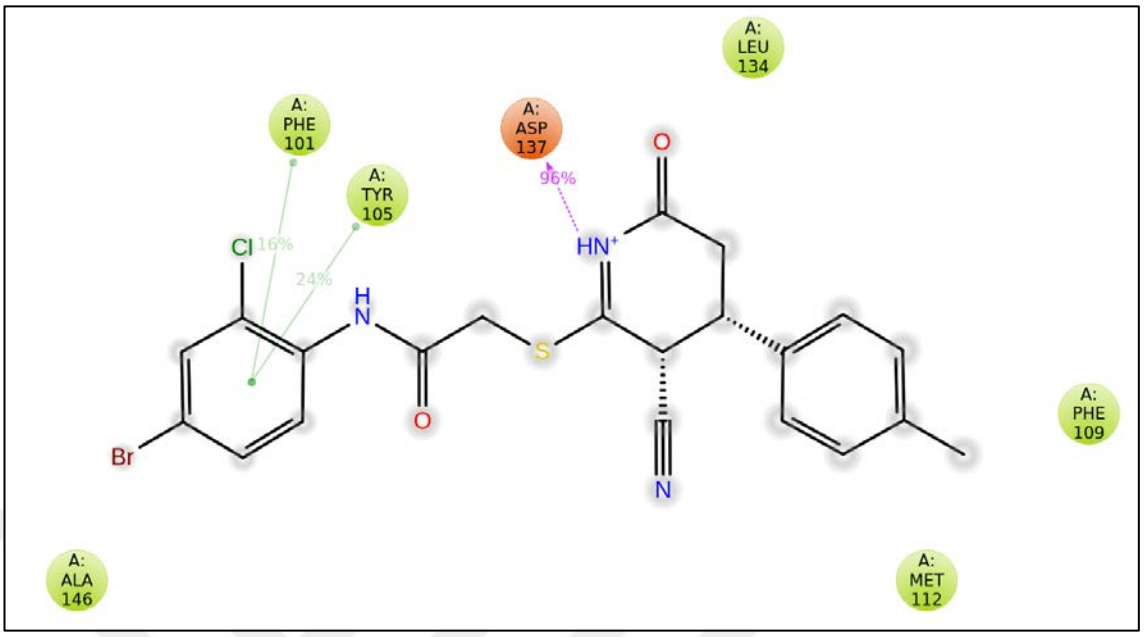
Following section shows 2D and 3D ligand interaction diagrams and crucial residue-ligant interactions for the selected hit molecules using top-docking poses and MD simulations.

Figure 4.2: 3D and 2D ligand interaction diagrams of top-scored IFD pose of compound 43.



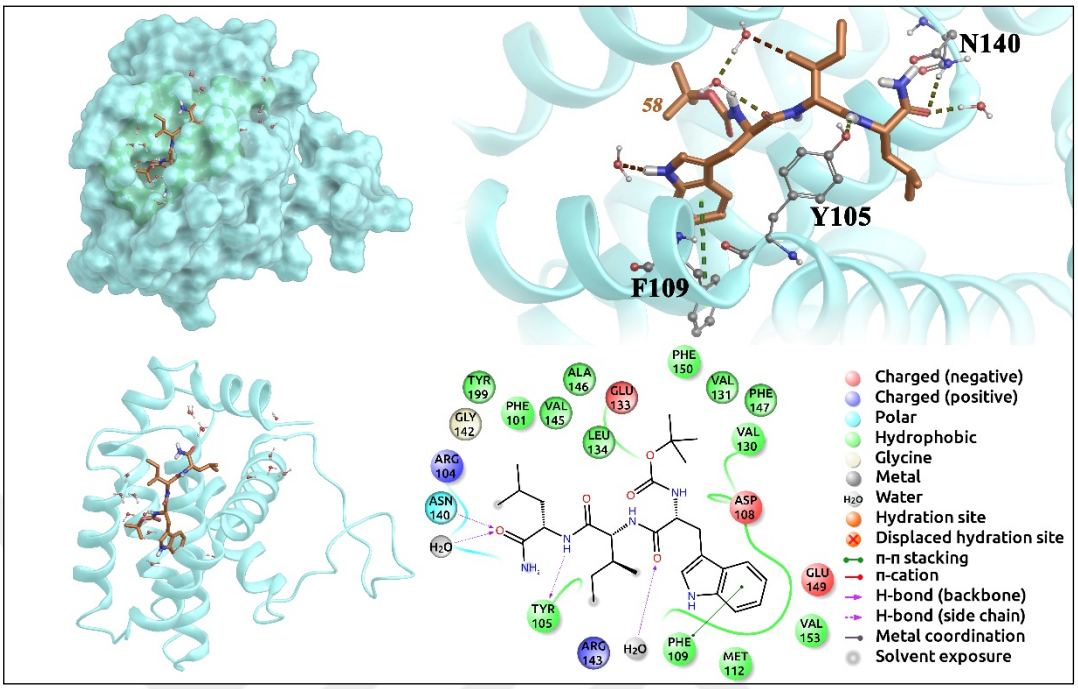
In the initial conformation, (i.e top docking pose from Glide/IFD) three π-π stacking interactions were observed between side chains of PHE101, ARG104 and PHE150 and molecule **43**. **(Figure 4.2)** Also three H-bonding network between crystal waters and molecule **43** was occured. ARG143 made a π-cation interaction between positively charged Nitrogen atom of the molecule **43** (Figure 4.2).

Figure 4.3: 2D ligand interactions diagram of compound 43 from MD stimulations.



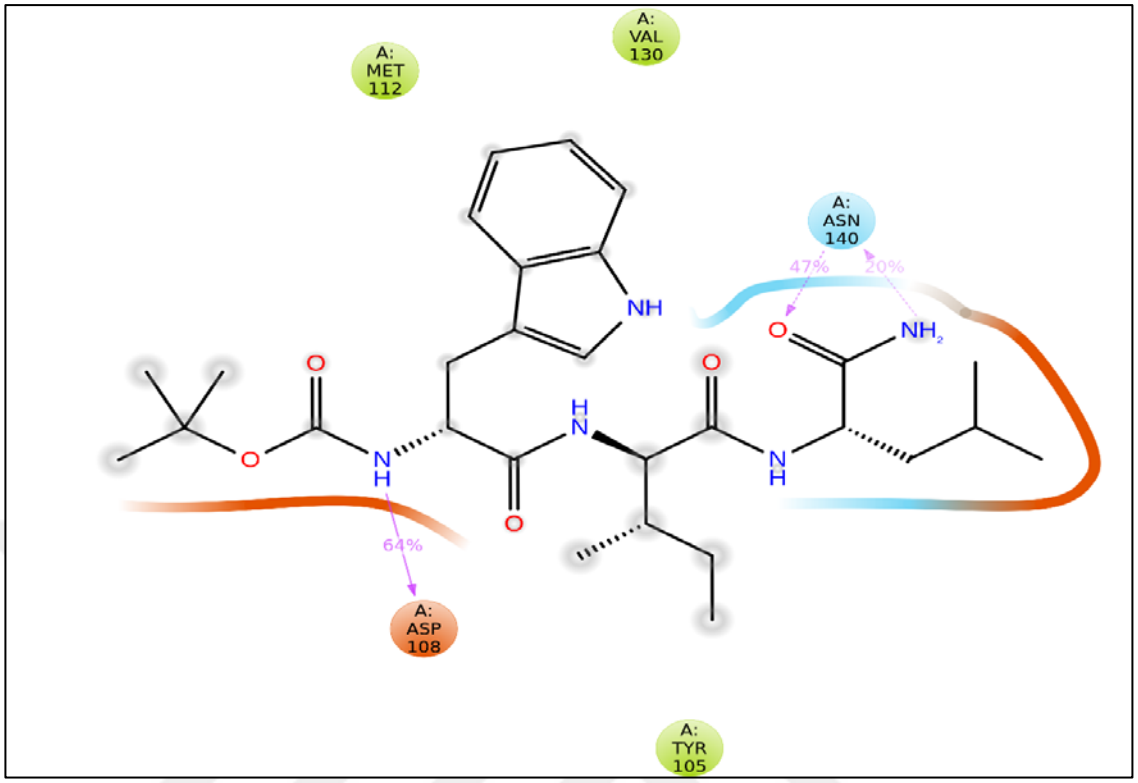
During MD simulations, π-π stacking interactions between PHE101 and the suggested inhibitor molecule is conserved **(Figure 4.3).** All three water H-bonding network, π-π and π-cation interactions between ARG104, PHE150, and ARG143 were not maintained, respectively. (Figure 4.3) However, a strong h-bond between ASP137 and molecule 43 is formed.

Figure 4.4: 3D and 2D ligand interaction diagrams of top-scored IFD pose of compound 58.



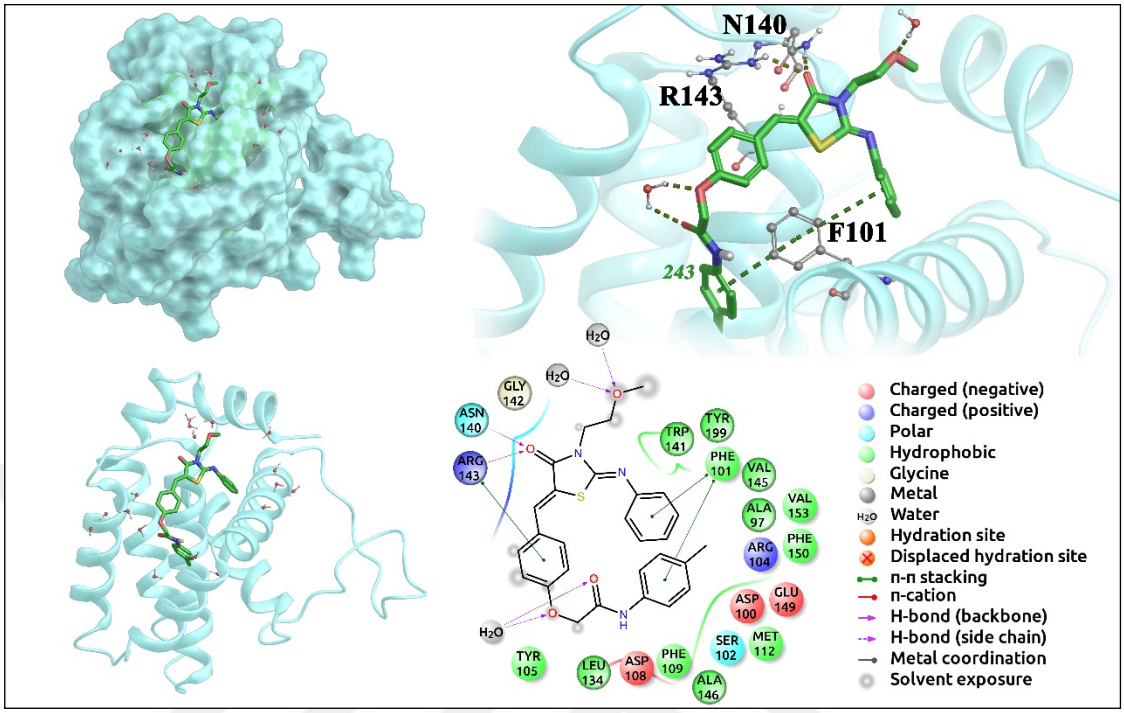
Top docking poses of **58** shows crucial interactions of ASN140 (H-bonding), PHE109 (π-π stacking), TYR105 (H-bonding). (Figure 4.4)

Figure 4.5: 2D ligand interactions diagram of compound 58 from MD stimulations.



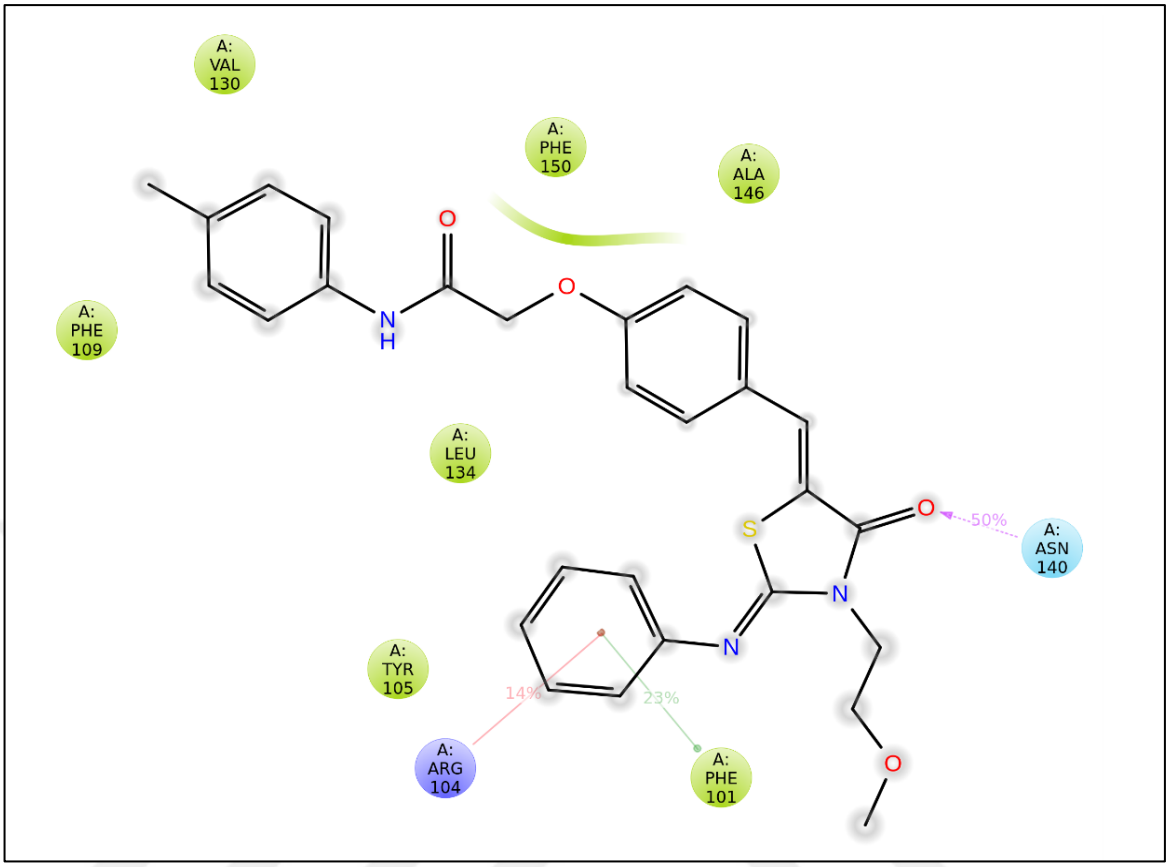
Throughout simulations hydrogen bonds between TYR105 and the ligand was not conserved, instead another hydrogen bonding interaction is constructed with ASP108 during 64% of the simulations time. Two hydrogen bonds were stable through MD simulations with the ASN140 and the terminal carbonyl and amino groups of the ligand. (Figure 4.5)

Figure 4.6: 3D and 2D ligand interaction diagrams of top-scored IFD pose of compound 243.



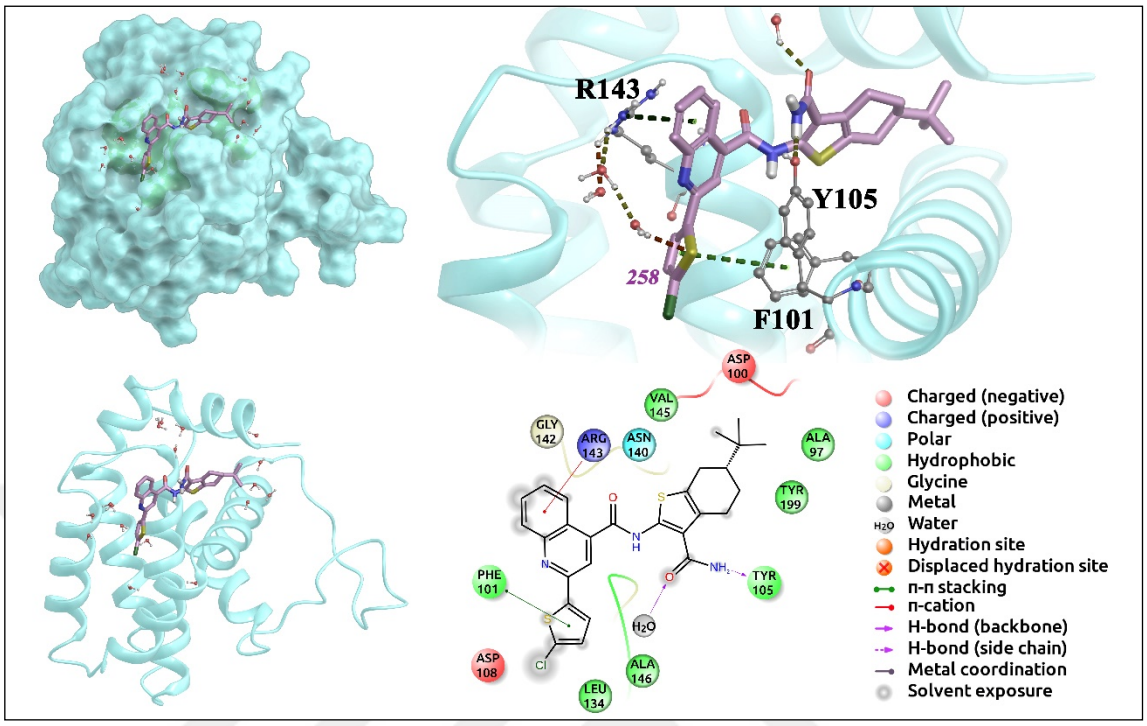
In the docking pose of compound **243**, PHE101 formed two π-π stacking interactions. ARG143 contributed to the binding of molecule **243**, via one π-π stacking interaction and one hydrogen bond. Also, one H-bond was observed between molecule **243** and ASN140. Three water molecules were contributed to the H-bond network via 4 H-bonds. (Figure 4.6)

Figure 4.7: 2D ligand interactions diagram of compound 243 from MD simulations.



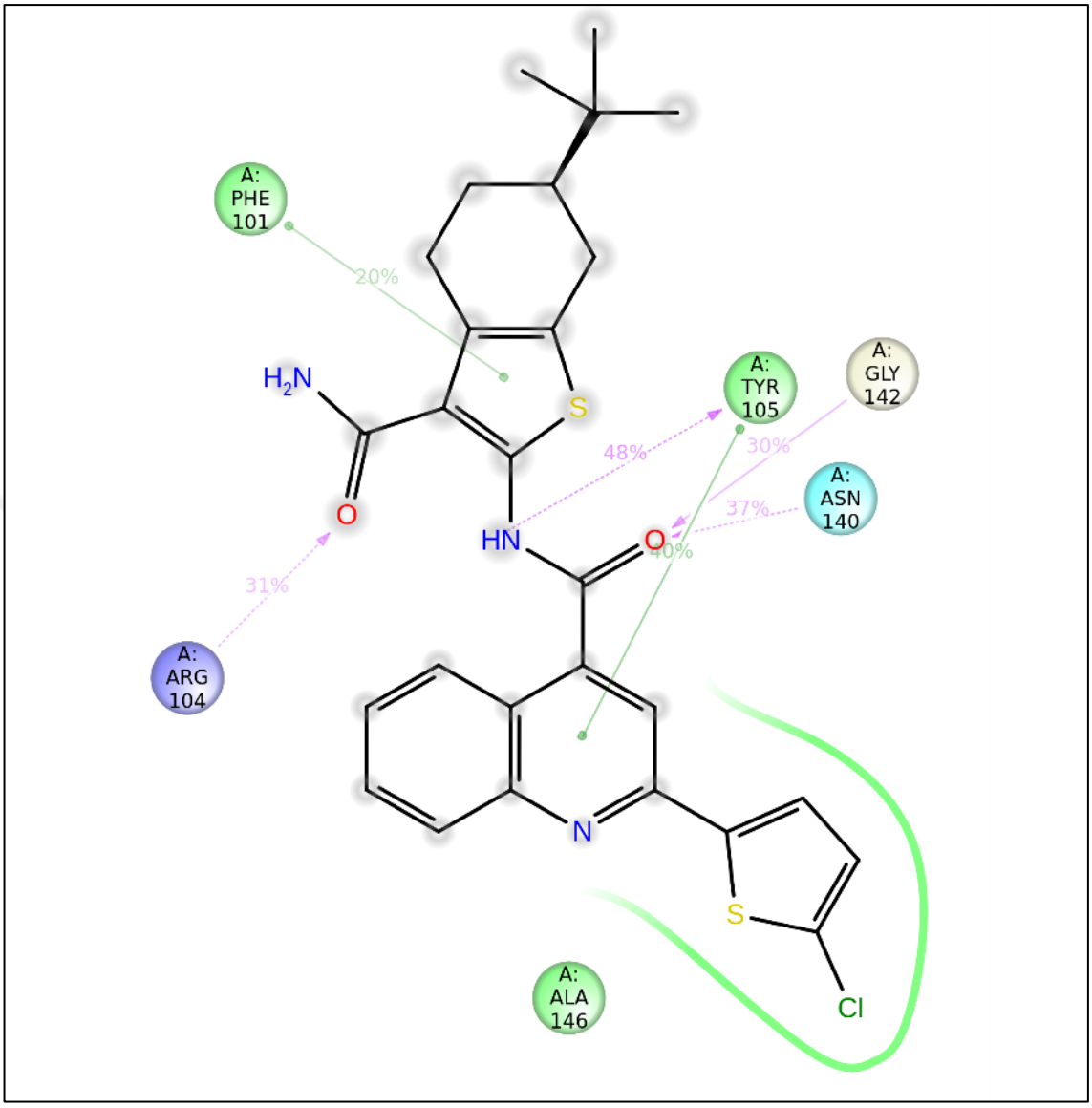
H-bond between ASN140 and π-π interaction between PHE101 and suggested inhibitor was maintained during MD simulations. However, a new weak π-cation interaction was formed between ARG104 and molecule **243** (conserved %14). All other unspecified interactions were not maintained in the course of MD simulations. (Figure 4.7)

Figure 4.8: 3D and 2D ligand interaction diagrams of top-scored IFD pose of compound 258.



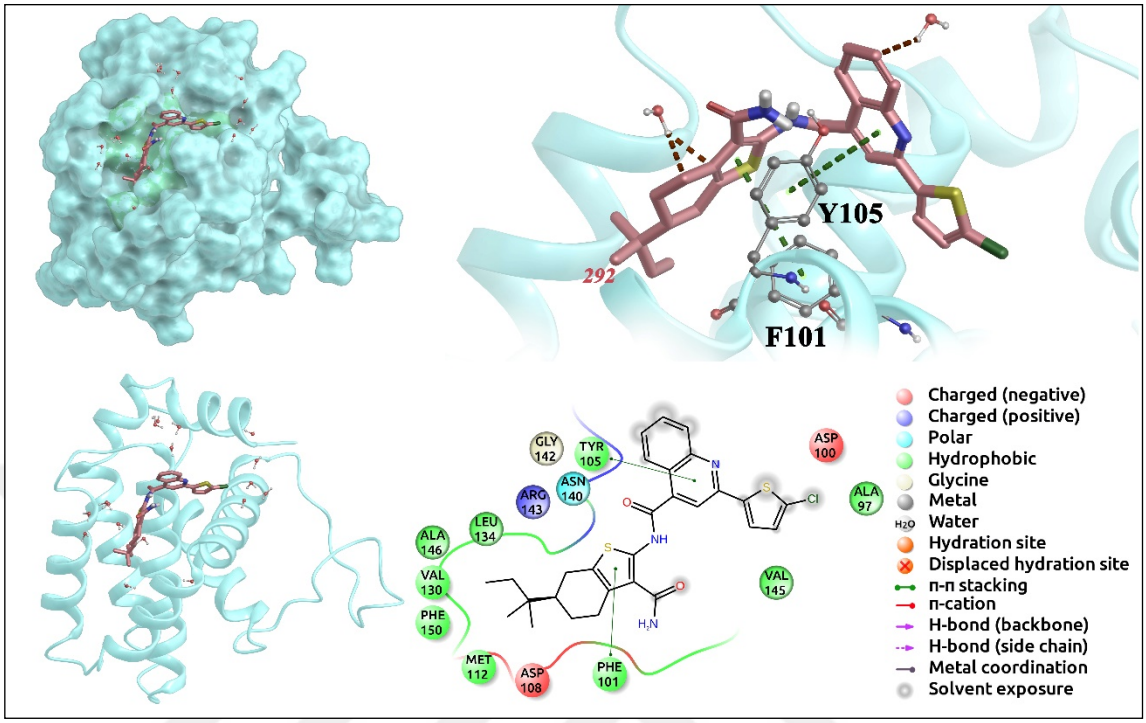
In the top-docking pose of molecule **258**, π-π stacking interactions between PHE101 and ARG143 were observed, and H-bonds were formed with following residues: TYR105, ARG143. (Figure 4.8)

Figure 4.9: 2D ligand interactions diagram of compound 258 from MD simulations.



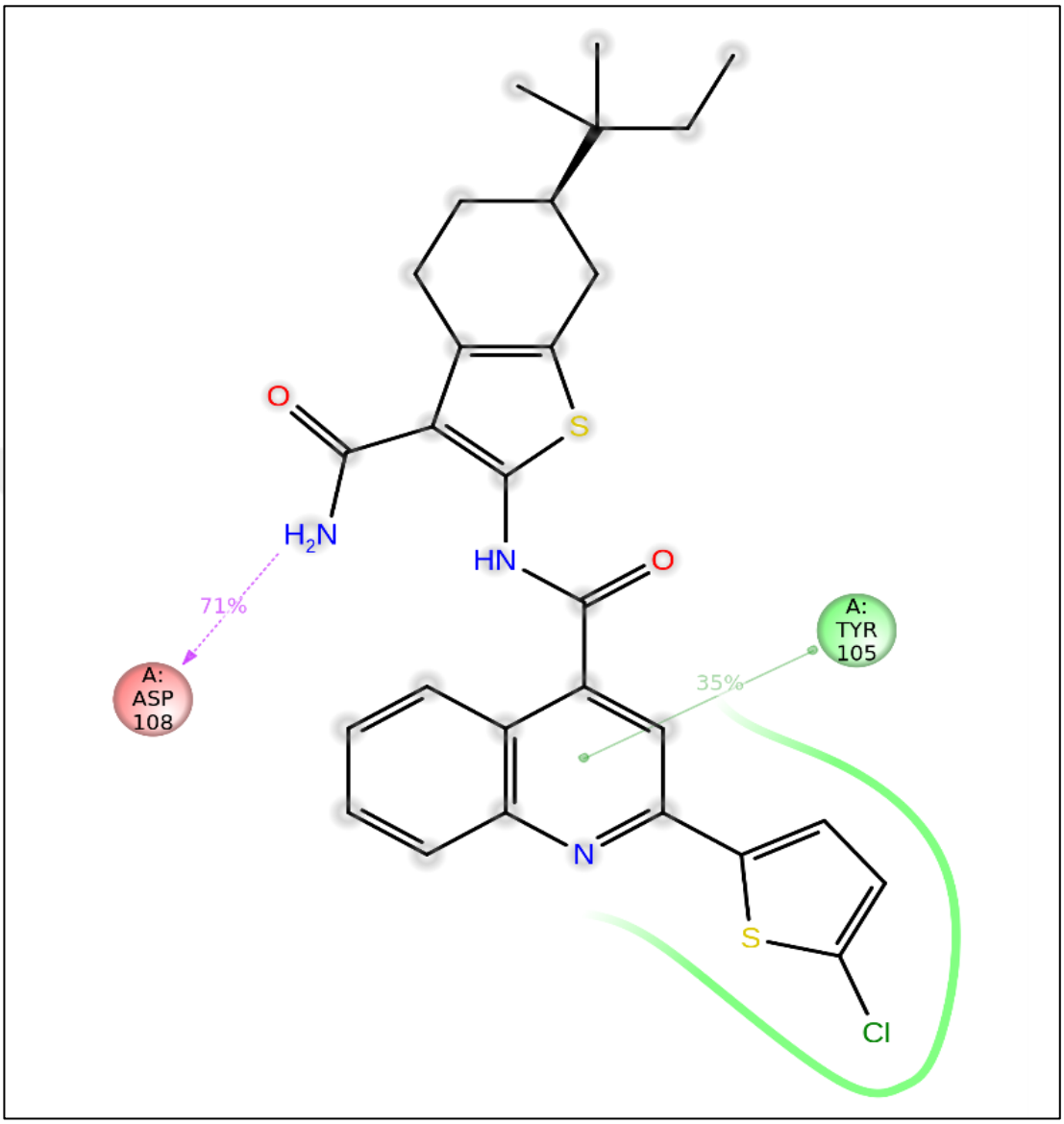
During MD simulations, only π-π stacking interaction between PHE101 and H-bond between ASN140 and molecule **258** were maintained, all other bonding interactions were lost. Interestingly, ARG104, TYR105 and GLY142 (which is critical in 4-(4-{[4-(4-chlorophenyl)-5,6-dihydro-2H-pyran-3-yl]methyl}piperazin-1-yl)-N-{[3-nitro-4-(tetrahydro-2H-pyran-4-ylamino)phenyl]sulfonyl}benzamide binding to BCL2 in 4LXD crystal structure) formed new three H-bonds and one π-π stacking interaction between molecule **258** and BCL2. (Figure 4.9)

Figure 4.10: 3D and 2D ligand interaction diagrams of top-scored IFD pose of compound 292.



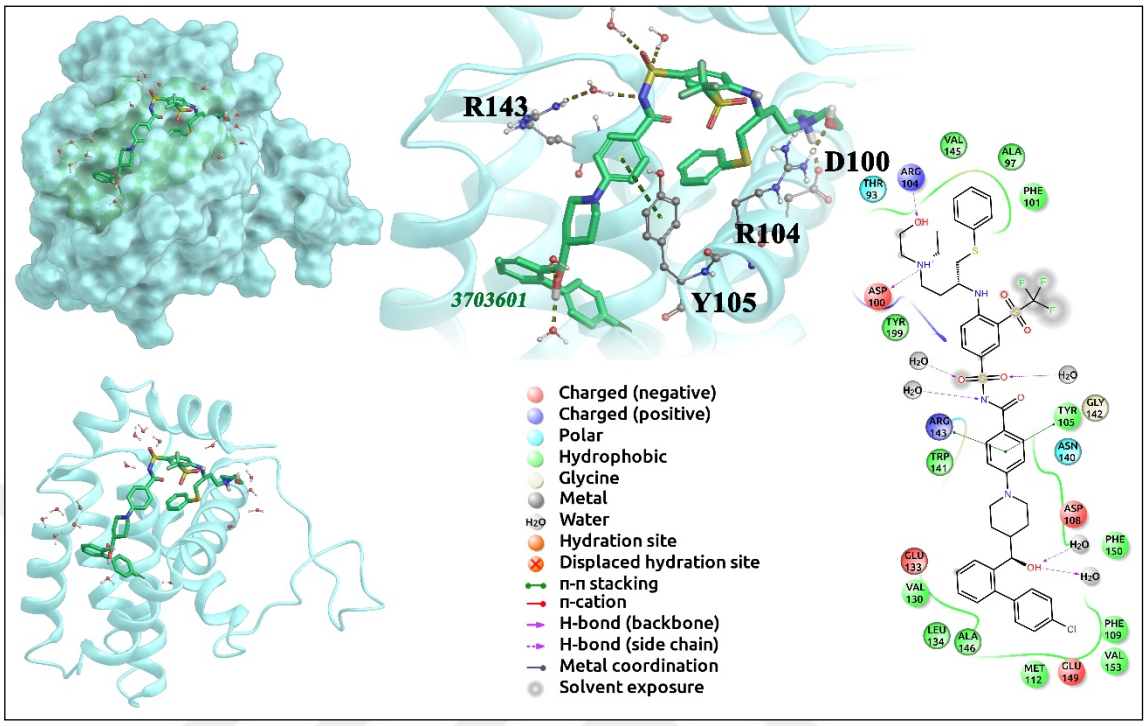
In the top-docking pose of molecule **292**, three H-bonds were observed between the suggested molecule and the protein. TYR105 was contributed to binding of molecule **292** via π-π stacking interactions. (Figure 4.10)

Figure 4.11: 2D ligand interactions diagram of compound 292 from MD simulations.



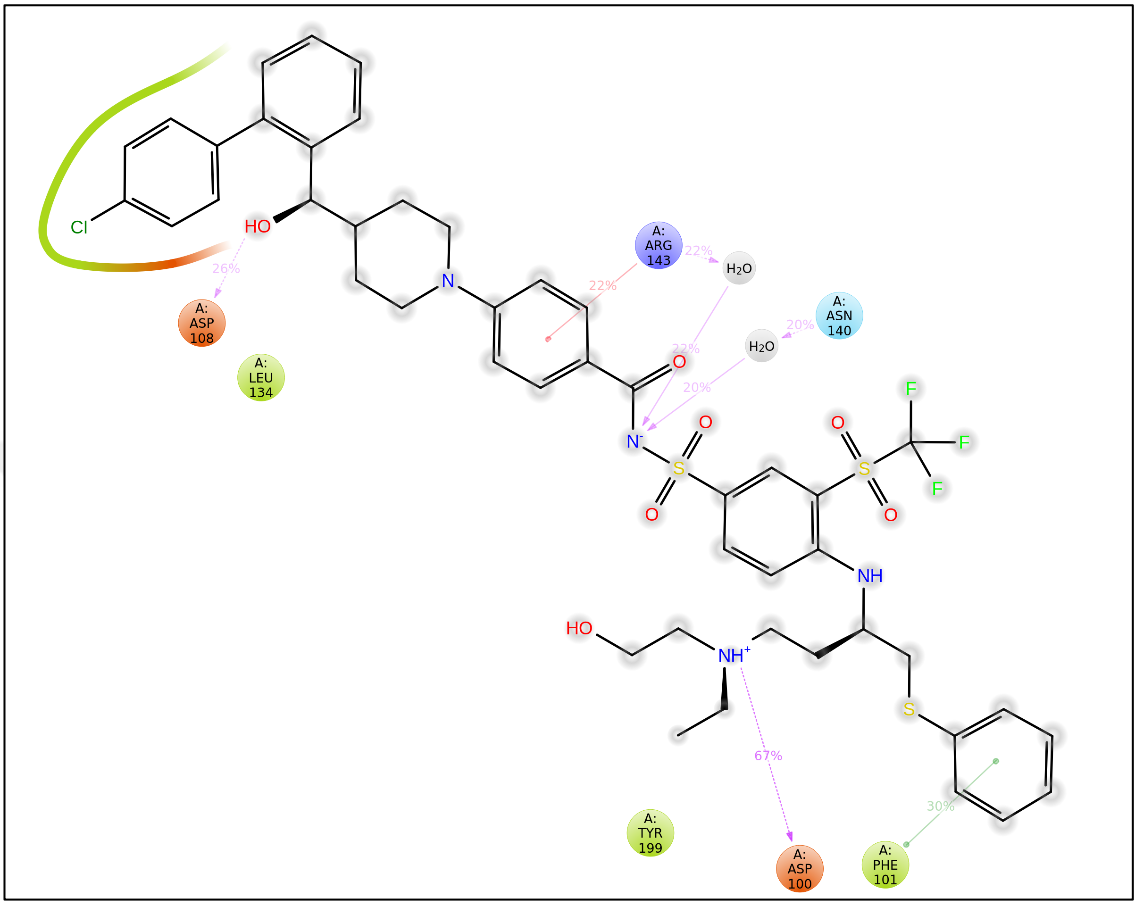
During MD simulations one strong H-bond was formed between ASP108 and BCL2. The π-π stacking interaction of TYR105 was maintained at the %35 of the time course of MD simulations. (Figure 4.11)

Figure 4.12: 3D and 2D ligand interaction diagrams of top-scored IFD pose of compound 3703601.



In the top-docking pose of the control molecule (known inhibitor - 3703601), two H-bond between ASP100 and ARG104 were formed with the inhibitor. Also, two π-π stacking interactions were observed between TYR105 and ARG143. Five different water molecules were contributed to h-bond network around the known inhibitor via 5 H-bonds. (Figure 4.12)

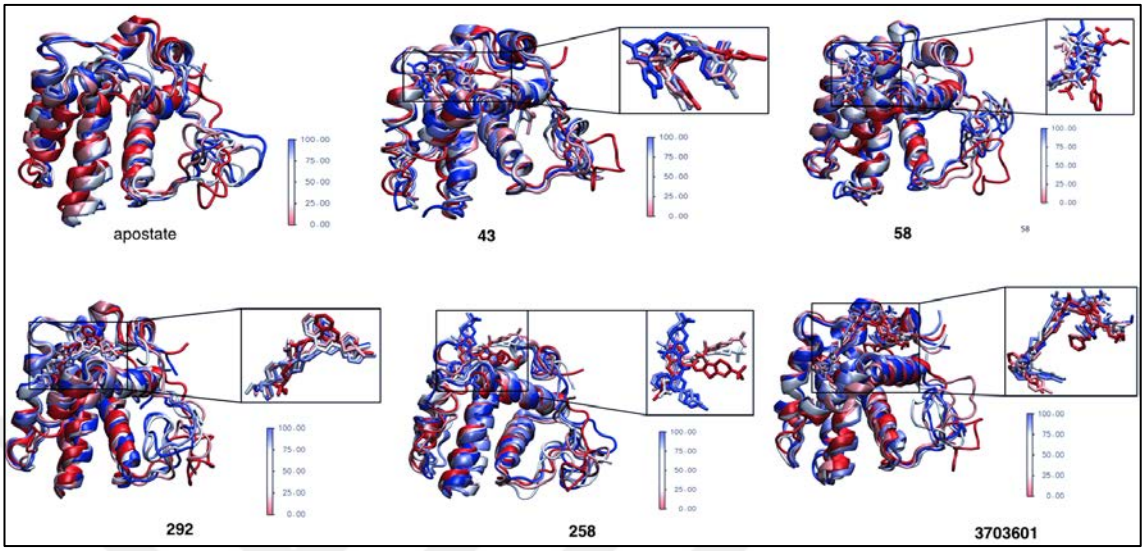
Figure 4.13: 2D ligand interactions diagram of compound reference molecule 3703601 from MD simulations.



During MD simulations, π-π stacking interactions between PHE101 and H-bond between ASP100 were conserved. Three new additional (one π-cation and two H-bonds) were formed between the control molecule and ARG143, ASP108 and ASN140, respectively. All 5 H-bonds were lost, but two new additional H-bonds were formed between the inhibitor and water molecules. (Figure 4.13)

In order to investigate the effect of the small molecule inhibitors to the target structure, ligand-bound complex structures were also compared with apo-state of the enzyme. Figure 4.14 represents color-scale changes (red initial of the simulation and blue end of the simulation) throughout the MD simulations.

Figure 4.14: Conformational changes of the apo-state of the enzyme, and hit ligands (43, 58, 292, and 258) as well as reference molecule (3703601) bound molecules throughout the MD simulations.



In the apo state as expected the higher fluctuations were observed at the mainly linker regions. Figure 4.14 also shows corresponding conformational changes when compound **43** is bound to the binding pocket. It is clear that together with binding fluctuation of the enzyme at the binding pocket as well as periphery of the target is decreased. Inhibitor keeps relatively similar conformations during the MD simulations. The most fluctuated regions of the ligand **43** is the terminal aromatic rings. Ligand shows higher fluctuations (both rotational and translational) compared to the **43** during the compound **58**-bound systems MD simulations. Figure 4.14 also shows which structural and dynamical changes happen throughout the MD simulations when compound **243** binds to the active site of the BCL-2. It shows structural stability at the binding pocket. The most fluctuated region of the hit molecule **243** is the terminal aromatic ring. Compounds **292** and reference molecule do not change their rotational and translational motions significantly at the binding pocket. However, compound **258** has very dramatic structural changes at the binding pocket throughout the MD simulations.

Root-Mean-Square-Deviation (RMSD) is a measurement for the average change in distance / displacement of the atoms of protein (selection of all-atoms, backbone, sidechain or alpha carbon atoms, etc.) or ligand for a particular trajectory frame with respect to a reference frame. For example, the RMSD for the frame x is:

$$RMSD_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(r_i'(t_x)\right) - r_i(t_{ref}))^2}$$

Where *N* is the number of selected atoms; t_{ref} is the reference time (typically the first frame is used, and it is regarded as time t=0); and r′ is the position of the selected atoms in frame x after superimposing on the reference frame, where frame x is recorded at time t_x. This procedure is repeated for each trajectory frames. The RMSD value is expressed in length, and the unit is mostly Ångstrom (Å). Figure 4.15 shows root-mean square deviation (RMSD) graphs of studied proteins during the simulations based on alpha carbons. As it can be seen in the Figure, RMSD plots do not change significantly after 20-ns. Additionally, we have checked the RMSD of the ligand molecules by considering two different fitting modes: 'fit on protein/profit' and 'fit on ligand/ligfit. While the first mode aligned indicates the structural stability of ligand with respect to protein i.e. its translational motion, the second mode shows the internal fluctuations of the ligand atoms in its binding pocket i.e. its rotational motion. The RMSD of C_α atoms for PR-ligand complexes could be found in Figure 4.16. As it can be seen from figures, the most stable structures (ligands do not diffuse at the binding pocket) were hit compounds **43**, **292**, and **243**. Reference molecule also states stable at the binding pocket.

Figure 4.15: Protein RMSD graph of the Ca atoms of the BCL-2 obtained from 100 ns MD simulations.

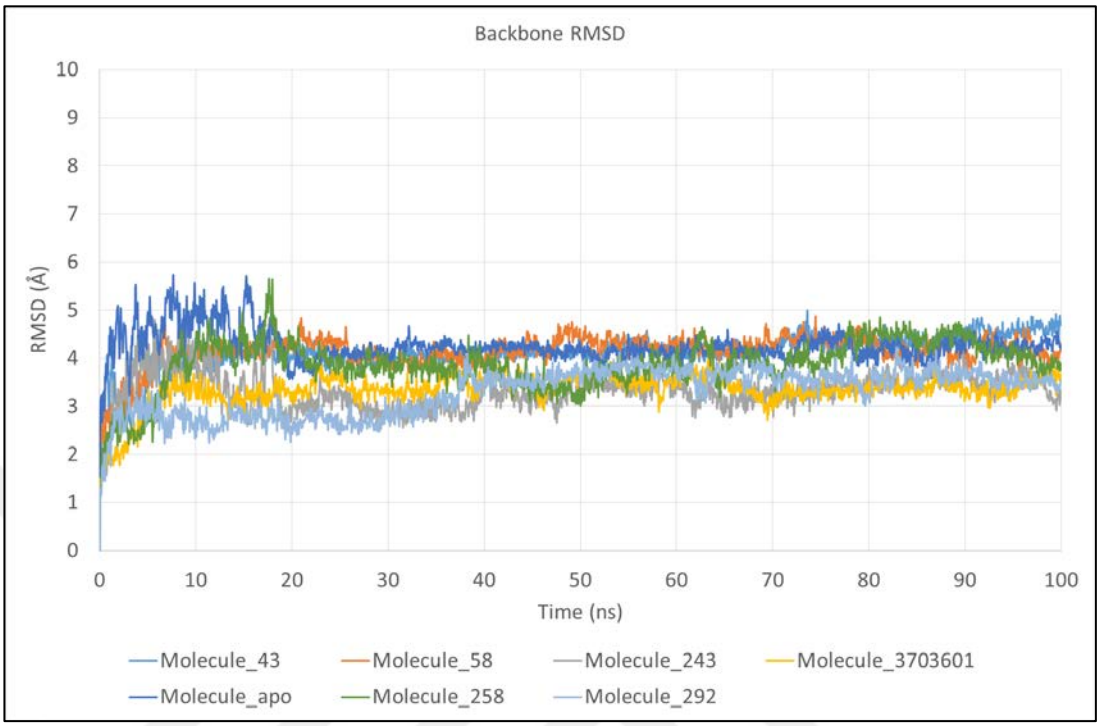


Figure 4.16: ProtFit RMSD graph of the hit and reference compounds.

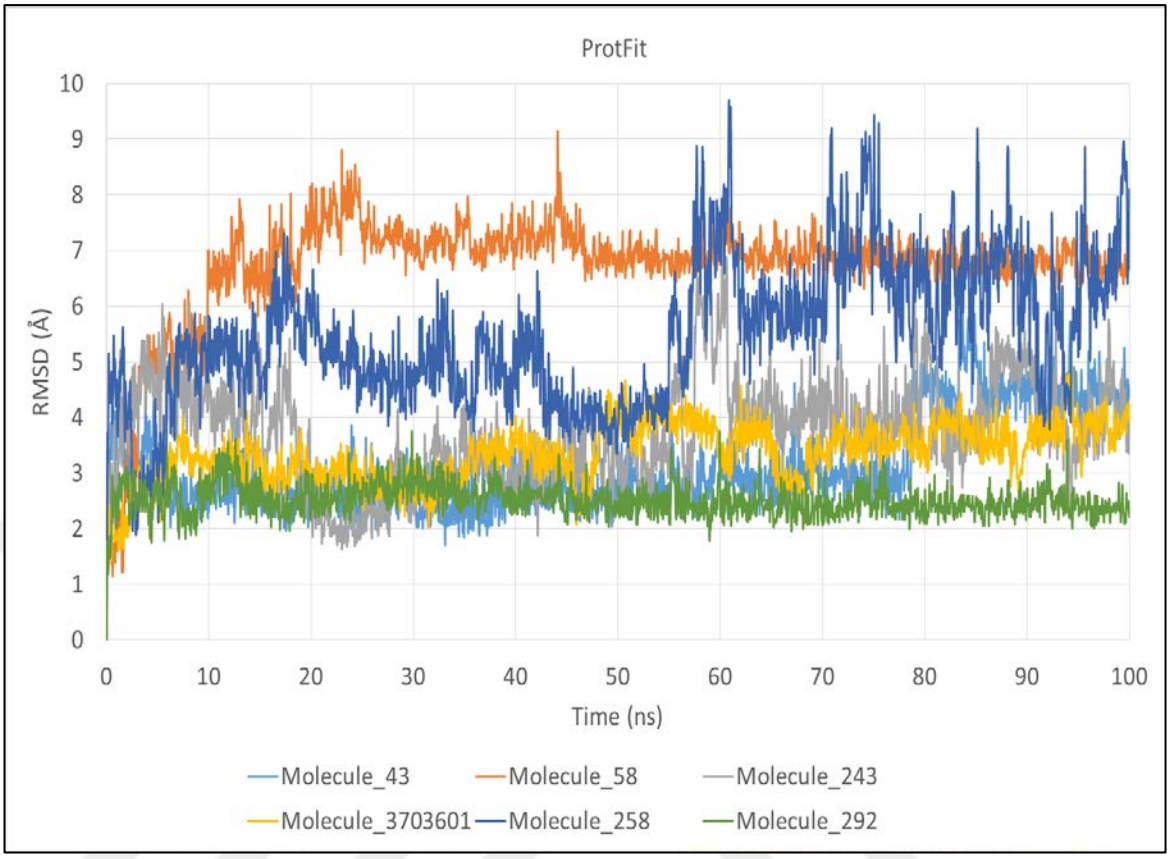


Figure 4.17: LigFit RMSD graph of the hit and reference compounds.

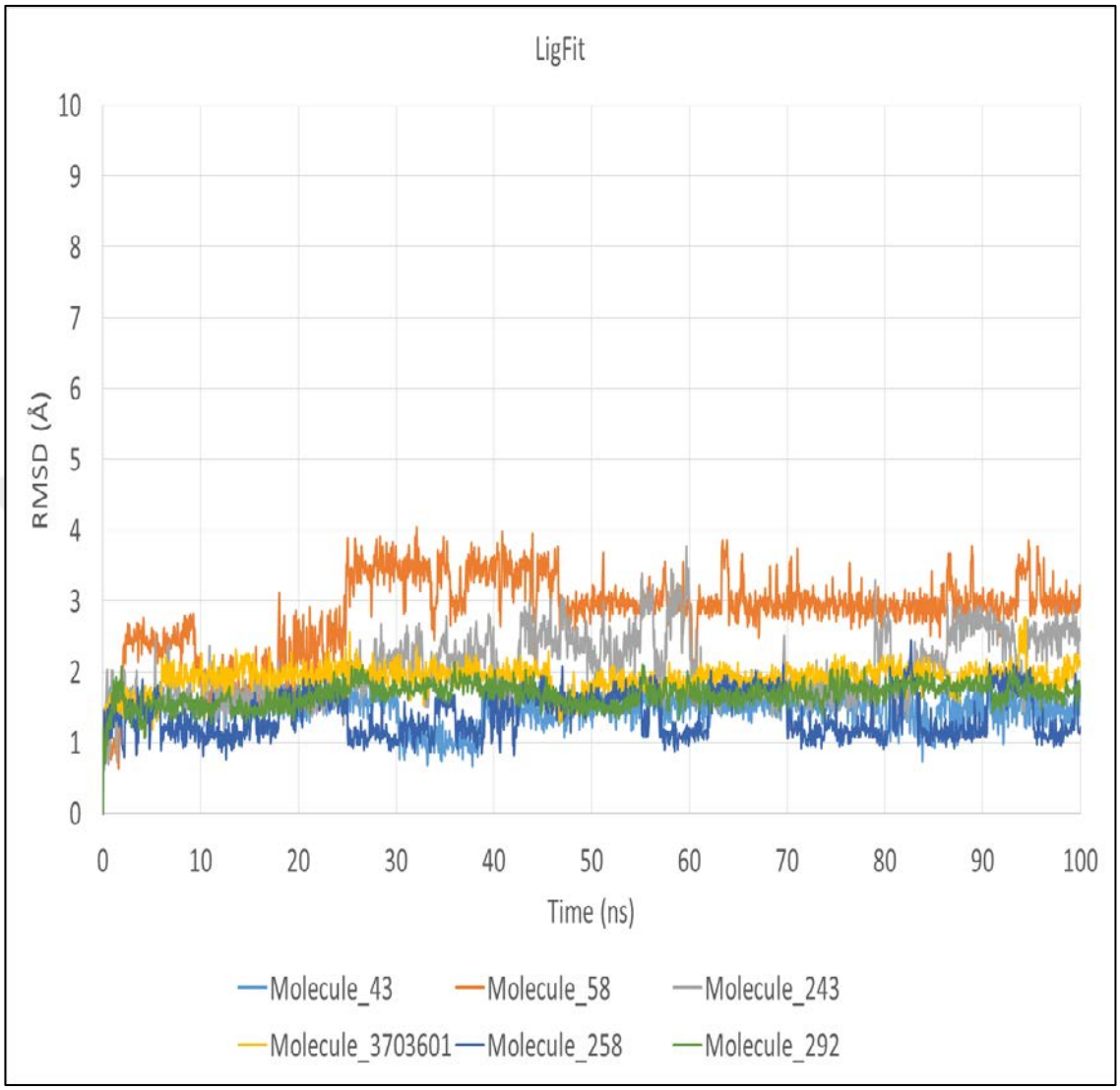


Figure 4.17 shows rotational motions of studied ligands at the binding pocket during the simulations. The most fluctuated molecules was **58** which has maximum RMSDs around 3-4 Å.

RMSD from the average over time can be referred to as the Root-Mean-Square-Fluctuation (RMSF) when a dynamical system fluctuates about some well-defined average position. The size of this fluctuation can be measured, and it is useful for characterizing local changes along the protein chain. The RMSF for residue *i* is:

$$RMSF_i = \sqrt{\frac{1}{T}\sum_{i=1}^{T} < (r_i'(t)) - r_i(t_{ref}))^2 >}$$

where T is the trajectory time over that the RMSF is calculated, t_{ref} is the reference time, r_i is the position of residue i, r_i' is the position of the atoms of residue i after superposition on the reference, and the angle brackets indicate the average of the square distance is taken over the selected atoms of the residue.

Figure 4.18: Protein RMSF graph.

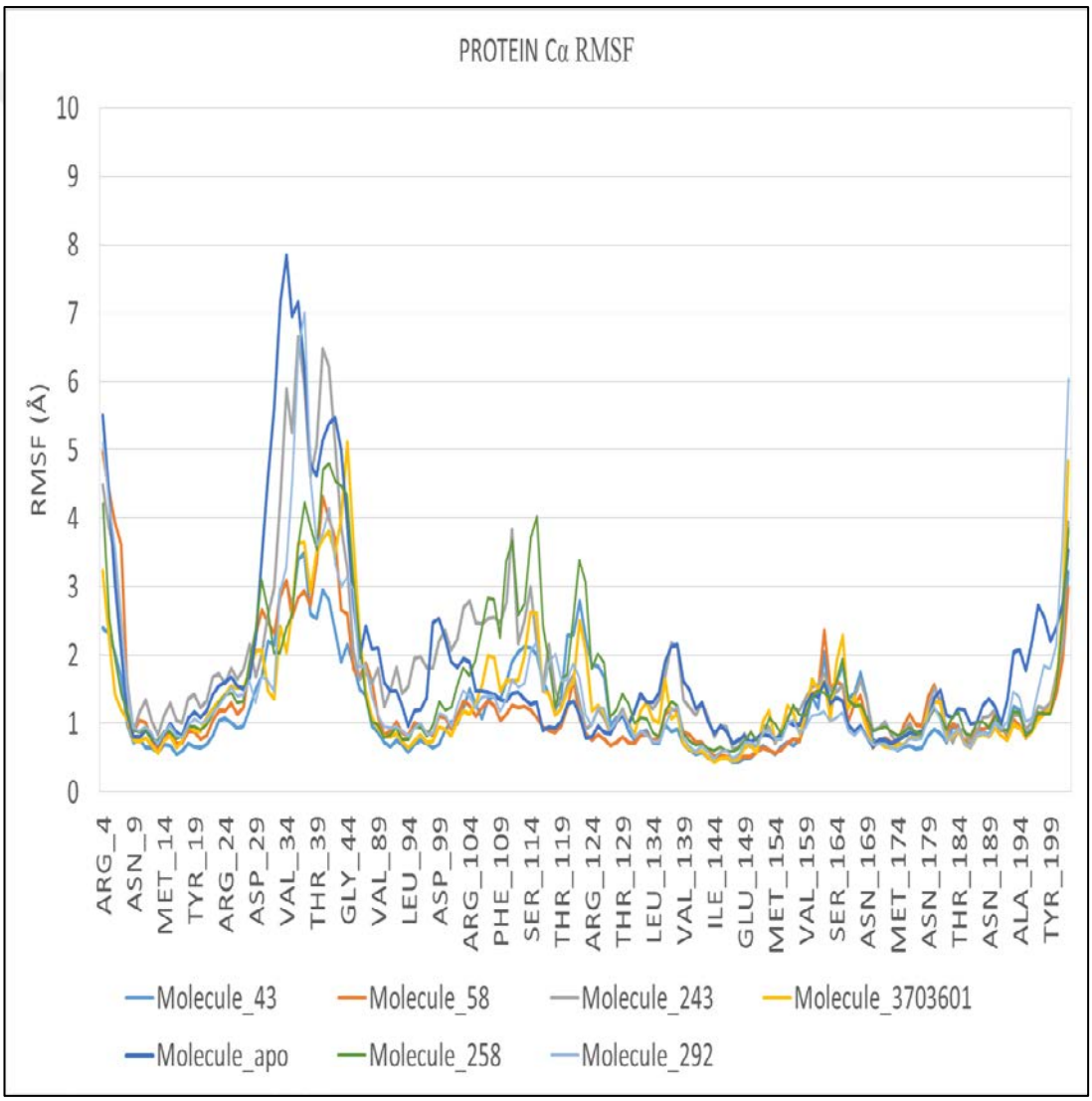


Figure 4.18 shows RMSF graphs of studied systems during the MD simulations. While secondary structural regions show fluctuations around 1-2 Å, other regions especially

residue numbers between 30-45 have larger fluctuations around 6-8 Å. N and C terminal residues also have higher fluctuations as expected.

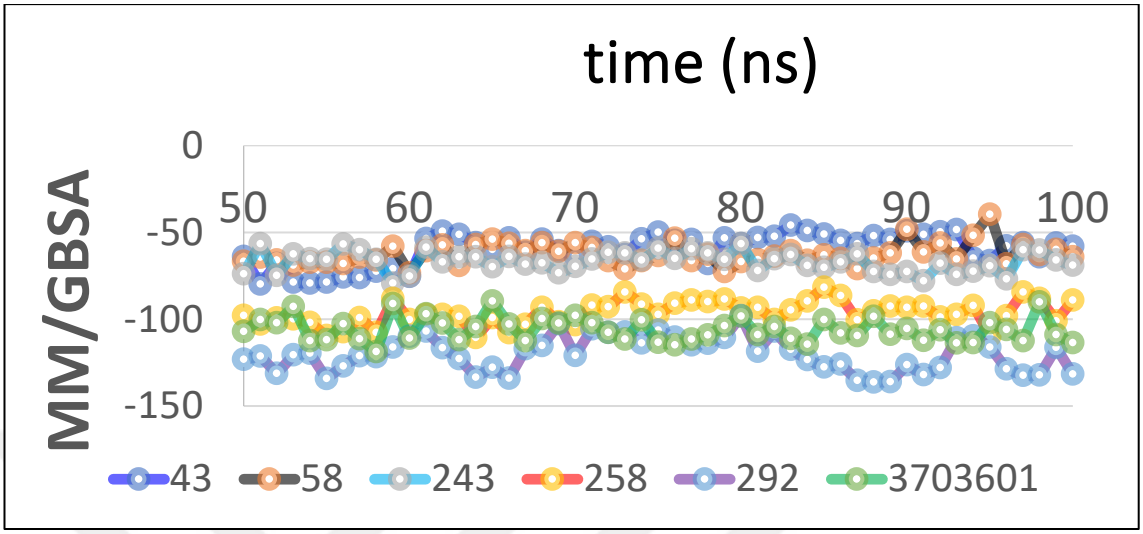
The trajectory frames obtained after MD simulations were subjected to post-process analysis. The MM/GBSA approach implemented in Prime module of Schrodinger was used to calculate the binding free energies. While Table 4.3 shows average binding free energies of studied ligands, Figure 4.19 shows changes of free energies during the simulations (last half of the simulations were considered).

Table 4.3: MM/GBSA average values of studied molecules.

Compounds	GBSA (kcal/mol)	Standart Deviation
43	**-59.59**	**9.44**
58	**-62.40**	**6.35**
243	**-66.79**	**5.82**
258	**-96.35**	**6.96**
292	**-120.06**	**9.96**
3703601	**-102.93**	**6.93**

MM/GBSA calculations were obtained for top-docking scored hit molecules.

Figure 4.19: MM/GBSA free energy analysis for the studied molecules at the binding pocket of BCL-2 throughout the last half of the MD simulations. (MM/GBSA binding energy values in kcal/mol)



5. CONCLUSIONS

In this thesis work, a molecular library (Specs-SC) composed of 212520 molecules, first filtered for their therapeutic effect against cancer, then obtained molecules again filtered to remove toxic compounds using MetaCore/MetaDrug web server. Identified 342 non-toxic and potent compounds (high therepeutic activity prediction against cancer) using MetaCore were then screened on target-driven approached using available BCL-2 structures (3 NMR and 1 X-ray structures were used). In order to compare the both structural and energetic results, a known BCL-2 inhibitor is used as positive control molecule and same computational protocols were applied for this compound. Identified hit molecules from both docking and short (i.e., 1-ns) MM/GBSA calculations that have similar/better binding energies compared to known-inhibitors, then subjected to longer (i.e., 100-ns) MD simulations. 5 molecules were identified as potent novel hit compounds: **43**, **58**, **243**, **258**, and **292**. Obtained MD trajectories were used in post MD MM/GBSA free energy calculations. Results of this thesis work, can be used in the discovery of novel BCL-2 family protein inhibitors.

REFERENCES

Periodicals

WHO (2012) Globocan 2012, Accessed 20 Dec 2016.

Acoca S, Cui Q, Shore GC, Purisima EO. 2011, Molecular dynamics study of small molecule inhibitors of the Bcl-2 family. *Proteins: Struct Funct Bioinform*. 79, 2624-36.

Alonso H, Bliznyuk AA, Gready JE. 2006, Combining docking and molecular dynamic simulations in drug design. *Med Res Rev*. 26, 531-68.

Arora K, Brooks Iii CL. 2009, Functionally important conformations of the Met20 loop in dihydrofolate reductase are populated by rapid thermal fluctuations. *J Am Chem Soc*. 131, 5642-7.

Azam SS, Abro A, Tanvir F, Parvaiz N. 2014, Identification of unique binding site and molecular docking studies for structurally diverse Bcl-xL inhibitors. *Med Chem Res*. 23, 3765-83.

Banks JL, Beard HS, Cao Y, Cho AE, Damm W, Farid R, et al. 2005, Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J Comput Chem*. 26, 1752-80.

Bártová I, Otyepka M, Kříž Z, Koča J. 2004, Activation and inhibition of cyclin-dependent kinase-2 by phosphorylation; a molecular dynamics study reveals the functional importance of the glycine-rich loop. *Protein Sci*. 13, 1449-57.

Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, et al. 2002, The protein data bank. *Acta Crystallogr Sect D Biol Crystallogr*. 58, 899-907.

Bernardo PH, Wan K-F, Sivaraman T, Xu J, Moore FK, Hung AW, et al. 2008, Structure− activity relationship studies of phenanthridine-based Bcl-XL inhibitors. *J Med Chem*. 51, 6699-710.

Billard C. 2013, BH3 mimetics: status of the field and new developments. *Mol Cancer Ther*. 12, 1691-700.

Blaineau SV, Aouacheria A. 2009, BCL2DB: moving 'helix-bundled'BCL-2 family members to their database. *Apoptosis*. 14, 923-5.

Bowers KJ, Chow DE, Xu H, Dror RO, Eastwood MP, Gregersen BA, et al., editors. Scalable algorithms for molecular dynamics simulations on commodity clusters.

SC'06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing; 2006: IEEE.

Bray F, Jemal A, Grey N, Ferlay J, Forman D. 2012, Global cancer transitions according to the Human Development Index (2008–2030): a population-based study. *The lancet oncology*. 13, 790-801.

Brown JM, Attardi LD. 2005, The role of apoptosis in cancer development and treatment response. *Nature reviews cancer*. 5, 231.

Bruncko M, Oost TK, Belli BA, Ding H, Joseph MK, Kunzer A, et al. 2007, Studies Leading to Potent, Dual Inhibitors of Bcl-2 and Bcl-xL. *J Med Chem*. 50, 641-62.

Buch I, Giorgino T, De Fabritiis G. 2011, Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proceedings of the National Academy of Sciences*. 108, 10184-9.

Chen G, Zhou D, Li XZ, Jiang Z, Tan C, Wei XY, et al. 2017, A natural chalcone induces apoptosis in lung cancer cells: 3D-QSAR, docking and an in vivo/vitro assay. *Sci Rep*. 7, 10729.

Chiba S, Ishida T, Ikeda K, Mochizuki M, Teramoto R, Taguchi Y, et al. 2017, An iterative compound screening contest method for identifying target protein inhibitors using the tyrosine-protein kinase Yes. *Sci Rep*. 7, 12038.

Chipuk JE, Moldoveanu T, Llambi F, Parsons MJ, Green DR. 2010, The BCL-2 family reunion. *Mol Cell*. 37, 299-310.

Chung C. 2018, Restoring the switch for cancer cell death: Targeting the apoptosis signaling pathway. *Am J Health Syst Pharm*. 75, 945-52.

Cory S, Huang DC, Adams JM. 2003, The Bcl-2 family: roles in cell survival and oncogenesis. *Oncogene*. 22, 8590.

Darden T, York D, Pedersen L. 1993, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*. 98, 10089-92.

Davis JM, Navolanic PM, Weinstein-Oppenheimer CR, Steelman LS, Hu W, Konopleva M, et al. 2003, Raf-1 and Bcl-2 induce distinct and common pathways that contribute to breast cancer drug resistance. *Clin Cancer Res*. 9, 1161-70.

Degterev A, Lugovskoy A, Cardone M, Mulley B, Wagner G, Mitchison T, et al. 2001, Identification of small-molecule inhibitors of interaction between the BH3 domain and Bcl-x L. *Nat Cell Biol*. 3, 173.

Delbridge AR, Grabow S, Strasser A, Vaux DL. 2016, Thirty years of BCL-2: translating cell death discoveries into novel cancer therapies. *Nat Rev Cancer*. 16, 99-109.

Dewson G, Kluck R. 2010, Bcl-2 family-regulated apoptosis in health and disease. *Cell Health and Cytoskeleton*. 2, 9-22.

Doruker P, Atilgan AR, Bahar I. 2000, Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: Application to α-amylase inhibitor. *Proteins: Structure, Function, and Bioinformatics*. 40, 512-24.

Durdagi S, Salmas RE, Stein M, Yurtsever M, Seeman P. 2016, Binding Interactions of Dopamine and Apomorphine in D2High and D2Low States of Human Dopamine D2 Receptor Using Computational and Experimental Techniques. *ACS Chem Neurosci*. 7, 185-95.

Ebrahim AS, Sabbagh H, Liddane A, Raufi A, Kandouz M, Al-Katib A. 2016, Hematologic malignancies: newer strategies to counter the BCL-2 protein. *J Cancer Res Clin Oncol*. 142, 2013-22.

Eimon PM, Ashkenazi A. 2010, The zebrafish as a model organism for the study of apoptosis. *Apoptosis*. 15, 331-49.

Elmore S. 2007, Apoptosis: a review of programmed cell death. *Toxicol Pathol*. 35, 495-516.

Enyedy IJ, Ling Y, Nacro K, Tomita Y, Wu X, Cao Y, et al. 2001, Discovery of small-molecule inhibitors of Bcl-2 through structure-based computer screening. *J Med Chem*. 44, 4313-24.

Ewing TJ, Makino S, Skillman AG, Kuntz ID. 2001, DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J Comput Aided Mol Des*. 15, 411-28.

Fesik SW. 2005, Promoting apoptosis as a strategy for cancer drug discovery. *Nature Reviews Cancer*. 5, 876.

Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. 2004, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem*. 47, 1739-49.

Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, et al. 2006, Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *J Med Chem*. 49, 6177-96.

Gandhi L, Camidge DR, De Oliveira MR, Bonomi P, Gandara D, Khaira D, et al. 2011, Phase I study of Navitoclax (ABT-263), a novel Bcl-2 family inhibitor, in patients with small-cell lung cancer and other solid tumors. *J Clin Oncol*. 29, 909.

Gapsys V, de Groot BL, Briones R. 2013, Computational analysis of local membrane properties. *J Comput Aided Mol Des*. 27, 845-58.

Güner OF. Pharmacophore perception, development, and use in drug design: Internat'l University Line; 2000.

Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, et al. 2004, Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J Med Chem*. 47, 1750-9.

Hallek M. 2015, Chronic lymphocytic leukemia: 2015 update on diagnosis, risk stratification, and treatment. *Am J Hematol*. 90, 446-60.

Hanahan D, Weinberg RA. 2000, The hallmarks of cancer. *Cell*. 100, 57-70.

Hanahan D, Weinberg RA. 2011, Hallmarks of cancer: the next generation. *Cell*. 144, 646-74.

Hardwick JM, Soane L. 2013, Multiple functions of BCL-2 family proteins. *Cold Spring Harb Perspect Biol*. 5,

Hayes RL, Noel JK, Mohanty U, Whitford PC, Hennelly SP, Onuchic JN, et al. 2012, Magnesium fluctuations modulate RNA dynamics in the SAM-I riboswitch. *J Am Chem Soc*. 134, 12043-53.

Hetts SW. 1998, To die or not to die: an overview of apoptosis and its role in disease. *JAMA*. 279, 300-7.

Hockenbery D, Nuñez G, Milliman C, Schreiber RD, Korsmeyer SJ. 1990, Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. *Nature*. 348, 334.

Hoover WG. 1985, Canonical dynamics: Equilibrium phase-space distributions. *Phys Rev A*. 31, 1695-7.

Igney FH, Krammer PH. 2002, Death and anti-death: tumour resistance to apoptosis. *Nature Reviews Cancer*. 2, 277.

Itchaki G, Brown JR. 2016, The potential of venetoclax (ABT-199) in chronic lymphocytic leukemia. *Therapeutic advances in hematology*. 7, 270-87.

Jacobson MP, Friesner RA, Xiang Z, Honig B. 2002, On the Role of the Crystal Environment in Determining Protein Side-chain Conformations. *J Mol Biol*. 320, 597-608.

Jacobson MP, Pincus DL, Rapp CS, Day TJF, Honig B, Shaw DE, et al. 2004, A hierarchical approach to all-atom protein loop prediction. *Proteins: Structure, Function, and Bioinformatics*. 55, 351-67.

Jagani H, Kasinathan N, Meka SR, Josyula VR. 2016, Antiapoptotic Bcl-2 protein as a potential target for cancer therapy: A mini review. *Artif Cells Nanomed Biotechnol*. 44, 1212-21.

Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. 2011, Global cancer statistics. *CA Cancer J Clin*. 61, 69-90.

Jones G, Willett P, Glen RC, Leach AR, Taylor R. 1997, Development and validation of a genetic algorithm for flexible docking. *J Mol Biol*. 267, 727-48.

Jones G, Willett P, Glen RC, Leach AR, Taylor R. 1997, Development and validation of a genetic algorithm for flexible docking11Edited by F. E. Cohen. *J Mol Biol*. 267, 727-48.

Kamath PR, Sunil D, Ajees AA, Pai K, Biswas S. 2016, N′-((2-(6-bromo-2-oxo-2H-chromen-3-yl)-1H-indol-3-yl) methylene) benzohydrazide as a probable Bcl-2/Bcl-xL inhibitor with apoptotic and anti-metastatic potential. *Eur J Med Chem*. 120, 134-47.

Kanan T, Kanan D, Erol I, Yazdi S, Stein M, Durdagi S. 2019, Targeting the NF-κB/IκBα complex via fragment-based E-Pharmacophore virtual screening and binary QSAR models. *J Mol Graphics Modell*. 86, 264-77.

Kang MH, Reynolds CP. 2009, Bcl-2 inhibitors: targeting mitochondrial apoptotic pathways in cancer therapy. *Clin Cancer Res*. 15, 1126-32.

Kitada S, Leone M, Sareth S, Zhai D, Reed JC, Pellecchia M. 2003, Discovery, characterization, and structure− activity relationships studies of proapoptotic polyphenols targeting B-cell lymphocyte/leukemia-2 proteins. *J Med Chem*. 46, 4259-64.

Kitchen DB, Decornez H, Furr JR, Bajorath J. 2004, Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature reviews Drug discovery*. 3, 935.

Kollek M, Müller A, Egle A, Erlacher M. 2016, Bcl-2 proteins in development, health, and disease of the hematopoietic system. *The FEBS journal*. 283, 2779-810.

Krammer PH, Behrmann I, Daniel P, Dhein J, Debatin K-M. 1994, Regulation of apoptosis in the immune system. *Curr Opin Immunol*. 6, 279-89.

Lama D, Sankararamakrishnan R. 2008, Anti-apoptotic Bcl-XL protein in complex with BH3 peptides of pro-apoptotic Bak, Bad, and Bim proteins: Comparative molecular dynamics simulations. *Proteins: Structure, Function, and Bioinformatics*. 73, 492-514.

Lama D, Sankararamakrishnan R. 2011, Molecular dynamics simulations of pro-apoptotic BH3 peptide helices in aqueous medium: relationship between helix stability and their binding affinities to the anti-apoptotic protein Bcl-X L. *J Comput Aided Mol Des*. 25, 413.

Leelananda SP, Lindert S. 2016, Computational methods in drug discovery. *Beilstein J Org Chem*. 12, 2694-718.

Lessene G, Czabotar PE, Colman PM. 2008, BCL-2 family antagonists for cancer therapy. *Nature reviews Drug discovery*. 7, 989.

Levine ZA, Venable RM, Watson MC, Lerner MG, Shea J-E, Pastor RW, et al. 2014, Determination of biomembrane bending moduli in fully atomistic simulations. *J Am Chem Soc*. 136, 13582-5.

Li H-q, Yang J, Ma S, Qiao C. 2012, Structure-based design of rhodanine-based acylsulfonamide derivatives as antagonists of the anti-apoptotic Bcl-2 protein. *Biorg Med Chem*. 20, 4194-200.

Li J, Abel R, Zhu K, Cao Y, Zhao S, Friesner RA. 2011, The VSGB 2.0 model: A next generation energy model for high resolution protein structure modeling. *Proteins: Structure, Function, and Bioinformatics*. 79, 2794-812.

Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. 2013, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des*. 27, 221-34.

Maity A, Yadav S, Verma CS, Dastidar SG. 2013, Dynamics of Bcl-xL in water and membrane: molecular simulations. *PLoS One*. 8, e76837.

Martyna GJ, Tobias DJ, Klein ML. 1994, Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*. 101, 4177-89.

Mavromoustakos T. 2011, Editorial [Hot Topic: Methodologies and Applied Strategies in the Rational Drug Design (Guest Editor: T. Mavromoustakos)]. *Curr Med Chem*. 18, 2516-.

Meier P, Finch A, Evan G. 2000, Apoptosis in development. *Nature*. 407, 796.

Mohamad Rosdi MN, Mohd Arif S, Abu Bakar MH, Razali SA, Mohamed Zulkifli R, Ya'akob H. 2018, Molecular docking studies of bioactive compounds from Annona muricata Linn as potential inhibitors for Bcl-2, Bcl-w and Mcl-1 antiapoptotic proteins. *Apoptosis*. 23, 27-40.

Nosé S. 1984, A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*. 81, 511-9.

Okada H, Mak TW. 2004, Pathways of apoptotic and non-apoptotic death in tumour cells. *Nature Reviews Cancer*. 4, 592.

Oltersdorf T, Elmore SW, Shoemaker AR, Armstrong RC, Augeri DJ, Belli BA, et al. 2005, An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature*. 435, 677.

Onufriev A, Bashford D, Case DA. 2004, Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Structure, Function, and Bioinformatics*. 55, 383-94.

Petros AM, Olejniczak ET, Fesik SW. 2004, Structural biology of the Bcl-2 family of proteins. *Biochim Biophys Acta*. 1644, 83-94.

Pinto M, del Mar Orzaez M, Delgado-Soler L, Perez JJ, Rubio-Martinez J. 2011, Rational design of new class of BH3-mimetics as inhibitors of the Bcl-xL protein. *J Chem Inf Model*. 51, 1249-58.

Pinto M, Perez JJ, Rubio-Martinez J. 2004, Molecular dynamics study of peptide segments of the BH3 domain of the proapoptotic proteins Bak, Bax, Bid and Hrk bound to the Bcl-x L and Bcl-2 proteins. *J Comput Aided Mol Des*. 18, 13-22.

Radha G, Raghavan SC. 2017, BCL2: A promising cancer therapeutic target. *Biochim Biophys Acta Rev Cancer*. 1868, 309-14.

Raghav PK, Verma YK, Gangenahalli GU. 2012, Molecular dynamics simulations of the Bcl-2 protein to predict the structure of its unordered flexible loop domain. *J Mol Model*. 18, 1885-906.

Rarey M, Kramer B, Lengauer T, Klebe G. 1996, A fast flexible docking method using an incremental construction algorithm. *J Mol Biol*. 261, 470-89.

Reed JC. 2002, Apoptosis-based therapies. *Nature reviews Drug discovery*. 1, 111.

Rodriguez JM, Ross NT, Katt WP, Dhar D, Lee Gi, Hamilton AD. 2009, Structure and Function of Benzoylurea-Derived α-Helix Mimetics Targeting the Bcl-xL/Bak Binding Interface. *ChemMedChem: Chemistry Enabling Drug Discovery*. 4, 649-56.

Rosas-Trigueros JL, Correa-Basurto J, Guadalupe Benítez-Cardoza C, Zamorano-Carrillo A. 2011, Insights into the structural stability of Bax from molecular dynamics simulations at high temperatures. *Protein Sci*. 20, 2035-46.

Rossé T, Olivier R, Monney L, Rager M, Conus S, Fellay I, et al. 1998, Bcl-2 prolongs cell survival after Bax-induced release of cytochrome c. *Nature*. 391, 496.

Ruefli-Brasse A, Reed JC. 2017, Therapeutics targeting Bcl-2 in hematological malignancies. *Biochem J*. 474, 3643-57.

Salmas RE, Seeman P, Aksoydan B, Erol I, Kantarcioglu I, Stein M, et al. 2017, Analysis of the Glutamate Agonist LY404,039 Binding to Nonstatic Dopamine Receptor D2 Dimer Structures and Consensus Docking. *ACS Chem Neurosci*. 8, 1404-15.

Sartorius UA, Krammer PH. 2002, Upregulation of bcl-2 is involved in the mediation of chemotherapy resistance in human small cell lung cancer cell lines. *Int J Cancer*. 97, 584-92.

Schmitt CA, Rosenthal CT, Lowe SW. 2000, Genetic analysis of chemoresistance in primary murine lymphomas. *Nat Med*. 6, 1029.

Sekijima M, Motono C, Yamasaki S, Kaneko K, Akiyama Y. 2003, Molecular dynamics simulation of dimeric and monomeric forms of human prion protein: insight into dynamics and properties. *Biophys J*. 85, 1176-85.

Shan Y, Kim ET, Eastwood MP, Dror RO, Seeliger MA, Shaw DE. 2011, How does a drug molecule find its target binding site?, *J Am Chem Soc*. 133, 9181-3.

Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M. 2007, Epik: a software program for pKaprediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des*. 21, 681-91.

Sherman W, Day T, Jacobson MP, Friesner RA, Farid R. 2006, Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *J Med Chem*. 49, 534-53.

Sodt AJ, Sandar ML, Gawrisch K, Pastor RW, Lyman E. 2014, The molecular structure of the liquid-ordered phase of lipid bilayers. *J Am Chem Soc*. 136, 725-32.

Solderquist RS, Eastman A. 2016, BCL2 Inhibitors as Anticancer Drugs: A Plethora of Misleading BH3 Mimetics. *Mol Cancer Ther*. 15, 2011-7.

Søndergaard CR, Olsson MHM, Rostkowski M, Jensen JH. 2011, Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J Chem Theory Comput*. 7, 2284-95.

Souers AJ, Leverson JD, Boghaert ER, Ackler SL, Catron ND, Chen J, et al. 2013, ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat Med*. 19, 202.

Tang G, Ding K, Nikolovska-Coleska Z, Yang C-Y, Qiu S, Shangary S, et al. 2007, Structure-based design of flavonoid compounds as a new class of small-molecule inhibitors of the anti-apoptotic Bcl-2 proteins. *J Med Chem*. 50, 3163-6.

Tang G, Nikolovska-Coleska Z, Qiu S, Yang C-Y, Guo J, Wang S. 2008, Acylpyrogallols as inhibitors of antiapoptotic Bcl-2 proteins. *J Med Chem*. 51, 717-20.

Thomas S, Quinn BA, Das SK, Dash R, Emdad L, Dasgupta S, et al. 2013, Targeting the Bcl-2 family for cancer therapy. *Expert Opin Ther Targets*. 17, 61-75.

Tropsha A. 2010, QSAR in drug discovery. *Drug Design: Structure-and Ligand-Based Approaches*. 151-64.

Tse C, Shoemaker AR, Adickes J, Anderson MG, Chen J, Jin S, et al. 2008, ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Res*. 68, 3421-8.

Tsujimoto Y, Cossman J, Jaffe E, Croce CM. 1985, Involvement of the bcl-2 gene in human follicular lymphoma. *Science*. 228, 1440-3.

Verdine GL, Walensky LD. 2007, The challenge of drugging undruggable targets in cancer: lessons learned from targeting BCL-2 family members. *Clin Cancer Res*. 13, 7264-70.

Verma J, Khedkar VM, Coutinho EC. 2010, 3D-QSAR in drug design-a review. *Curr Top Med Chem*. 10, 95-115.

Verma S, Singh A, Mishra A. 2015, Complex disruption effect of natural polyphenols on Bcl-2-Bax: molecular dynamics simulation and essential dynamics study. *J Biomol Struct Dyn*. 33, 1094-106.

Wakui N, Yoshino R, Yasuo N, Ohue M, Sekijima M. 2018, Exploring the selectivity of inhibitor complexes with Bcl-2 and Bcl-XL: A molecular dynamics simulation approach. *J Mol Graph Model*. 79, 166-74.

Wang G, Nikolovska-Coleska Z, Yang C-Y, Wang R, Tang G, Guo J, et al. 2006, Structure-based design of potent small-molecule inhibitors of anti-apoptotic Bcl-2 proteins. *J Med Chem*. 49, 6139-42.

Wang S, Yang D, Lippman ME, editors. Targeting Bcl-2 and Bcl-XL with nonpeptidic small-molecule antagonists. Semin Oncol; 2003: Elsevier.

Wu H, Medeiros LJ, Young KH. 2018, Apoptosis signaling and BCL-2 pathways provide opportunities for novel targeted therapeutic strategies in hematologic malignances. *Blood Rev*. 32, 8-28.

Xing C, Wang L, Tang X, Sham YY. 2007, Development of selective inhibitors for anti-apoptotic Bcl-2 proteins from BHI-1. *Biorg Med Chem*. 15, 2167-76.

Yildirim A, Sharma M, Varner BM, Fang L, Feig M. 2014, Conformational preferences of DNA in reduced dielectric environments. *The Journal of Physical Chemistry B*. 118, 10874-81.

Yin H, Lee G-i, Sedey KA, Rodriguez JM, Wang H-G, Sebti SM, et al. 2005, Terephthalamide derivatives as mimetics of helical peptides: disruption of the Bcl-xL/Bak interaction. *J Am Chem Soc*. 127, 5463-8.

Yip K, Reed J. 2008, Bcl-2 family proteins and cancer. *Oncogene*. 27, 6398.

Yoshino R, Yasuo N, Hagiwara Y, Ishida T, Inaoka DK, Amano Y, et al. 2017, In silico, in vitro, X-ray crystallography, and integrated strategies for discovering spermidine synthase inhibitors for Chagas disease. *Sci Rep*. 7, 6666.

Yoshino R, Yasuo N, Inaoka DK, Hagiwara Y, Ohno K, Orita M, et al. 2015, Pharmacophore modeling for anti-chagas drug design using the fragment molecular orbital method. *PLoS One*. 10, e0125829.

Youle RJ, Strasser A. 2008, The BCL-2 protein family: opposing activities that mediate cell death. *Nature Reviews Molecular Cell Biology*. 9, 47.

Zhang H, Nimmer P, Tahir S, Chen J, Fryer R, Hahn K, et al. 2007, Bcl-2 family proteins are essential for platelet survival. *Cell Death Differ*. 14, 943.

Ziedan NI, Hamdy R, Cavaliere A, Kourti M, Prencipe F, Brancale A, et al. 2017, Virtual screening, SAR, and discovery of 5-(indole-3-yl)-2-[(2-nitrophenyl) amino][1, 3, 4]-oxadiazole as a novel Bcl-2 inhibitor. *Chem Biol Drug Des*. 90, 147-55.