Virtual Screening and Elucidation of Putative Binding Mode for Small Molecule Antagonist of BCL2 BH4 Domain

JOEL Ireoluwa Yinka[1], **ADIGUN** Temidayo Olamide[1], **AJIBOLA** Ahmeedah Ololade[1], **BANKOLE** Olukayode Olusola[1,2], **OZOJIOFOR** Ugochukwu Okechukwu[3], **REMI-ESAN** Ifelolu Adeseye[2], **SULAIMON** Lateef Adegboyega[4,5]

1. Department of Biochemistry, University of Ilorin, Ilorin Kwara state
2. Department of Science Laboratory Technology, School of Pure & Applied Sciences, Federal Polytechnic Ilaro, Ogun state
3. Department of Biotechnology, Nigerian Defense Academy, Kaduna, Kaduna State
4. Department of Chemical Sciences, Crescent University Abeokuta, Ogun State.
5. Department of Biochemistry, University of lagos

Keywords: BCL2, BH4, Induced-Fit Docking, QM-polarized docking, QM/MM optimization, and Computer-Aided Drug Design

Corresponding Author: JOEL Ireoluwa Yinka, Department of Biochemistry, University of Ilorin, Nigeria; joelireoluwa@gmail.com or 17-68ez012pg@students.unilorin.edu.ng

Abstract

Evading apoptosis is one of the hallmarks of cancer cells, therefore a lot of therapeutic strategies have been developed to induce cell death in these cells. BCL2 family protein governs the intrinsic pathway of cell death. The BCL2 family protein possess both pro and anti-apoptosis members. BCL2 protein, a pioneering member of the anti-apoptosis protein has been implicated in several cancerous cells; with dozens of small-molecule inhibitors developed to inhibit its activity. BCL2 family protein processes BH domains (1-4) for heterodimerization and homodimerization amidst it family members and targeting the BH4 domain is an established new strategy for cancer therapy; however, only BDA366 has been reported so far as a BH4 binding molecule. Using Computer-aided drug discovery techniques (Molecular docking, QM-polarized docking, Induced-fit docking, QM-MM optimization, and QM electronic descriptors) we screened M sample (~ 99,000 compounds) of the SCUBIDOO database to find ligands that interacts with the BH4 domain of BCL2 protein with high affinity, we identified 11 putative BH4 specific small molecules with binding affinity ranging from ~ -84kcal/mol to -64kcal/mol with a putative binding hypothesis with BH4 amino acid residues: ASP10, ARG12, GLU13, MET16, LYS17, and HIS20

Introduction

The uncontrolled proliferation of cancer cells is not only driven by oncogenes but also results from defective apoptotic machinery or a combination of both [1][2] as well as the development of resistance by cancer cells to chemotherapeutics-induced apoptosis. These culminate in poor prognosis with an 18% overall survival rate in cancer patients [1]. Acquired resistance to apoptotic cell death is largely due to an over-expression of anti-apoptotic genes, and down-regulation or mutation of pro-apoptotic genes [3]. Therefore, overcoming resistance to apoptosis by activating apoptosis pathways has been a major focus in the development of therapeutic strategies for cancer treatment.

Apoptosis is an evolutionarily preserved mechanism of controlled cell deletion in nature playing a critical role such as the deletion of redundant or damaged cells in diverse fundamental body processes of multicellular organisms [4]. Its machinery consists of two major inextricably linked activation pathways: the extrinsic pathway via interaction of transmembrane receptors and death ligands (such as Fas, TNF-α, and tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)) [5][6] and the intrinsic pathway with its consequential leakage of multiple pro-apoptotic proteins including cytochrome c and Smac/DIABLO from the mitochondria to the cytosol due to mitochondrial membrane potential loss [7] as well as a third, though less well-known, initiation pathway known as the intrinsic endoplasmic reticulum pathway[6].

The BCL2 family of proteins play major roles in the intrinsic pathway of apoptosis; the family is subdivided into two: pro-apoptosis (BH123: BAK, BAX and BH3-only proteins: BID, BIM, BAD, BIK, PUMA, and NOXA) and anti-apoptosis (BCL2 protein, BCL-xL, MCL1, BCL-W, BCL-B, and BCL2a1) [8]

Briefly, the BCL2 family of protein regulates apoptosis in the following ways: firstly in the absence of an apoptosis stimulus, BCL2 protein (anti-apoptosis) binds with BH123 proteins: BAK and BAX (pro-apoptosis), thereby preventing activation i.e. formation of homodimers on the mitochondrion membrane causing the release of caspase and activating the intrinsic apoptosis pathway [9]; however, when an apoptosis stimulus is received BH3-only proteins binds with BCL2, preventing it from interacting with BH123 protein, this process, therefore, activate the pro-apoptotic protein BAK and BAX [5][9].

The complex binding and interactions within the BCL2 family occur via their BCL2 alpha-helical homology (BH) domains [20]. There exist four BH domains (BH 1- 4) of which the BH3 domain is responsible for activation and inactivation interactions of pro-apoptosis and anti-apoptosis BCL2 family proteins. Since overexpression of BCL2 protein has implicated in several cancerous tumors [10], BH3 domain has been targeted for developing small molecule inhibitors of BCL2 protein (anti-apoptosis) to induced cell death in cancerous cells; inhibitors such as ABT-737, ABT-263, ABT-199, Disarib, etc.[8] have been developed with varying degree of success [5][6]. However, none has been approved for use; this is primarily due to the highly conserved nature of BH3 domains among the BCL2 family [10]; the search for small molecules inhibitors or antagonists of BCL2 protein has thus continued.

Several studies have suggested developing small molecule inhibitors targeting the BH4 domain of BCL2 protein as a potential therapeutic strategy [9][11][12]. Han *et al* [13] recently used BDA366 as proof of concept to show the therapeutic potential in targeting the BH4 domain of BCL2 protein. The BH4 domain in BCL2 protein is important in the regulation of the anti-apoptotic activity of BCL2 protein; it is also required in the interaction (heterodimerization) of BCL2 protein with pro-apoptosis protein BAX [10]. Furthermore, mutant BCL2 protein without the BH4 domain has been shown to promote apoptosis instead of inhibiting apoptosis [10][12].

At the time of writing, BDA366 is the only small molecule targeting the BH4 domain that has been reported. We, therefore, sought to virtually screen for potential BH4 binding small molecules and investigate a putative binding mode for the identified lead compounds.

Materials and Methods

2.1. Small molecules Library

The M sample of the SCUBIDOO database [14] was downloaded for this study. The M sample which consists of ~ 99,000 compounds is a sample representation of the whole database, which consists of ~21,000,000 compounds.

2.2. Molecular Docking

2.2.1. Protein Preparation

Schrodinger's Protein preparation wizard [15] was used to import the BCL2 protein from PDB (PDB ID:1GM), and prepossessed: Hydrogen atoms were added, disulfide bonds were created, missing loops and side chains were added using Prime [16] Termini was capped, water molecules beyond 5Å from het groups were deleted, finally, het sates at pH 7.0 +/- 2.0 was generated using Epik [4]. Hydrogen bonds in the protein structure were optimized, water molecule with less than 3 Hydrogen bonds to non-water residues/molecule was deleted, and finally, the whole protein structure was minimized converging heavy atoms to RSMD: 0.3Å.

2.2.1. Ligand Preparation

The Small molecule library was prepared using Ligprep [15]: the molecule was ionized generating all possible states at pH 7.0 +/- 2.0, the ligand was desalted and tautomers was generated. Possible stereoisomers were also generated (at most 32 per ligand). A total of ~ 200,000 structures were thereafter generated.

2.2.2. Active site Grid generation

Using the receptor grid generation grid of Schrodinger suit, the docking grid file was generated. However, since there is still no co-crystallized ligand in the BH4 domain of BCL2, we manually inputted the amino acid residues for the active sites based on information from the literature (AA 6-31) [10,13].

2.2.3. High throughput virtual screening

Using Schrodinger High throughput virtual screening workflow, the small molecule library was docked into the BH4 domain of BCL2 protein. The workflow included frittering the library based on drug and lead-likeness criteria, the workflow the docking stage in the workflow docked the compounds using the 3 different glide docking protocol (HTVS, SP, XP), and finally post-possessed using Molecular Mechanics-Generalized Born Surface Area (MM-GBSA).

2.2.4. Molecular Mechanics-Generalized Born Surface Area

Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) evaluated the binding free energies (binding affinity) and implemented geometric minimization of the docked protein-ligand complex [18]. The binding free energy was calculated using VSGB 2.0 implicit solvation model and OPLS-2005 via Prime [19]. The binding free energy is calculated using Eq.1

$$\Delta G_{bind} = G_{complex} - G_{protein} - G_{ligand} \quad (1)$$

Where $G_{complex}$, $G_{protein,}$ and G_{ligand} represents the free binding energy of the protein-ligand complex, protein only, and ligand only respectively [18].

2.3. QM-Polarized Docking

QM-Polarized docking (qpld) module [20] of Schrodinger suits were used to implement the qpld experiment. The module used the grid file generated during molecular docking after an initial docking was performed; the docking used Glide SP [21] protocol, ligand van der Waals scaling was set to 0.8 and a maximum of 5 poses was generated per ligand. After the initial docking, Quantum mechanics (QM) charges for the generated ligand poses were calculated using a semi-empirical method with charge type set to Coulson. With this generated QM charges the ligand poses were redocked using Glide SP protocol and a maximum of 10 poses per ligand was generated.

2.4. Induced-fit Docking

In other to predict conformation changes that could occur as a result of the ligand binding, the Schrodinger Induced-fit docking module [22] was used to model the conformation changes. The module docked the ligand flexibly into the active site Using Glide SP docking protocol [21] with ring conformation sampled at an energy window of 2.5kcal/mol; the side chain of the active site residue was trimmed, with receptor and ligand van der waals scaling at 0.80. The maximum number of poses per ligand was set to 20. The generated poses were refined using Prime [16]: the residues within 5.0Å of ligand pose were refined and side chains optimized. The refined structures within 30.0kcal/mol of the best structure and the top 20 structure of overall were redocked using the Glide SP docking protocol.

2.5. QM/MM Optimization

The Induced-fit docking poses were optimized using QM/MM optimization [23] protocol implemented via the Schrodinger Qsite module. The ligand and side-chain residues interacting with the ligand were treated as the QM region, while the protein was treated as MM region. DFT-B3LYP and basics set 631G** level was used for the QM calculations. The MM region was treated using OPLS2005 and energy was minimized using the Truncated Newton Algorithm. Convergence interaction was set at energy and gradient.

2.6. Electron Descriptors

The following Quantum mechanics (QM) electronic descriptors (Highest occupied molecular orbital, lowest unoccupied molecular orbital, Electrostatic surface potential) were calculated using the Schrodinger Jaguar Single point energy module [41]. The module used hybrid Density flow theory (DFT) with Becke‘s three-parameter exchange potential, Lee-Yang-Parr correlation functional (B3LYP), and basis set 631G ** level [24][25]. From this aforementioned descriptors the following descriptors were calculated [26];

- HOMO-LUMO gap= $E_{LUMO} – E_{HOMO}$,
- ionization energy (I) = $-E_{HOMO}$,
- electron affinity (A) = $-E_{LUMO}$,
- global hardness (η) = $(-E_{HOMO} + E_{LUMO})/2$,
- chemical potential (μ) = $(E_{HOMO} + E_{LUMO})/2$.
- global electrophilicity power of a ligand as $\omega = \mu 2/2\eta$ (proposed by Parr et al. [27]).

Results

1.1. Molecular docking Virtual screening

Molecular docking is by far the most used technique in computer-aided drug discovery. It's a molecular simulation method used in predicting energetically favorable binding conformation of small molecules (ligands) in the protein active site [28]; using a scoring function, the binding affinity of the ligand is determined or inferred [29]. This, therefore, enables a large number of small molecules to be screened computationally and ranked according to their binding affinity. This technique was therefore used to screen 99,000 small molecules of the M sample of SCUBIDOO database [14] (see methods) in other to find ligands interacting with the BH4 domain of BCL2 protein with high binding affinity. BDA366 which has been established as a selective BH4 interacting ligand [30] was used as the standard for this study. After docking and scoring, the binding affinity of the ligands was rescored and geometrically minimized using MM-GBSA (see methods).

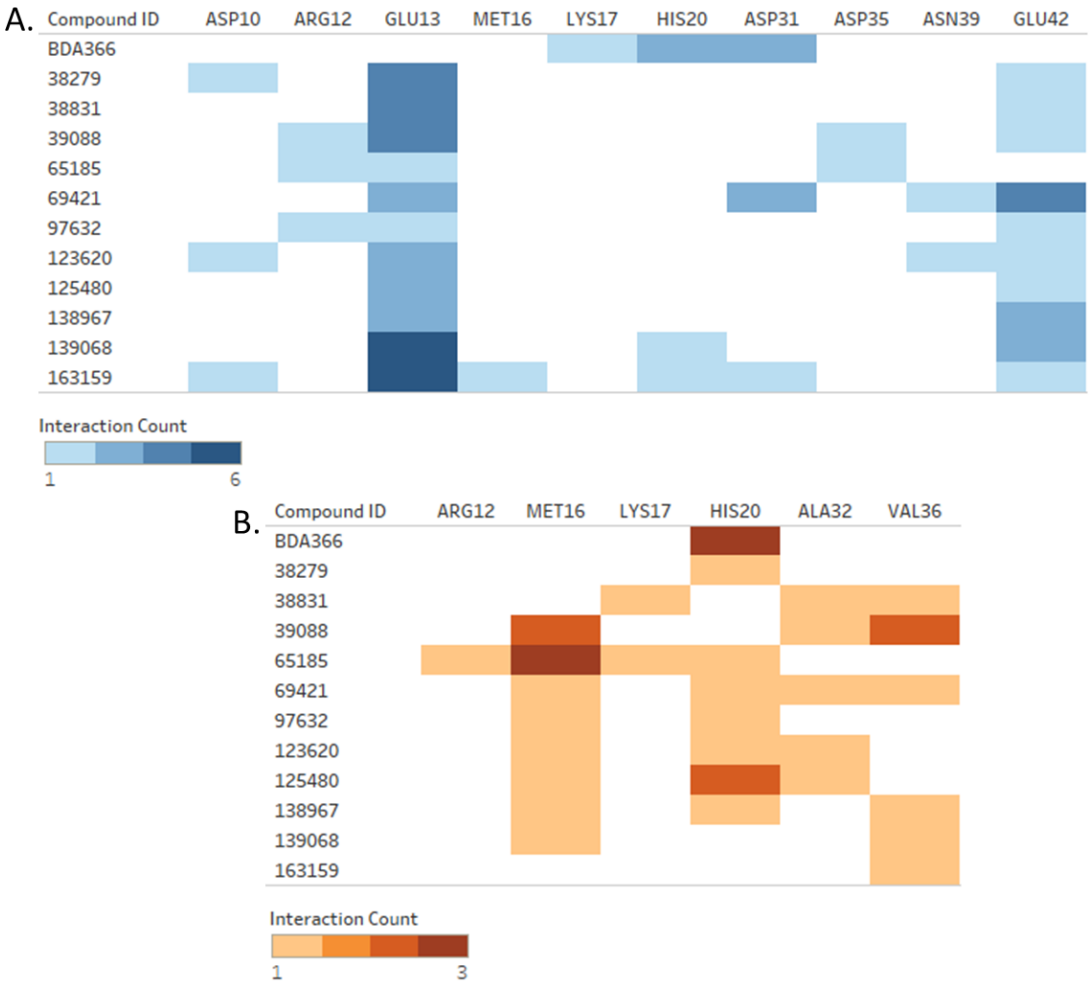

Figure 1: Molecular Docking binding interaction analysis: *a) Hydrogen bond and electrostatic interaction count b) Hydrophobic Interaction Count*

The binding affinity of the ligands ranged from ~ -74kcal/mol to -24kcal/mol, BDA366 had a binding affinity of -45.39kcal/mol. The top-ranked 11 compounds binding interactions (Hydrogen bond, electrostatic interactions, and hydrophobic interactions) were analyzed (Figure 1). The results showed different compound interactions within BH4 (AA: 6-31) and interaction with amino acid residues outside of the BH4 domain (figure 1).

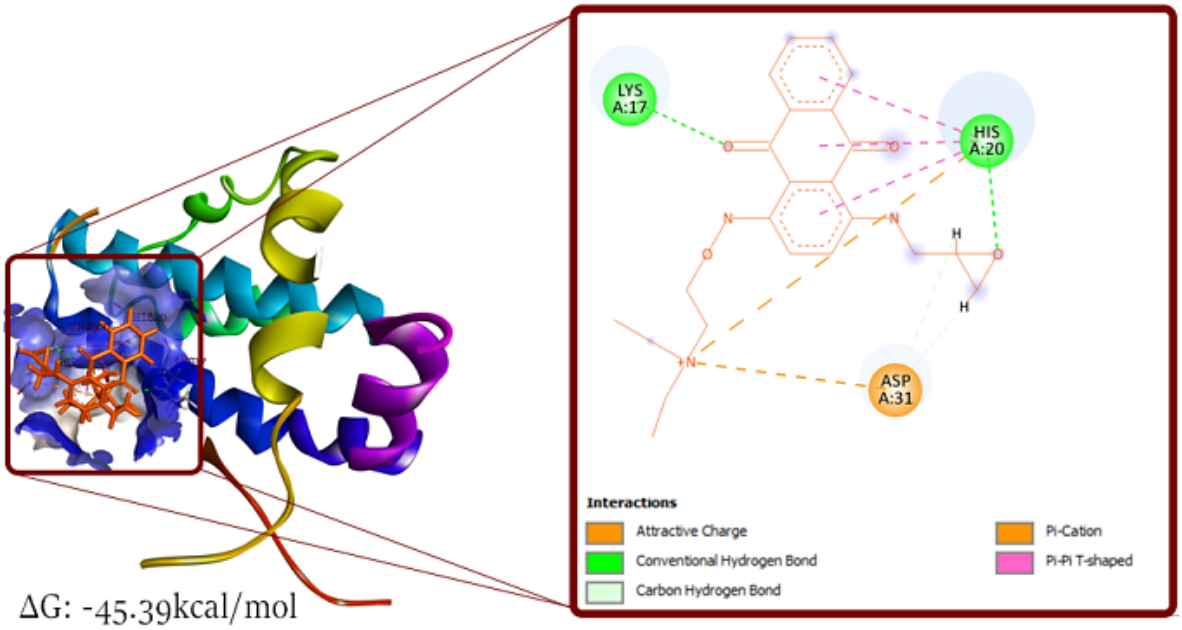

Figure 2: Molecular docking pose visualization: BDA366

Visual inspection of BDA366 showed Hydrogen Bond (H-bond) formations with LYS17 and HIS20 and hydrophobic interactions with HIS20 and ASP31 (Figure 2). The binding conformation of the top 3 ranked compounds (-74 to -64kcal/mol) was inspected (figure 3). Interaction with different BH4 residues was observed (AA: 10, 13, 16, 20, and 31), and non-BH4 residue interactions (AA: 32, 36, and 42) were observed (figure 3). Overall, interaction with GLU13, HIS20, VAL36, and GLU42 was the most observed interactions among the top 11 compounds (Figure 1, 2, and 3). Using molecular docking we were able to virtually screen ~99,000 compounds and identified 11 lead compounds with binding affinity ranging from -74 to -63kcal/mol all having higher binding affinity than BDA366 (-45.39kcal/mol). However, molecular docking is not the most accurate method in elucidating a binding mode and determining binding affinity, the lead compounds were hence subjected to more accurate binding simulation protocols.

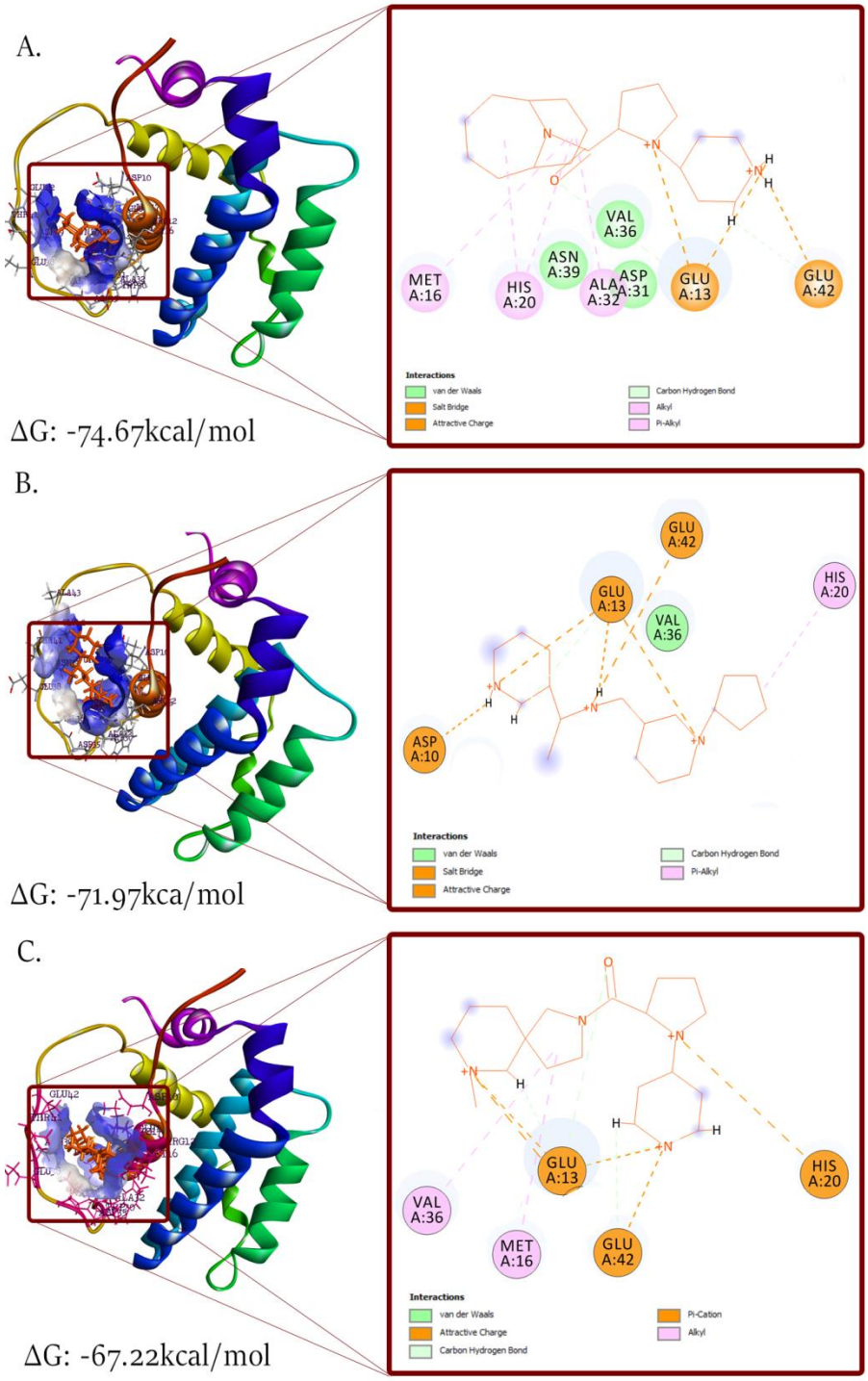

Figure 2: Molecular Docking pose visualization: *a) 125480 b) 38279 c) 139068*

1.2. QM-Polarized Docking

In finding an energetically favorable binding conformation in a docking simulation, electrostatic charges on the ligand atoms influence the accuracy of the simulation (binding affinity and docking pose) [28]. Various models for calculating these atomic charges have been developed and compared (force-fields, *ab intio*, empirical, and semi-empirical) with various levels of accuracy improvement [44]. Apart from accurately calculating the atomic charges, calculating the charges induced on the ligands by the active site residues have been shown to greatly improve docking simulation accuracy (binding affinity and docking confirmation) when compared with experimental data [20][44][31]. QM-Polarized Docking (qpld) is an improvement on classical docking protocol; wherein ligands are docked initially (using atomic charges calculated via force field models) into the active site. Using the docking poses generated Quantum mechanics calculations are used to determine the charges induced on the ligands by the active site residues, these new charges acquired is used in the next run of the docking simulation. Using Schrodinger QM-Polarized Docking module (see methods), qpld was implemented thereafter; the docking conformations were post-processed using MM-GBSA (see methods) and binding affinity rescored.

Using BDA366 as the control for this docking run (figure 3) we docked the 11 lead compounds; binding interaction was analyzed (figure 4) and the docking pose of the top 3 ranked compounds (binding affinity) was visualized (figure 5).

The binding interaction analysis showed BDA366 to have acquired new binding interactions with BH4 residues: SER24, TYR21, ILE14, and non-BH4 residues: HIS94 and LEU95, the interactions with LYS17 and HIS20 were retained (figure 3 and 4), however, HIS20 lost the 3 electrostatic interactions (pi-cation) while LYS17 gained additional two electrostatic interactions (pi-cation), attractive charge interaction with ASP31 was lost (figure 2 and 3). An increase in the binding affinity (-45.39kcal/mol to -49.89kcal/mol) of BDA366 was observed.

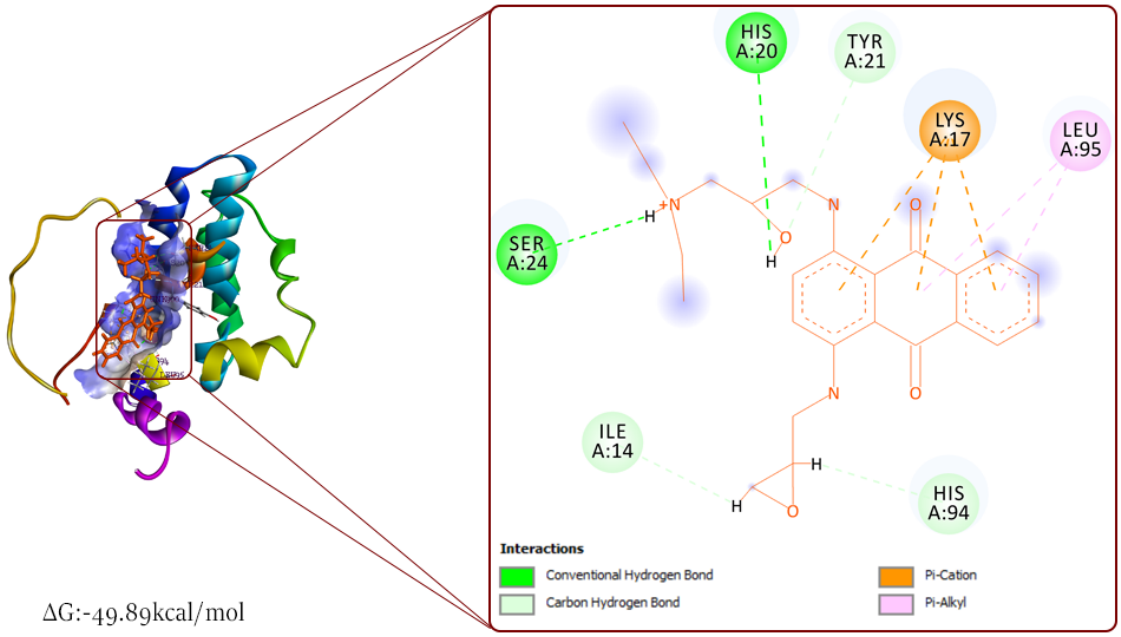

Figure 3: QM-Polarized Docking pose visualization: BDA366

Analysis of the lead compounds binding interactions showed the various interaction with BH4 residues (AA: 10-31) and also with non-BH4 residues (AA: 32, 35, 36, 39, 94, and 95) (figure 4). The most common BH4 residue interactions were with GLY13, MET16, and HIS20 while non-BH4 residues: ALA32 and GLU42 (figure 4). The binding affinity ranking of the lead compounds changed with compound 125480 still retain its position as the highest binding affinity with no significant change in its binding affinity, however compound 38831 had the second-highest binding affinity: -70.95kcal/mol (an increase from molecular docking binding affinity: -63.81kcal/mol). Compound 38279 had a binding affinity of -67.76kcal/mol which represents a decrease in the binding affinity (-71.97kcal/mol) observed in the previous docking run (figure 4 and 5).

Figure 4: QM-Polarized Docking binding interaction analysis: *a) Hydrogen bond and electrostatic interaction count b) Hydrophobic Interaction Count*

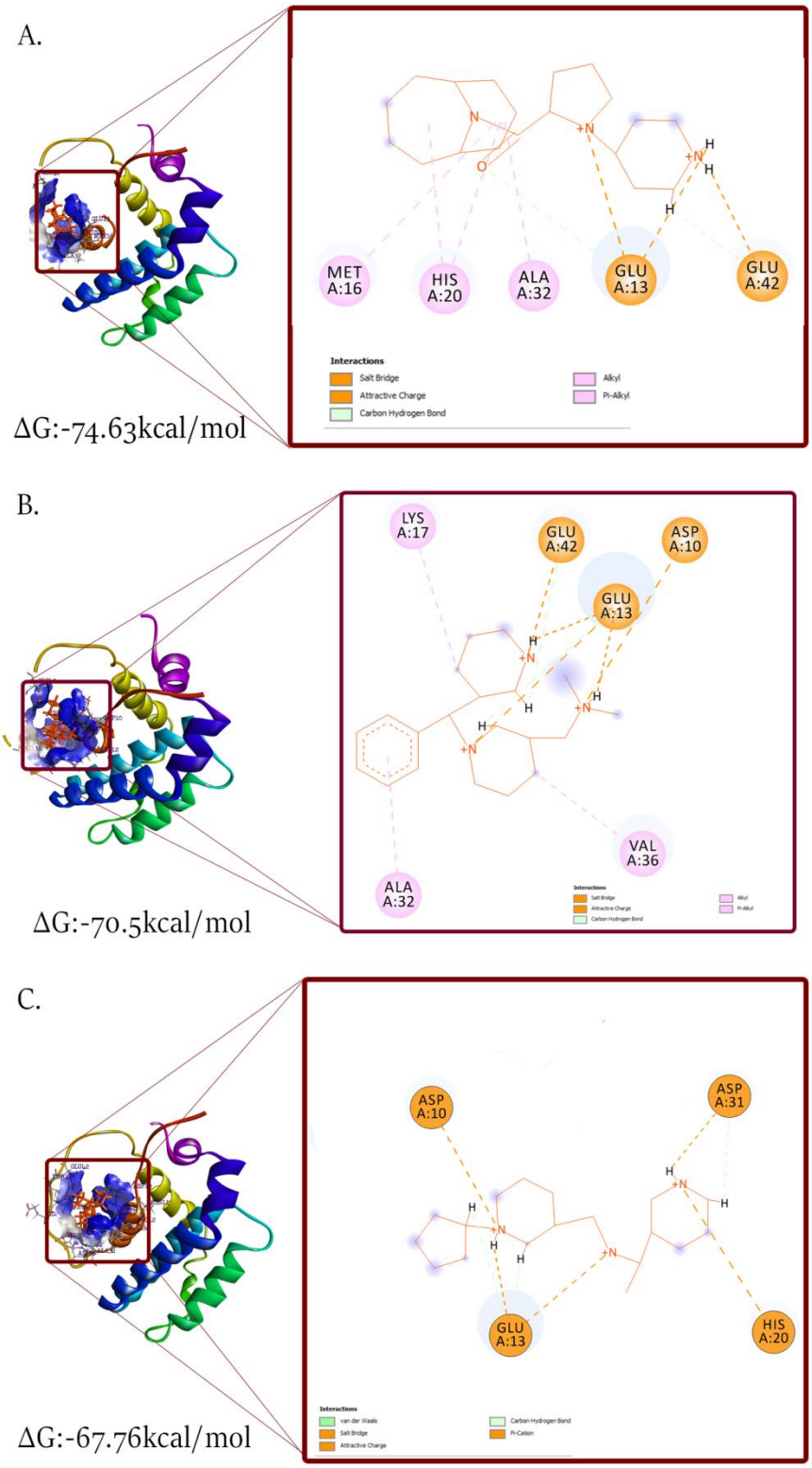

Figure 5: QM-Polarized Docking pose visualization: *a) 125480 b) 38831 c) 38279*

1.3. Induced-Fit Docking

As stated earlier that the primary aim of docking simulation is to search for binding conformation that is energetically favorable, in the process of searching the ligand is held flexible and different conformers are exhaustively searched till a conformer that binds favorably is achieved. This approach has a major limitation; in a biological system, both ligand and protein are flexible. The binding of a ligand to a protein active site induces series of conformation changes in the protein, whereby the protein 'shifts' to accommodate the ligand causing a better binding fit [32][22].

This is not accounted for in classical docking simulations. Advances in docking technique have, however, allowed for the combination of rigid receptor-ligand docking protocol and protein structure prediction protocol; the ligand is docked into the receptor rigidly and the protein structure prediction protocol predicts conformation changes that could occur as the ligand binds [22]. This protocol has been used to accurately predict and provide insights into the protein-ligand binding conformation of various protein targets [39][47][33].

Using Schrodinger's induced-fit docking module (see methods), the lead compounds (BDA366 as control) were subjected to Induced fit docking to provide insights into induced conformational changes that might occur and how this 'fits' might affect the binding affinity of the ligands previously observed.

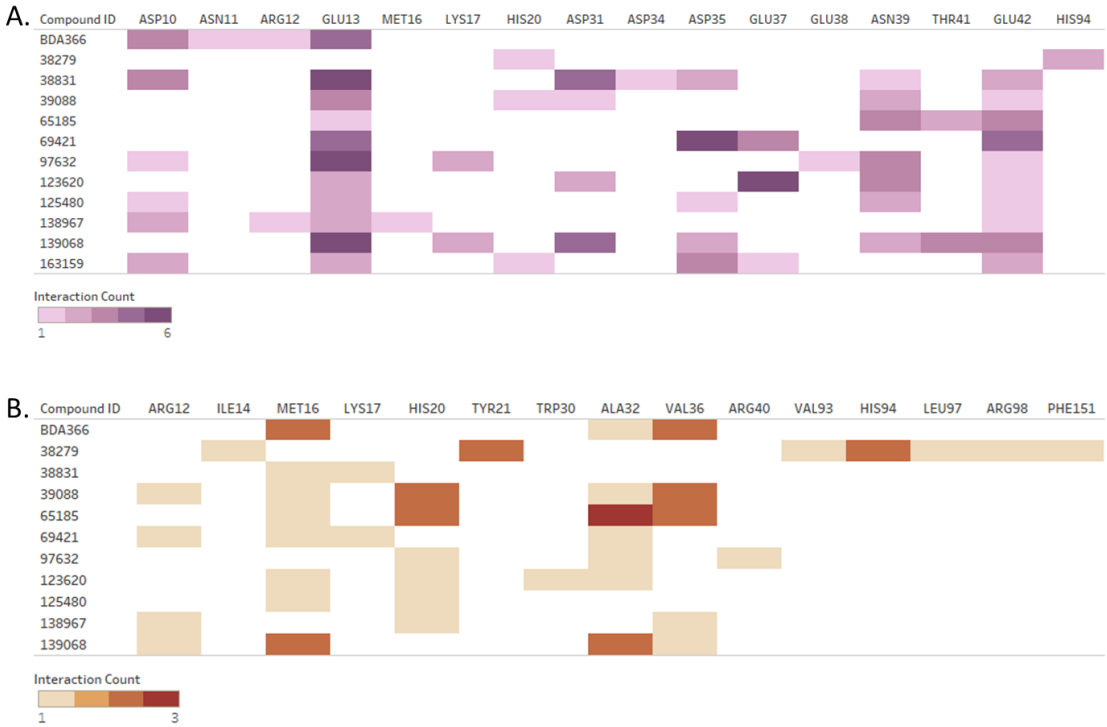

Figure 6: Induced-Fit Docking binding interaction analysis: *a) Hydrogen bond and electrostatic interaction count b) Hydrophobic Interaction Count*

Binding interaction analysis showed an increase in the number of interactions, with binding affinity increasing significantly (figure 5 and 8). BDA366 formed H-bond and electrostatic interaction with ASP10, ASN11, ARG12, GLU13, MET16, and hydrophobic bonds with ALA32 and VAL36 (figure 5 and 6). This result is consistent with the BH4 residues suggested by Han *et. al* [30]. It was suggested that BDA366 interacts with ASP10, ASN11, ARG12, and GLU13 in other to regulate BCL2. The interaction with these residues induces a conformational change in BCL2 by exposing its BH3 domain; converting BCL2 protein from a survival molecule (anti-apoptosis) to a killer molecule (pro-apoptosis) [30]. BDA366 had a binding affinity of -64.90kcal/mol.

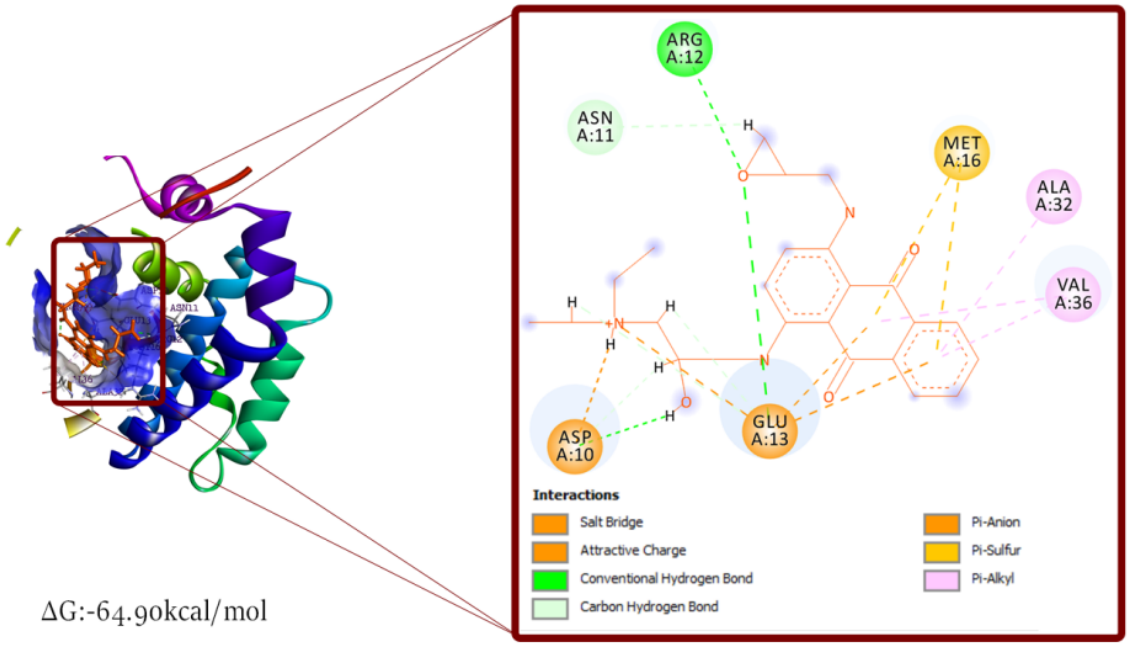

Figure 7: Induced-Fit Docking pose visualization: BDA366

The lead compounds had their binding affinity ranging from -100 to -85kcal/mol. The top 3 lead compounds were 139068, 138967, and 38831 (-100.78, -95.27 and -93.20kcal/mol respectively.). Each of these compounds interacted with at least two of the four BH4 residues by which BDA366 might regulate BCL2 alongside other BH4 domain residues and BH4 hydrophobic residues (figure 7). Compound 139068 interacted with BH4 residues AA: 12, 13, 16, 17, and 31. Non-BH4 interactions included AA: 32, 32, 36, 41, and 42. Compound 138967 interacted with BH4 residues AA: 10, 12, 13, 16, 20, and non-BH4 residues AA: 36 and 42. Compound 38831 interacted with BH4 domain residues AA: 10, 13, 16, 17, and 31 and non-BH4 domain residues AA: 35, 34, 39, and 42 (figure 7).

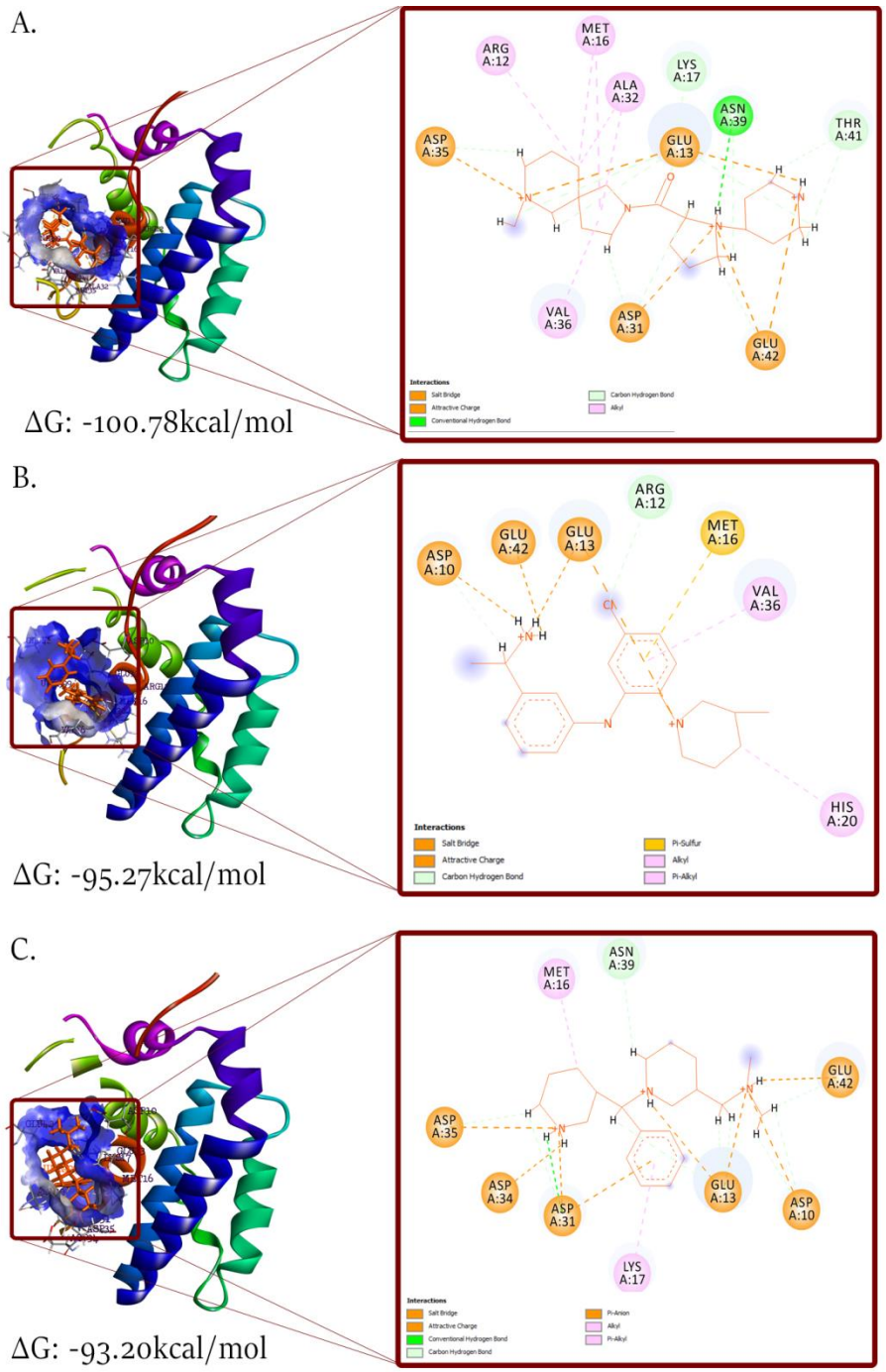

Figure 8: Induced Fit Docking pose visualization: *a) 139068 b) 138967 c) 38831*

So far, three different binding affinities have been elucidated from the three different docking protocols. In other to consolidate these binding affinities a weighted average was calculated. The weights used were calculated by adding a point (1) for every interaction with a BH4 residue ("reward") and subtracting 0.5 for every interaction with non-BH4 residues ("punish"). We infer that ligand interactions with BH4 domain residues should be giving more priority and "rewarded" than non-BH4 residue interactions.

Compound ID	Molecular Docking	Qm-polarized Docking	Induced-Fit Docking	Weighted Average
BDA366	-45.39	-49.89	-64.90	-53.40
38279	-71.97	-67.76	-85.82	-70.60
38831	-63.81	-70.95	-93.20	-83.88
39088	-66.25	-67.24	-88.43	-76.16
65185	-65.68	-65.78	-87.21	-67.26
69421	-67.10	-71.73	-94.47	-74.18
97632	-64.52	-65.83	-90.49	-75.00
123620	-66.31	-66.39	-91.18	-71.75
125480	-74.67	-74.63	-88.60	-78.53
138967	-62.90	-63.90	-95.27	-80.49
139068	-67.22	-69.30	-100.78	-84.46
163159	-64.35	-64.75	-94.47	-68.82

ΔG (kcal/mol)

-100.78 -45.39

Figure 9: Ligand binding affinity for different Docking Protocol

Using this protocol we calculated the weighted average binding affinity for the 11 lead compounds including BDA366 (control) (Figure 9). Compound 139068, 138967, and 38831 were the top 3 compounds with the binding affinity of -84.46kcal/mol, -83.88kcal/mol, and -80.49kcal/mol respectively.

1.4. QM-MM Optimization

The induced-fit docking (IFD) complex of the top three compounds (139068, 138967, and 38831) was optimized using Quantum mechanics - Molecular mechanics (QM-MM) calculations and compared. The ligand and interacting side-chain residues were treated as the QM region while the protein as MM region. This calculation aimed to geometrically optimize the docking complex thereby, validating bond interactions observed and to inquire how this interaction contributes to the calculated binding affinity. Using Qsite (see methods) the QM-MM optimization was implemented, after which the binding affinity of the optimized structure was calculated using MM-GBSA (see methods). The pre-optimized pose (IFD pose) and post-optimized pose (QM-MM pose) alongside their corresponding binding affinity were thereafter compared.

The optimized complex of compound 139068 maintained all interactions formed (when compared with pre-optimized pose) and a slight increase in calculated binding affinity (-102.15kcal/mol) (figure 10). Two new interactions were formed with Compound 138967: H-bond with GLU37 and electrostatic interactions with GLU38. H-bond with ARG12 was changed to hydrophobic interactions. This observed interaction changes resulted in the increase in binding affinity from -95.27 to -99.20kcal/mol (figure 10). Compound 38831 formed a new H-bond with THR41 and additional electrostatic interaction with ASP31. This new interaction however did not result in an increase in binding affinity but a reduction: -93.20 to -80.69kcal/mol (figure 10).

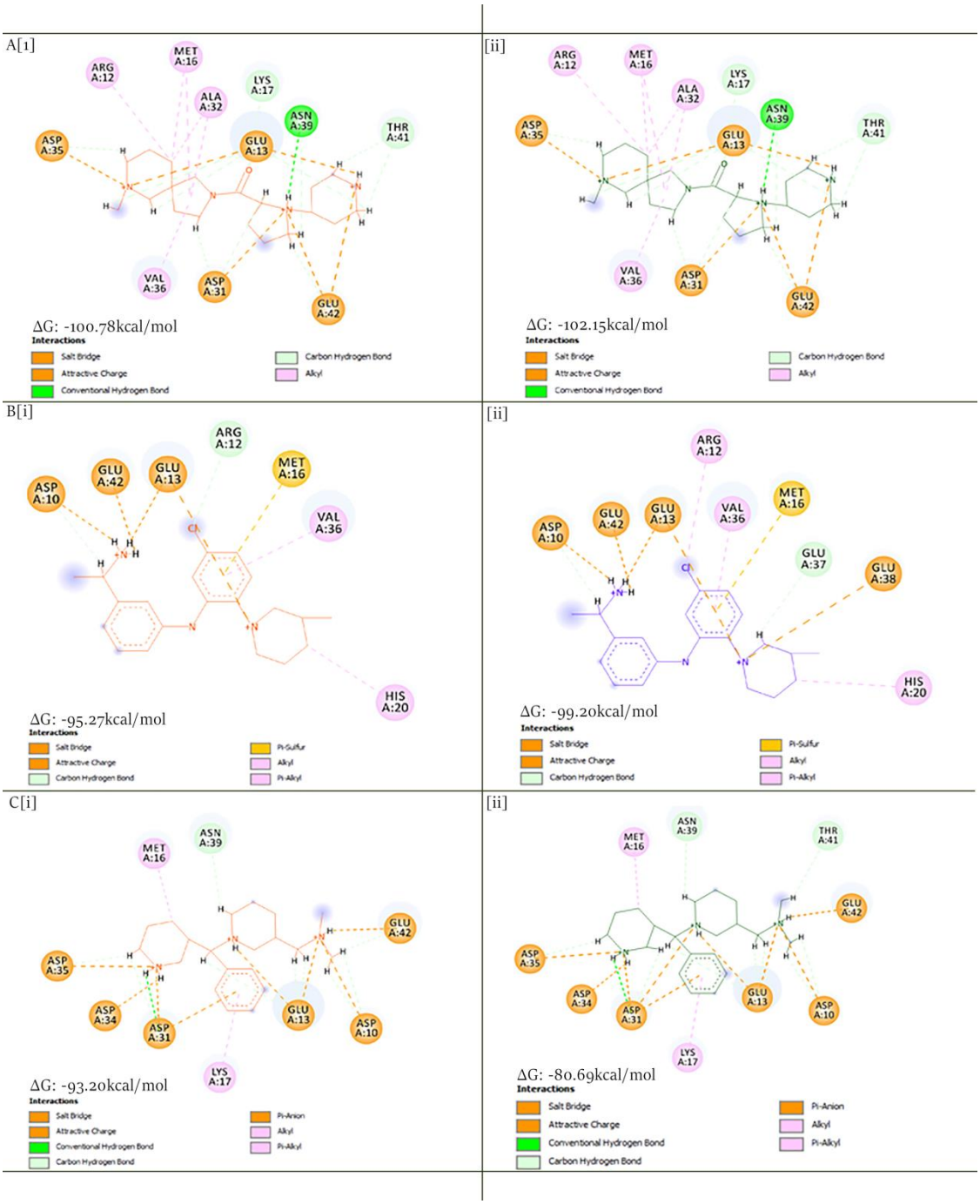

Figure 10: Comparison of pre (i) and post (ii) optimized induced-fit binding pose: *a) 139068 b) 138967 c) 38831*

1.5. Electronic Descriptor calculations

The chemical reactivity of a ligand provides insight into how a ligand might exercise its biological activity and provides insight on how to optimize for better biological activity. Using Single point energy module (see methods) Quantum electronic descriptors were calculated to show the mechanics of interaction. The descriptors calculated include HOMO, LUMO, and MESP; from these descriptors, the following were extrapolated (see methods): HOMO-LUMO gap, Ionization energy, Electron affinity, Chemical potential, Global hardness, and Global electrophilicity.

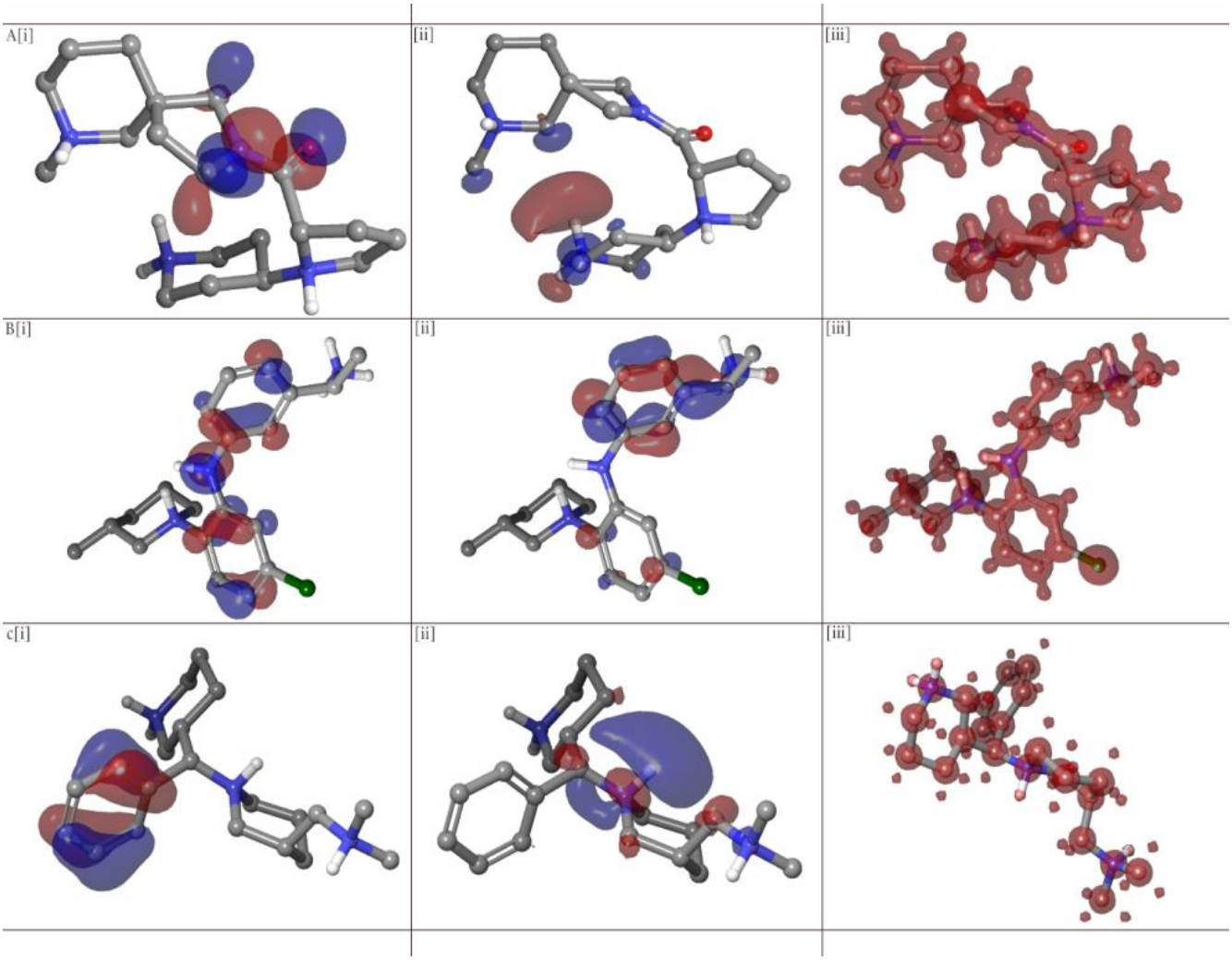

Figure 11: HOMO, LUMO, and MESP for*: a) 139068 b) 138967 c) 38831*

Highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO) gives information about the orbitals in a compound; calculating the HOMO-LUMO gap shows how easy it will be for an electron to move from a region of LUMO to HOMO and stability of the compounds; a small HOMO-LUMO gap could indicate a more reactive compound but less stable compound [34][35]. Ionization energy refers to the energy required to remove an electron

from a gaseous atom (it gives information about the energy of the orbital it originated from) while electron affinity is the energy change/released when an electron is added to a gaseous atom. Generally, electrons move from regions of high chemical potential to regions of low chemical potentials; therefore, a compound with low chemical potential indicates an electrophile. The Global hardness of a compound is the HOMO-LUMO gap of the compound [49]; compounds that are "extremely" hard are generally poor electrophiles. Global electrophilicity indicates the general reactivity of the compound. These descriptors are shown in Table 1.

Table 1: Quantum electronic descriptors

Compound Id	HOMO (eV)	LUMO (eV)	HOMO-LUMO GAP	Ionization energy (I)	Electron Affinity (A)	Global Hardness (η)	Chemical Potential (μ)	Global Electrophilicity (ω)
BDA366	-0.26	-0.17	0.09	0.26	0.17	0.09	-0.22	-2.39
139068	-0.57	-0.36	0.21	0.57	0.36	0.21	-0.47	-2.21
138967	-0.43	-0.25	0.18	0.43	0.25	0.18	-0.34	-1.89
163159	-0.6	-0.32	0.28	0.6	0.32	0.28	-0.46	-1.64
38831	-0.54	-0.33	0.21	0.54	0.33	0.21	-0.44	-2.07
123620	-0.39	-0.24	0.15	0.39	0.24	0.15	-0.32	-2.10
39088	-0.53	-0.34	0.19	0.53	0.34	0.19	-0.44	-2.29
69421	-0.53	-0.38	0.15	0.53	0.38	0.15	-0.46	-3.03
97632	-0.41	-0.26	0.15	0.41	0.26	0.15	-0.34	-2.23
125480	-0.44	-0.25	0.19	0.44	0.25	0.19	-0.35	-1.82
65185	-0.34	-0.12	0.22	0.34	0.12	0.22	-0.23	-1.05
38279	-0.57	-0.33	0.24	0.57	0.33	0.24	-0.45	-1.88

Compound 123620, 69421, and 97632 appeared to be the strongest electrophiles (among the 11 lead compounds) with a global hardness of 0.15, however, compound 69421 was the most reactive (Global electrophilicity: -3.03). Of the top three leads (139068, 138967, and 38831), compound 139068 and 38831 appeared to be the most reactive (Global electrophilicity: 0.21); however, Compound 138967 was the strongest electrophile (chemical potential: -0.34; HOMO-LUMO gap: 0.18; Global Hardness: 0.18).

Molecular electrostatic potential (MESP) of the Compounds were calculated; it refers to the work done in bringing a unit positive charge from infinity to a point [36]; it shows the surface charge distribution on the ligand thereby identify regions that might be involved in the electrophilic reaction (positively charged) and regions involved in the nucleophilic reaction (negatively charged) [36]. The MESP of the top 3 was visualized and showed all atoms of ligands to be positively charged (figure 11), this, therefore, suggests that the ligands interact with the BH4 domain via electrophilic reactions.

1.6. BH4 specificity investigation

One of the downsides of small molecule inhibitors that have been developed to targeting the BH3 homology binding domain is that the small molecules also bind with other proteins in the BCL2 families (BCL-XL, MCI-1, BFI-1. etc.) due to the highly conserved nature of the BH3 domains. Since the primary aim of this study is to identify BH4 specific binding small molecules, we, therefore, docked these compounds into the BH3 domain of BCL2, BCL-Xl, and MCL-1 and compare their binding affinity against BDA366.

Using BDA366 as a negative control (since BDA366 has been shown experimentally to be highly selective for BH4), Venetoclax and ABT 737 [8] as a positive control for BCL2 and BCL-XL BH3 docking, and 2,5-substituted benzoic acid[] as a positive control for MCL-1 BH3 domain docking, the lead compound was docked and binding affinity calculated.

Table 2: BH4 specificity docking

Compound ID	**BCL2(BH3) (kcal/mol)**	**BCL-XL(BH3) (kcal/mol)**	**MCL-1(BH3) (kcal/mol)**
Venetoclax	-103.14	-86.17	-
ABT737	-115.44	-80.24	-
BDA366	-70.57	-55.99	-84.24
2,5-substituted benzoic acid	-	-	-124.33
38279	-66.93	-48.52	-80.02
38831	-70.96	-43.66	-62.46
39088	-58.2	-41.09	-57.68
65185	-53.01	-44.73	-67.47
69421	-78.4	-47.01	-58.52
97632	-53.25	-39.65	-59.28
123620	-75.49	-39.53	-77.62
125480	-62.98	-40.37	-69.24
138967	-67.68	-54.66	-83.84
139068	-70.01	-47.03	-69.32
163159	-67.07	-44.79	-55.19

The results showed Compounds 123620 and 69421 had a higher binding affinity (-75.49 and -78.40kcal/mol respectively) for BCL2- BH3 domain than BDA366 (-70.57kcal/mol) (Table 2) suggesting that this compounds might also bind and interact with BCL2 BH3 domains. However, docking with other BCL2 family protein none of the lead compounds bind to the BH3 domain with a binding affinity greater than BDA366 (Table 2

Discussion

BCL2 family protein is responsible for the intrinsic pathway for cellular apoptosis; each member of the family has a homology binding site where heterodimerization and homodimerization among the BCL2 family protein occurs. Except for BH3 only protein all possesses BH 1-4 domains [10][8]. Most small molecules (ABT737, ABT376, Venetoclax, Disarib), developed so far have targeted the BH3 domain, however, due to the highly conserved nature of BH3 among the BCL2 family, non-specificity of this compounds occurs leading to toxicity and loss of efficacy [8]. The BH4 domain of BCL2 protein is not conserved in another BCL2 family (not even it is closet protein: BCL-xl)[13]; also BCL2 protein wherein the hydrophobic residues of the BH4 domain were mutated resulted in the 'conversion' of BCL2 protein from anti-apoptosis to pro-apoptosis protein [9]. Using BDA366 as an antagonist, the BH4 domain of BCL2 was targeted and used to show proof of concept [7].

However, BDA366 is the only specific BH4 binding small molecule that has been developed so far. Docking ~99,000 small molecules 11 lead compounds where identified to show a higher binding affinity for the BH4 domain when compared with BDA366. In other to elucidate a binding mode, QM-polarized and induced-fit docking protocol was implemented and top 3 induced-fit docking pose geometrically optimized using QM/MM optimization. Han *et al* identified four amino acid: Asp10, ASN11, ARG12, GLU13 by which BDA366 carries out its antagonist activity [7], aside for these four, interactions with MET16, LYS17, HIS20, and Asp31 were also observed; also interactions with hydrophobic residues of BH4 domain has been shown to cause the 'conversion' from pro-apoptosis to anti-apoptosis. With this as a guide binding interaction (Hydrogen bonds, electrostatic interaction, and hydrophobic interactions) analysis were carried out.

Interactions with GLU13, MET16, and HIS20 were consistently observed, aside these, interactions with ASP10, ARG12, and LYS17 were common. From these, interactions with ASP10, ARG12, GLU13, MET16, LYS17, and HIS20 are suggested as the binding mode for the 11 lead compounds observed (it should be noted that not all this amino acid interactions were observed at the same time).

Three different binding affinities were calculated (Molecular docking, QM-polarized docking, and Induced-fit docking). As a way to rank the lead compounds and consolidate the various binding affinities, a weighted average was calculated. This average rewarded every interaction with BH4 domain residues and punished for every non-BH4 residue. When compared with BDA366 all the lead binding affinity was significantly higher. We suggest that these high binding affinities observed could be as a result of interacting with other non-BH4 residues (especially GLU 42 which was consistently observed); recently BDA366 was shown to interact with PRO127 a non-BH4 residue when subjected to molecular dynamics simulation [37]; this, therefore, suggests that interaction with non-BH4 residues plays a role in regulating BCL2 protein, and in this study possibly an increase in binding affinity.

Quantum mechanics descriptors calculated to explain chemical reactivity (electronic properties) of the compounds showed the compounds to be electrophiles with various degrees of reactivity (Global electrophilicity, HOMO-LUMO gap), electrophilic strength (Global hardness, Chemical potential), and kinetic stability (HOMO-LUMO gap). Compound 69421 was the most reactive and among the strongest electrophile.

Docking simulations with closely related anti-apoptosis BCL2 protein (BCL-Xl, MCL-1) and BCL2 protein in their BH3 domains, showed the leads not to have binding affinity higher than BDA366 (BDA366 was used as a negative control), this data, potentially suggest that this leads might be specific for BH4 domain of BCL2 protein, however there is still need for experimental validation.

One limitation of this study is the absence of experimental data to fully validate the biological activity for this lead compound, also the use of Molecular dynamics simulations might also validate the binding mode for the lead compounds we suggested. However this study has suggested set of compounds that can be further investigated and optimized as BH4 binding small molecules.

Conclusion

Targeting the BH4 domain in BCL2 protein is a promising strategy in ‘converting’ survival proteins to death proteins in cancer cells; identifying and developing BH4 small molecule is therefore important. Using computer simulations, 11 diverse small molecules have been identified; a binding mode has been elucidated and chemical reactivity of the compounds explained. However, experimental data are still needed to validate the antagonistic activity of these lead compounds.

References

[1] Brahmer J, Reckamp KL, Baas P, Crinò L, Eberhardt WEE, Poddubskaya E, et al. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. N Engl J Med 2015;373:123–35. https://doi.org/10.1056/NEJMoa1504627.

[2] Carrizosa DR, Gold KA. New strategies in immunotherapy for non-small cell lung cancer. Transl Lung Cancer Res 2015;4:553–9. https://doi.org/10.3978/j.issn.2218-6751.2015.06.05.

[3] Reed JC. Apoptosis-based therapies. Nat Rev Drug Discov 2002;1:111–21. https://doi.org/10.1038/nrd726.

[4] Hu Y, Zhou L, Zhu X, Dai D, Bao Y, Qiu Y. Pharmacophore modeling , multiple docking , and molecular dynamics studies on Wee1 kinase inhibitors. J Biomol Struct Dyn 2018;0:13. https://doi.org/10.1080/07391102.2018.1495576.

[5] Wang L, Xiang S, Williams KA, Dong H, Bai W, Nicosia S V., et al. Depletion of HDAC6 Enhances Cisplatin-Induced DNA Damage and Apoptosis in Non-Small Cell Lung Cancer Cells. PLoS One 2012;7. https://doi.org/10.1371/journal.pone.0044265.

[6] Haggarty SJ, Koeller KM, Wong JC, Grozinger CM, Schreiber SL. Domain-selective small-molecule inhibitor of histone deacetylase 6 (HDAC6)-mediated tubulin deacetylation. Proc Natl Acad Sci U S A 2003;100:4389–94. https://doi.org/10.1073/pnas.0430973100.

[7] Yang TM, Barbone D, Broaddus VC. Bortezomib-induced apoptosis in the treatment of non-small cell lung cancer. J Intern Med Taiwan 2009;20:106–19. https://doi.org/10.6314/JIMT.2009.20(2).02.

[8] Radha G, Raghavan SC. BCL2: A promising cancer therapeutic target. Biochim Biophys Acta - Rev Cancer 2017;1868:309–14. https://doi.org/10.1016/j.bbcan.2017.06.004.

[9] Hirotani M, Zhang Y, Fujita N, Naito M, Tsuruo T. NH2-terminal BH4 domain of Bcl-2 is functional for heterodimerization with Bax and inhibition of apoptosis. J Biol Chem 1999;274:20415–20. https://doi.org/10.1074/jbc.274.29.20415.

[10] Liu Z, Wild C, Ding Y, Ye N, Chen H, Wold EA, et al. BH4 domain of Bcl-2 as a novel target for cancer therapy. Drug Discov Today 2016;21:989–96. https://doi.org/10.1016/j.drudis.2015.11.008.

[11] Rong YP, Barr P, Yee VC, Distelhorst CW. Targeting Bcl-2 based on the interaction of its BH4 domain with the inositol 1,4,5-trisphosphate receptor. Biochim Biophys Acta - Mol Cell Res 2009;1793:971–8. https://doi.org/10.1016/j.bbamcr.2008.10.015.

[12] Rong YP, Bultynck G, Aromolaran AS, Zhong F, Parys JB, De Smedt H, et al. The BH4 domain of Bcl-2 inhibits ER calcium release and apoptosis by binding the regulatory and coupling domain of the IP3 receptor. Proc Natl Acad Sci U S A 2009;106:14397–402. https://doi.org/10.1073/pnas.0907555106.

[13] Han B, Park D, Li R, Xie M, Owonikoko TK, Zhang G, et al. Small-Molecule Bcl2 BH4 Antagonist for Lung Cancer Therapy. Cancer Cell 2015;27:852–63. https://doi.org/10.1016/j.ccell.2015.04.010.

[14] Chevillard F, Kolb P. SCUBIDOO: A Large yet Screenable and Easily Searchable Database of Computationally Created Chemical Compounds Optimized toward High Likelihood of Synthetic Tractability. J Chem Inf Model 2015;55:1824–35. https://doi.org/10.1021/acs.jcim.5b00203.

[15] Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des 2013;27:221–34. https://doi.org/10.1007/s10822-013-9644-8.

[16] Jacobson MP, Pincus DL, Rapp CS, Day TJF, Honig B, Shaw DE, et al. A Hierarchical Approach to All-Atom Protein Loop Prediction. Proteins Struct Funct Genet 2004;55:351–67. https://doi.org/10.1002/prot.10613.

[17] Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M. Epik: A software program for pKa prediction and protonation state generation for drug-like molecules. J Comput Aided Mol Des 2007;21:681–91. https://doi.org/10.1007/s10822-007-9133-z.

[18] Singh N, Villoutreix BO, Ecker GF. Rigorous sampling of docking poses unveils binding hypothesis for the halogenated ligands of L-type Amino acid Transporter 1 (LAT1). Sci Rep 2019;9:1–20. https://doi.org/10.1038/s41598-019-51455-8.

[19] Lyne PD, Lamb ML, Saeh JC. Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. J Med Chem 2006;49:4805–8. https://doi.org/10.1021/jm060522a.

[20] Kombo DC, Grinevich VP, Hauser TA, Sidach S, Bencherif M. QM-polarized ligand docking accurately predicts the trend in binding affinity of a series of arylmethylene quinuclidine-like derivatives at the α4β2 and α3β4 nicotinic acetylcholine receptors (nAChRs). Bioorganic Med Chem Lett 2013;23:4842–7. https://doi.org/10.1016/j.bmcl.2013.06.094.

[21] Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. J Med Chem 2004;47:1750–9. https://doi.org/10.1021/jm030644s.

[22] Sherman W, Day T, Jacobson MP, Friesner RA, Farid R. Novel procedure for modeling ligand/receptor induced fit effects. J Med Chem 2006;49:534–53. https://doi.org/10.1021/jm050540c.

[23] Murphy RB, Philipp DM, Friesner RA. A mixed quantum mechanics/molecular mechanics (QM/MM) method for large-scale modeling of chemistry in protein environments. J Comput Chem 2000;21:1442–57. https://doi.org/10.1002/1096-987X(200012)21:163.0.CO;2-O.

[24] Bochevarov AD, Harder E, Hughes TF, Greenwood JR, Braden DA, Philipp DM, et al. Jaguar : A High-Performance Quantum Chemistry Software Program with Strengths in Life and Materials Sciences. Int J Quantum Chem 2013, 2013:2110–2142. https://doi.org/10.1002/qua.24481.

[25] Gupta VP. Density Functional Theory (DFT) and Time Dependent DFT (TDDFT). 2016. https://doi.org/10.1016/b978-0-12-803478-1.00005-4.

[26] Pavitha P, Prashanth J, Ramu G, Ramesh G, Mamatha K, Venkatram Reddy B. Synthesis, structural, spectroscopic, anti-cancer and molecular docking studies on novel 2-[(Anthracene-9-ylmethylene)amino]-2-methylpropane-1,3-diol using XRD, FTIR, NMR, UV–Vis spectra and DFT. J Mol Struct 2017;1147:406–26. https://doi.org/10.1016/j.molstruc.2017.06.095.

[27] Parr RG, Szentpály L V., Liu S. Electrophilicity index. J Am Chem Soc 1999;121:1922–4. https://doi.org/10.1021/ja983494x.

[28] Tsai K, Wang S, Hsiao N, Li M, Wang B. Bioorganic & Medicinal Chemistry Letters The effect of different electrostatic potentials on docking accuracy : A case study using DOCK5 . 4. Bioorg Med Chem Lett 2008;18:3509–12. https://doi.org/10.1016/j.bmcl.2008.05.026.

[29] Chen YC. Beware of docking! Trends Pharmacol Sci 2015;36:78–95. https://doi.org/10.1016/j.tips.2014.12.001.

[30] Han B, Park D, Li R, Xie M, Owonikoko TK, Zhang G, et al. Small-Molecule Bcl2 BH4 Antagonist for Lung Cancer Therapy. Cancer Cell 2015;27:852–63. https://doi.org/10.1016/j.ccell.2015.04.010.

[31] Kurczab R. halogen bonding The evaluation of QM / MM-driven molecular docking combined with MM / GBSA calculations as a halogen-bond scoring strategy. Strutural Sci Cryst Eng Mater 2017:1–7. https://doi.org/10.1107/S205252061700138X.

[32] A. Sotriffer C. Accounting for Induced-Fit Effects in Docking: What is Possible and What is Not? Curr Top Med Chem 2012;11:179–91. https://doi.org/10.2174/156802611794863544.

[33] Barreca ML, Iraci N, De Luca L, Chimirri A. Induced-fit docking approach provides insight into the binding mode and mechanism of action of HIV-1 integrase inhibitors. ChemMedChem 2009;4:1446–56. https://doi.org/10.1002/cmdc.200900166.

[34] Prashanth J, Ramesh G, Laxman Naik J, Kishan Ojha J, Venkatram Reddy B. Molecular geometry, NBO analysis, Hyperpolarizability and HOMO-LUMO energies of 2-azido-1-phenylethanone using Quantum chemical calculations. Mater. Today Proc., vol. 3, Elsevier Ltd; 2016, p. 3761–9. https://doi.org/10.1016/j.matpr.2016.11.025.

[35] Aihara J. Reduced HOMO−LUMO Gap as an Index of Kinetic Stability for Polycyclic Aromatic Hydrocarbons.pdf n.d.

[36] Gupta VP. Principles and Applications of Quantum Chemistry. n.d.

[37] Li T, Cui Y, Wu B. Molecular dynamics investigations of structural and functional changes in Bcl-2 induced by the novel antagonist BDA-366. J Biomol Struct Dyn 2019;37:2527–37. https://doi.org/10.1080/07391102.2018.1491424.