BIRZEIT UNIVERSITY

FACULTY OF GRADUATE STUDIES

Kinase Inhibitors as Potential Drugs: A Molecular Dynamics Simulation Study

This thesis is submitted in partial fulfillment of the requirements for the Master's Degree in

Applied Chemistry from the Faculty of Graduate Studies at Birzeit University, Palestine

Iman Hammad

Supervisor: Dr. Mazen Hamed

2017

Kinase Inhibitors as Potential Drugs: A Molecular Dynamics Simulation Study

By Iman Esam Hammad

Accepted by the Faculty of Graduate studies, Birzeit University, in partial fulfillment of the requirements of the degree of Master of Applied Chemistry

Committee Members	Signature
Mazen Hamed, Ph. D (Principal Advisor)	……
Wael Karain, Ph. D (Member)	……..
Abdallah Sayyed-Ahmad, Ph. D (Member)	…….

ACKNOWLEDGEMENTS

I would like to extend my sincere appreciation to my supervisor Dr. Mazen Hamed for his support, guidance and motivation throughout the project. I would also like to express my gratitude to my family and husband for providing me with veracious support and continuous motivation throughout this project and my years of study.

TABLE OF CONTENTS

ABSTRACT

Molecular dynamics simulation and binding free energy (ΔG_{bind}) calculations were done to inspect the interaction between five inhibitors and PDK-1 kinase. The free energy (ΔG_{bind}) values were computed using MM-GBSA and MM-PBSA free energy calculation methods.

The entropic contribution of the binding free energy ΔS was computed using normal mode (NMODE) method. The change of enthalpy (ΔH) was calculated using the equation ΔG=ΔH-T.ΔS.

There is a noticeable difference in the values of ΔG depending on the calculation method whether MM-PBSA or MM-GBSA, and this is due to the calculation different approach in each case.

PDK-1 kinase is a well validated anticancer target. The results gave the binding modes between PDK-1 kinase and the five inhibitors, which can be used in the future in the drug design processes for cancer treatment. The placement of water molecules in the binding sites are known. This can be used to design better inhibitors through adding substituents to the inhibitor to replace a water molecule that binds kinase in the active site based on the creation of an inhibitor that includes a structural water mimic.

Through Molecular dynamics simulation, we identify potency PDK-1 inhibitor (5) that have unique binding to the inactive kinase conformation (DFG-out). On the other hand, inhibitors (1-4) are consider as classical ATP-competitive kinase inhibitors (Type I) which are bind to the active conformation DFG-in.

It was reported that **type I kinase inhibitors** form water-mediated hydrogen bond networks (both water molecules W1 and W2 are commonly observed) and the ligand does not extend to the water-filled cavity. These two features distinguish type I from type II inhibitors and these two features were obvious in our study in the binding modes of inhibitors (1-4) with the PDK-1 kinase.

LIST OF FIGURES

LIST OF TABLES

ملخص

في هذه الدراسة قمنا باستخدام التمثيل الجزيئي من اجل حساب قيم طاقة الربط لخمس مثبطات لانزيم الكاينيز PDK 1 –باستخدام الطريقتين (Molecular Mechanics/Generlized-Born surface area) (Molecular Mechanics/Poisson Boltzmann surface area) . وتم حساب قيم مساهمة الفوضى T.ΔS باستخدام طريقة (Nmode)وتم حساب التغير في الطاقة HΔ باستخدام هذه المعادلة ΔG = ΔH – T.ΔS.

هناك فرق ملحوظ في قيم التغير في طاقة الربط اعتمادا على طريقة الحساب سواء باستخدام (MM-PBSA or MM-GBSA) بحيث كانت نتائج MM-GBSA افضل من نتائج MM-PBSA وهذا يرجع الي الاختلاف في طريقة الحساب التي يتم اعتمداها من قبل MM-PBSA and MM-GBSA.

كذلك وضحت نتائج هذه الدراسة كيفية ارتباط هذه المثبطات بانزيم الكاينيز PDK-1 والتي يمكن استخدامها لعلاج السرطان, بحيث أن موقع جزيئات الماء كانت موجودة في مواقع الربط ويمكن استخدام هذه النتائج في تصميم مثبطات أفضل من خلال إضافة ذرات مثل الأوكسجين والهيدروجين للمثبط لكي يكون قادرا على تكوين روابط هيدروجينية لكي يحل مكان جزيئات الماء التي كانت موجودة في مواقع الربط من خلال الاستناد على مبدا تكوين مثبطات جديدة محاكية لجزيئات الماء.

من خلال محاكاة الديناميات الجزيئية قمنا بتحديد مثبط (المركب 5) والذي يرتبط بشكل فريد بأنزيم الكاينيز PDK-1 لتكوين كاينيز غير نشيط على نقيض المركبات الأخرى(1-4) والتي يتم اعتبراها مثبطات كاينيز كلاسيكية منافسة ل ATP . ومن الجدير بالذكر ان المثبط (المركب 5) هو دواء فعال وذلك بسبب قيمة طاقة الربط الكبيرة ΔG = -52.3 Kcal/mol, هذه القيمة الكبيرة ناتجة من تكوين خمس روابط هيدروجينية قوية بين هذا المثبط وأنزيم الكاينيز PDK-1.

وفقا لما ذكرته التقارير فان النوع الأول من مثبطات الكاينيز (مثبطات كلاسيكية منافسة ل ATP) تتميز عن النوع الثاني من المثبطات من خلال تشكيل شبكات من الروابط الهيدروجينية بواسطة جزيئين من الماء وهذا يتوافق مع نتائج هذه الدراسة بحيث انه تم ملاحظة هذه الميزة في المثبطات (المركب4) و (المركب 3).

Chapter 1

1. INTRODUCTION

In this work, structure-based drug design that employs molecular dynamics, binding free energy calculations is used to investigate anti-cancer inhibitors of 3-phosphoinositide dependent kinase-1 (PDK-1)[1]. This computational drug design approach has been successfully applied to both lead optimization and hit identification against PDK-1[1].

Fast expansion in this area has been made possible by advances in software and hardware computational power and sophistication, identification of molecular targets, and an increasing database of publicly available target protein structures. CADDD is being utilized to identify hits (active drug candidates), select leads (most likely candidates for further evaluation), and optimize leads i.e. transform biologically active compounds into suitable drugs by improving their physicochemical, pharmaceutical, ADMET/PK (pharmacokinetic) properties.[2]

The purpose of usages of computational tools as we hope to improve effectiveness and efficiency of drug discovery and development process, decrease use of animals, and increase predictability.[2]

1.1 Computer-aided drug design

A general strategy for drug discovery efforts can be summarized by the following steps: Identifying a target, screening for inhibitors, isolating hits, optimization and selection of a candidate molecule for clinical studies[1]. Drug discovery is a very long and expensive process in which twelve to twenty-four years are needed to discover and develop a new drug. In addition, the average cost to develop a new drug into markets is more than $1 billion[1].

Historically, Paul Ehrlich[3] was the first person to postulate on the existence of chemoreceptors that can be exploited therapeutically. It was reported that initial stages in chemotherapy was focused on the isolation and purification of active ingredients from natural products such as plants. This was followed by rational approach which was based on understanding the mechanisms of action and the drug-receptor interactions[3]. Despite using the newer approach, it was reported that only 18 new chemical entities (NCE) were approved in the year 2005 and 2006[4]. For this reason, a newer approach was developed to increase the efficiency of drug discovery process.

Computer-aided drug design (CADD) is an example on the newer approaches that is using computer based techniques to analyze molecules and molecular systems to predict their biological properties.[5] This approach helps in the identification and optimization of new potential drugs[5]. Computational drug design has played an

important role in the successful development of marketed drugs such as saquinavir, ritonavir, and indinavir were utilized in the treatment of human immunodeficiency virus (HIV)[6].

There are two broad strategies in computational drug design:[5]

1. Using Ligands for drug design

2. Receptor-based drug design Related to structure

The first strategy is a ligand-based drug design (LBDD) that is usually applied if a number of biologically active compounds are characterized and the target 3D structure is unknown[5]. By analyzing the physico-chemical properties of these active molecules, LBDD aims at predicting new chemical structures that are likely to have better biological properties[6]. There are many methods that use active known molecules to predict new ligands including Quantitative structure-activity relationship, Pharmacophore modeling and shape-based screening methods[7]. For example, Quantitative structure-activity relationship (QSAR) method is based on the regression analysis of relationship between biological activity of set of homologous compounds and their various physico-chemical descriptors such as hydrophobic properties, electrostatic properties, steric factors, donor-acceptor[5].

Another example is the Pharmacophore modeling which based on accounting for direct protein–ligand binding, and ignores other interactions outside the pharmacophore region[7]. A Pharmacophore is defined as the three-dimensional representation of active chemical features of active compounds[7].

The second strategy is Structure-based drug design methods are usually used when the X-ray crystal/NMR structure of the target protein is available[5]. The core strategy of this approach is based on analyzing the active ligand interaction with the binding site on the target protein. This means that ligands that exhibit similar interactions to the active ones will have similar biological effects[6].

Ligand Docking and de novo drug design are two examples on the SBDD methods. Docking methods require the structure of the target protein to estimate the binding energy of a number of ligands and rank them according to their estimated binding free energies. In comparison, De novo drug design require the structure of the active site as starting point[5].

1.2 Protein Kinases

There are more than 500 protein kinases known in the human genome. They are the second largest group of currently investigated drug targets [8,9]. Protein kinases main biological function is to catalyze the transfer of phosphoryl group of ATP to a hydroxyl group of threonine, tyrosine or serine residues[10]. Signal-transduction

pathways are activated by the phosphorylation of certain proteins which are responsible for the transition of cellular signals throughout the cells and to the nucleus[10].

The deregulation of phosphorylation reaction due to mutations in kinase genes are known to cause 218 diseases. For this reason, protein kinases are considered as an important therapeutic target in different diseases and viral infections. Therefore, the protein kinases are considered as important effectors in human pathology[10] and thereby a highly attractive therapeutic target in drug discovery[9].

1.3 Role of the kinase enzymes in cancer

It was reported that abnormality in protein kinases can lead to the development of several reported disorders and major diseases such as, endocrine disorders, cardiovascular disease are due to malfunction of phosphorylation process[11].

There are 500 genes that encode kinases are involved in cancer, while other oncogenes activate kinases or are phosphorylated by other kinases. This what makes kinases as potential targets for drug development[11].

There are three specific sites in protein kinase that are involved in phosphorylation: an ATP binding site, a domain catalyzing the transfer of phosporyl group from ATP (phosphate pocket) and a substrate- binding site (PIF-pocket).[12]

Perturbed signal transduction provokes deregulation of different processes in cell migration, which can lead to malignant phenotype. 50% or more of receptor tyrosine kinase (RTK) and several serine/threonine kinases have been perturbed in different human malignancies. Irregular activity of a protein kinase which occur by genomic rearrangements result in hybrid proteins with catalytic domains of a protein kinase and another unrelated protein.[12] A second mechanism that damages the normal function of protein kinases is the mutations responsible for kinase constitutive activity. The third mechanism is explained by increasing expression of protein kinases. Finally, deregulation of kinase activity by activation of oncogenes can also contribute to tumorigenesis.

The chemotherapy treatment of breast cancer by using cyclophosphamide causes weight gain, ovarian failure, cardiac toxicity and Probability of developing a second cancers[13]. Studies estimated that in 2050 the global cancer will increase to 27million new cases[14]. Therefore, there are urgent need to discover a potent and selective cancer drug with no side effects.

1.4 3-Phosphoinositide dependent kinase-1 (PDK-1)

PDK-1 is a 556 amino acids serine/threonine kinase that belongs to the AGC protein kinase family. It plays an important role in the phosphorylation and activation of a number of proteins such as protein kinase B (PKB), protein kinase

C isoforms, the p70 ribosomal S6 kinase and serum and glucocorticoid-induced kinase[15].

Structure of PDK-1 is consisting of two lobes: C-terminal lobe and N-terminal lobe and is similar in overall structure to PKA[16]. C-terminal pleckstrin homology domain (PH) is essential for interaction of PDK-1 with the cell membrane because it binds with phosphoinositide lipids of the plasma membrane[15].

PDK-1 residues Val124, Val127 on the α-helix, Lys115, Ile119, Ile118 on the B-helix, and Leu155 on B-sheet form a hydrophobic pocket (PIF) pocket (Fig 1.1) [16]. Since Leu155 is presented at the center of this pocket, whereas the other residues form a lining of the inside wall of the pocket[17].

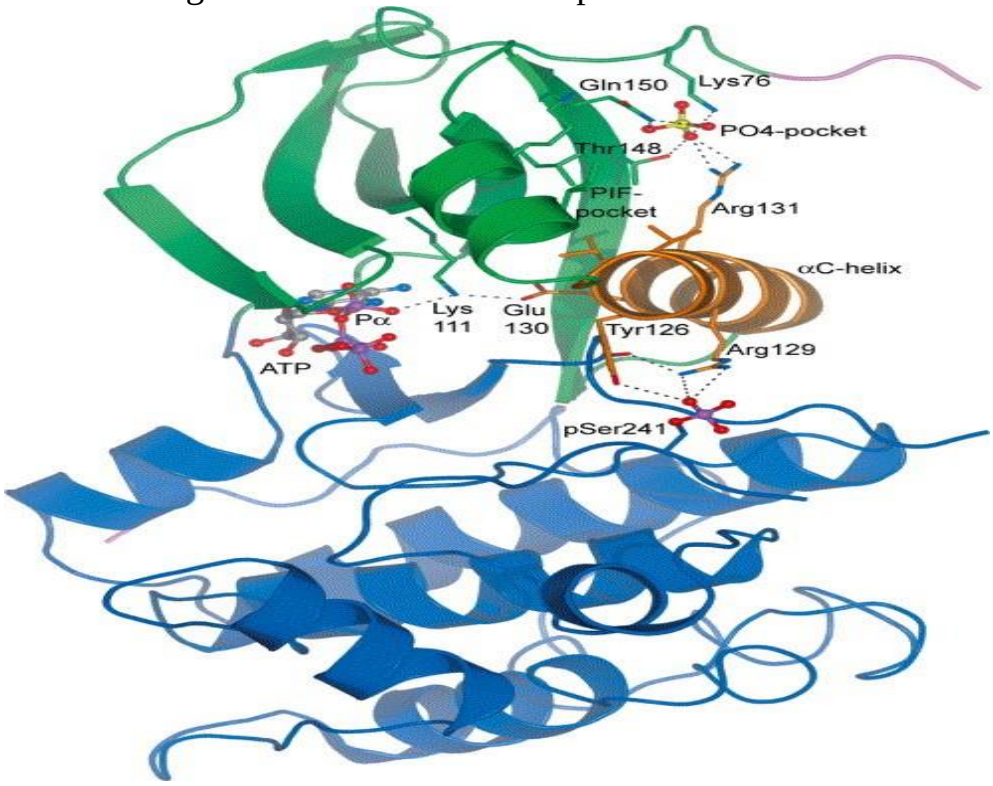


Figure 1.1: Structure of PDK-1 kinase domain with ATP molecule. The C-terminal lobe (in blue), the C-helix in green, the N-terminal lobe (in green), and the pSer241 in the T-loop (in purple/red spheres)[16]

S6K1 substrate interacts with the PIF-pocket of PDK-1 with higher affinity when it is phosphorylated at its hydrophobic motif. This indicated that the phosphate-binding site may be located close to the PIF-pocket.[17]

It was reported that the mutation of Leu155 to Glu canceled the ability of PDK-1 to interact with a peptide (PIFtide) substrates such as PRK2, S6K1 and SGK1.[17] Whereas mutation of Ile119, Lys115, Leu155, and Glu150 to Ala decreased the affinity of PDK-1 to PIFtide binding substrates but did not abolish the ability of PDK-1 for phosphorylation and activation of PIFtide substrates such as S6K1 and SGK1.[17]

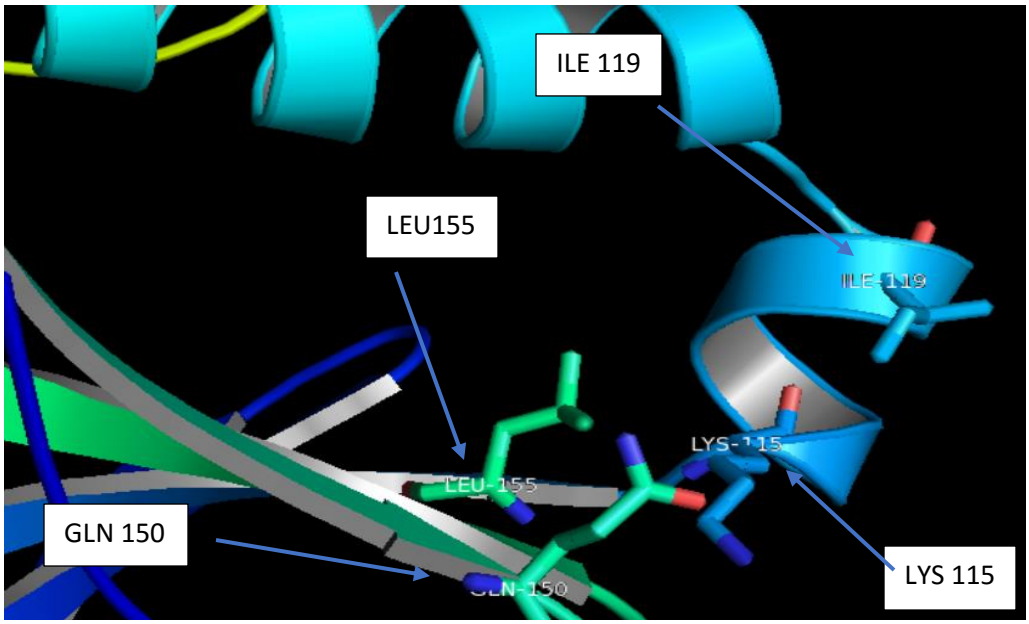


Figure 1.2: PIF-binding pocket of PDK-1 kinase

Phosphate-docking site is another small pocket lined with basic residues. This pocket is located in close vicinity to the PIF pocket (Fig 1.2). In the crystal structure

shown in figure 1.1, this pocket was occupied by a sulfate-anion that interacts with four residues lining the phosphate pocket, namely Gln150, Arg131, Lys76, and Thr148.[16],[17]

The αC-helix (residues 129–131) is an important element in the core of PDK-1 structure formed from residues 124–136. It links both the N-terminal lobe and the C-terminal lobe with the active site as well as the phosphopeptide pocket with the phosphoserine in the T-loop. In particular, Val127 and Val124 are involved in formation of the hydrophobic pocket (PIF-pocket).

Arg129 and Arg131 form two hydrogen bonds with the phosphorylated Ser241 and sulfate in the phosphate pocket, respectively (Fig. 1.1). In addition, each of Glu130 and Lys111 forms a hydrogen bond with the phosphate of bound ATP that are crucial for kinase activation. Finally, Tyr126 forms a hydrogen bond with the phosphorylated Ser241.[16]

1.5 Identification of residues in the ATP pocket

The ATP binding pocket as described in Figure 1.3 consists of multiple regions[18] as described below:

1) The Adenine region which is a conserved hydrophobic region. It is made up of residues at positions P2 (residue 88), P10 (residue 96), P13 (residue 109), P17 (residue142), P35 (residue 212). The adenine ring of ATP makes hydrophobic contacts with these five residues. In addition, it makes two hydrogen bonds with the

backbone of the hinge region residues (P20-P27). A third hydrogen bond occurs between two C-H groups of pyrimidiene ring with the carbonyl group of P23. The adenine-binding region is not characterized by large variability of amino acids, as a result of this it is not a good site for high degree of specificity.[18]

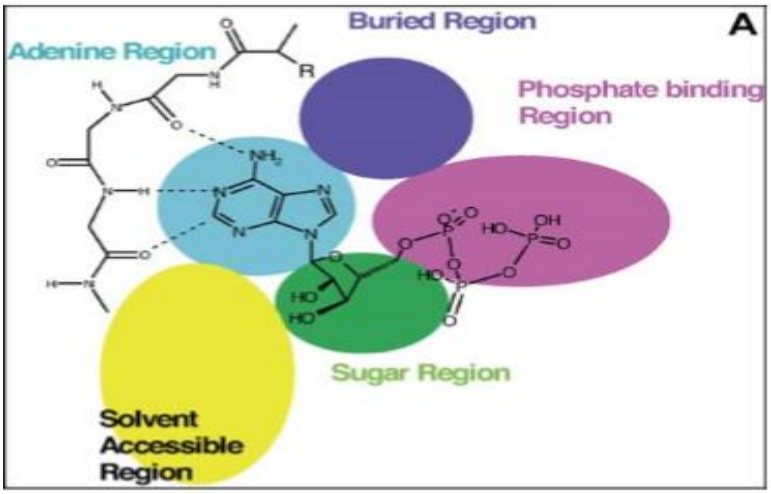


Figure 1.3: ATP binding pocket region: phosphate region (in magenta); sugar region (in green); Adenine region (in cyan); buried region (in violet) and solvent accessible region in (yellow)[18]

and P27 (residue 164).[18] In 80.7% of protein kinases P27 residue is a serine, a glutamate, an aspartate or glutamine. The variability in P27 allows for the development of selective and potent inhibitors as demonstrated in the EGFR family of kinases where a unique cysteine placed in P27 position.[18]

3) Phosphate region which contains many highly polar residues and it consists of two parts: (a) glycine-rich loop (GXGXXGXV: P3 -P10) lies on the N-terminal lobe. It is the only one that shows significant conformational flexibility.[18]

(b) alpha-helix which consists an essential and conserved residues: which is made up of residues at position P14 (residue 111), P15 (residue 129), P33 (residue 209), P37 (residue 223) and P38 (residue 225).[1] P15 (residue 129) with three-dimensional location that makes an essential indicator of the active state of any kinase. P37 (residue 223) and P38 (residue 225) are conserved in all protein kinases and these are essential for the transfer of phosphate group from ATP to the substrate. This part gives an indication whether the kinases are in their active or inactive conformations.[18]

4) Buried region: the largest sequence diversity in the ATP pocket residues are found in this region, this region is not occupied by ATP, which is made up of residues at position P16 (residue 133), P17(residue 142), P18 (residue 144), P19 (residue 156), P20 (residue 158), P36 (residue 222).[18] The residue in position P20 is important in determining the size of this specific region in the ATP binding pocket. P20 is often a bulky amino acid (40% methionine, 15% phenylalanine). It acts as a “molecular gate” to the buried ATP binding region. The introduction of a group to the buried region increases potency and, increases selectivity compared to that of kinases when this region is smaller.[18]

5) Solvent accessible regions: this region is important in exploited to increasing the binding affinity and to modulate ADME (toxicity) properties of ligands. The major difference in shape of solvent accessible area is contributed to the presence or

absence of glycine residue in positon P26. Often the NH of glycine forms intramolecular hydrogen bond with the carbonyl of P23 residue.[18]

1.6 Inhibition of PDK-1 Kinase enzyme

It was reported that overexpression of PDK-1 resulted in vitro and the PDK-1 phosphorylation was reported to suffer a high elevated levels in vivo breast cancers.[19] This explained that there are a strong relationship between PDK-1 and malignant phenotype.

The main strategy of developing kinase inhibitors is to reduce ATP binding and/ or inhibit kinase activity[12]. ATP and PDK-1 inhibitors compete in binding to the PDK-1 active site. When the PDK-1 inhibitors bind to PDK-1 active site they act to stop the transmission of phosporyl group from ATP to different amino acids. As a result, PDK-1 signal transduction is blocked. Development of PDK-1 inhibitors could lead to development of better treatment options for cancer.

specificity would not be a challenge if the target protein has unique catalytic functions and active site structures.[20] All 500 protein kinases encoded in the human genome have similar ATP-binding site structure.[21] In the last decade, more than 50 patents of PDK-1 inhibitors were published in which the ATP-pocket within the kinase domain was the target.[22]

Therefore, it is easy to establish the reason of kinase inhibitors being very specific and why the off-target effects are inevitable.[20] Nonetheless, Off-target effects are sometimes advantageous in clinical drugs. For example, Gleevec (Imatinib) was developed for treatment of chronic leukemia as an oral inhibitor of BCR-Abl. Currently, it has been approved for treatment of gastrointestinal stromal tumor (GIST).[20] This work demonstrates that the context of cells determine specificities of chemical inhibitors in vivo conditions.

Figure 1.4 shows that the specificities of inhibitors depend on cellular context. For example, in vitro standardized conditions were employed for enzyme-substrate and ATP concentrations, in addition to 'standard temperature and pressure' in physical chemistry. These conditions do not reflect the situation in living cells.[20] This reflect the facts that even we discovered the excellent drug in vitro condition, it may be not become an excellent one in vivo condition.

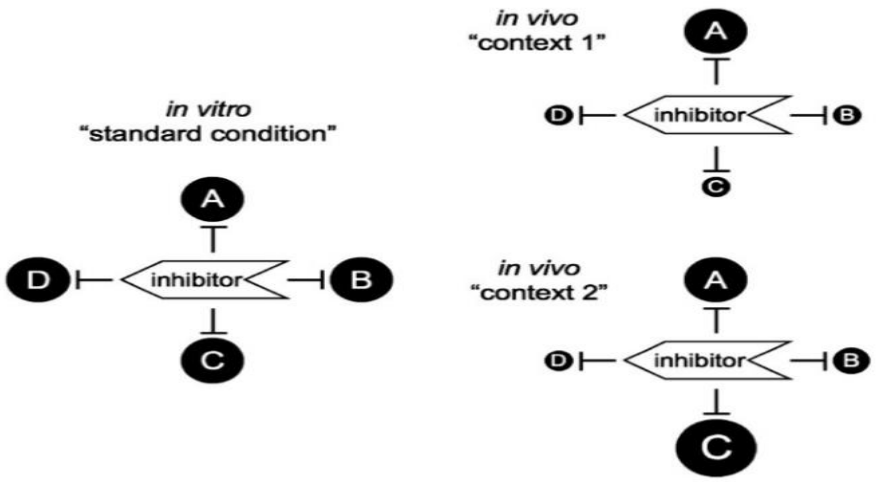


Figure 1.4: Context of target cells determined target specificities of inhibitors[20]

1.7 Known potent drugs for cancer diseases

BX-320, BX-795 and BX-912 (Figure 1.5) are considered as potent and selective competitive inhibitors of PDK-1 enzyme activity with respect to its substrate (ATP). BX-320 which inhibit the PDK-1 signaling pathway in different cancer cell lines including MDA-453 (breast), U87-MG (glioblastoma), PC-3 (prostate), HCT-116 (colon), MiaPaCa (pancreatic) and LOX (melanoma)cells. BX-795 and BX-912 potently inhibited the growth of PC-3, U87-MG and MDA-453 cancer cell lines only.[23]

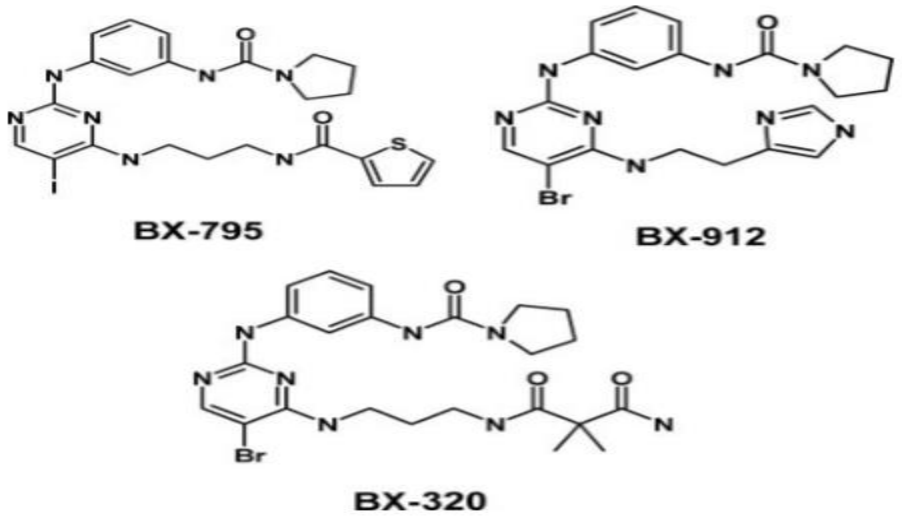


Figure 1.5: Example on potent inhibitors of PDK-1: BX-795, BX-912 and BX-320, respectively[23]

The high potency and selectivity of BX-320 is due to the formation of two hydrogen bonds between two nitrogens of amino-pyrimidine group with Ala162, which lies in the hinge region of the PDK-1 (Figure 1.6) .[23]

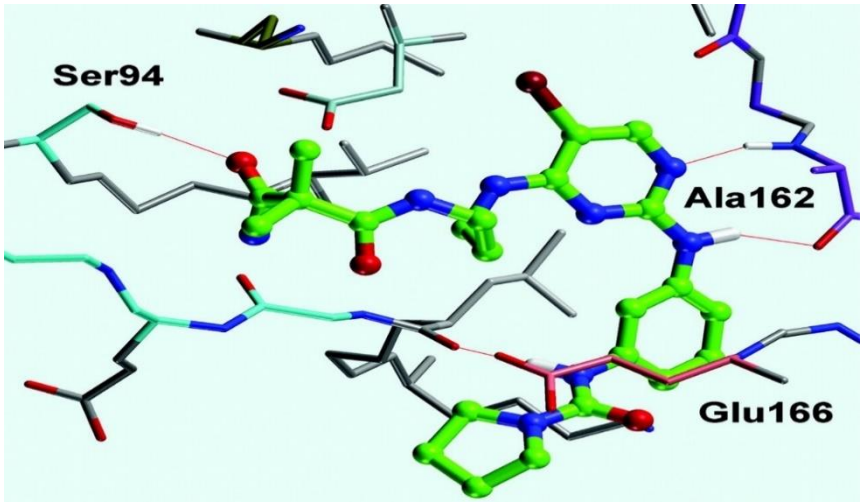
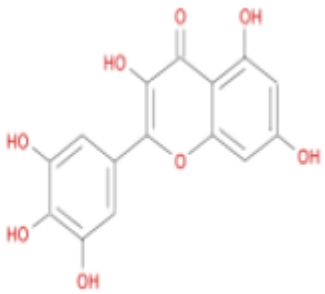


Figure 1.6: Structure of BX-320 bound to the ATP binding pocket of PDK-1[23]

Singh et al[24] reported using molecular docking that myricetin (3,5,7-Trihydroxy-2-(3,4,5-trihydroxyphenyl)-4-chromenone) acts as a probable anti-cancer agent. Myricetin is naturally occurring flavanol and it is a polyphenolic compound.[24]

Myricetin is considered as a potent PDK-1 inhibitor because of: (1) negative Docking energy of Myricetin-PDK-1 complex (-41 Kcal/mol), which indicates a favorable binding of Myricetin at the binding site of the PDK-1 kinase. (2) Formation of the most essential type of interaction between PDK-1 receptor and myricetin molecule (hydrogen bonding). The residues involved in formation of hydrogen bonds were Thr 222, Ala 162, Lys111, Asp 223, Ser 160, and Glu 130. Ala 162 and Ser 160 among these amino acid residues lie in the Hinge of PDK-1 protein. This type of interaction confirms that Myricetin fits into the active pocket of PDK-1receptor tightly (Table 1.1).[2]

Table 1.1: Molecular docking analysis of Myricetin[24]

Myricetin	Docking energy (Kcal/mol)	Hydrogen bonding residues	Hydrogen bond distances (Å)
	-41	O5: H-Lys[111]	2.44
		O5: H-Lys[111]	2.19
		O8: H-Ala[162]	2.34
		H28: O-Glu[130]	2.31
		H30: O-Asp[223]	2.09
		H23: O-Ser[162]	2.14
		H33: O-Ser[162]	1.98

Ong et all[25] reported that Myricetin possesses both antioxidant properties and prooxidant properties, it also has a therapeutic potential in cancer treatment, cardiovascular diseases and diabetes mellitus. Benzo(a)- pyrenes cause cancer of the skin and lungs. Myricetin reduces the risk of skin cancer caused by polycyclic aromatic hydrocarbons.[25] Polycyclic aromatic hydrocarbons when metabolized produce carcinogenic metabolites.[25] Myricetin was found to inhibit the hydroxylation of benzo(a)pyr- ene in the human liver microsomes.[25]

Virtual screening, NMR-based fragment screening, and ultrahigh throughput screening (UHTS) led to the identification of diverse chemicals as PDK-1 inhibitors which bind the PDK-1 kinase in the ATP-site with at least one H-bond towards the

hinge region. The first four inhibitors that are used in this investigation were identified using a combined screening method (HTS and virtual screening).[22]

The inhibitors studied in this work are 6-methoxy-2-(1H-pyrazol-5-yl)-1H-benzimidazole (inhibitor 1), 4-dicarboxylicacid diamide (inhibitor 2), 4-butyl-6-(1H-pyrrolo[2,3-b]pyridin-3-yl)pyrimidin- 2-amine (inhibitor 3), 4-ethyl-6-[5-(1H-pyrazol-4-yl)-1H-pyrrolo[2,3- b]pyridin-3-yl]pyrimidin-2-amine (inhibitor 4)[22] and 1-(3,4-difluorobenzyl)-2-oxo-N-{(1R)-2-[(2- oxo-2,3-dihydro-1H-benzimidazol-5-yl)oxy]- 1-phenylethyl}-1,2-dihydropyridine-3-carboxamide (inhibitor 5) which are shown in figure 1.7.

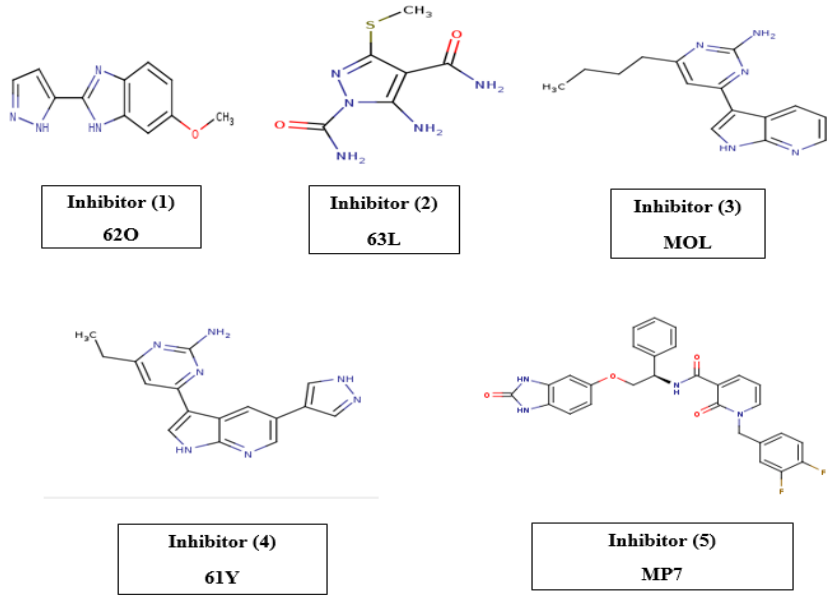


Figure 1.7: The five PDK-1 inhibitors that used in this investigation[22]

Inhibitor (5) is an example of pyridinonyl-based PDK-1 inhibitors described by Sunesis and Biogen Idec.[26] The concept of these inhibitors is mainly dependent on

the presence of a flexible linker to a hinge binding moiety (HBM) bearing neighboring H bond donor (HD) and H bond acceptor (HA) groups.[26]

1.8 Classification of inhibitors

Protein kinase inhibitors are classified according to their binding modes as follows:

Type I inhibitors: which are classical ATP-competitive and bind the ATP-binding site when the protein kinase is an activated state. They bind the hinge region with at least one hydrogen bond. FDA recently approved Type I anticancer kinase inhibitors: gefitinib, dasatinib, sunitinib, lapatinib, ruxolitinib, pazopanib, vemurafenib, crizotinib, erlotinib, and bosutinib.[27]

Type II inhibitors: are also ATP-competitive with binding to the extended ATP-binding site of protein kinase in an inactive state. If a significant change in the protein conformation occurs, it means that the inhibitor belongs to type II. Conformational changes in the protein kinase structure open a new hydrophobic pocket in the back of the protein that is called the *Deep Pocket* (also called the Phe pocket or allosteric pocket). These inhibitors, usually are hydrogen bonded to the hinge region but this is not a requirement for their action.[28]

Type III inhibitors (Allosteric binders): are ligands that target allosteric binding sites of protein kinase, therefore they are non-ATP-competitive. As

allosteric binding sites are highly specific for a protein kinase, this means a high degree of selectivity can be achieved. A specific feature that characterized this type of inhibitors no hydrogen bond in the hinge region. At present no drug on the market belongs to this type.[28] Traxler has developed a pharmacophore model for ATP-competitive inhibitors (type I) that identifies five different regions within the ATP-binding site.[28]

1.9 Drug design and drug properties

The drug-likeness of oral small molecules were evaluated by several guidelines. Up to 2015, a total of 28 small molecule kinase inhibitors SMKIs are FDA approved for the treatment of human cancer. Lipinski's Rule of Five (ROF) has been used as a rule of thumb to evaluate their absorption, permeability and solubility of drugs. The Veber Rules (Number of rotatable bonds (NRB)≤10 and polar surface area (PSA)≤140 $Å^2$) to evaluate oral bioavailability. Analysis of the number of rings (NOR) as included in the MDDR Rule (NOR ≥3).[29]

The physicochemical properties of more than 2,000 drugs and candidate drugs in clinical trials were analyzed by Lipinski: A compound has drug-like properties if it matches the following criteria (The Lipinski rule of five).[30]

●Its molecular weight (M.W) < 500.

●The lipophilicity property of compound (logP) which is the logarithm of the partition coefficient between 1-octanol and water[31] ≤5.

●The number of atoms in the molecule that donate hydrogen atoms to form hydrogen bonds (OH & NH) ≤ 5

●The number of atoms that can accept hydrogen atoms to form hydrogen bonds

(O & N) ≤10

Poor absorption or permeability is possible when the compound properties not obey the rule of five.[32]

Analysis of 28 FDA approved SMKIs revealed that 28 SMKIs were fitted well with ROF (HBD ≤5 and HBA≤ 10) and the molecular weight of 28 SMKIs is in the range 400 to 600. It is worth noting that with the exception of exitinib all inhibitors have at least six HBAs this reflect the fact that nitrogen and oxygen atoms are beneficial for kinase inhibitors. The Veber Rule[33] are abided by all inhibitors except dabrafenib. The analysis of number of rings as included in MDDR Rule[34] showed good adherence, NOR was no more than five for all SMKIs.[29]

1.10 Computational approach for binding free energy calculation using MM-GBSA (or MM-PBSA)

Several computational methods are available for calculating the binding free energy of protein-protein interactions, protein-DNA interactions and ligand-protein

interactions.[35] Some methods are more accurate but computationally intensive such as the thermodynamic integration (TI) and the free energy perturbation (FEP) methods.[36] On the other hand, less accurate methods such as molecular mechanics/Poisson Boltzmann surface area (MM/PBSA) and molecular mechanics/Generalized Born surface area (MM/GBSA) are less time-consuming methods.[37] MM/PBSA and MM/GBSA methods are end point methods because they calculate the binding affinity through simulations of only two end states (unbound and bound states of a ligand and its protein target).[37]

1.10.1 Binding free energy of ligand-protein complex using MM-GBSA

In the MM-GBSA formulation, the binding free energy of a ligand to a protein is calculated as the difference between the free energy of protein-ligand complex and the sum of the free energies of protein and ligand separately as follows.[38]

$$\Delta G^0_{\text{binding, solvated}} = G^0_{\text{complex, solvated}} - [G^0_{\text{receptor, solvated}} + G^0_{\text{ligand, solvated}}] \quad (1.1)$$

From the thermodynamic cycle shown in Figure 1.8. The binding free energy calculated as illustrated in equation (1.2).

$$\Delta G^0_{\text{binding, solvated}} = \Delta G^0_{\text{binding, vacuum}} + \Delta G^0_{\text{solv, complex}} - (\Delta G^0_{\text{solv, receptor}} + \Delta G^0_{\text{solv, ligand}}) \quad (1.2)$$

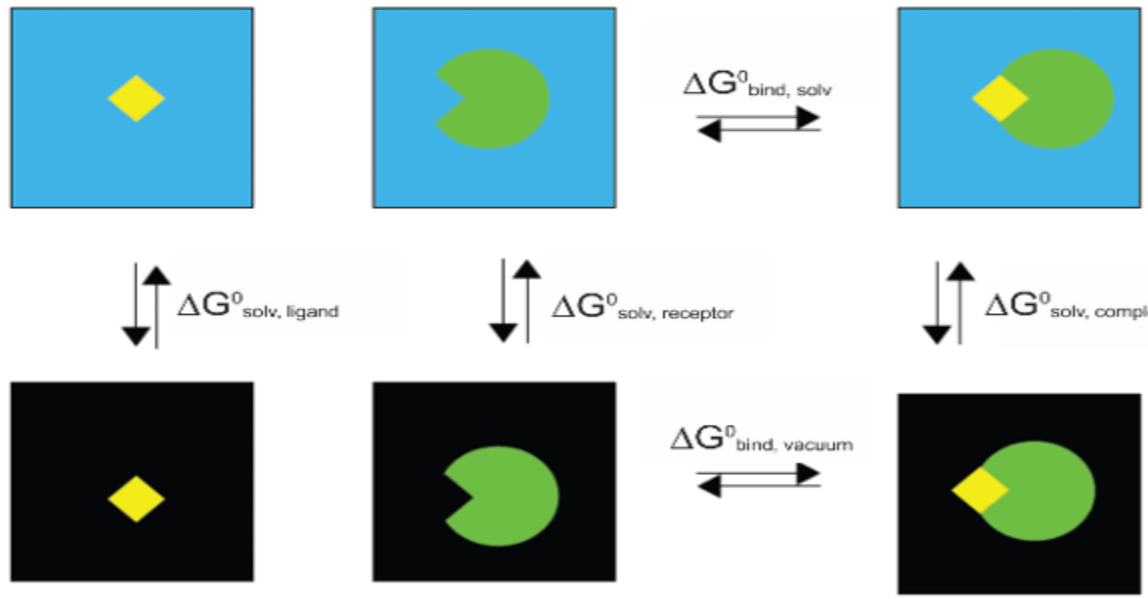


Figure 1.8: Thermodynamic cycles for binding free energy calculations for complex solvated systems (in blue boxes), whereas systems in the gas phase (in white boxes)

In the calculation of the solvation free energy term polar and nonpolar contributions are considered. For the polar contribution, the change in the free energy resulting from transfer of a charged molecule from gas-phase (modeled as a homogeneous medium with dielectric constant=1) to solvent (modeled as a homogeneous medium with=80), equation (1.4) γ and β values are constants dependent on the applied method.[39,40]

$$\Delta G^0_{sol} = \Delta G^0_{polar} + \Delta G^0 \quad (1.3)$$

$$\Delta G^0_{sol(nonpolar)} = \gamma(SASA) + \beta \quad (1.4)$$

$$\Delta G^0_{sol(polar)} = G^0_{electrostatic,\varepsilon=80} - G^0_{electrostatic,\varepsilon=1} \quad (1.5)$$

The approximation formula of the electrostatic contribution appears in equation (1.6), but extended Generalized Born model consists of a set of radii (ai) and charges contributions from Equation 1.7 for each particle.[41]

$$\Delta G_{elec} = -\frac{q^2}{2a}\left(1 - \frac{1}{\varepsilon}\right) \tag{1.6}$$

$$\Delta G_{elec} = -\left(1 - \frac{1}{\epsilon}\right)\sum_{i=1}\sum_{i=i+1}^{N}\frac{q_i q_j}{r_{ij}} - \frac{1}{2}\left(1 - \frac{1}{\varepsilon}\right)\sum_{i=1}^{N}\frac{q_i^2}{a_i} \tag{1.7}$$

$$\Delta G^0{}_{\text{bind, vacuum}} = \Delta E^0{}_{\text{MM}} - T.\Delta S^0{}_{\text{Nmode}} \tag{1.8}$$

$$E_{MM} = E_{bond} + E_{Angle} + E_{Torsion} + E_{Van\ der\ Waals} + E_{electrostatioc} \tag{1.9}$$

Protein -inhibitor average binding energy estimated by in the gas phase by molecular mechanics. Two types of energy are involved: first are non-covalent energies consisting of van der Waals energy and electrostatic energy. The second type are the covalent energies represented by bonds, angles and dihedral energies.[39]

Another way to calculate binding affinity is by molecular mechanics-Poisson Boltzmann surface area(MM-PBSA). Both MM-GBSA and MM-PBSA use the same previous equations to calculate the binding free energy, but the difference in the calculation of the electrostatic solvation energy G_{sol} (polar contribution).[42]

$$\Delta G_{sol} = \Delta G_{PB/GB} + \Delta G_{SA} \tag{1.10}$$

Where ΔG_{sol} is the sum of electrostatic solvation energy (polar contribution), $\Delta G_{PB/GB}$, and the nonelectrostatic solvation component (nonpolar contribution), ΔG_{SA}.[42] The electrostatic energy (ΔG_{PB}) is calculated by solving the Poisson-Boltzmann numerically. By combining Poisson's equation (1.11) for the electrostatic potential with Boltzmann's equation (1.12) that gives the charge distribution, you end up with the Poisson-Boltzmann equation (1.13).[41]

$$\nabla^2 \phi(\boldsymbol{r}) = -\frac{4\pi\rho(\boldsymbol{r})}{\epsilon} \tag{1.11}$$

$$n(r) = Ne^{-\frac{V(r)}{k_b T}} \tag{1.12}$$

$$\Delta G\ solv = \frac{1}{2}\sum_i qi\ (\emptyset i^{\varepsilon=80}\ -\ \emptyset i^{\varepsilon=1}\ ) \tag{1.13}$$

Chapter 2

Computational Methods

2.1 Protein- Inhibitor structures

The crystal structure of PDK-1 complex with five inhibitors were taken from the Brookhaven Protein Data Bank. The PDB codes of the PDK-1 with inhibitors 6-methoxy-2-(1H-pyrazol-5-yl)-1H-benzimidazole (inhibitor 1), 4-dicarboxylicacid diamide (inhibitor 2), 4-butyl-6-(1H-pyrrolo[2,3-b]pyridin-3-yl)pyrimidin- 2-amine (inhibitor3),4-ethyl-6-[5-(1H-pyrazol-4-yl)-1H-pyrrolo[2,3-b]pyridin-3-yl]pyrimidin-2-amine (inhibitor 4), and 1-(3,4-difluorobenzyl)-2-oxo-N-{(1R)-2-[(2-oxo-2,3-dihydro-1H-benzimidazol-5-yl)oxy]-1-phenylethyl}-1,2-dihydropyridine-3-carboxamide (inhibitor 5) are 5HNG, 5HO7, 5HO8, 5HKM, and 3NAX respectively. Water molecules and two sulfate groups were removed from the PDB files.

It was reported that part of N-terminal lobe of PDK-1 (residues 1-50) interact with Ralguanine nucleotide exchange factors. This region was not present in the PDB file of PDK-1 structure, because this region assumed a unique conformation in PDK-1.[17] PDK-1 protein consists of 556 amino acids, the Phosphoserine residue (SEP) is in position 241 is linking L-peptide ($C_3 H_8 N O_6 P$).

It was reported that 3-Phosphoinositide-dependent protein kinase-1 (PDK-1) expressed in 293 cells was phosphorylated at Ser25, Ser241, Ser393, Ser396 and Ser410. Mutation of Ser241 to Ala canceled PDK-1 activity, whereas mutation of the other phosphorylation sites individually to Ala did not aff ect PDK-1 activity. Also it was reported that PDK-1 can phosphorylate itself at Ser241, leading to its own activation.[43] The pdb files of PDK-1 structure in inhibitors (1-4) show a phosphorylated T-loop in Ser241 therefore, it is in an active state.[17]

2.2 Equilibration of the solvated system

Four steps were used to equilibrate the system: minimization, heating, density equilibration and unrestrained equilibration.

(a) Relaxation of the solvated system.

We used ***sander*** to minimize our system in order to remove any bad contacts as a results of the hydrogenation steps in ***xLeap***. Minimization (imin=1) was done in two steps: The first step involves the relaxation of water molecules only, whereas protein and inhibitor atoms were fixed by using a harmonic restraint (restraint_wt=2.0). The second step involves the minimization of the whole system using Sander. The input file min.in was used to perform the first step and min_all.in was used to perform the second step of minimization (see Appendix B). Minimization was performed using 500 steps of the steepest descent method, then

switching to conjugate gradient algorithm for the remaining steps (maxcyc=1000, ncyc=500). Constant volume periodicity was applied (ntb=1).

(b) Heating the solvated system.

The system was then heated (imin=0) using langevin thermostat (ntt=3) to maintain the temperature of our system to 300K, with a collision frequency 2 ps^{-1}. This method is more efficient than Berendsen method (ntt=1) due to hot solvent, cold solute phenomena.[44] The file titled heat.in (see Appendix B) was used to perform the heating process.

(c) Density equilibration.

The file titled density.in (see Appendix B) was used to perform this step. The system was equilibrated at 300 K with constant pressure periodic boundary (ntp=1) using Particle mesh Ewald (PME) method and positional restrains of 2 Kcal/mol. $Å^2$ was applied.

(d) Unrestrained equilibration

The file titled equil.in (see Appendix B) was used to perform this step. The unrestrained system was equilibrated at 300 K with constant pressure periodic boundary (ntp=1). The ***SHAKE*** method[45] was applied (ntc=2, ntf=2) to hold all covalent bonds containing hydrogen atoms.

2.3 Production step of the solvated system

The production simulation time is 2ns that run using the same conditions as in the final phase of equilibration to prevent any sudden jump in the potential energy due to a change in simulation conditions. The production run was carried out over four sequential steps using the input file prod.in (see Appendix B). During all the MD simulations, the Particle Mesh Ewald (PME) method was utilized with 10Å cutoff for long-range interactions.

2.4 Calculating the binding free energy of the protein-inhibitor complex

We carried out the binding free energy calculation using both the MM-GBSA method and the MM-PBSA method for comparison. This is achieved using input file for mmpbsa.in (see Appendix B).

2.5 Calculating the entropic contribution

Normal Mode Analysis ***(Nmode)*** was used to calculate the entropic contribution.[46]

The file titled mmpbsa_nm.in (see Appendix B) was used to do this step.

Chapter 3

Results and discussion

3.1 Analysis of simulations

The system reached state of equilibrium after different stages of simulation. This was checked by monitoring of different properties during the simulation. The system properties were extracted from the output files, and were ploted versus time. Figures 3.1 to 3.4 show plots for inhibitor 1-protein complex.

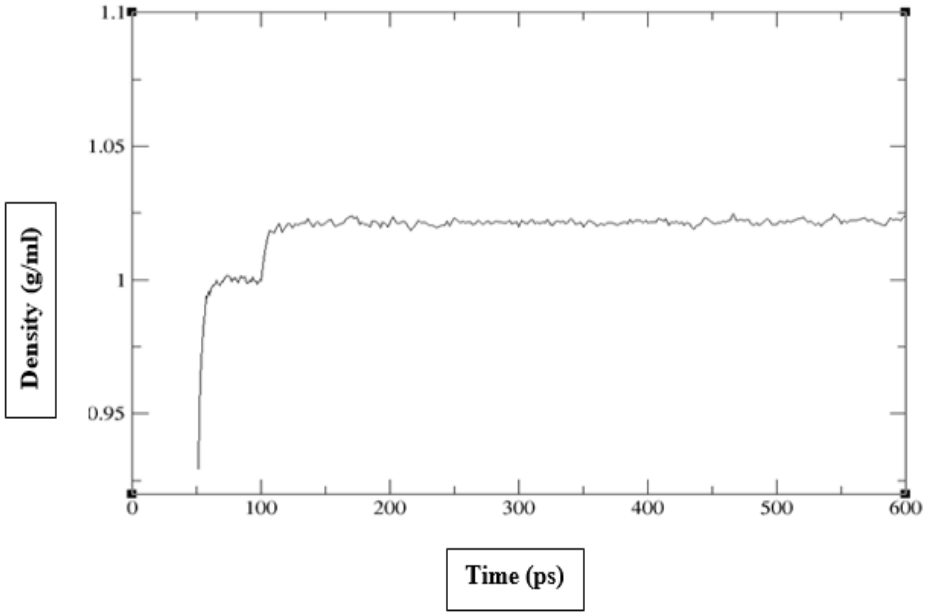


Figure 3.1: Density of protein-inhibitor (1) complex system during equilibration

As shown in Figure 3.1, the first 50 ps of the simulation represents the heating stage. There was no density data recorded due to the constant volume condition that took place until 50 ps. After that the density increased up to 1.02 g/ml and stayed around this number until the last 550 ps. This is reasonable because the density of pure water at 300 K is 1.00 g/ml, so adding inhibitor 1-protein complex lead to a rise in density of the system by 4%.[47]

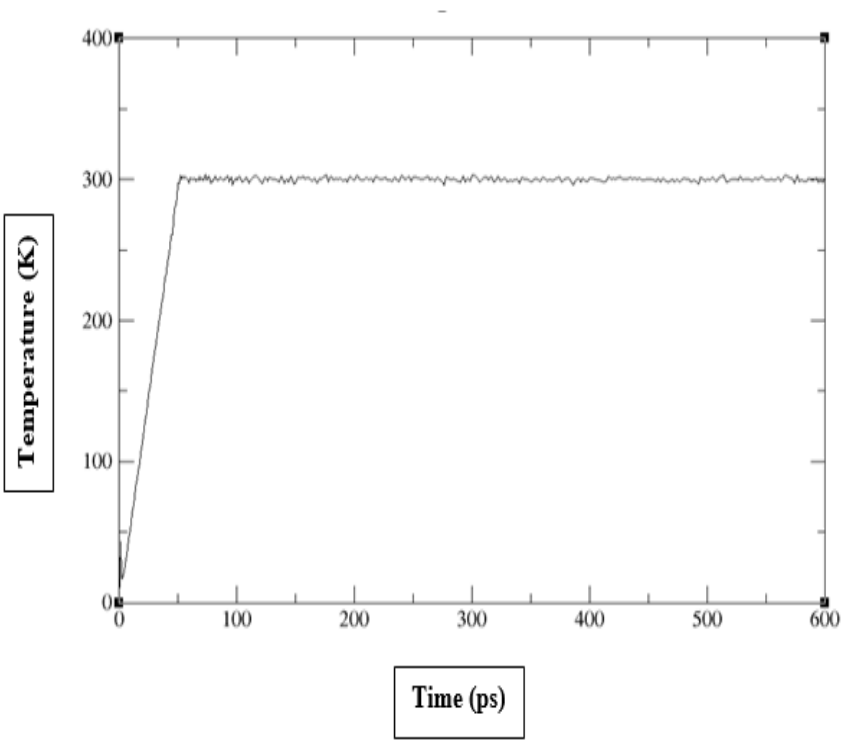


Figure 3.2: Temperature of protein-inhibitor (1) complex system during equilibration runs

In Figure 3.2, the temperature rises regularly from 0 K to 300 K. After that the temperature of the system reached an equilibrium value of 300 K over the last

stage of simulations, indicating that Langenvin dynamics applied successfully in this case.[47]

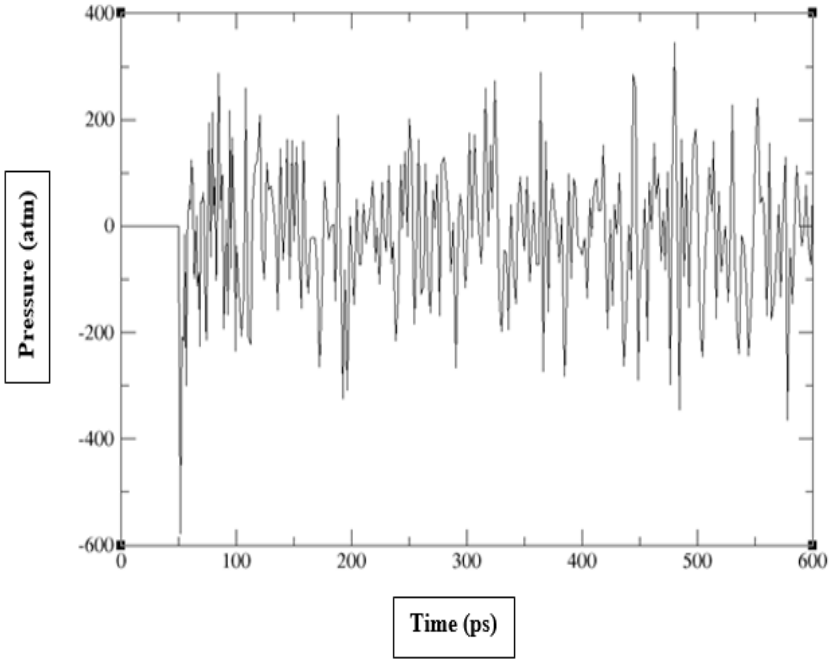


Figure 3.3: Pressure of protein-inhibitor (1) complex system during equilibration

In the pressure versus time plot (Figure 3.3). It shows that in the time interval between 0-50 ps the pressure was zero, because it was running at constant volume. At 50 ps the system changed to constant pressure, the volume of the box changed and the pressure dropped sharply becoming negative.[48] Positive values of pressure reflect a force trying to make the water box larger, whereas negative pressure values reflect a force trying to reduce the volume of the water box.[47] While the pressure plot shows that the pressure fluctuated during the simulation, pressure stabilized at1 atm, this indicates a successful equilibration.

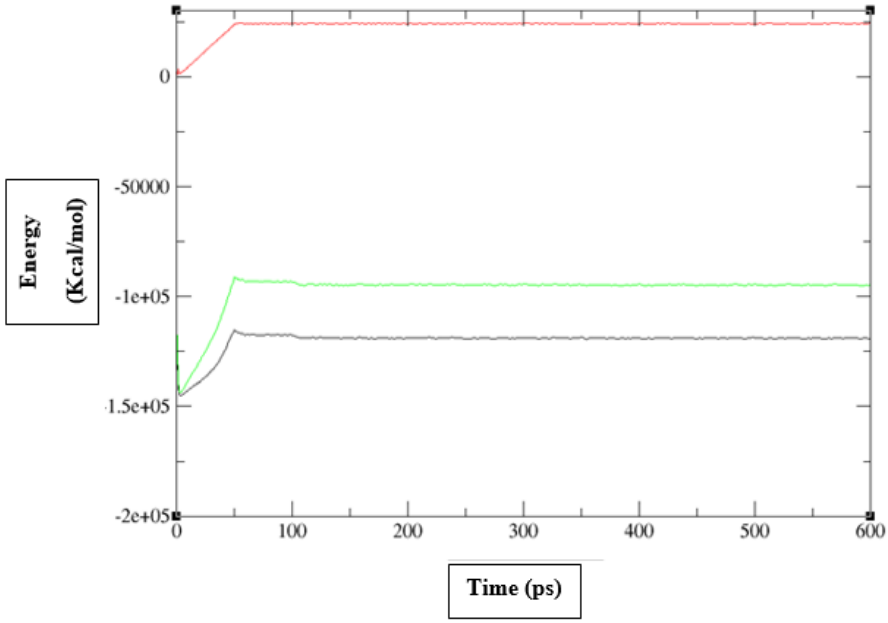


Figure 3.4: Total, kinetic and potential energy of protein-inhibitor (1) complex system during equilibration runs. The kientic energy (in red line); the potential energy (in black line); the total energy (in green line) which is the sum of kinteic and potential energy

According to the energy plot versus time (Figure 3.4) the first 50 ps of simulation there was An increased in all energies corresponding to heating from 0 K to 300 K. The kinetic energy remained constant in the last stages indicating a successful performance of temperature thermostat.[47]

The potential energy and the total energy initially increased, then during the constant volume stage (0 to 20 ps) there was a plateau, then at 20 to 40 ps there was a decrease in the energy values because, at this stage we switched off the protein-ligand restraints and moved to constant pressure. After that the potential energy

leveled off for the reminder of our simulation indicating stability and a relaxed system.[47]

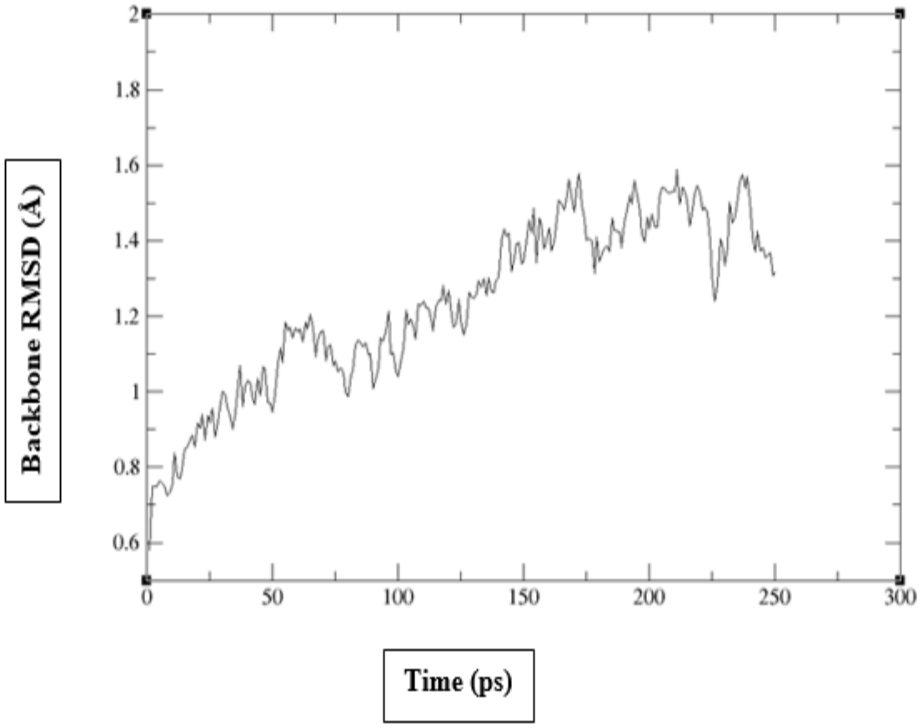


Figure 3.5: RMSD of protein backbone during unrestrained equilibration run of protein-inhibitor (1) complex

In order to quantify the similarity between a native inhibitor1-protein complex (com_wat.inpcrd) and a generated inhibitor1-protein complex (equil.mdcrd), the mass weighted RMSD (Root-mean square deviation) can be calculated between these two structures. Figure 3.5 shows that the root mean square deviation (RMSD) values increased rapidly in the first 75 ps, then it fluctuated around a value of 1.4 Å until the last 250 ps, which is an acceptable value. RMSD

values were around 1.4 Å, this reflects an acceptable conformational changes in the protein backbone .

3.2 Study of binding energies of kinase – inhibitors

3.2.1 Binding free energies of protein-inhibitor complexes: MM-PBSA versus MM-GBSA

MM-PBSA and MM-GBSA are direct methods for the quantitative prediction of binding free energy of ligand-protein complex.[49]

Both methods are used in this work to calculate the binding free energy of PDK-1 kinase with five inhibitors (Table 3.1). As shown in Figure 3.6 there is a good correlation between binding free energies were calculated by MM-GBSA, and experimental values of binding free energies which are derived from the experimental reported IC_{50} values (R^2 =0.54).

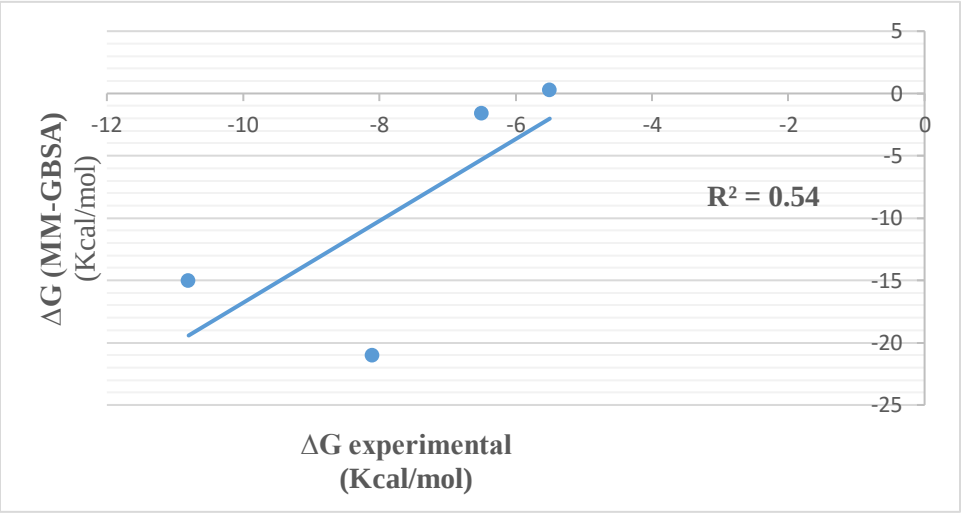


Figure 3.6: Correlation between ΔG calculated by MM-GBSA and ΔG experimental values

To the contrary, the correlation between the binding free energies were calculated by MM-PBSA and experimental values of binding affinity (Fig 3.7) which are derived from the experimental reported IC_{50} values is weaker (R^2 =0.06).

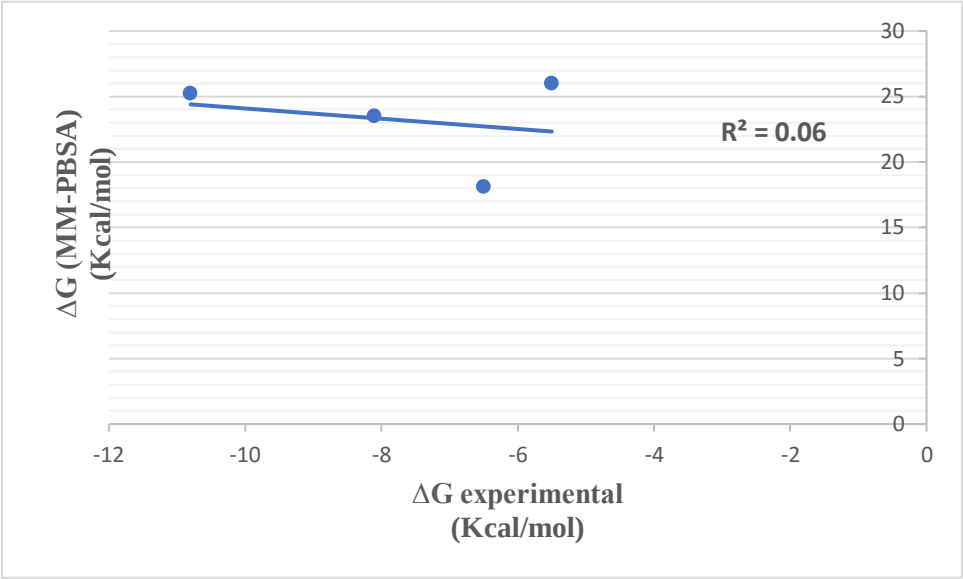
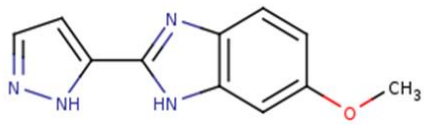
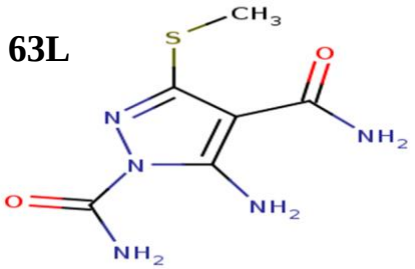
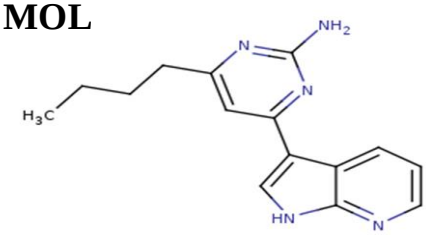
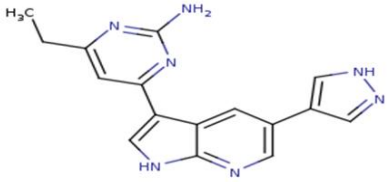
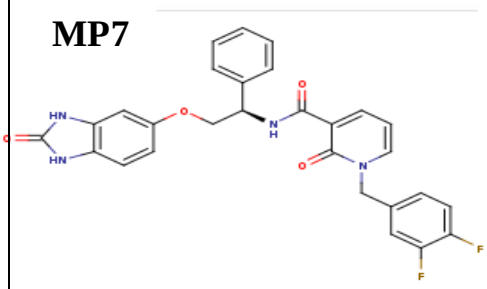


Figure 3.7: Correlation between ΔG calculated by MM-PBSA and ΔG experimental values

It is observed that the calculated binding free energies using MM-GBSA method were closer to the experimental values than those calculated using MM-PBSA method. In principle, PB is more theoretically rigorous than GB, but it does not mean that MM/PBSA can give better predictions than MM/ GBSA.[50] Our result agrees with some of the reports that MM-GBSA based on GB^{OBCI} is considered a better approach than the MM-PBSA in calculating the binding free energies when heterocyclic and aromatic system is present .[51]

Table 3.1: Binding free energies (kcal/mol) calculated at T= 300 K and P= 1 atm for PDK-1 binding with the four inhibitors

	Experimental		Calculated	
Inhibitor	**IC_{50}** (µM)	**ΔG_{exp}** [22] (Kcal/mole)	**ΔG_{calc}** **(MM-GBSA)** (Kcal/mole)	**ΔG_{calc}** **(MM-PBSA)** (Kcal/mole)
62O 	93	-5.5	0.3 ± 1.6	26.0 ± 1.7
63L 	17	-6.5	-1.6 ± 2.0	18.1 ± 2.0
MOL 	1.1	-8.1	-21.0 ± 1.5	23.5 ± 2.3
61Y 	0.013	-10.8	-15.0 ± 1.8	25.2 ± 2.6
MP7 	-	-	-52.3 ± 2.8	-11.4 ± 3.0

The experimental binding free energies (ΔG_{bind}) were calculated from the experimental values of IC_{50}, by using this equation:

$$\Delta G_{bind} = RT\ln K_D = RT\ln IC_{50} \quad (3.1)$$

The reported IC_{50} values are concentrations at which the PDK-1 kinase activity is inhibited by 50% of the initial concentration.[52] The kinetic study of enzyme-inhibitor reaction in the absence of inhibitor follows a simple Michaelis-Menten equation (3.2).[53] The following equation assumes that the concentration of enzyme is sufficiently low (neglected).

$$V0 = \frac{V\max\ S}{K\mathrm{m} + S} \quad (3.2)$$

$$VI = \frac{Vmax\ \ S}{Km\ \left(1 + \frac{I}{KI}\right) + S} \quad (3.3)$$

Where, V_{max} = maximum velocity; V_0 = velocity in the absence of the inhibitor; K_m =Michaelis constant of the substrate; V_I = velocity in the presence of inhibitor; I = concentration of inhibitor; S= substrate concentration; K_I= dissociation constant of enzyme-inhibitor complex (EI).

When $I = I_{50}$, $V_0 = 2V_I$ then[53]

$$\frac{2\,Vmax\;\;S}{Km\;\left(1+\frac{I50}{KI}\right)+S}=\frac{Vmax\;\;\;S}{Km+S} \qquad (3.4)$$

By rearranging equation 3.4:

$$I50\;=KI\;\;(1+\frac{S}{Km}) \qquad (3.5)$$

In the case of a competitive inhibitor, S << Km, then Ki ~ IC_{50}.

3. 3 Analysis of the binding mode of inhibitor (1)-PDK-1 complex

In this section we discussed the binding mode of inhibitor (1) with PDK-1 kinase complex (Figure 3.8(a) and (b)). The non-covalent interaction of inhibitor to the proteins is governed by different interactions including van der Waal and hydrogen bond interaction.[24]

Inhibitor (1) makes two strong hydrogen bond interactions with the backbone oxygen atom of Ser^{90} in the adenine region of the kinase with a distance of 2.1 Å and 1.8 Å (Table 3.2). The benzimidazole ring is in a buried region. It is surrounded by residue Thr^{152} (4.6 Å not considered as hydrogen bond).

Other weak interactions were formed between the inhibitor (1) and the PDK-1 (Figure 3.8(c)), C-H ….. π interaction[54] between carbon hydrogen atom of Leu^{18} and the center of benzimidazole ring of inhibitor (3.1 Å is the average distance from the hydrogen atom to the center of ring). The same type of interaction was made by Ala^{92} in the hinge region, but the average distance is 3.9 Å.

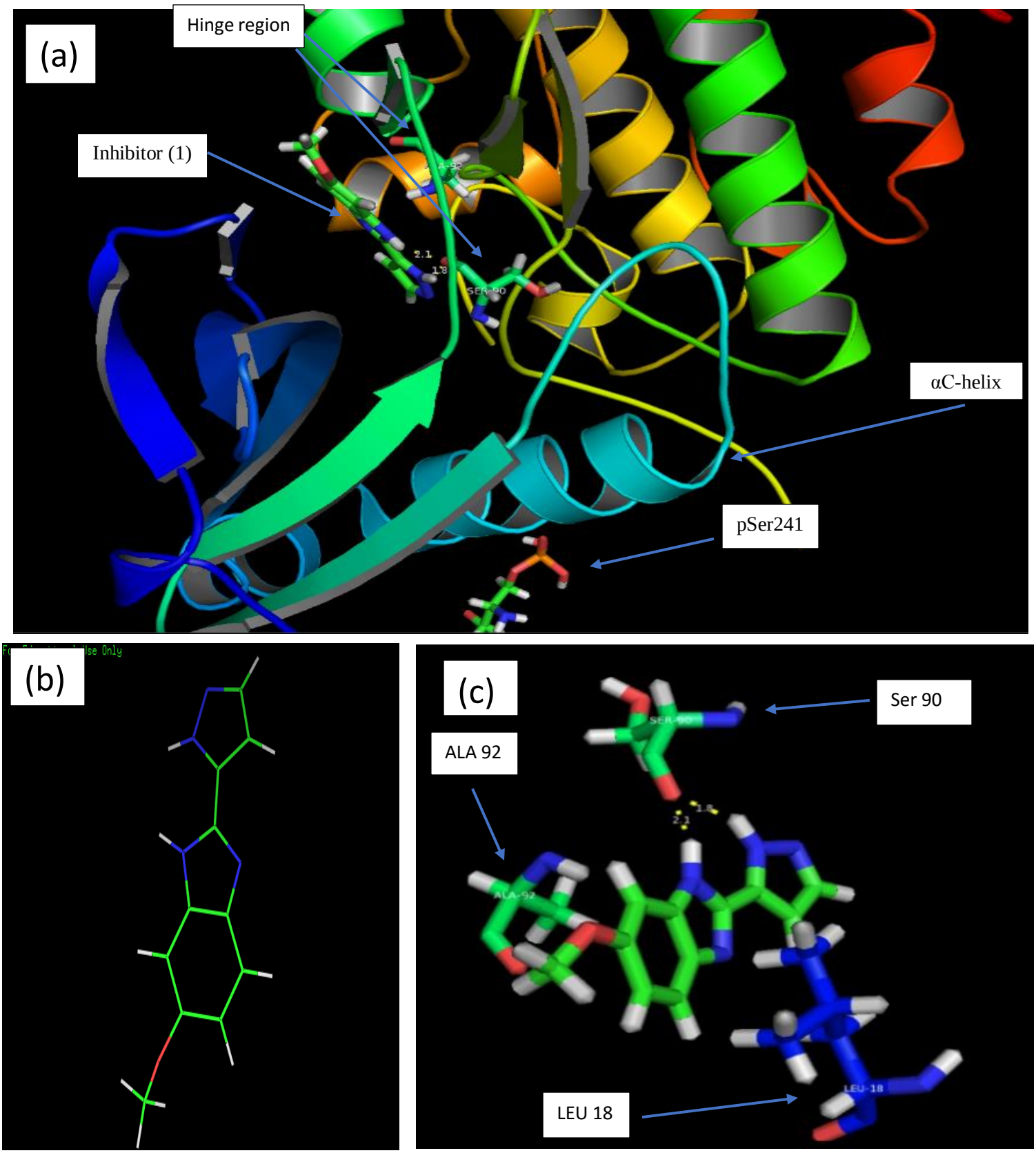
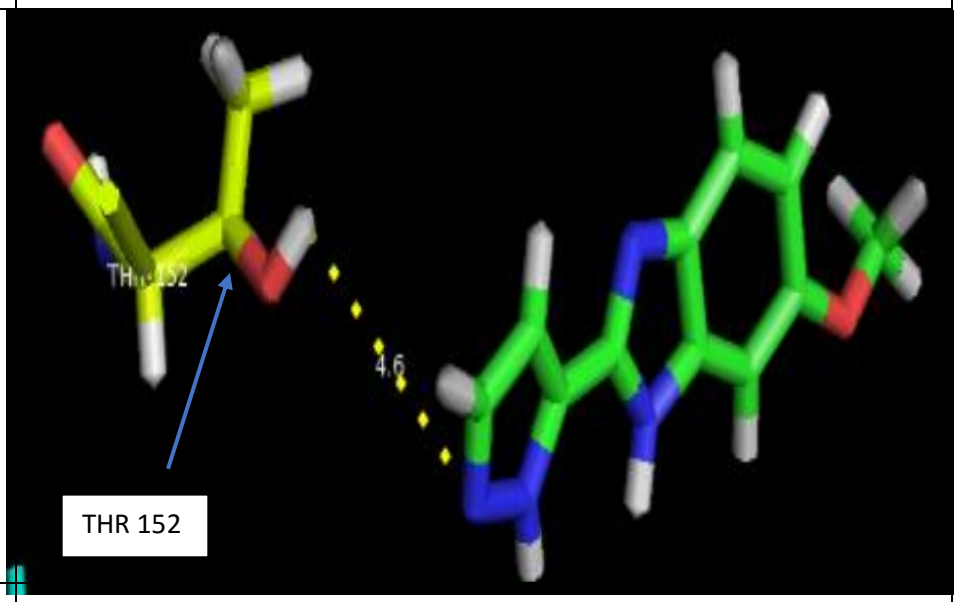
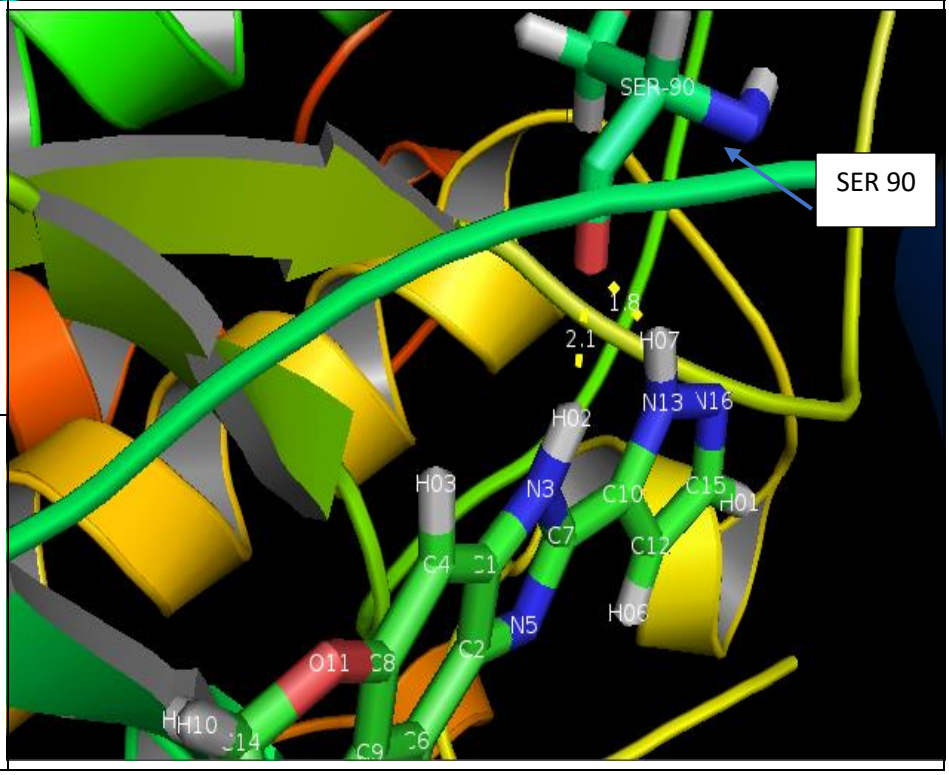


Figure 3.8: (a) Inhibitor (1)-PDK-1 complex, (b) graphical representation of inhibitor (1) and (c) other weak interactions between inhibitor (1) and kinase

Table 3.2: Hydrogen bond analysis of inhibitor (1) in 62O-PDK-1complex

Atom of inhibitor	Atom of protein/H_2O	Comment
N16	H-Thr[152] no considerable hydrogen bonds	
HO2	O-Ser[90]	
HO7	O-Ser[90]	

3.4 Analysis of the binding mode of inhibitor (2)-PDK-1 complex

In this section we discussed the binding mode of inhibitor (2) with PDK-1 kinase complex (Figure 3.9(a) and (b)). Inhibitor (2) has hydrogen bond interaction with the backbone carbonyl group of Ser^{90} with 2.1 Å distance, and another two hydrogen bonds are formed with Ala^{92} at 3.2 Å and 2.1 Å distances in the hydrophobic adenine pocket (Table 3.3).

A strong hydrogen bond interaction with the carbonyl group of Thr^{152} in the buried region (length = 1.8 Å). It is worth noting that this inhibitor has an intramolecular hydrogen bond between hydrogen atom of amino group (HO2) and oxygen of carbonyl group (O14) as shown in table 3.3.

The highest frequency of intramolecular hydrogen bonds for planer, six membered rings stabilized by conjugation with a π-system. The formation of an intramolecular hydrogen bond result in an increased lipophilicity and membrane permeability accompanied by reduced aqueous solubility. These are due to the removal of one donor and one acceptor function from the surface of a molecule.[55]

Replacing real rings by such pseudo rings to form pseudo six-membered ring is a new and non-conventional strategy and the new classes of kinase inhibitors follow this approach.[56]

We noted that inhibitor (2) interacts with the active site in water-mediated hydrogen bonds with active-site residues. A water-mediated network of hydrogen bonds is formed by 2 water molecules to inhibitor (2) as shown in table 3.3.

In addition to all of these interactions, other weak interactions were formed between inhibitor (2) and PDK-1 (Figure 3.11(c)), C-H…...C=O interaction[54] between hydrogen atom of Tyr^{91} and the carbonyl group of inhibitor (2.6 Å).

Another weak interaction was formed of the type C-H…….π interaction[54] between carbon hydrogen atom of Val^{26} with the center of pyrazole ring (3.7 Å is average distance between the center of the pyrazole ring and the hydrogen atom) as shown in figure 3.9(c).

It is worth noting that Leu^{18} is close to inhibitor due to the C-H…….C=O weak interaction between hydrogen atom of Leu^{18} and the carbonyl group of Val^{26} (2.1 Å) as shown in figure 3.9(c).

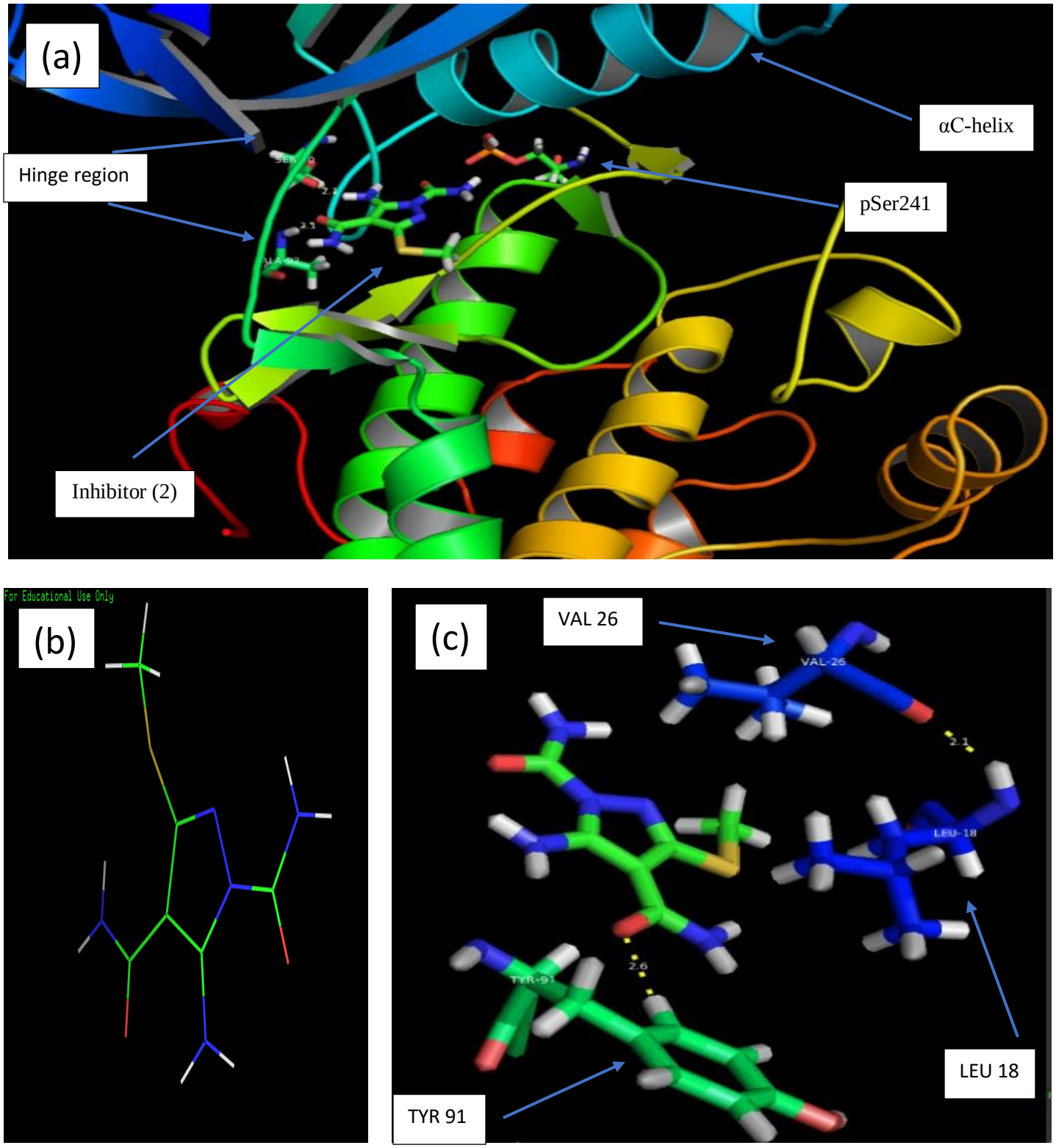


Figure 3.9: (a) Inhibitor (2)-PDK-1 complex, (b) graphical representation of inhibitor (2) and (c) other weak interactions between inhibitor (2) and kinase

Table 3.3: Hydrogen bond analysis of inhibitor (2) in 63L-PDK-1complex

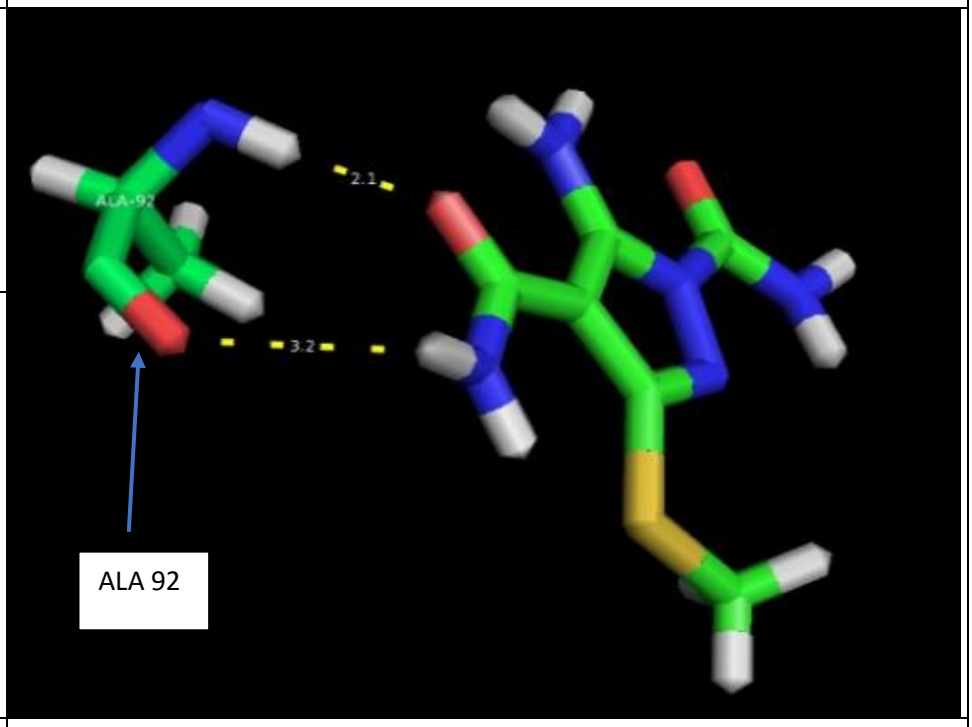
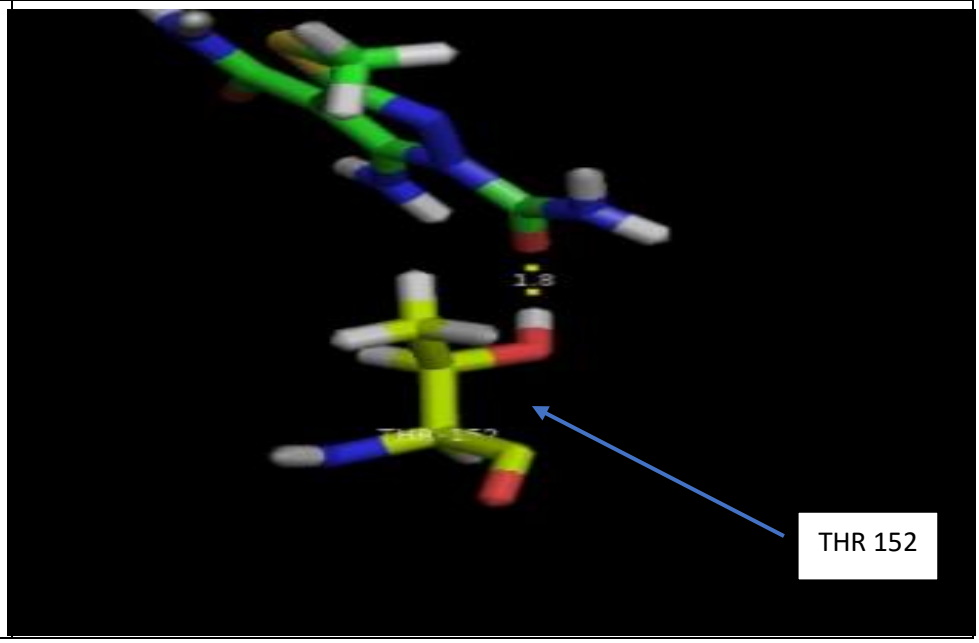
Atom of inhibitor	Atom of protein/H_2O	Comment
O10	H-Ala^{92}	
HO7	O-Ala^{92}	
O14	H-Thr^{152}	

Table 3.3: Hydrogen bond analysis of inhibitor (2) in 63L-PDK-1complex

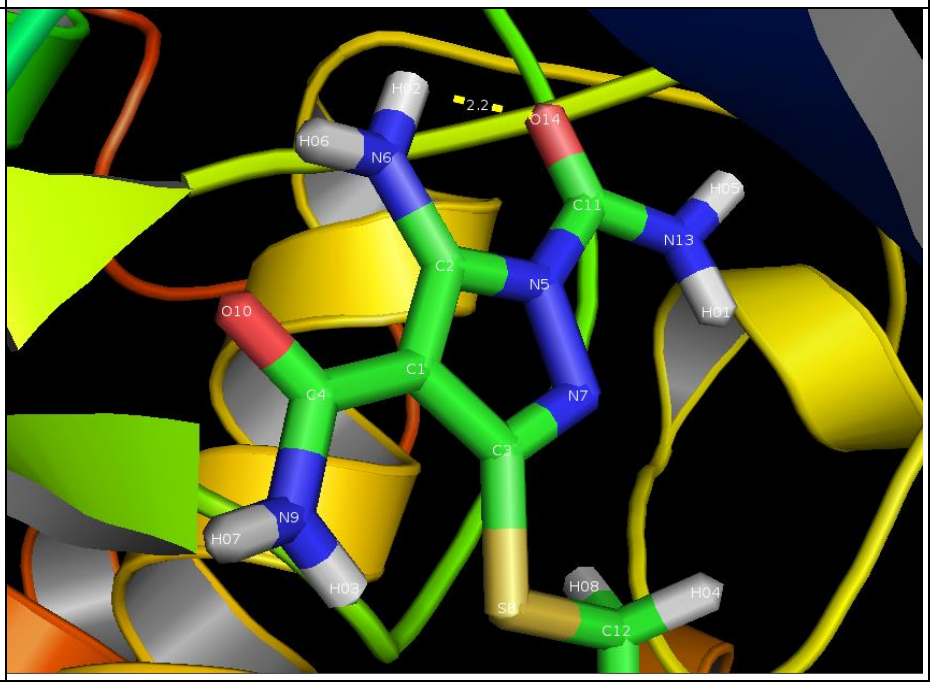
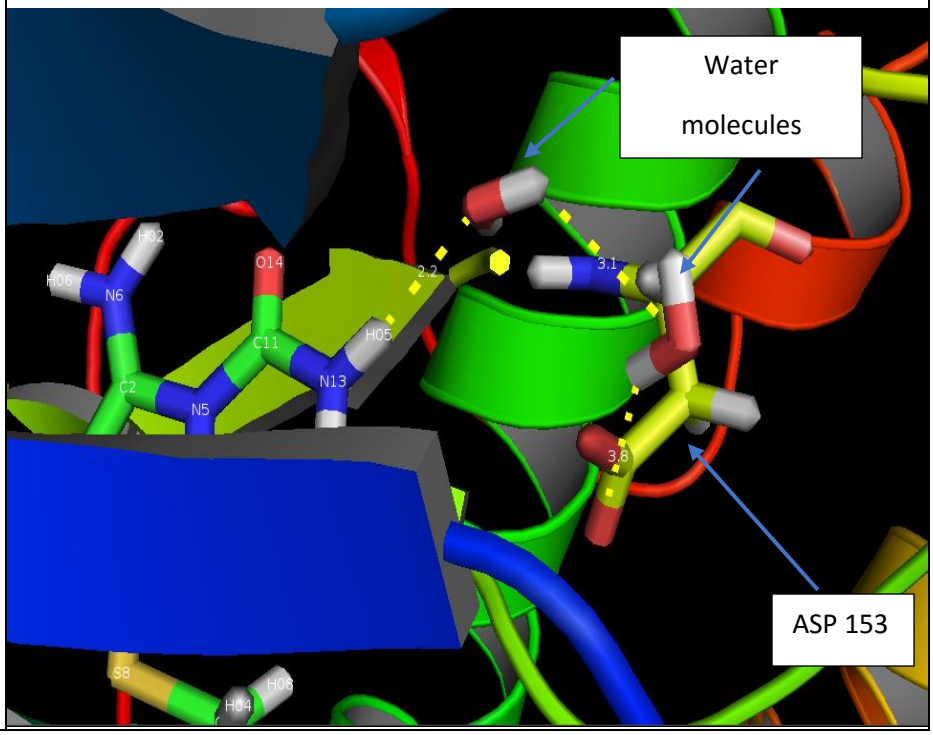
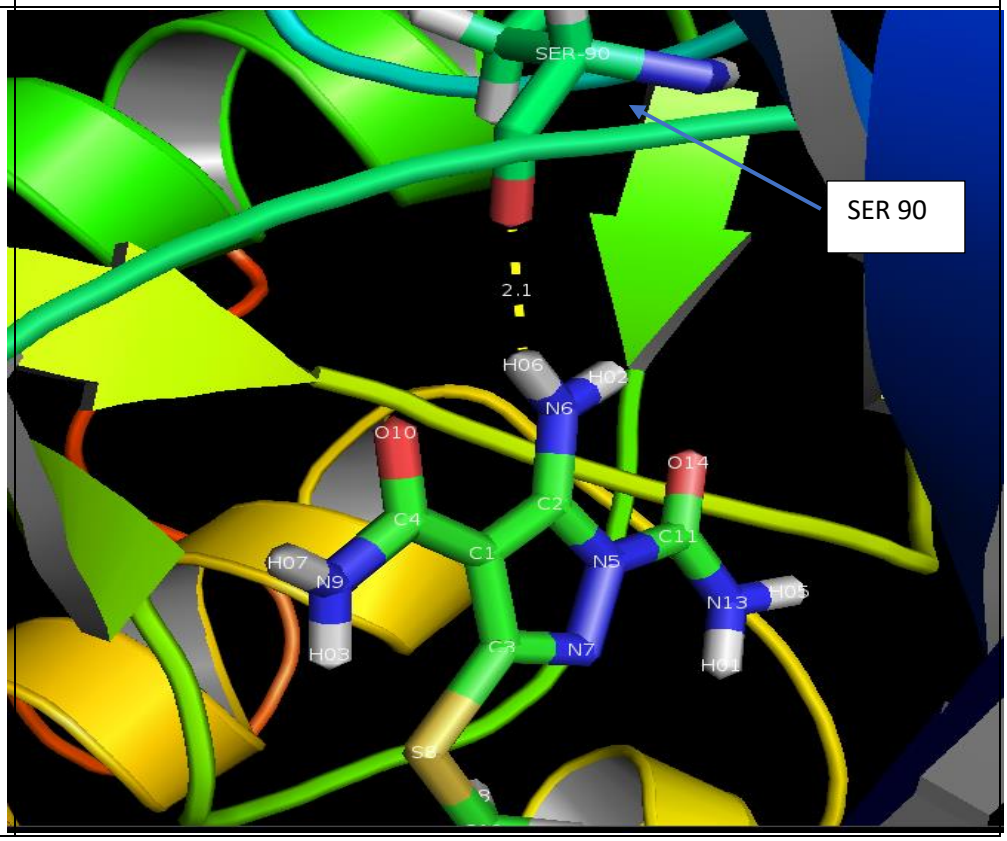
Atom of inhibitor	Atom of protein/H_2O	Comment
O14	H-63L[282] (Intramolecular hydrogen bond)	
HO5	O-Asp[153] (Water mediated)	

Table 3.3: Hydrogen bond analysis of inhibitor (2) in 63L-PDK-1complex

Atom of inhibitor	Atom of protein/H_2O	Comment
HO6	O-Ser^{90}	

3.5 Analysis of the binding mode of inhibitor (3)-PDK-1 complex

In this section we discussed the binding mode of inhibitor (3) with PDK-1 kinase complex (Figure 3.10(a) and (b)). Inhibitor (3) is located in the ATP-binding site (which lies between the N-terminal and C-terminal lobes of kinases).[57] The 7-azaindole ring mimics the interactions of the adenine base in ATP with the protein backbone, where two conserved hydrogen bonds are formed between the 7-azaindole nitrogen N3 in inhibitor (3) and the backbone-hydrogen of Ala^{92}, and the 7-azaindole hydrogen HO5 and the backbone-oxygen of Ser^{90} as shown in table 3.4.

In addition to the presence of direct hydrogen bonds, there are water-mediated hydrogen bond interactions. The water mediated hydrogen bond interaction occur between inhibitor (3) and Lys^{41} in the phosphate region, and Thr^{152} in the buried region as shown in table 3.4.

It was reported that the discovery of aminoindazole ring and the addition of one heterocyclic ring which is involved by using its nitrogen atoms in the hydrogen bond interaction with inhibitor. This is critical for binding. Overall cumulative data confirm that each nitrogen in the aminoindazole positively contributes to PDK-1 binding and inhibition activity.[58]

Other weak interactions were formed between inhibitor (3) and the PDK-1,

C-H….π interaction[54] between carbon hydrogen atom of Leu[18] and the center of 7-azaindole ring (3.3 Å is average distance between the hydrogen atom and the center of the ring) as shown in figure 3.10(c).

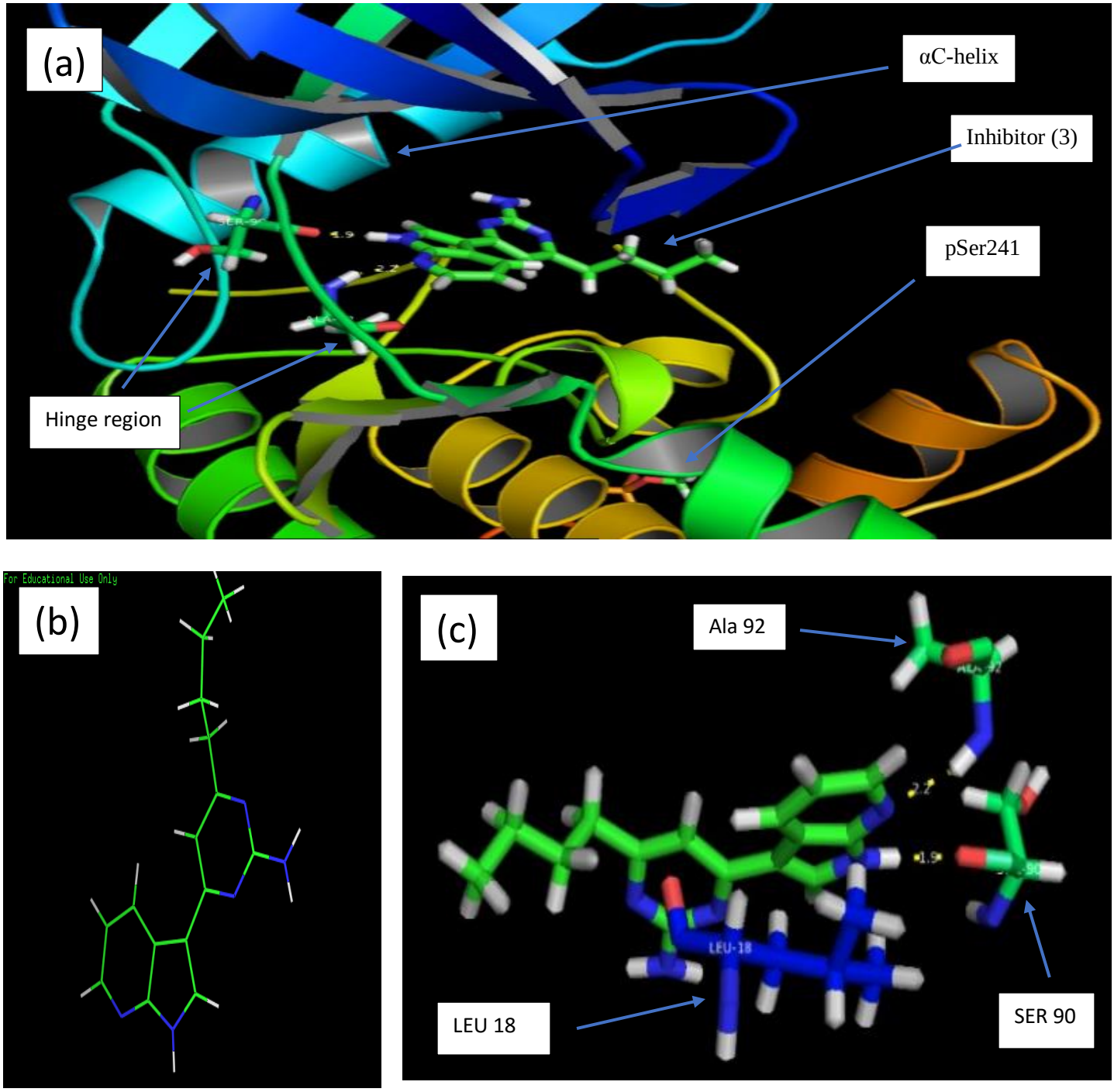


Figure 3.10: (a) Inhibitor (3)-PDK-1 complex, (b) graphical representation of inhibitor (3) and (c) other weak interactions between inhibitor (3) and kinase

Table 3.4: Hydrogen bond analysis of inhibitor (3) in MOL-PDK-1complex

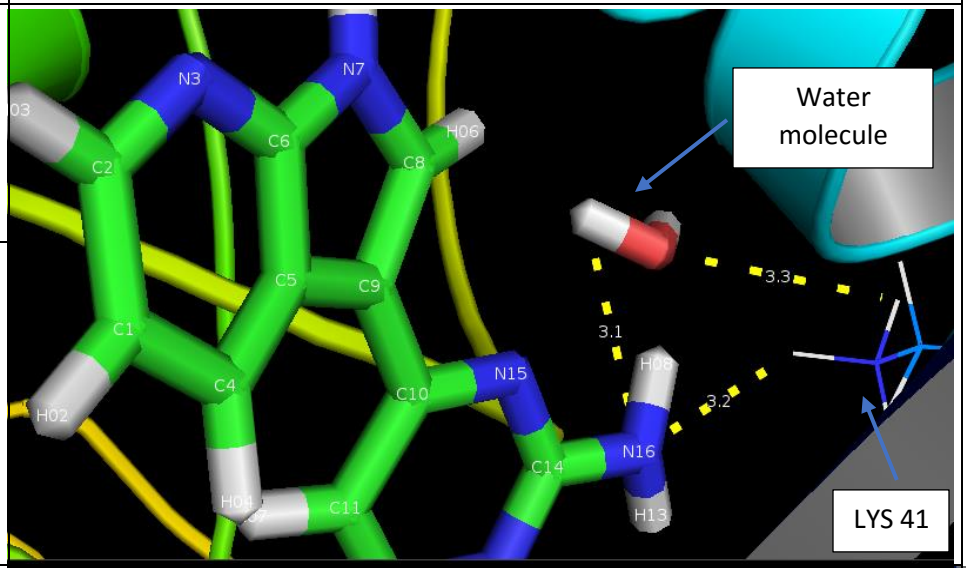
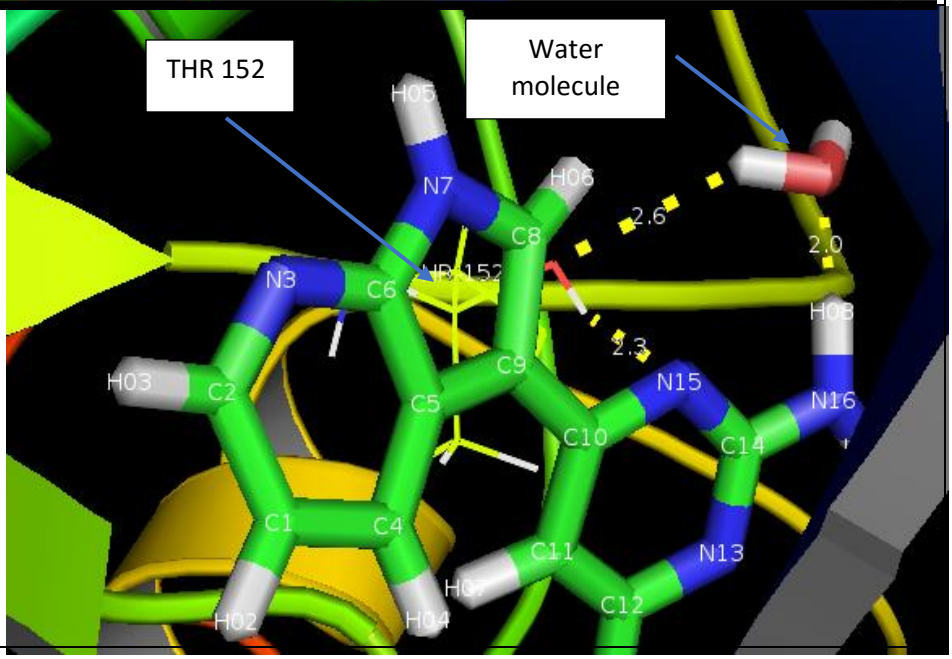
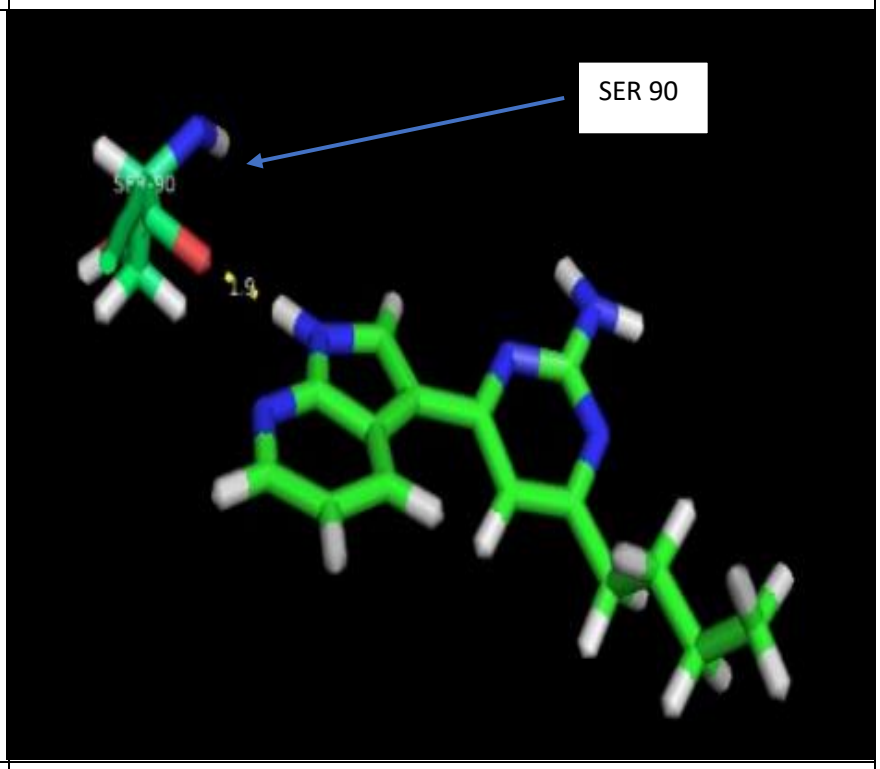
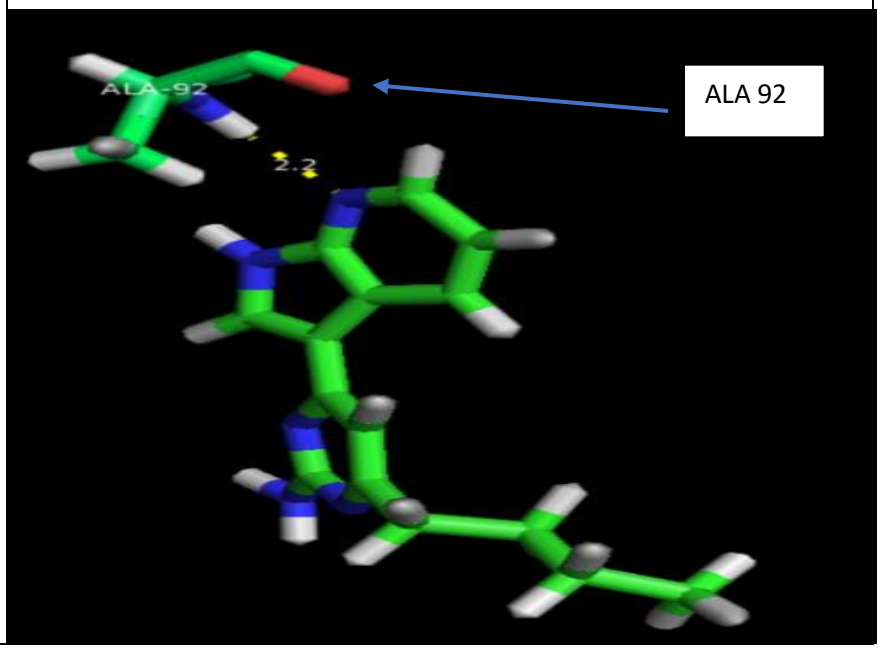
Atom of inhibitor	Atom of protein/H_2O	Comment
N16	H-Lys[41] (Water mediated)	
N16	H-Lys[41]	
N15	H-Thr[152]	
HO8	O-Thr[152] (Water mediated)	

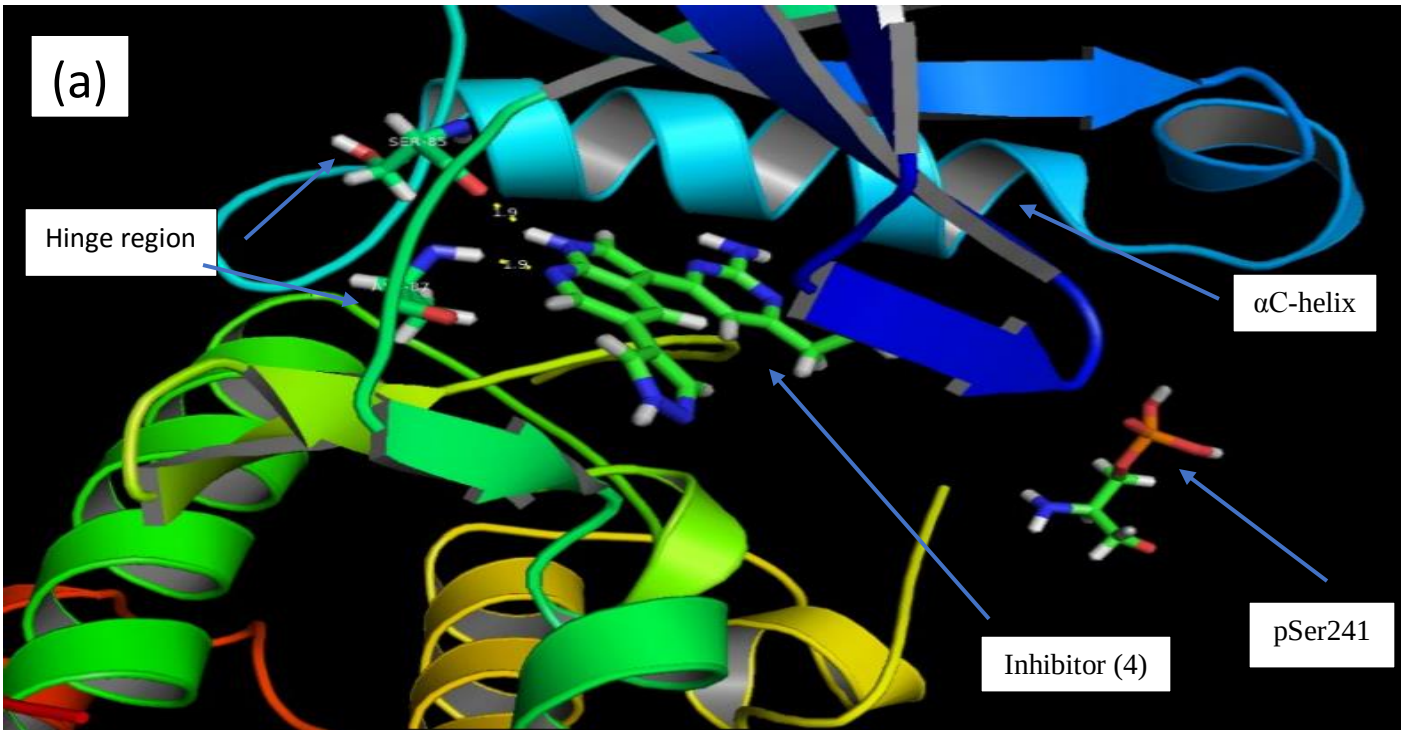
Table 3.4: Hydrogen bond analysis of inhibitor (3) in MOL-PDK-1complex

Atom of inhibitor	Atom of protein/H_2O	Comment
HO5	O-Ser^{90}	
N3	H-Ala^{92}	

3.6 Analysis of the binding mode of inhibitor (4)-PDK-1 complex

In this section we discussed the binding mode of inhibitor (4) with PDK-1 kinase complex (Figure 3.11(a) and (b)). Inhibitor (4) forms four hydrogen bonds with the protein. The N12 atom binds the hydroxyl group hydrogen of Thr[147] in the buried region, N4 atom binds hydroxyl group hydrogen of Ala[87] in the adenine region, N23 atom binds the hydrogen atom of Lys[94], N23 atom binds hydroxyl group hydrogen atom of Glu[91] in the sugar region and HO5 hydrogen atom binds to the backbone carbonyl group of Ser[85] in the adenine region (Table 3.5).

Other weak interaction was formed between inhibitor (4) and the PDK-1, C-H…. π interaction[54] between carbon hydrogen atom of Leu[13] and the center of 7-azaindole ring as shown in figure 3.11(c).



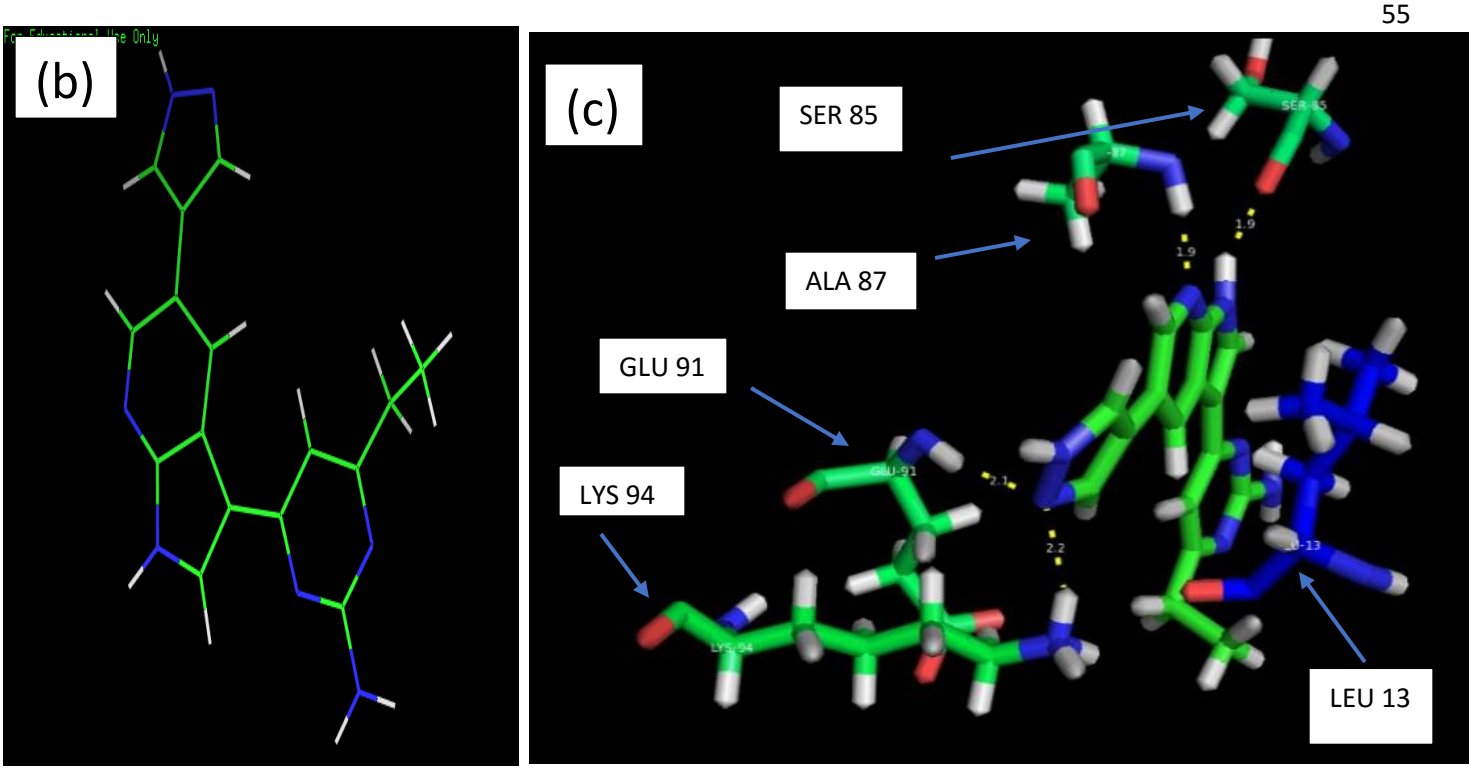


Figure 3.11: (a) Inhibitor (4)-PDK-1 complex, (b) graphical representation of inhibitor (4) and (c) other weak interactions between inhibitor (4) and kinase

It was reported that **type I kinase inhibitors** form water-mediated hydrogen bond networks (both water molecules W1 and W2 are commonly observed) and the ligand does not extend to the water-filled cavity. These two features distinguish type I from type II inhibitors.[59]

Figure 3.12 illustrates the typical distribution of ligand-W1 hydrogen bond distances for 180 ATP-binding site ligands.[59] According to inhibitor (4), the ligand-W1 hydrogen bonds was 3.0 Å which is agree with the typical distribution of hydrogen bonds.

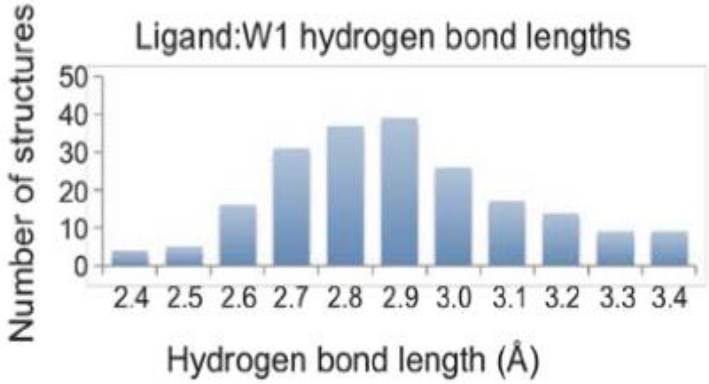


Figure 3.12: The distribution of hydrogen bond lengths for 180 ATP-binding site ligand-W1 hydrogen bonds[59]

Table 3.5: Hydrogen bond analysis of inhibitor (4) in 61Y-PDK-1complex

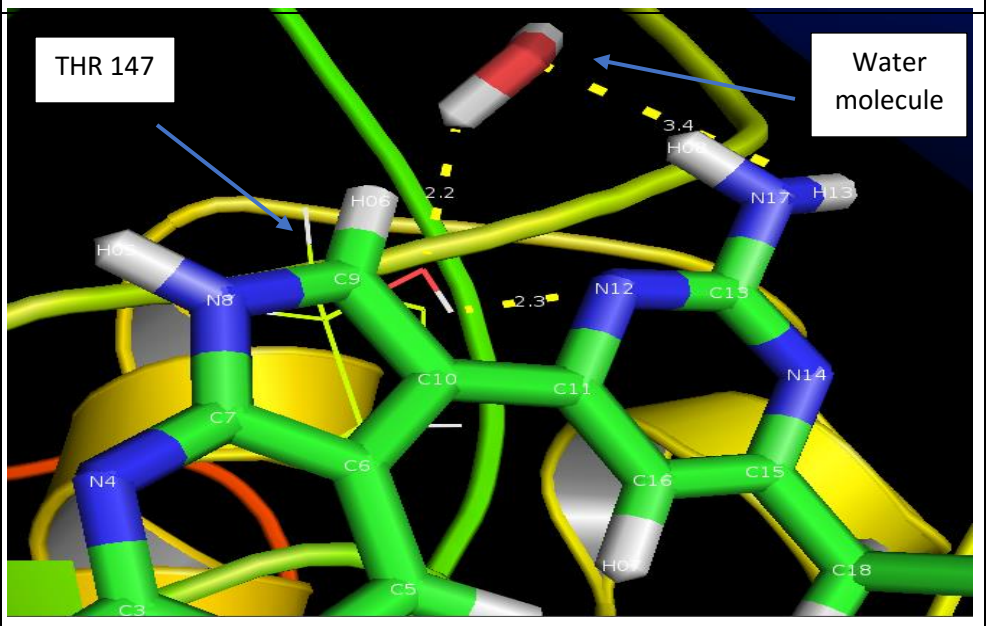
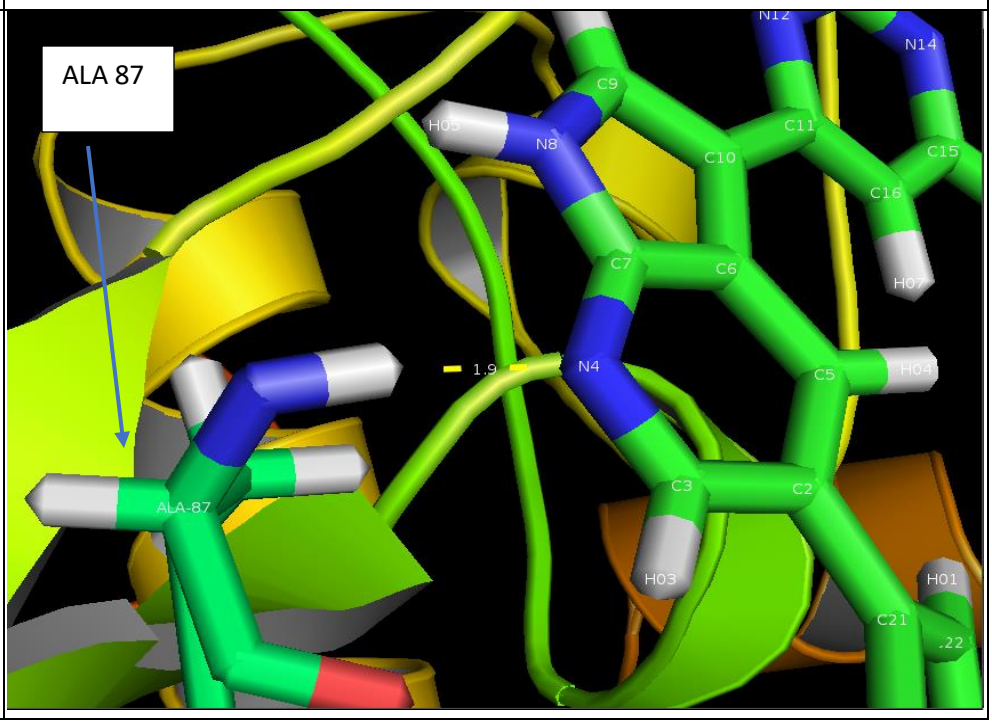
Atom of inhibitor	Atom of protein/H_2O	Comment
N12	H-Thr[147]	
H13	O-Thr[147] (Water mediated)	
N4	H-Ala[87]	

Table 3.5: Hydrogen bond analysis of inhibitor (4) in 61Y-PDK-1complex

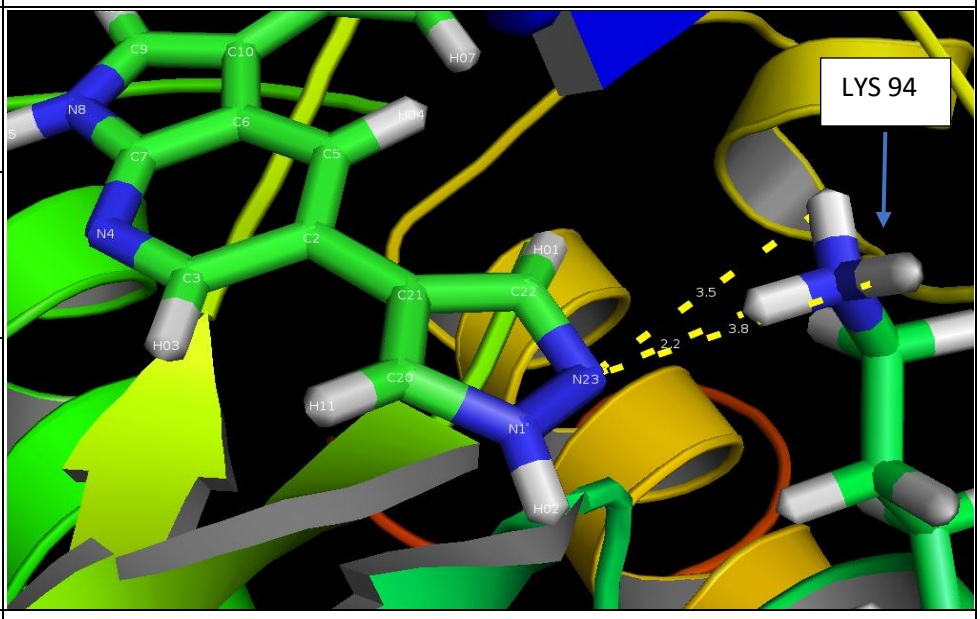
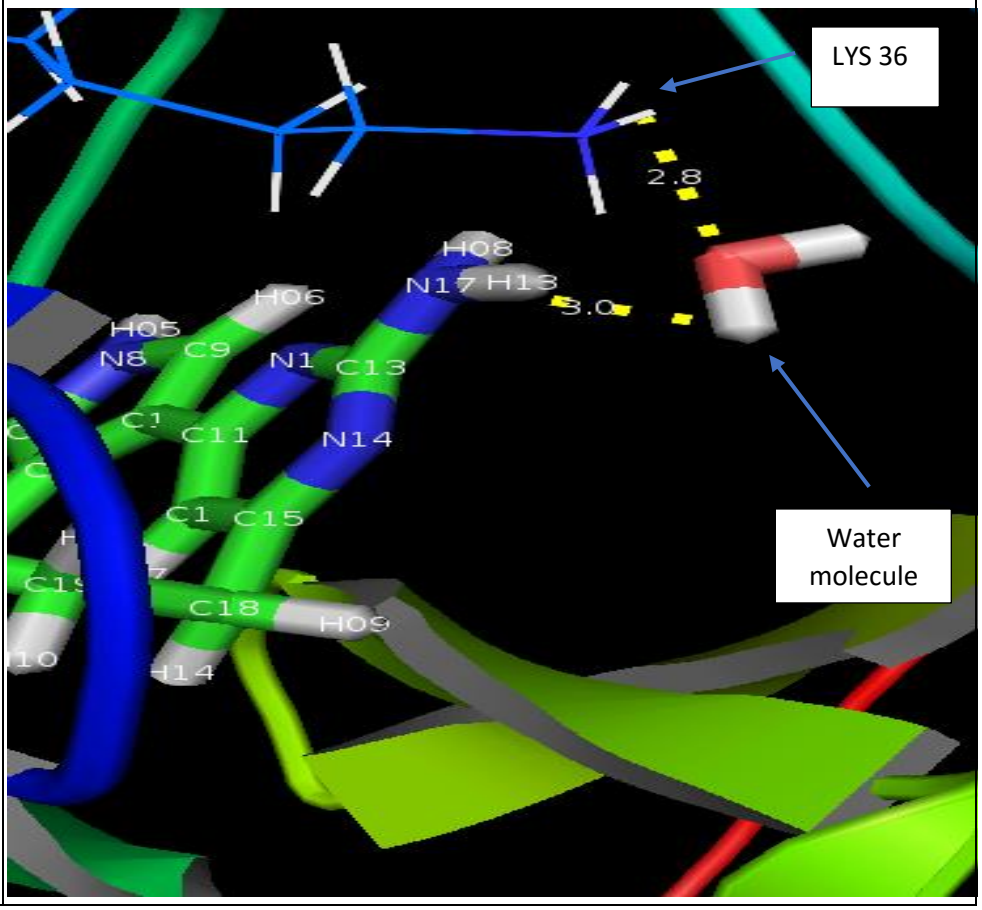
Atom of inhibitor	Atom of protein/H_2O	Comment
N23	H-Lys[94]	
N23	H-Lys[94]	
N23	H-Lys[94]	
N17	H-Lys[36] (Water mediated)	

Table 3.5: Hydrogen bond analysis of inhibitor (4) in 61Y-PDK-1complex

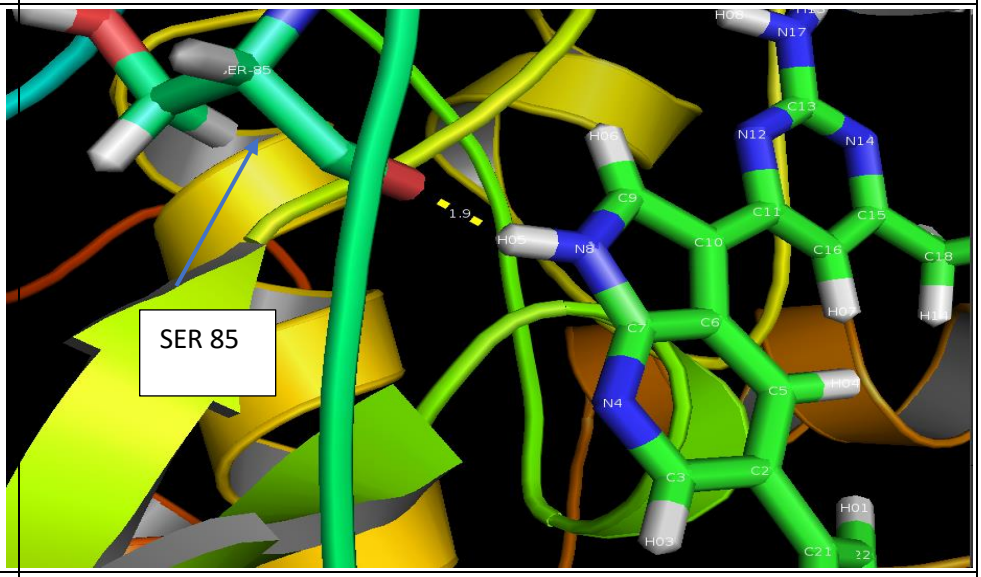
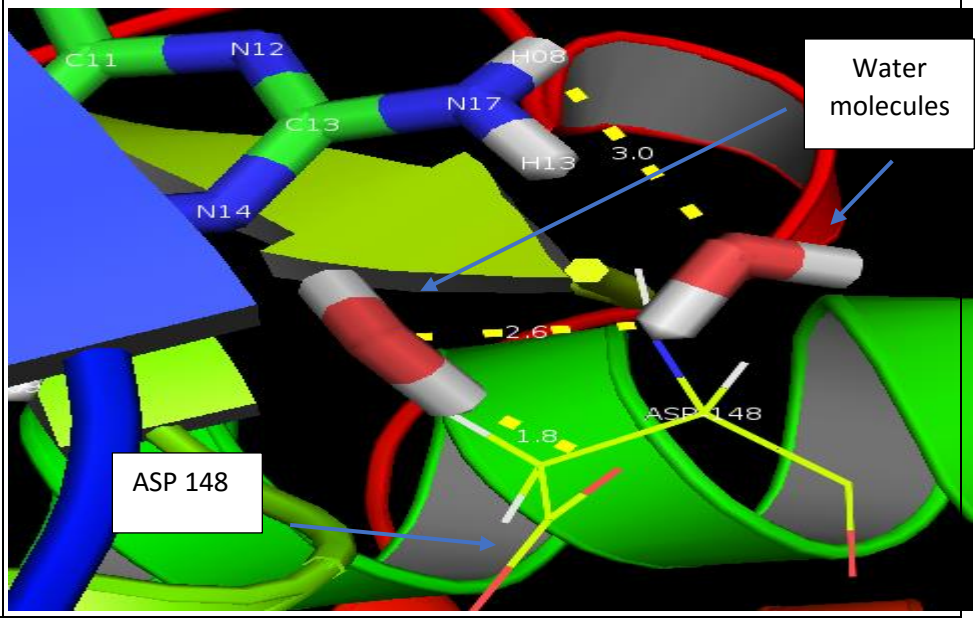
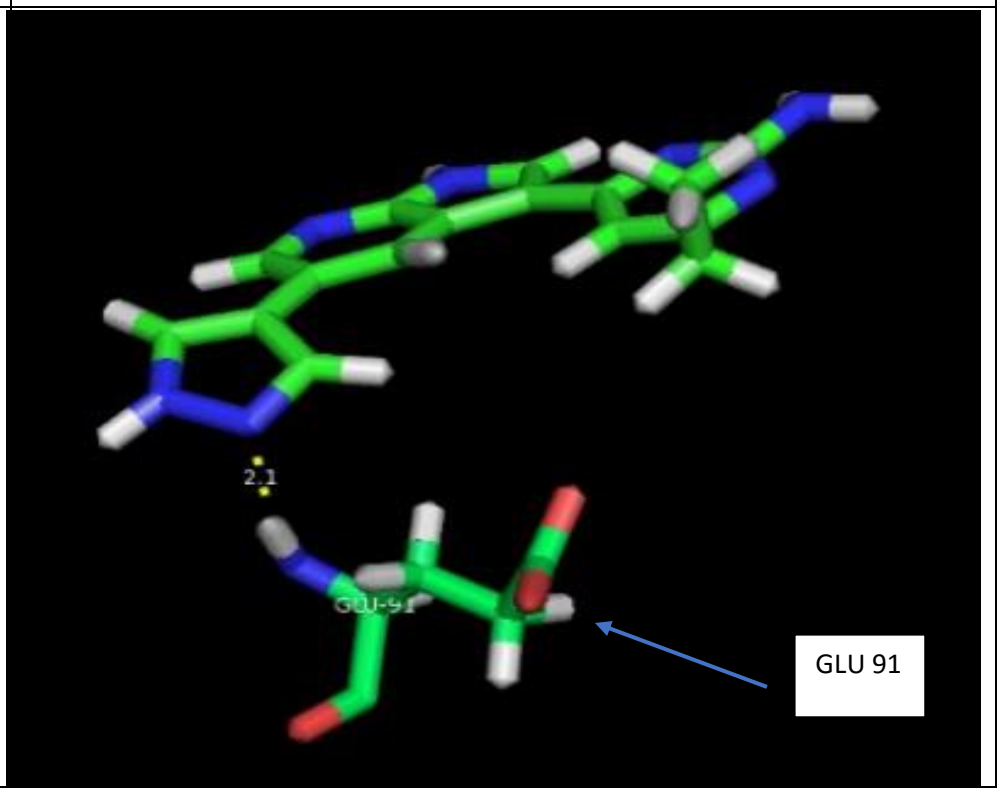
Atom of inhibitor	Atom of protein/H_2O	Comment
HO5	O-Ser[85]	
HO8	O-Asp[148] (Water mediated)	

Table 3.5: Hydrogen bond analysis of inhibitor (4) in 61Y-PDK-1complex

Atom of inhibitor	Atom of protein/H_2O	Comment
N23	H-Glu^{91}	

Dunitz et. al[60] reported that the entropic gain of releasing a bound water molecule from the binding site of protein can be 7 cal/mol. K, corresponding to an energy gain of 2.1 kcal/mol at 300 K.[60] The entropic contribution to binding affinity is observed upon removing water molecules from the binding sites of protein molecules, and is an essential part which cannot be ignored in drug design.[61]

The position of water molecules in the binding sites can be used to design better inhibitors in which the principle lies in the fact that a substituent is added to the inhibitor that replaced a water molecule that bounded to kinase (design inhibitor that includes a structural water mimic).[62]

An increase in ligand affinity can result if the contribution of substituent is greater than free energy cost which results from displacing solvent molecules. This is an easy process because the ligand already has paid the energy cost as translational and rotational entropy.[61]

3.7 Analysis of the binding mode of inhibitor (5)-PDK-1 complex

In this section we discussed the binding mode of inhibitor (5) with PDK-1 kinase complex (Figure 3.13(a) and (b)). Inhibitor (5) occupied allosteric site of the protein kinase PDK-1 called the PDK1 -interacting fragment (PIF)tide-binding site, or PIF pocket. This inhibitor was occupied PIF/Phosphate pocket which was determined by Lys115, Ile118, Ile119, Val124, Leu155 residues.

Inhibitor (5) binds to the inactive kinase conformation (DFG-out) in the PIF/Phosphate pocket of PDK-1 kinase, so this inhibitor considered as Type II (Deep pocket binder) inhibitor. It is worth noting that this is the first reported example of Type II (DFG-out) kinase inhibitor for AGC kinase.[63]

Inhibitor (5) consists three molecular fragments: a hinge binding group, a linker, and a hydrophobic moiety.[63] The hydrophobic moiety interacts with the phosphate pocket through four strong hydrogen bond interactions. Three hydrogen bonds are formed between O34 atom and hydrogen atoms (HZ1, HZ2, and HZ3) of the amino group of Lys^{36}. The fourth strong hydrogen bond is formed between the carbonyl group of inhibitor and the amino group hydrogen atom of Asp^{148} as shown in table 3.6.

This inhibitor interacts with the hinge region through Ser^{85}, and Ala^{87}. The first interaction is between H22 amino group hydrogen atom of inhibitor and the

carbonyl group of Ser[85]. The second interaction is between carbonyl group of inhibitor and amino group hydrogen atom of backbone Ala[87] (Table 3.6).

It is worth noting that a strong intramolecular hydrogen bond interaction is present. This intramolecular interaction is between the carbonyl group and the amino group hydrogen atom (HO1). This type of interaction is like inhibitor (2) between hydrogen atom of amino group (HO2) and oxygen of carbonyl group (O14) as shown in table 3.6.

As previously stated: replacing real rings by pseudo rings to form pseudo six-membered ring is a new and non-conventional strategy and the new classes of kinase inhibitors follow this approach.[56]

In addition to the strong interactions, other weak interactions were formed between inhibitor (5) and the PDK-1 (Figure 3.13(c)), C-H….C=O interaction[54] between the carbonyl group of Phe[149] and the carbon hydrogen atom of inhibitor (3.0 Å distance), between the carbon hydrogen atom of Leu[84] and the carbonyl group of inhibitor (3.0 Å distance) and between the carbon hydrogen atom of Tyr[86] and the carbonyl group of inhibitor (3.0 Å distance) as shown in figure 3.14(c).

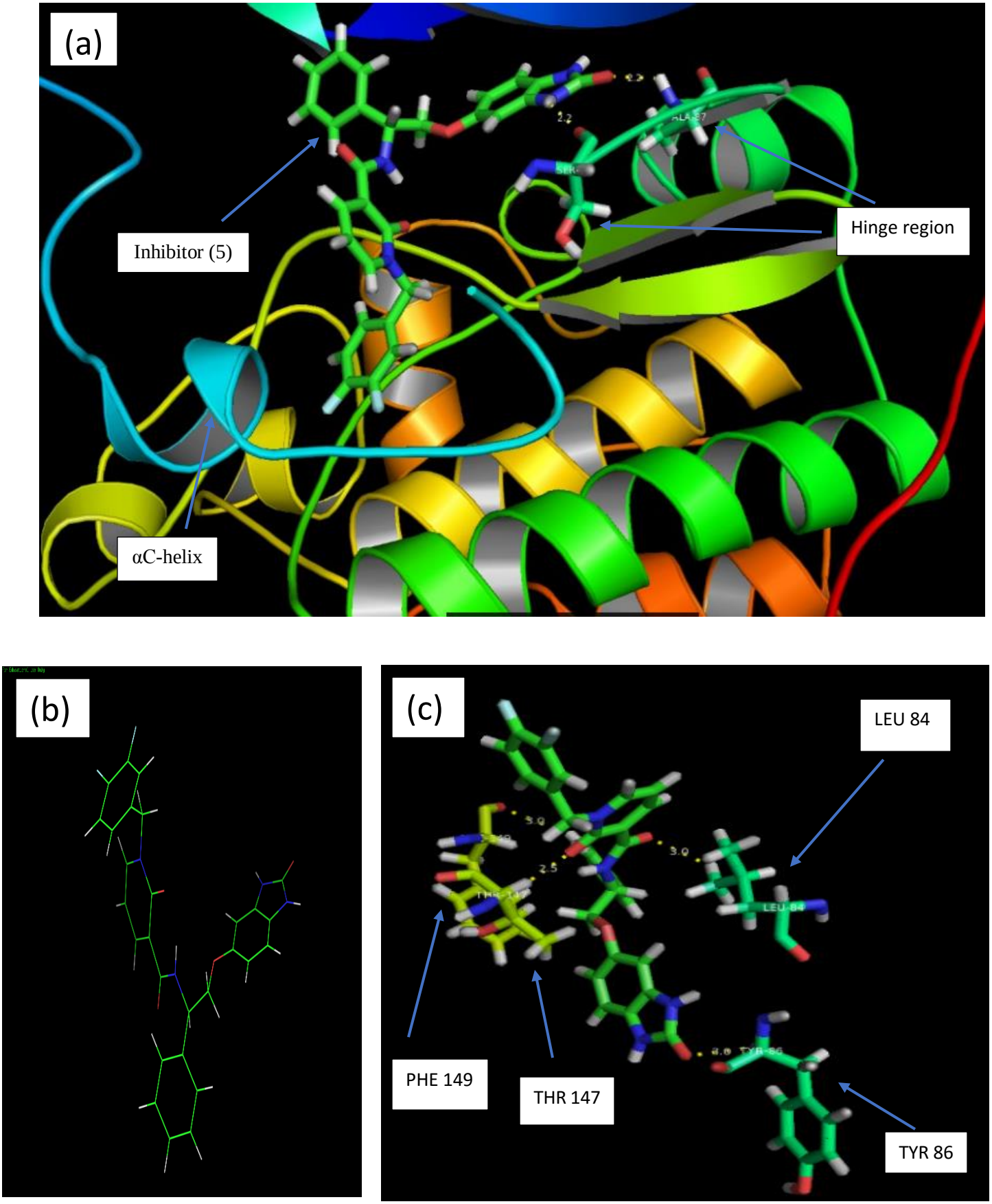


Figure 3.13: (a) Inhibitor (5)-PDK-1 complex, (b) graphical representation of inhibitor (5) and (c) other weak interactions between inhibitor (5) and kinase

Table 3.6: Hydrogen bond analysis of inhibitor (5) in MP7-PDK-1complex

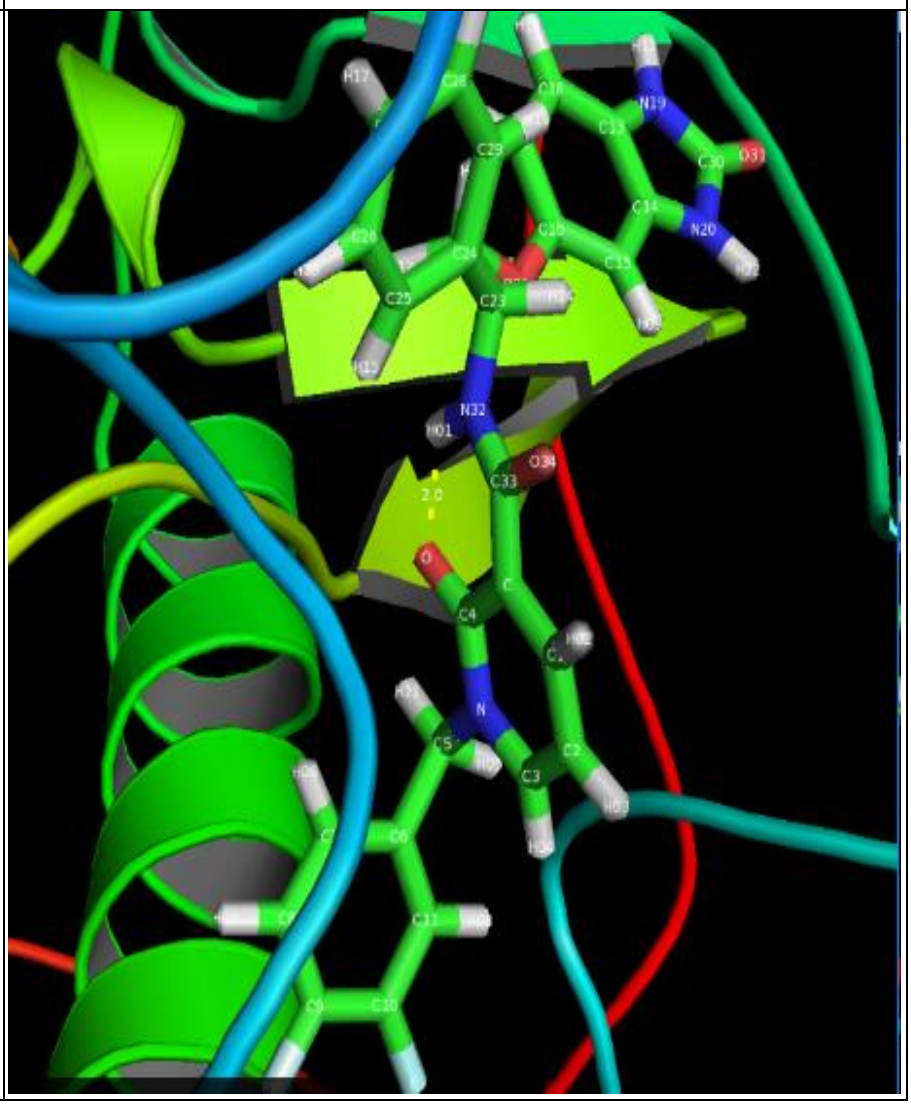
Atom of inhibitor	Atom of protein/H_2O	Comment
O	H-MP7[279] (Intramolecular hydrogen bond)	

Table 3.6: Hydrogen bond analysis of inhibitor (5) in MP7-PDK-1complex

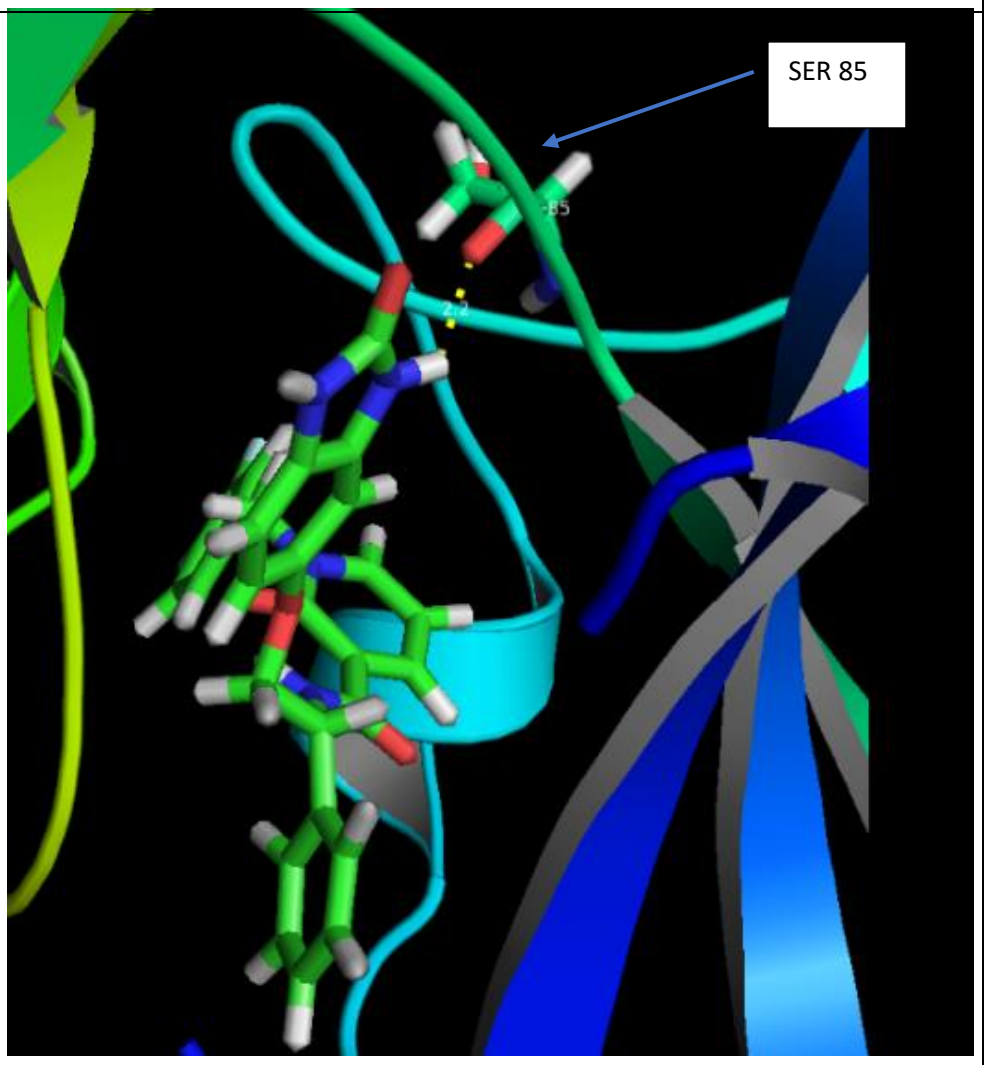
Atom of inhibitor	Atom of protein/H_2O	Comment
H22	O-Ser^{85}	

Table 3.6: Hydrogen bond analysis of inhibitor (5) in MP7-PDK-1complex

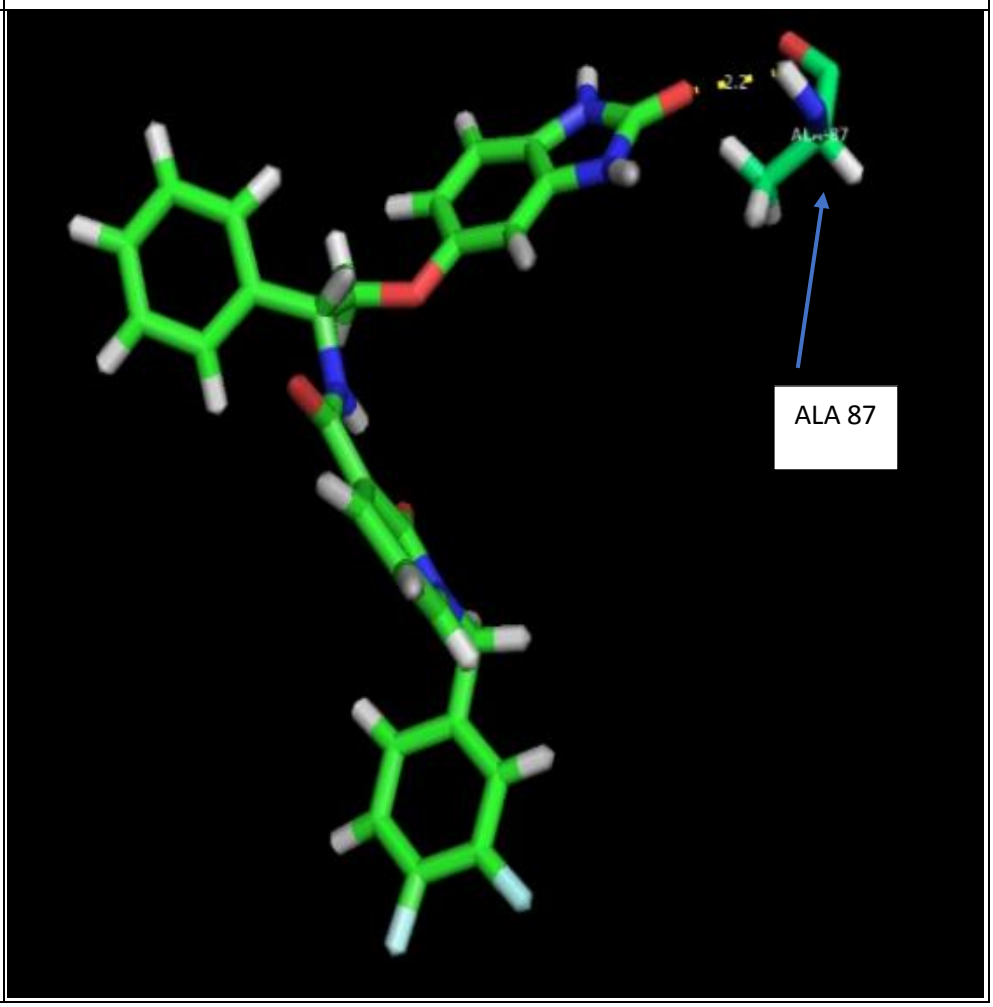
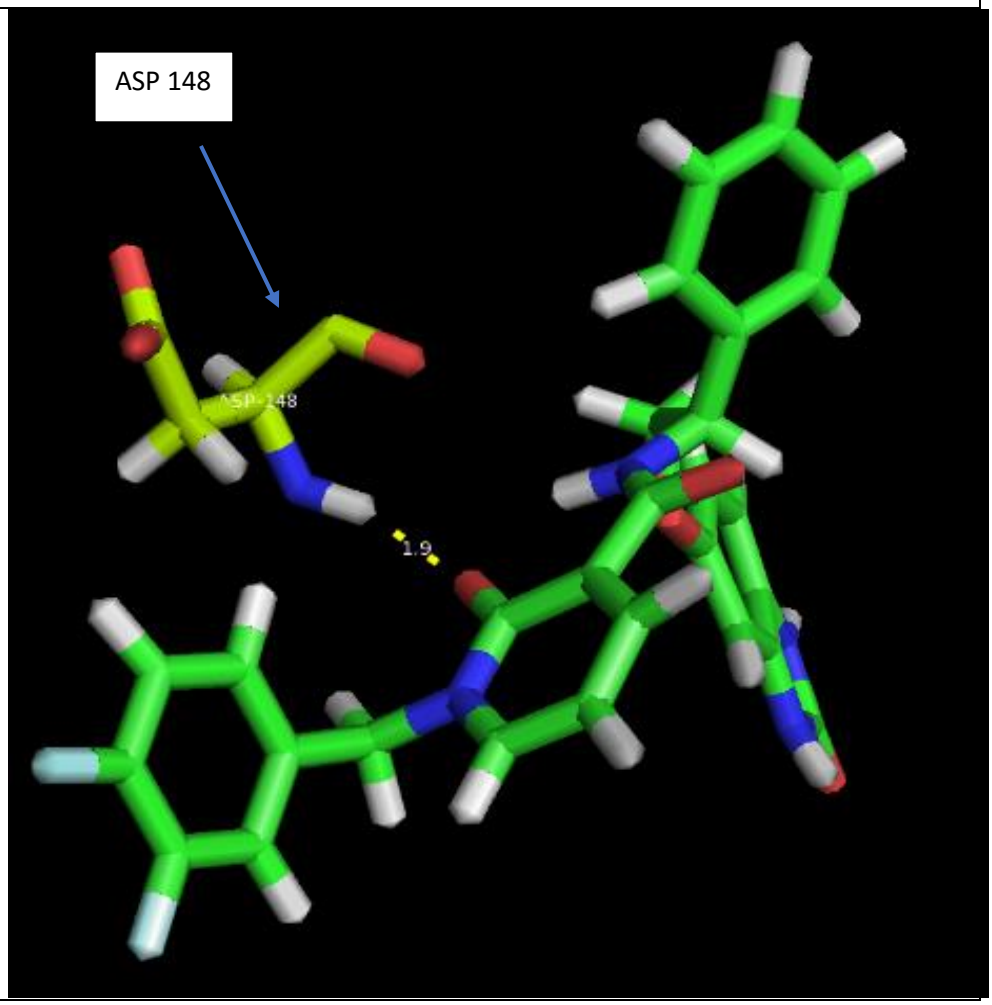
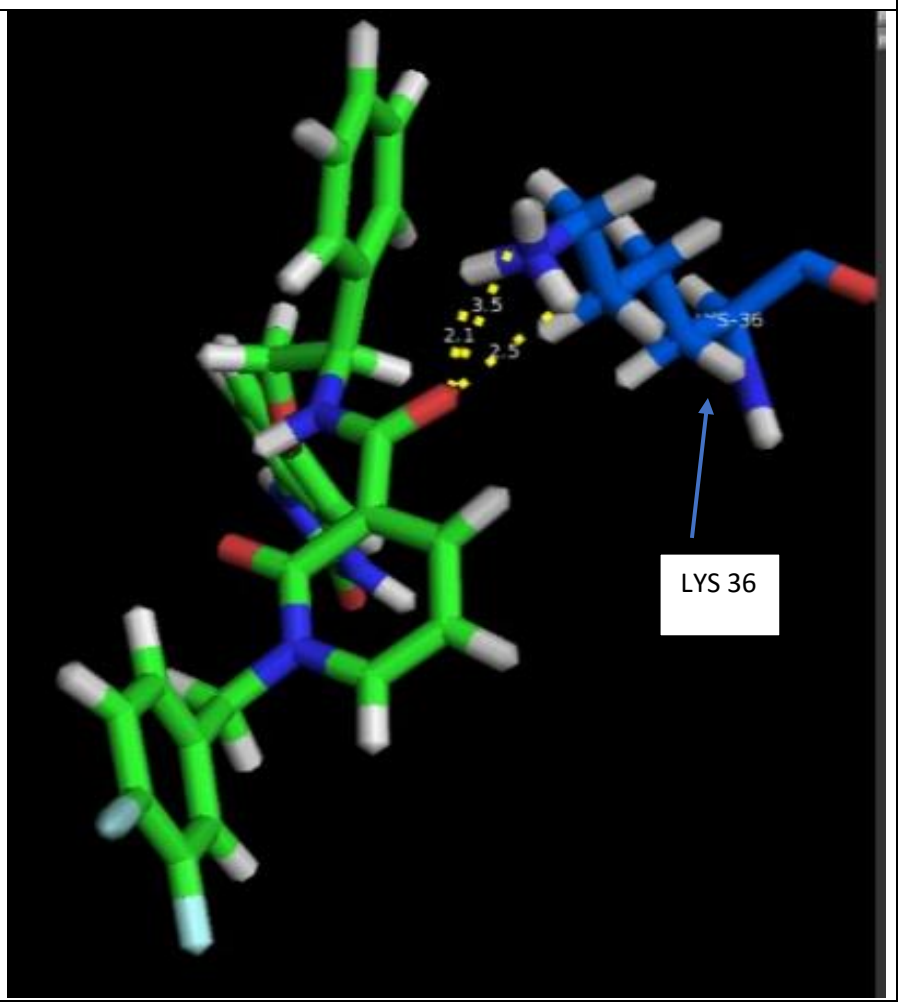
Atom of inhibitor	Atom of protein/H_2O	Comment
O31	H-Ala^{87}	

Table 3.6: Hydrogen bond analysis of inhibitor (5) in MP7-PDK-1complex

Atom of inhibitor	Atom of protein/H_2O	Comment
O	H-Asp^{148}	

Atom of inhibitor	Atom of protein/H_2O	Comment
O	H-Lys[36]	

A significant difference between the classical ATP-competitive inhibitors and Type II (Deep pocket binder) inhibitors were firstly, αC-helix of PDK-1 kinase was distorted. The conformational change in this helix is due to displacing of Glu-130 residue from the active site[60], this disruption was observed in our study.[64]

Secondly, there was a hydrogen bond interaction between Arg^{131} with PS^{48} in the classical ATP-competitive inhibitors, but this is replaced by Arg^{131} with Glu^{130} in the inactive conformation.[64] This is consistent with our results.

It was reported that the distortion of DFG motif is the most obvious in Type II (Deep pocket binder) inhibitor. Rotation about ϕ main chain torsion angle of Asp^{381}, as a result of this rotation, Phe^{382} removed from ATP pocket and Asp^{381} becomes to the back pocket[65].

This conformational change creates an inactive state of the kinase because the flipped-out phenylalanine blocks ATP-binding site.[65] This distortion was not observed in our study because Asp^{381} and Phe^{382} residues were not present in the original PDB files of inhibitor-protein complex.

3.8 Effect of Thermodynamic parameters on the protein-inhibitor complexes

The unfavorable interactions between studied inhibitors (1-5) and PDK-1is reflected in the negative value of entropic contribution ($T\Delta S$). This is due to the release of the ordered H_2O molecules in addition to the conformational change, which is typically negative as the association of a ligand with its target results in the loss of conformational freedom for one or both molecules.[66]

In other words, the negative entropic contribution resulted from “freezing out” of translational, rotational and internal degrees of freedom of the ligand on

binding.[67] As illustrates in figure 3.14, the coefficient of determination value is R^2 =0.12. This reflects that entropic contribution is not a driving force of binding affinity. Whereas, as appears in figure 3.15. The difference of enthalpy is considered as a driving force of binding free energy, due to the high value of coefficient of determination (R^2 = 0.98).

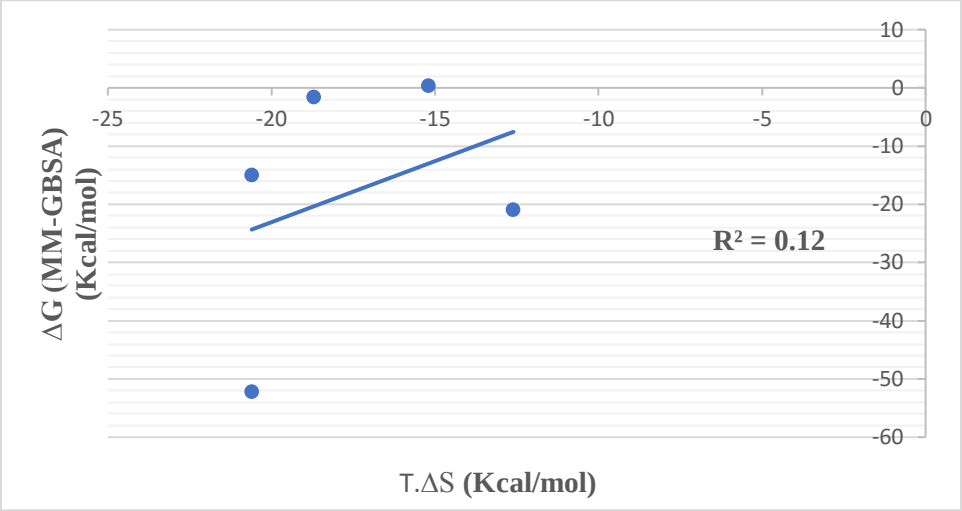


Figure 3.14: Relationship between the difference of Gibbs free energy (ΔG) and the entropic contribution (T.ΔS)

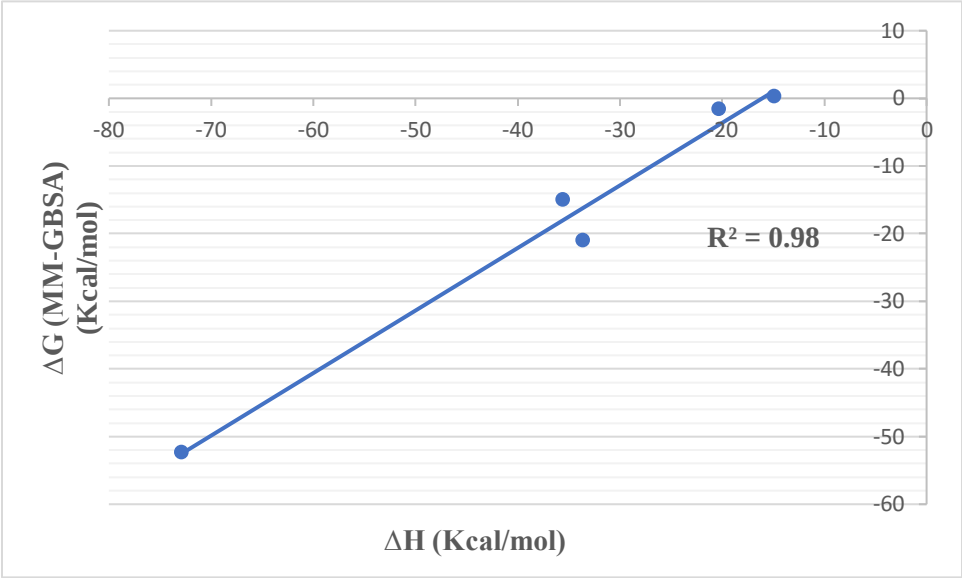


Figure 3.15: Relationship between the difference of Gibbs free energy (ΔG) and the difference of enthalpy (ΔH)

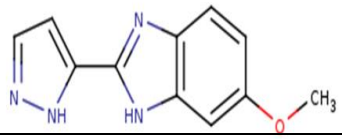
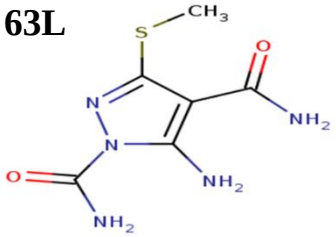
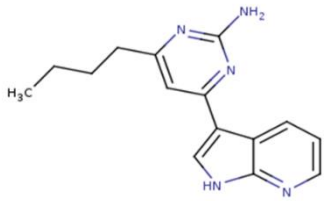
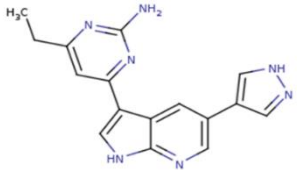
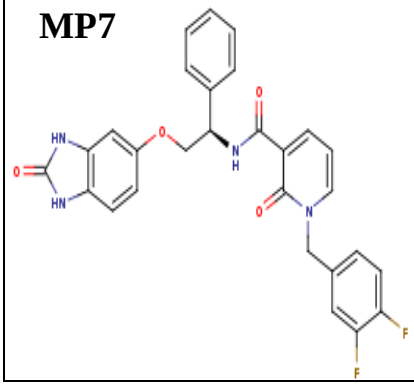
Positive contributions from both enthalpy and entropy are requirements for high affinity of binding. These two aspects of drug association should be optimized in a challenging and perplexing process because enthalpy optimization can frequently be offset by a loss in entropy. Maximizing the enthalpy contribution is difficult due to the formation of favorable H-bonds and van der Waal contacts and this is opposed by the cost of desolvation of incorrectly positioned polar moieties within a molecule.[66]

The previous statement is clearly observed when we try to maximize the number of hydrogen bond interactions in inhibitor (4)-protein complex, conversely, we faced positive contribution of entropy.

The binding free energy of a ligand for its target is a function of enthalpic and entropic contributions as defined by the Gibbs free energy change. This can be parsed into individual contributions of intermolecular van der Waal attractive forces, H-bonding interactions, and repulsive forces like the hydrophobic effect that drive a ligand out of water and into the hydrophobic cavity of a protein.[66]

The negative binding free energy (ΔG_{bind}) of all complexes reflects the favorable interaction between inhibitor-protein complexes in pure water except in the first complex where it gave positive value.

Table 3.7: Thermodynamic parameters of the five protein-inhibitor complexes that calculated at T= 300 K and P= 1 atm

	Experimental		Calculated				
Inhibitor	**IC_{50}** (μM)	**ΔG_{exp}** (Kcal/mole)	**ΔE_{MM}**	**ΔG_{sol}**	**ΔG_{calc}** **(MM-GBSA)** (Kcal/mole)	**$T\Delta S$** (Kcal/mole)	**ΔH** (Kcal/mole)
62O 	93	-5.5	-35.9	21.0	0.3 ± 1.6	-15.2 ± 3.2	-14.9
63L 	17	-6.5	-56.7	36.4	-1.6 ± 2.0	-18.7 ± 2.6	-20.3
MOL 	1.1	-8.1	-65.8	32.2	-21.0 ± 1.5	-12.6 ± 7.2	-33.6
61Y 	0.013	-10.8	-58.7	23.1	-15.0 ± 1.8	-20.6 ± 3.6	-35.6
MP7 	-	-	-132.0	59.1	-52.3 ± 2.8	-20.6± 3.6	-72.9

3.9 Correlation between IC_{50} and binding free energy

The acceptable IC_{50} value for a ligand to possess a drug-like property is (1-10) nM. According to this parameter which is most critical in determining the drug candidate, we conclude that inhibitor (1), inhibitor (2), inhibitor (3), and inhibitor (4) do not possess a drug-like property (IC_{50} = 93 µM, 17 µM, 1.1 µM, 0.013 µM), respectively, because they dissociate and do not stay bound to the enzyme.

These values are correlated with binding free energies of these complexes that calculated by MM-GBSA which are 0.3 K cal/mol, -1.6 Kcal/mol, -21.0 Kcal/mol, -15.0 Kcal/mol. We noted that IC_{50} becomes lower, binding free energy become larger in negative sign, and become more druggable property.

IC_{50} of Inhibitor (1) is equals 93 µM, this inhibitor does not possess drug-like property, and this result agree with what we computed, ΔG_{bind} of this inhibitor with PDK-1 kinase is 0.3 Kcal/mol. The positive sign of this value resulted from the formation of only one hydrogen bond interaction.

But if we look about inhibitor (5), we noted that this inhibitor has high value of ΔG_{bind} equals -52.3 Kcal/mol, this high negative value resulted from the formation of six strong hydrogen bond interactions. with no water mediated was present.

But if we compared inhibitor (3) and inhibitor (4) we see the IC_{50} values of inhibitor (3) and inhibitor (4) is 1.1 µM and 0.013 µM respectively. But the binding affinities of these inhibitors are -21.0 Kcal/mol and -15.0 Kcal/mol respectively. This is explained by the fact that the stronger the hydrogen bonds formed between the water molecule and the binding site, more favorable enthalpic contribution occur, and at the same time water molecules become less disordered and more highly restricted (less favorable entropic contribution).[65]

In inhibitor (4)-protein complex there were three water molecules instated in inhibitor (3)-protein complex there were only two water molecules.

As appear in figure 3.16, the IC_{50} values are well correlated with the binding affinities that calculated by MM-GBSA. The coefficient of determination in this case equals 0.55.

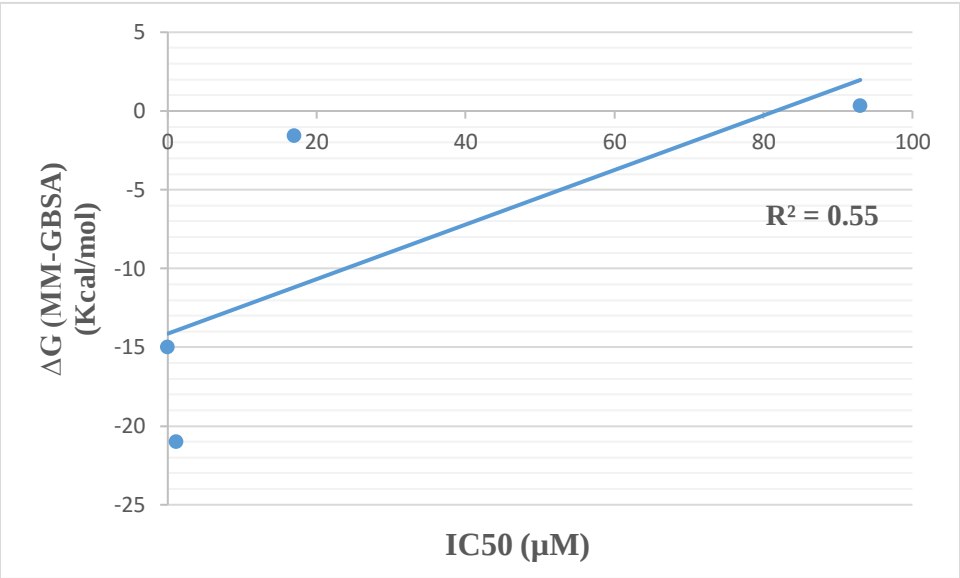


Figure 3.16: Correlation between the IC_{50} values and binding free energies that calculated by MM-GBSA

Whereas, the correlation between the IC_{50} and the binding free energy that calculated by MM-PBSA, is weaker (Figure 3.17). The coefficient of determination in this case equals 0.14.

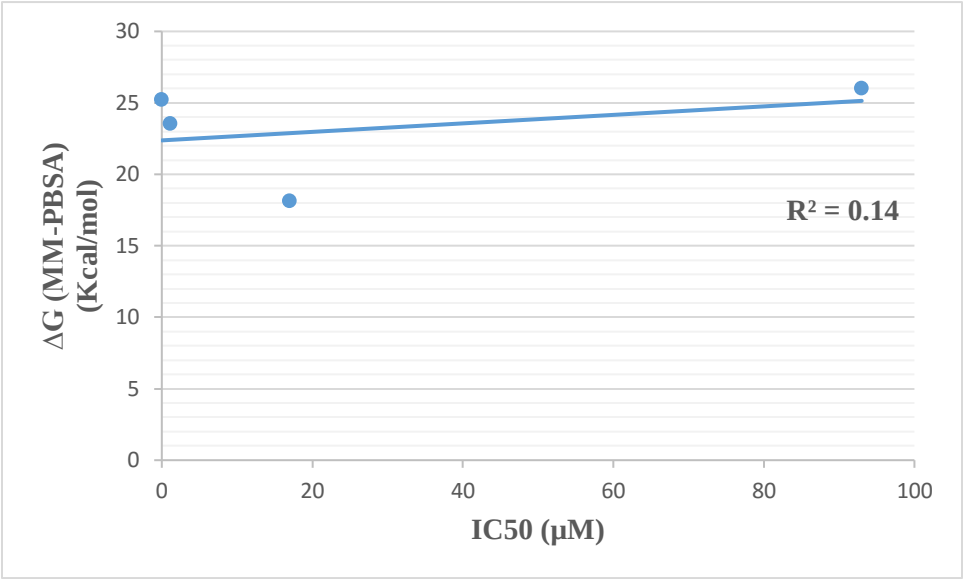


Figure 3.17: Correlation between the IC_{50} values and binding free energies that calculated by MM-PBSA

It is worth noting, that the coefficient of determination is the same with correlation between IC_{50} values and the experimental binding free energies equals 0.55 (Figure 3.18).

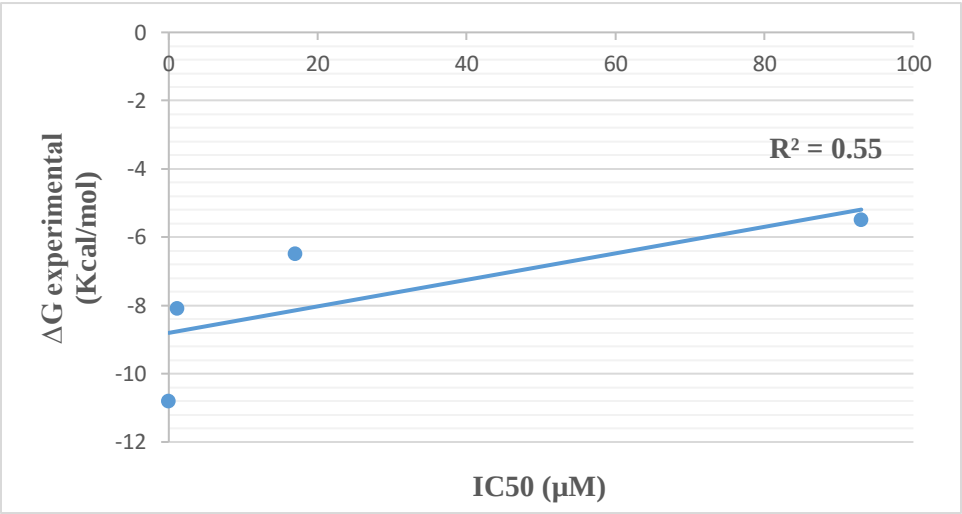


Figure 3.18: Correlation between the IC_{50} values and the experimental binding free energies

3.10 Energies calculated by MM-GBSA and contributing energies

All inhibitors studied in this work except inhibitor (3) share the most prominent binding contributions from the van der Waals (VDW) interactions. VDW values were -42.8 Kcal/mol, -30.7 Kcal/mol, -40.3 Kcal/mol, -76.4 Kcal/mol for inhibitor (1)-kinase complex, inhibitor (2)-kinase complex, inhibitor (4)-kinase complex, and inhibitor (5)-protein complex, respectively (Table 3.8, Figure 3.19).

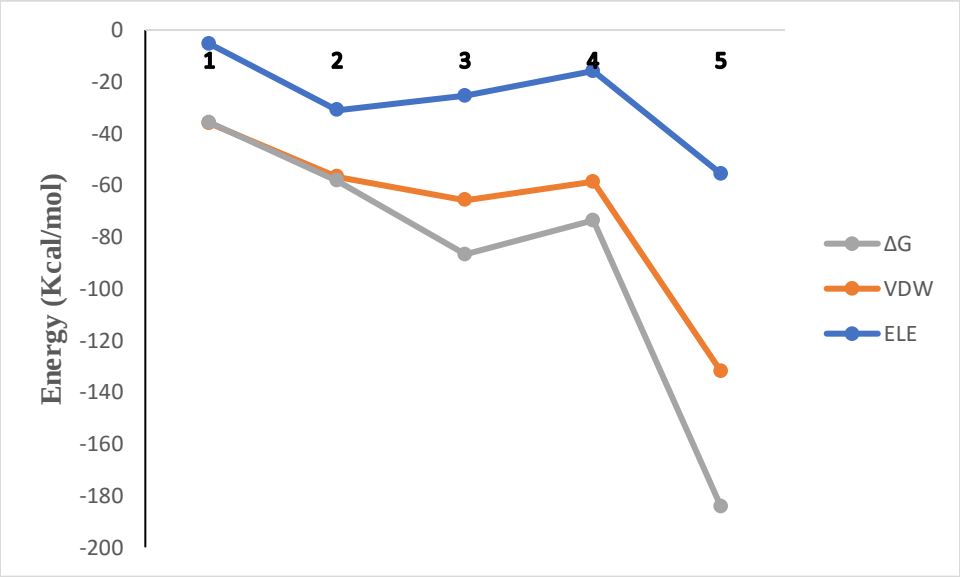
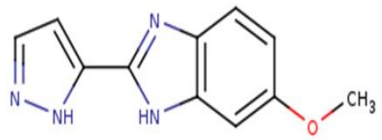
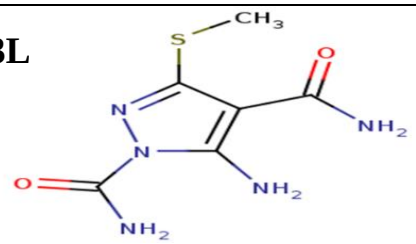
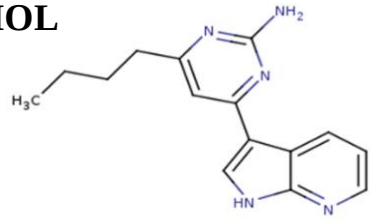
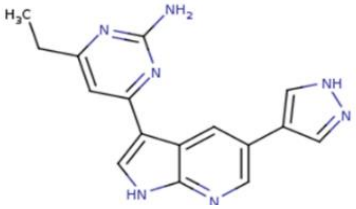
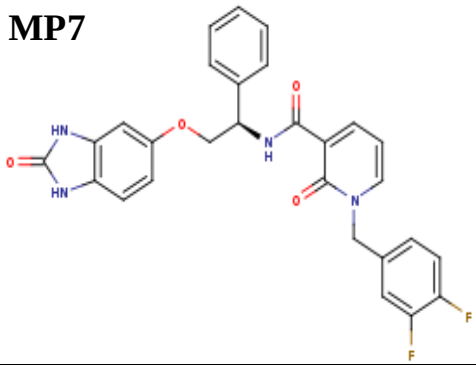


Figure 3.19: Contributions of electrostatic energy and van der Waals energies to the difference of Gibbs free energy (ΔG) of different inhibitor-PDK-1 kinase complexes

Table 3.8: Total binding free energy (ΔGtotal), van der Waals energy (VDW), electrostatic energy (ELE), solvation free energy (ΔGsol), and binding free energy of inhibitor-protein complex (ΔGbind) that calculated by MM-GBSA. All energies are in unit kcal/mol

Inhibitor	ΔG_{total}	VDW	ELE	ΔG_{sol}	ΔG_{bind} (MM-GBSA)
62O 	-15.0	-30.7	-5.3	21.0	0.3 ± 1.6
63L 	-20.3	-25.7	-31.0	36.4	-1.6 ± 2.0
MOL 	-33.6	-40.3	-25.5	32.2	-21.0 ± 1.5
61Y 	-35.6	-42.8	-15.9	23.1	-15.0 ± 1.8
MP7 	-72.9	-76.4	-55.6	59.1	-52.3 ± 2.8

The binding free energy values are well correlated with van der Waals Energy calculated by molecular mechanics (VDW). The coefficient of determination value in this case equals 0.96 (Figure 3.20).

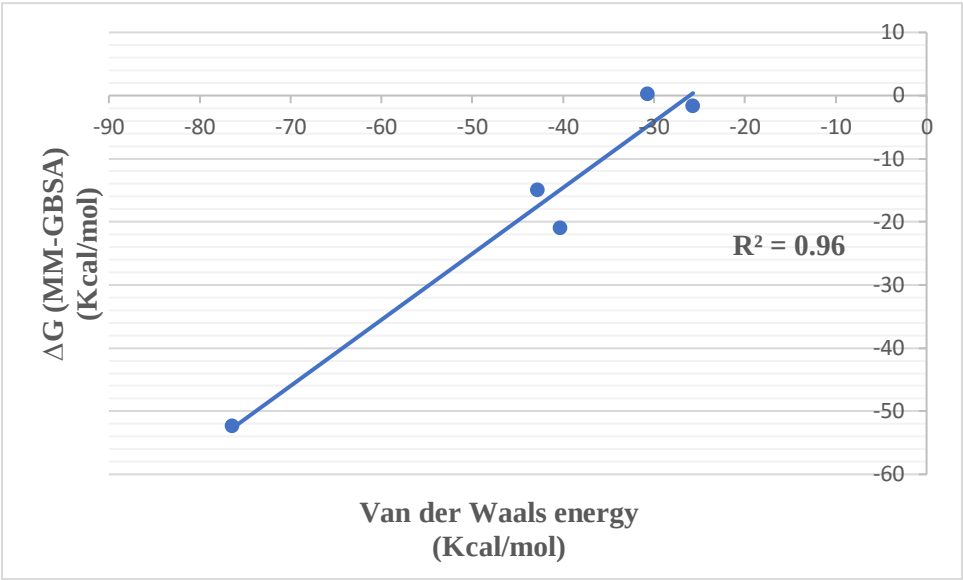


Figure 3.20: Relationship between the difference of Gibbs free energy (ΔG) and the Van der Waals energy

To the contrary, the correlation between the electrostatic energy and the binding free energy, is weaker (Figure 3.21). The coefficient of determination in this case equals 0.70.

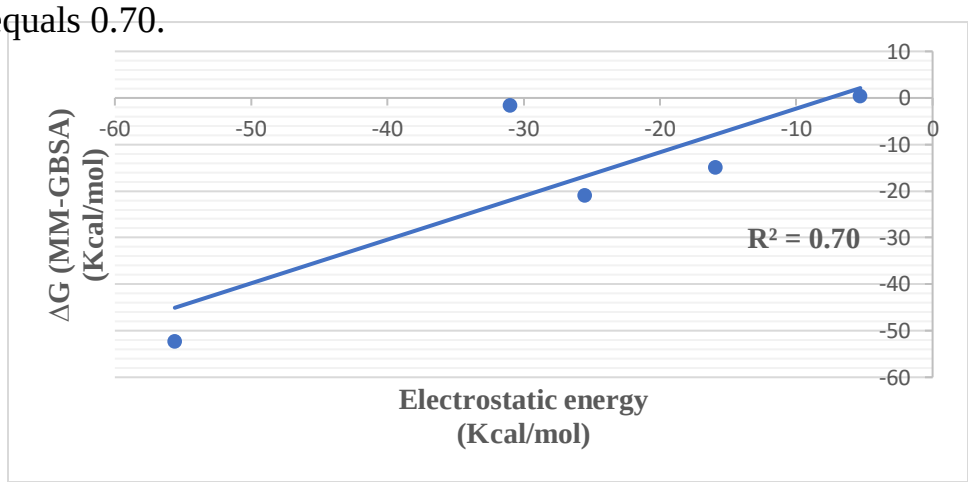


Figure 3.231: Relationship between the difference of Gibbs free energy (ΔG) and the electrostatic energy

3.11 Classification of inhibitors studied

The structure of all inhibitor-protein complexes except inhibitor (5) show a phosphorylated T-loop and are therefore, assumed to be in an active state.[17]According to Traxler model, inhibitors (1-4) bind to the activated (phosphorylated) form of the protein kinase and occupy the ATP binding site with at least a formation of one hydrogen bond with the hinge region. This reflects the fact that these inhibitors belong to traditional pharmacophore model (type I) or classical ATP-competitive inhibitor.[68]

It is worth noting that these inhibitors are reversible ATP-competitive because the type of interaction is mostly hydrogen bonding interaction and no irreversible covalent bond formation.[69]

Inhibitor (5) binds to the inactive kinase conformation (DFG-out) in the PIF/Phosphate pocket of PDK-1 kinase, so this inhibitor is considered as Type II (Deep pocket binder) inhibitor. It is worth noting that this is the first reported example of Type II (DFG-out) kinase inhibitor for AGC kinase.[63]

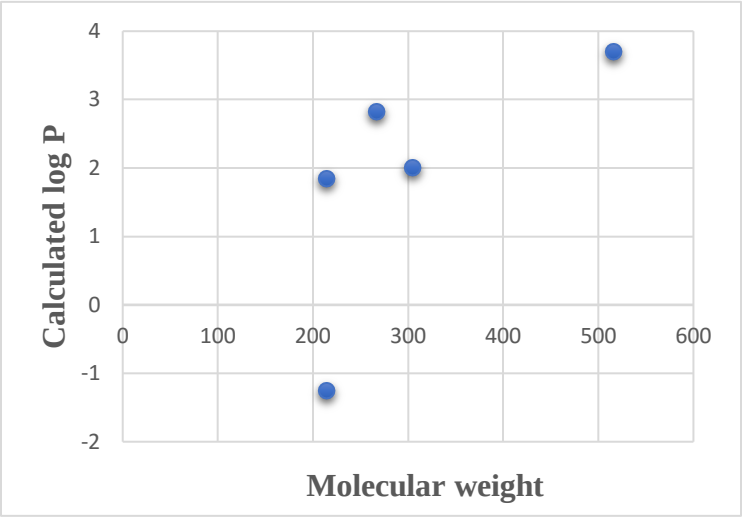
Another evidence proved that inhibitor (5) is considered as deep-pocket binder inhibitor Type (II) is the absence of water-mediated hydrogen bond interactions. The presence of water molecules in the binding sites is consider a feature that distinguish Type (I) from Type (II) inhibitors.

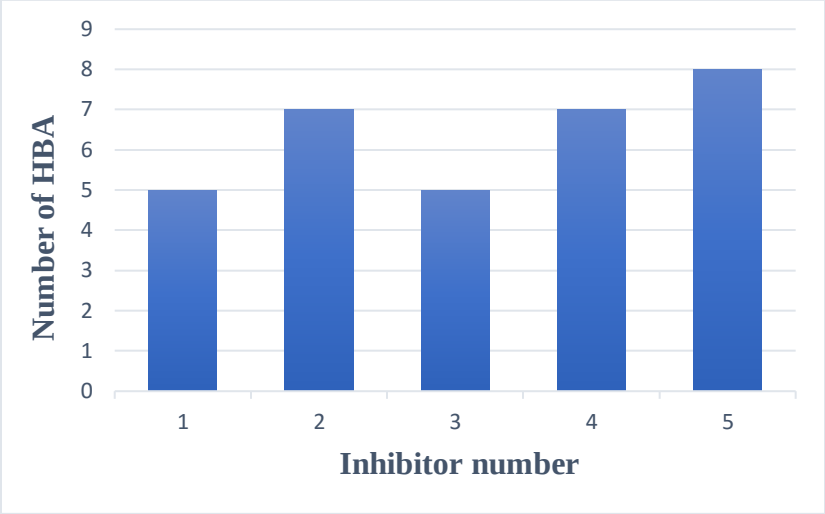
3.12 Analysis of the inhibitors according to Lipinski's Rule of five, Veber Rule and MDDR Rule

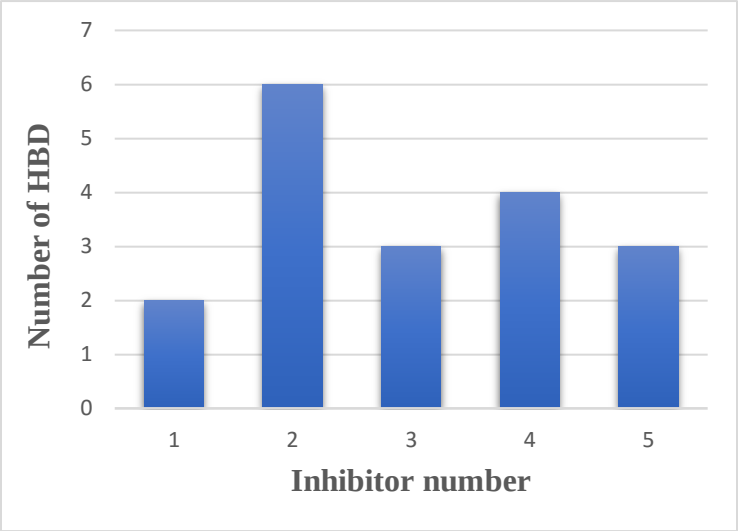
The fragment -based approach to calculate the polar surface area (PSA) descriptor, is a free software package. In this study, ***Molinspirstion*** was used to calculate PSA (with other useful molecular descriptors). ***SMILES*** files are required to process the values.[70]

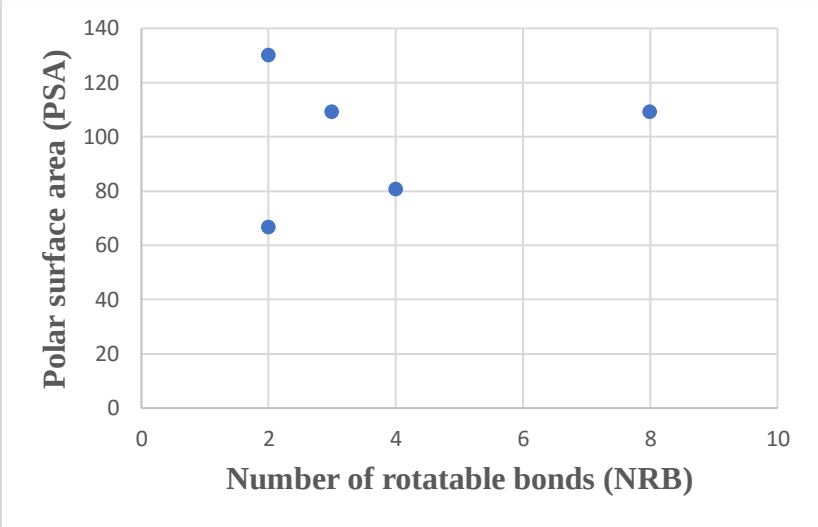
All inhibitors in this study agreed with the Lipinski's rule of five (ROF), except for inhibitor (2). The number of atoms that donate hydrogen atoms to form hydrogen bonds (HBD) was 6 which is higher than the acceptable value (Fig 3.22).

According to Veber's and MDDR Rules, all values for inhibitors in this study are consistent with these Rules except for inhibitor (2) which has NOR value of one (Fig 3.22). This value is lower than the acceptable value (NOR=3).









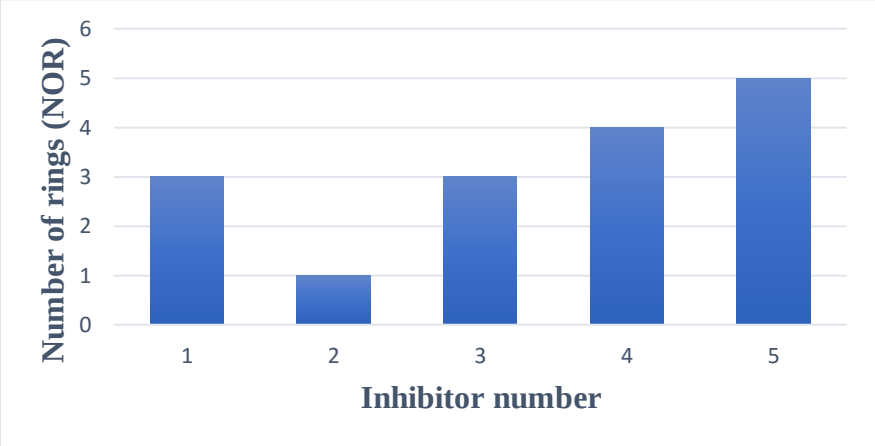
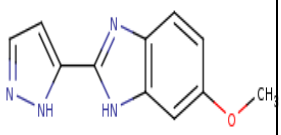
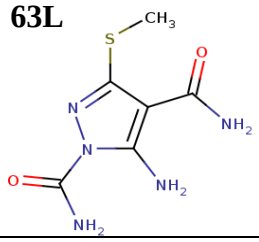
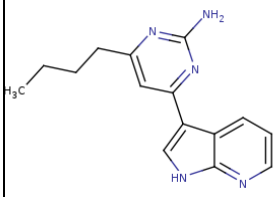
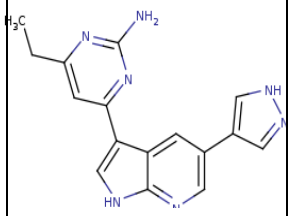
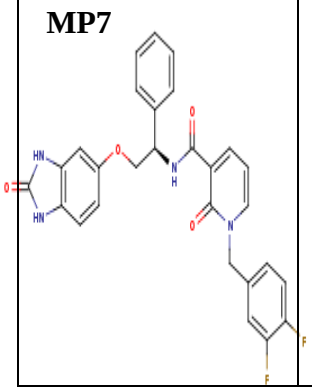


Figure 3.22: Analysis of the inhibitors according to Lipinski's Rule of five, Veber and MDDR Rules

Table 3.9: **Analysis of the inhibitors according to Lipinski's Rule of five, Veber and MDDR Rules**

Inhibitor	Molecular weight (g/mol)	A^{70} logP	PSA^{70} ($Å^2$)	Num-H acceptor Atoms	Num-H donar Atoms	Num-of rings	Number of Rotatable Bond	Ligand efficiency
62O 	214.23	1.84	66.60	5	2	3	2	0.33
63L 	215.24	-1.26	130.04	7	6	1	2	0.45
MOL 	267.34	2.81	80.49	5	3	3	4	0.45
61 	305.35	2.00	109.17	7	4	4	3	0.47
MP7 	516.50	3.70	108.99	8	3	5	8	0.32

In addition to the previous parameters that determine the drug-like properties, the ligand efficiency is an essential common metric to assess the drug-like quality of a compound[71]. This is estimated by relating binding free energy to the number of heavy atoms in a molecule **(LE= -ΔG/ HA).**[72] Therefore, the resulting ligand efficiency tends to be maximal for small molecules (e.g. fragments) and then steadily decreases as heavier atoms are added. The LE value for a small molecule that inhibit protein-protein interaction is a round 0.24, whereas LE is equal to 0.3 or higher is a desired value.

It is observed that all ligand efficiency values of our inhibitors were higher than 0.3. So we can conclude that all our inhibitors except inhibitor (2) have the drug-like properties when appling Lipinski's Rule of five, Veber Rule and MDDR Rule.

There is a good correlation between the difference of Gibbs free energy (ΔG) calculated by MM-GBSA and the molecular weight, the value of correlation coefficient of determination is $R^2 = 0.94$ (Figure 3.23).

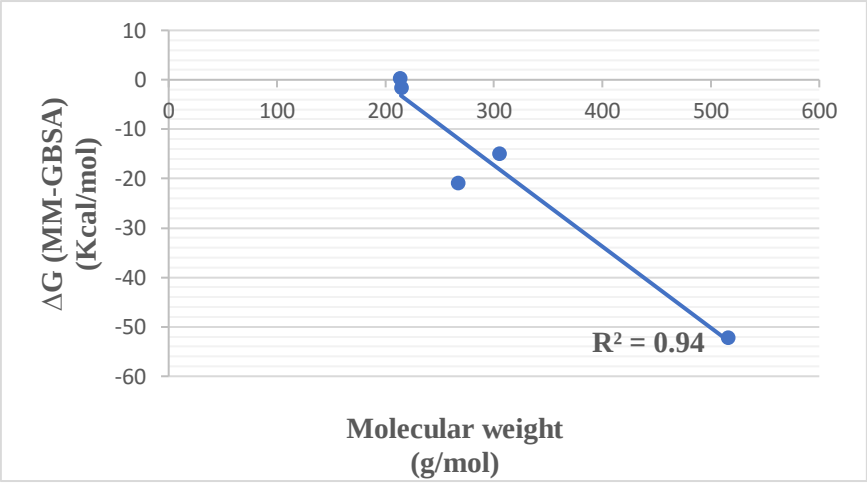


Figure 3.23: Relationship between the difference of Gibbs free energy (ΔG) calculated by MM-GBSA and the molecular weight

Also a good correlation between the difference of Gibbs free energy (ΔG) calculated by MM-GBSA and the lipophilicity property of inhibitor (logP), the value of coefficient of determination is $R^2 = 0.54$ (Figure 3.24).

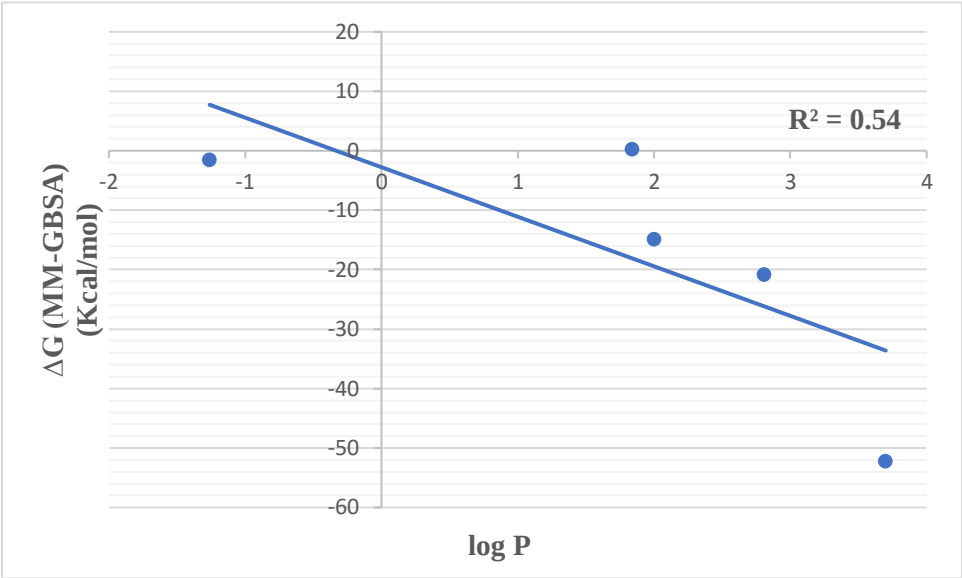


Figure 3.24: Relationship between the difference of Gibbs free energy (ΔG) calculated by MM-GBSA and log P

To the contrary, the correlation between the binding free energy (ΔG) calculated by MM-GBSA and the polar surace area (PSA), is weaker (Figure 3.25). The coefficient of determination in this case equals 0.02.

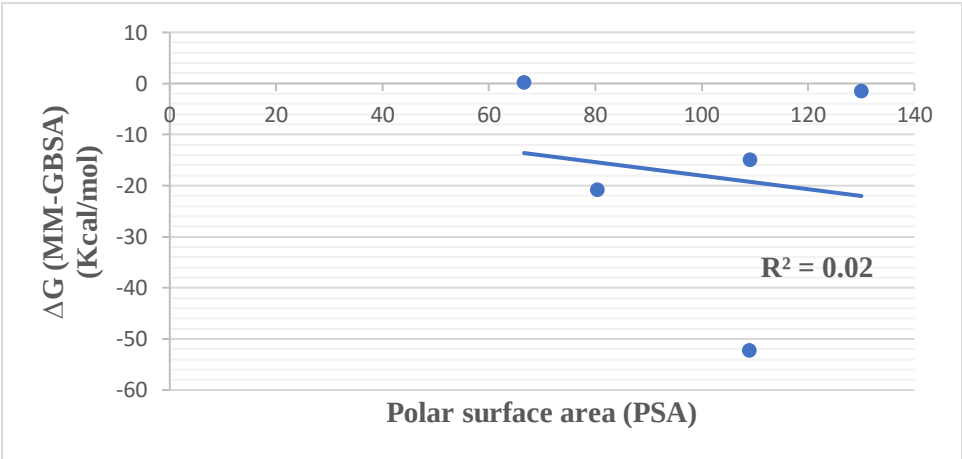


Figure 3.25: Relationship between the difference of Gibbs free energy (ΔG) calculated by MM-GBSA and PSA

3.13 Potency and selectivity of Inhibitor (5)

Inhibitor (5) made six strong hydrogen bonds with the PIF/Phosphate pocket of PDK-1 kinase with no water mediated hydrogen bond interaction; this interaction reflects the high affinity (ΔG_{bind} = -52.3 Kcal/mol), which indicates high potency of this inhibitor.

The high sequence similarity in the ATP binding pocket between different kinases is a major challenge for developing inhibitors that are specific for one or a small number of kinases.

Inhibitor (5) is consider as type II inhibitors, which occupied PIF/phosphate pocket of PDK-1 kinase. This allosteric site in protein kinase are especially used for developing more selective inhibitors[73]. This offers a possibility to develop more compounds with higher selectivity more than in the case of classical ATP-competitive inhibitors.[64]

Deep-pocket binder molecules, when used as single substrate, can be classified as substrate-selective PDK-1 inhibitors. When used in combination with ATP-competitive inhibitors they tend to suppress the activation of the downstream kinases.[73]

4. CONCLUSION

Molecular dynamics simulation was employed to identify an exquisitely potent PDK-1 inhibitor 5 (1-(3,4-difluorobenzyl)-2-oxo-N-{(1R)-2-[(2- oxo-2,3 dihydro-1H-benzimidazol-5-yl)oxy]-1-phenylethyl}-1,2-dihydropyridine carboxamide) that uniquely binds to the inactive kinase conformation.

This inhibitor is tightly bound to PDK-1 through five strong hydrogen bonds with the PIF/Phosphate pocket of PDK-1 kinase with no water mediated hydrogen bond interactions. This interaction reflects the high affinity of drug to receptor (ΔG_{bind} = -52.3 Kcal/mol).

In contrast to compounds 1-4, which are classical ATP-competitive kinase inhibitors (DFG-in) which are 6-methoxy-2-(1H-pyrazol-5-yl)-1H-benzimidazole (inhibitor 1), 4-dicarboxylicacid diamide (inhibitor 2), 4-butyl-6-(1H-pyrrolo[2,3-b]pyridin-3-yl)pyrimidin- 2-amine (inhibitor 3), 4-ethyl-6-[5-(1H-pyrazol-4-yl)-1H-pyrrolo[2,3- b]pyridin-3-yl]pyrimidin-2-amine (inhibitor 4).

MM-PBSA and MM-GBSA both methods are used in this work to calculate the binding free energies of PDK-1 kinase with four inhibitors. There is a good correlation between binding free energy which was calculated by MM-GBSA and experimental values of binding free energy which are derived from the experimental

reported IC_{50} values (R^2 =0.55). To the contrary, the correlation between the binding free energy was calculated by MM-PBSA and experimental values of binding free energy which are derived from the experimental reported IC_{50} values is weaker (R^2 =0.14). This result agrees with some of the reports that MM-GBSA is considered a better approach than the MM-PBSA in calculating the binding free energies when metals are not involved.

According to IC_{50} values we conclude that inhibitor (1), inhibitor (2), inhibitor (3), and inhibitor (4) do not possess a drug-like property (IC_{50} = 93 µM, 17 µM, 1.1 µM, 0.013 µM), respectively, because they dissociate and do not stay bound to the enzyme.

And these values are correlated with binding free energies of these complexes that calculated by MM-GBSA which are 0.3 K cal/mol, -1.6 Kcal/mol, -21.0 Kcal/mol, -15.0 Kcal/mol, -52.3 Kcal/mol. We noted that IC_{50} become lower, and binding free energy become larger in negative sign, and become more druggable property.

As PDK-1 is a well validated anticancer target, the final results reveal the binding modes between PDK-1 kinase and the five inhibitors which can be used in the future in drug design for cancer treatment. The position of water molecules in the binding sites of inhibitor (2)-kinase and inhibitor (3)-kinase

complexes can be used to design better inhibitors in which the principle lies in the fact that a substituent is added to the ligand that displaces a bound water molecule based on the creation of new inhibitor that includes a structural water mimic.

References

[1] Hughes, J. P.; Rees, S. S.; Kalindjian, S. B.; Philpott, K. L. Br. *J. Pharmacol*. **2011**, *162* (6), 1239–1249.

[2] Kapetanovic. I. Chem Biol Interact. **2008**, 171(2), 165–176.

[3] Drews, J. *Science*. **2000**, *287* (5460), 1960–1964.

[4] Owens, *J. Nat. Rev*. *Drug Discov.* **2007**, *6* (2), 99–101.

[5] Veselovsky, V; Ivanov, S. *Curr*. *Drug Targets. Infect. Disord*. **2003**, *3* (1), 33–40.

[6] Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe Jr., E. W. *Pharmacol Rev* **2014**, *66* (1), 334–395.

[7] Wilson, G. L.; Lill, M. A. *Future Med. Chem*. **2011**, *3* (6), 735–750.

[8] Manning, G. *Science* **2002**, *298* (5600), 1912–1934.

[9] Hirpara, K. V; Aggarwal, P.; Mukherjee, A. J.; Joshi, N.; Burman, A. C. *Anticancer*. *Agents Med. Chem***. 2009**, *9* (2), 138–161.

[10] Shchemelinin, I.; Sefc, L.; Necas, E. *Folia Biol. (Praha).* **2006**, *52* (3), 81–101.

[11] Giamas, G.; Stebbing, J.; Vorgias, C. E.; Knippschild, U. *Pharmacogenomics* **2007**, *8*, 1005–1016.

[12] Mikalsen, T.; Gerits, N.; Moens, U. *Biotechnol*. *Annu. Rev*. **2006**, *12* (6), 153–223.

[13] Shapiro. C.; Recht. A. *N Engl J Med* **2001**, *344* (26), 1997-2008.

[14] Garcia, M.; Jemal, A.; Ward, E. M.; Center, M. M.; Hao, Y.; Siegel, R. L.; Thun, M. *American Cancer Society*, **2007**, *18*

[15] Scheid, M. P.; Parsons, M.; Woodgett, J. R. *Mol Cell Biol* **2005**, *25* (6), 2347–2363.

[16] Mora, A.; Komander, D.; Van Aalten, D. M. F.; Alessi, D. R. *Semin. Cell Dev. Biol*. **2004**, *15* (2), 161–170.

[17] Biondi, R. M.; Komander, D.; Thomas, C. C.; Lizcano, J. M.; Deak, M.; Alessi, D. R.; Van Aalten, D. M. F. *EMBO J*. **2002**, *21* (16), 4219–4228.

[18] Vulpetti, A.; Bosotti, R. *Farmaco* **2004**, *59* (10), 759–765.

[19] Zeng, H.; Cao, R.; Zhang, H. J. *Chem. Inf. Model* **2008**, *48* (9), 1760–1772.

[20] Miyazawa, K. J. *Biochem*. **2011**, *150* (1), 1–3.

[21] Bain, J.; Plater, L.; Elliott, M.; Shpiro, N.; Hastie, C. J.; Mclauchlan, H.; Klevernic, I.; Arthur, J. S. C.; Alessi, D. R.; Cohen, P. *Biochem. J* **2007**, *408* (3), 297–315.

[22] Wucherer-Plietker, M.; Merkul, E.; Müller, T. J. J.; Esdar, C.; Knöchel, T.; Heinrich, T.; Buchstaller, H. P.; Greiner, H.; Dorsch, D.; Finsinger, D.; Calderini, M.; Bruge, D.; Grädler, U. *Bioorganic Med. Chem. Lett*. **2016**, *26* (13), 3073–3080.

[23] Feldman, R. I. *J. Biol. Chem***. 2005**, *280* (20), 19867–19874.

[24] Singh, S.; Srivastava, P. *Comput Mol Biosci* **2015**, *5*, 20–33.

[25] Ong, K. and Khoo, H. *Gen. Pharmac* **1997**, *29* (2), 121-126.

[26] Peifer, C.; Alessi, D. R. *ChemMedChem* **2008**, *3* (12), 1810–1838.

[27] Blanc, J.; Geney, R.; Menet, C. *Anticancer Agents Med Chem*. **2013**, *13*, 731.

[28] Backes, A. C.; Zech, B.; Felber, B.; Klebl, B.; Müller, G. *Expert Opin. Drug Discov*. **2008**, *3*, 1409-1425.

[29] Wu, P.; Nielsen, T. E.; Clausen, M. H. *Drug Discov. Today* **2016**, *21* (1), 5–10.

[30] Leeson, P.; Oncology, H.; Hospital, C. *Nature* **1912**, *481*, 455–456.

[31] Nguyen, T.; Lee, S.; Wang, H.-K.; Chen, H.-Y.; Wu, Y.-T.; Lin, S.; Kim, D.-W.; Kim, D. *Molecules* **2013**, *18* (12), 15600–15612.

[32] Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev*. **1997**, *23*, 3–25.

[33] Veber, D. F. *J. Med. Chem*. **2002**, *45*, 2615–2623.

[34] Oprea, T. I. *J. Comput. Aided. Mol. Des*. **2000**, *14* (3), 251–264.

[35] Swanson, J. M. J.; Henchman, R. H.; McCammon, J. A. *Biophys. J* **2004**, *86* (1), 67–74.

[36] Lu, N.; Kofke, D. A. *J. Chem. Phys*. **2001**, *115* (15), 6866–6875.

[37] Synthetic, S. O. F. Molecular Dynamics - Studies of Synthetic and Biological Macromolecules; **2012**.

[38] Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. *J. Chem. Theory Comput*. **2012**, *8* (9), 3314–3321.

[39] Homeyer, N.; Gohlke, H. *Mol. Inform*. **2012**, *31* (2), 114–122.

[40] Honig, B.; Nicholls, A. *Science* **1995**, *268* (5214), 1144–1149.

[41] Hermansson, A. Calculating Ligand-Protein Binding Energies from Molecular Dynamics Simulations - Thesis in Physical Chemistry; **2015**.

[42] Hou, T.; Wang, J.; Li, Y.; Wang, W. *J. Chem. Inf. Model*. **2011**, *51*, 69–82.

[43] Casamayor, A. Morrice, A. Alessi, R. *Biochem. J* **1999**, *342* (2), 287.

[44] Lingenheil, M.; Denschlag, R.; Reichold, R.; Tavan, P. *J. Chem. Theory Comput* **2008**, *4*, 1293–1306.

[45] Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys* **1977**, *23* (3), 327–341.

[46] Genheden, S.; Kuhn, O.; Mikulskis, P.; Ryde, U. *J. Chem. Inf. Model* **2012**, *52*, 2079-2088.

[47] Nurisso, A.; Daina, A.; Walker, R. C. A practical introduction to molecular dynamics simulations: applications to homology modeling; **2012***; 857*.

[48] Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graph. Model*. **2006**, *25* (2), 247–260.

[49] Costanzi, S.; Tikhonova, I. G.; Harden, T. K.; Jacobson, K. A. *J. Comput. Aided. Mol. Des.* **2009**, *23* (11), 747–754.

[50] Hou, T.; Wang, J.; Li, Y.; Wang, W. *J. Chem. Inf. Model* **2011**, *51*, 69–82.

[51] Liu, J.; He, X.; Zhang, J. Z. H. *J. Mol. Model*. **2014**, *20* (10), 2451.

[52] Zhu, J.; Huang, J. W.; Tseng, P. H.; Yang, Y. T.; Fowble, J.; Shiau, C. W.; Shaw, Y. J.; Kulp, S. K.; Chen, C. S. *Cancer Res*. **2004**, *64* (12), 4309–4318.

[53] Cheng, Y.; Prusoff, W. H. *Biochem. Pharmacol* **1973**, *22* (23), 3099–3108.

[54] Alzate-Morales, J. H.; Contreras, R.; Soriano, A.; Tuñon, I.; Silla, E. *Biophys. J*. **2007**, *92* (2), 430–439.

[55] Kuhn, B.; Mohr, P.; Stahl, M. *J. Med. Chem* **2010**, *53* (6), 2601–2611.

[56] Furet, P.; Caravatti, G.; Guagnano, V.; Lang, M.; Meyer, T.; Schoepfer, *J. Bioorganic Med. Chem. Lett.* **2008**, *18* (3), 897–900.

[57] Zhao, B.; Lehr, R.; Smallwood, A. M.; Ho, T. F.; Maley, K.; Randall, T.; Head, M. S.; Koretke, K. K.; Schnackenberg, C. G. *Protein Sci.* **2007**, *16* (12), 2761–2769.

[58] Medina, J. R.; Blackledge, C. W.; Heerding, D. A.; Campobasso, N.; Ward, P.; Briand, J.; Wright, L.; Axten, J. M. *ACS Med. Chem. Lett.* **2010**, *1* (8), 439–442.

[59] Levinson, N. M.; Boxer, S. G. *Nat. Chem. Biol*. **2013**, *10* (2), 127–132.

[60] Dunitz, J. D. *Science*, **1994**, *264*, 670.

[61] Hamelberg, D.; McCammon, J. A. *J. Am. Chem. Soc*. **2004**, *126* (24), 7683–7689.

[62] Lam, P. Y. S.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; Chang, C.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R.; Erickson-viitanen, S. *Science*, **1994**, *263*, 380–384.

[63] Nagashima, K.; Shumway, S. D.; Sathyanarayanan, S.; Chen, A. H.; Dolinski, B.; Xu, Y.; Keilhack, H.; Nguyen, T.; Wiznerowicz, M.; Li, L.; Lutterbach, B. A.; Chi, A.; Paweletz, C.; Allison, T.; Yan, Y.; Munshi, S. K.; Klippel, A.; Kraus, M.; Bobkova, E. V.; Deshmukh, S.; Xu, Z.; Mueller, U.; Szewczak, A. A.; Pan, B. S.; Richon, V.; Pollock, R.; Blume-Jensen, P.; Northrup, A.; Andersen, J. N. *J. Biol. Chem.* **2011**, *286* (8), 6433–6448.

[64] Kappe, C. O.; Dallinger, D.; Mannhold, E. R.; Kubinyi, H.; Folkers, G.; Allerton, D. a; Walker, H.; Don, K.; Lira, M. J.; Leurs, S. a. Protein–Protein Interactions in drug discovery. **2012**.

[65] Stroud, R. M.; Finer-Moore. J. Computational and Structural Approaches to Drug Discovery Ligand–Protein Interactions. **2008**.

[66] Meanwell, N. A. *Chem. Res. Toxicol*. **2016**, *29*, 564–616.

[67] Barratt, E.; Bingham, R. J.; Warner, D. J.; Laughton, C. A.; Phillips, S. E. V; Homans, S. W. *J. Am. Chem. Soc*. **2005**, *127*, 11827–11834.

[68] Backes. A.; Zech. B.; Felber. B.; Klebl. B.; Müller. G. *Expert Opin. Drug Discov*. **2008,** *3*, 1409-1425

[69] Blanc. J.; Geney. R.; Menet. C. *Anti-Cancer Agents in Medicinal Chemistry*, **2013**, *13*, 731-747.

[70] Ertl, P.; Rohde, B.; Selzer, P. *J. Med. Chem***. 2000**, *43*, 3714-3717.

[71] Arkin, M.; Tang, Y.; Wells, *J. Chem. Biol*. **2014**, *21*, 1102.

[72] Schultes. S.; Graaf. C.; Haaksma. E.; Esch. I.; Leurs. R.; Kramer. O. *Drug Discov. Today. Technol*. **2010**, *7*, 157.

[73] Rettenmaiera, T.; Sadowskyb, J.; Thomsenb, N.; Chenc, S.; Doakb, A.; Arkinc, M.; Wellsb, J. *Proc. Natl. Acad. Sci*. **2014**, *111*, 18590.

[74] Wang.J.; Wolf. R.; Caldwell. J.; Kollman. P.; Case. D. *J Comput Chem* **2004**, *25*, 1157-1174.

[75] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys* **1983**, *79* (2), 926–935.

APPENDICES

APPENDIX A: SEP NON-STANDARD RESIDUE AND INHIBITOR FILES

Input file for force filed modification of SEP residue

File 1: SEP_leap.frcmod

```
From VanBeek et al. Biophys J. (2007) 92, 4168-4178
MASS


BOND


ANGLE
OH-P-OH    45.000   109.500


DIHE


IMPROPER

NONBON
```

Input file for identification atom types and atom charges of inhibitor (1)

File 2: Inhibitor1.mol2

```
62O

  26    28     1     0     0

SMALL

bcc

@<TRIPOS>ATOM

     1 C1        31.4290   24.4900    5.9150 ca        1 62O     -0.035200

     2 C2        32.8340   24.3570    5.9830 ca        1 62O      0.031400

     3 N3        31.1770   25.7430    5.4030 na        1 62O     -0.283300

     4 C4        30.6050   23.4550    6.3540 ca        1 62O     -0.221000

     5 N5        33.3580   25.5140    5.4990 nc        1 62O     -0.525100

     6 C6        33.3920   23.1590    6.4750 ca        1 62O     -0.035000

     7 C7        32.3840   26.3250    5.1600 cd        1 62O      0.468400

     8 C8        31.1650   22.2790    6.8370 ca        1 62O      0.150100
```

Input file for force filed modification of inhibitor (1)

File 3: Inhibitor1.frcmod

```
remark goes here
MASS
BOND
ANGLE
DIHE
IMPROPER
ca-ca-ca-na    1.1    180.0    2.0   Using default value
ca-ca-ca-nc    1.1    180.0    2.0   Using default value
ca-cd-na-hn    1.1    180.0    2.0   General improper torsional angle (2 general atom types)
ca-ca-ca-ha    1.1    180.0    2.0   General improper torsional angle (2 general atom types)
cd-na-cd-nc    1.1    180.0    2.0   Using default value
ca-ca-ca-os    1.1    180.0    2.0   Using default value
cc-cd-cd-na    1.1    180.0    2.0   Using default value
cc-cd-cc-ha    1.1    180.0    2.0   Using default value
cd-hn-na-nd    1.1    180.0    2.0   General improper torsional angle (2 general atom types)
cc-h4-cc-nd    1.1    180.0    2.0   Using default value
NONBON
```

APPENDIX B: INPUT FILES FOR SIMULATION

Input file for minimization of water and ion molecules in inhibitor (1) -protein complex

File 1: min.in

```
Minimization of water
&cntrl
  imin=1,maxcyc=1000,ncyc=500,
  cut=10.0,ntb=1,
  ntc=2,ntf=2,
  ntpr=100,
  ntr=1, restraintmask=':1-282',
  restraint_wt=2.0
 /
Hold protein and ligand fixed
10.0
RES 1-283
END
END
```

Input file for minimization of the whole complex

File 2: min_all.in

```
Minimization of the whole system
 &cntrl
  imin=1,maxcyc=1000,ncyc=500,
  cut=10.0,ntb=1,
  ntc=2,ntf=2,
  ntpr=100,
  ntr=0,
 /
END
```

Input file for heating the inhibitor (1) -protein complex from 0K to 300K

File 3: eat.in

```
Heating from 0K to 300K with weak restraints
 &cntrl
  imin=0,irest=0,ntx=1,
  nstlim=25000,dt=0.002,
  ntc=2,ntf=2,
  cut=10.0, ntb=1,
  ntpr=500, ntwx=500,
  ntt=3, gamma_ln=2.0,
  tempi=0.0, temp0=300.0, ig=-1,
  ntr=1, restraintmask=':1-282',
  restraint_wt=2.0,
  nmropt=1
 /
 &wt TYPE='TEMP0', istep1=0, istep2=25000,
  value1=0.1, value2=300.0, /
 &wt TYPE='END' /
```

Input file for density equilibration of inhibitor (1) -protein complex

File 4: density.in

```
&cntrl
 imin=0,irest=1,ntx=5,
 nstlim=25000,dt=0.002,
 ntc=2,ntf=2,
 cut=10.0, ntb=2, ntp=1, taup=1.0,
 ntpr=500, ntwx=500,
 ntt=3, gamma_ln=2.0,
 temp0=300.0, ig=-1,
 ntr=1, restraintmask=':1-282',
 restraint_wt=2.0,
/
```

Input file for unrestrained equilibration of inhibitor (1) -protein complex

File 5: equil.in

```
&cntrl
 imin=0,irest=1,ntx=5,
 nstlim=250000,dt=0.002,
 ntc=2,ntf=2,
 cut=10.0, ntb=2, ntp=1, taup=2.0,
 ntpr=1000, ntwx=1000,
 ntt=3, gamma_ln=2.0,
 temp0=300.0, ig=-1,
/
```

Input file for unrestrained production of inhibitor (1) -protein complex

File 6: prod.in

```
&cntrl
 imin=0,irest=1,ntx=5,
 nstlim=250000,dt=0.002,
 ntc=2,ntf=2,
 cut=10.0, ntb=2, ntp=1, taup=2.0,
 ntpr=5000, ntwx=5000,
 ntt=3, gamma_ln=2.0,
 temp0=300.0, ig=-1,
/
```

Input file for running MM-PBSA and MM-GBSA

File 7: mmpbsa.in

```
Input file for running PB and GB
&general
  endframe=50, verbose=1,
#  entropy=1,
/
&gb
 igb=2, saltcon=0.100
/
&pb
 istrng=0.100,
/
```

Input file for running entropy calculations using Nmode

File 8: mmpbsa_nm.in

```
Input file for running entropy calculations using NMode
&general
  endframe=50, keep_files=2,
/
&nmode
  nmstartframe=5, nmendframe=45,
  nminterval=5, nmode_igb=1, nmode_istrng=0.1,
/
```

Input file for running mass-weighted RMSD measurements

File 9: mesure_equil_rmsd.ptraj

```
trajin equil.mdcrd

reference com_wat.inpcrd

rms reference out equil.rmsd @CA,C,N 0.1
```

Input file for running conversion from mdcrd file to binpos file

File 10: mdcrd_to_binpos.ptraj

```
trajin prod1.mdcrd
trajin prod2.mdcrd
trajin prod3.mdcrd
trajin prod4.mdcrd
trajout prod.binpos binpos
```

Input file for hydrogen bonding analysis of inhibitor (1) -protein complex

File 11: analyse_hbond.ptraj

```
trajin prod.binpos

hbond :1-274 out nhb.dat avgout avghb.dat
```

APPENDIX C: OUTPUT FILES

Output file of ΔG_{bind} for inhibitor (1)-protein complex that resulted from MM-GBSA

File 1: MM-GBSA for ΔG_{bind} protein-inhibitor (1) complex

```
GENERALIZED BORN:

Complex:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2387.1522              18.1991              2.5737
EEL                      -19906.0747              28.5435              4.0367
EGB                       -4475.4622              20.5100              2.9006
ESURF                        88.0450               0.4429              0.0626

G gas                    -22293.2270              34.4401              4.8706
G solv                    -4387.4172              20.4611              2.8936

TOTAL                    -26680.6441              32.3518              4.5752

Receptor:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2354.6887              18.1946              2.5731
EEL                      -19888.1698              28.4029              4.0168
EGB                       -4477.9034              20.2762              2.8675
ESURF                        89.4199               0.4413              0.0624

G gas                    -22242.8585              34.5269              4.8828
G solv                    -4388.4835              20.2199              2.8595

TOTAL                    -26631.3420              32.4518              4.5894


Ligand:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                      -1.8136               0.2344              0.0331
EEL                         -12.6183               0.5191              0.0734
EGB                         -22.0954               0.4232              0.0598
ESURF                         2.1777               0.0136              0.0019

G gas                       -14.4319               0.4936              0.0698
G solv                      -19.9177               0.4177              0.0591

TOTAL                       -34.3496               0.5901              0.0835


Differences (Complex - Receptor - Ligand):
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                     -30.6500               1.0904              0.1542
EEL                          -5.2866               1.6542              0.2339
EGB                          24.5367               1.0683              0.1511
ESURF                        -3.5526               0.0346              0.0049

DELTA G gas                 -35.9366               1.8382              0.2600
DELTA G solv                 20.9841               1.0689              0.1512

DELTA TOTAL                 -14.9525               1.6290              0.2304
```

Output file of ΔG_{bind} for inhibitor (2)-protein complex that resulted from MM-PBSA

File 2: MM-PBSA for ΔG_{bind} protein-inhibitor (1) complex

```
POISSON BOLTZMANN:

Complex:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2387.1522              18.1991              2.5737
EEL                      -19906.0747              28.5435              4.0367
EPB                       -4075.2673              19.1133              2.7030
ENPOLAR                    2236.9806               3.0541              0.4319
EDISPER                   -1309.5224               2.5485              0.3604

G gas                    -22293.2270              34.4401              4.8706
G solv                    -3147.8090              19.3236              2.7328

TOTAL                    -25441.0360              31.8761              4.5080
Receptor:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2354.6887              18.1946              2.5731
EEL                      -19888.1698              28.4029              4.0168
EPB                       -4083.8299              19.0226              2.6902
ENPOLAR                    2231.3477               3.0596              0.4327
EDISPER                   -1318.4636               2.5071              0.3546

G gas                    -22242.8585              34.5269              4.8828
G solv                    -3170.9459              19.3747              2.7400

TOTAL                    -25413.8044              31.9662              4.5207
Ligand:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                      -1.8136               0.2344              0.0331
EEL                         -12.6183               0.5191              0.0734
EPB                         -21.0535               0.3862              0.0546
ENPOLAR                      25.0435               0.1278              0.0181
EDISPER                     -27.6378               0.1729              0.0245

G gas                       -14.4319               0.4936              0.0698
G solv                      -23.6478               0.4276              0.0605

TOTAL                       -38.0796               0.5793              0.0819
Differences (Complex - Receptor - Ligand):
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                     -30.6500               1.0904              0.1542
EEL                          -5.2866               1.6542              0.2339
EPB                          29.6162               1.0069              0.1424
ENPOLAR                     -19.4106               0.2120              0.0300
EDISPER                      36.5790               0.3302              0.0467

DELTA G gas                 -35.9366               1.8382              0.2600
DELTA G solv                 46.7846               1.1997              0.1697

DELTA TOTAL                  10.8480               1.6722              0.2365
```

Output file of ΔG_{bind} for inhibitor (2)-protein complex that resulted from MM-GBSA

File 3: MM-GBSA for ΔG_{bind} protein-inhibitor (2) complex

```
GENERALIZED BORN:

Complex:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2382.0840              19.2872              2.7276
EEL                      -19827.2949              40.0446              5.6632
EGB                       -4408.0595              25.1368              3.5549
ESURF                        87.6976               0.4609              0.0652

G gas                    -22209.3789              35.4489              5.0132
G solv                    -4320.3619              24.9850              3.5334

TOTAL                    -26529.7408              27.4538              3.8825

Receptor:
Energy Component            Average             Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2357.5888              19.7084              2.7872
EEL                      -19954.9192              40.0431              5.6629
EGB                       -4425.5261              25.1074              3.5507
ESURF                        89.3861               0.4567              0.0646

G gas                    -22312.5080              35.2631              4.9869
G solv                    -4336.1399              24.9597              3.5298

TOTAL                    -26648.6480              27.4911              3.8878


Ligand:
Energy Component            Average             Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                       1.2240               1.0341              0.1463
EEL                         158.6175               2.1640              0.3060
EGB                         -23.0458               0.9649              0.1365
ESURF                         2.4317               0.0103              0.0015

G gas                       159.8415               2.6065              0.3686
G solv                      -20.6142               0.9607              0.1359

TOTAL                       139.2273               2.6542              0.3754


Differences (Complex - Receptor - Ligand):
Energy Component            Average             Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                     -25.7192               2.1653              0.3062
EEL                         -30.9932               3.1003              0.4384
EGB                          40.5125               2.1040              0.2976
ESURF                        -4.1202               0.0338              0.0048

DELTA G gas                 -56.7124               2.9103              0.4116
DELTA G solv                 36.3922               2.0901              0.2956

DELTA TOTAL                 -20.3201               1.9515              0.2760
```

Output file of ΔG_{bind} for inhibitor (2)-protein complex that resulted from MM-PBSA

File 4: MM-PBSA for ΔG_{bind} protein-inhibitor (2) complex

```
POISSON BOLTZMANN:

Complex:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2382.0840              19.2872              2.7276
EEL                      -19827.2949              40.0446              5.6632
EPB                       -4033.4629              22.8552              3.2322
ENPOLAR                    2225.5414               3.0962              0.4379
EDISPER                   -1297.4909               2.5382              0.3590

G gas                    -22209.3789              35.4489              5.0132
G solv                    -3105.4124              22.5362              3.1871

TOTAL                    -25314.7913              26.9355              3.8092

Receptor:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2357.5888              19.7084              2.7872
EEL                      -19954.9192              40.0431              5.6629
EPB                       -4053.8735              22.9975              3.2523
ENPOLAR                    2223.0370               3.0854              0.4363
EDISPER                   -1307.2618               2.5773              0.3645

G gas                    -22312.5080              35.2631              4.9869
G solv                    -3138.0983              22.6852              3.2082

TOTAL                    -25450.6063              27.0122              3.8201

Ligand:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                       1.2240               1.0341              0.1463
EEL                         158.6175               2.1640              0.3060
EPB                         -21.0045               0.8966              0.1268
ENPOLAR                      21.9338               0.1046              0.0148
EDISPER                     -24.3504               0.1600              0.0226

G gas                       159.8415               2.6065              0.3686
G solv                      -23.4211               0.9001              0.1273

TOTAL                       136.4204               2.5309              0.3579

Differences (Complex - Receptor - Ligand):
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                     -25.7192               2.1653              0.3062
EEL                         -30.9932               3.1003              0.4384
EPB                          41.4151               2.2681              0.3208
ENPOLAR                     -19.4294               0.2129              0.0301
EDISPER                      34.1212               0.3239              0.0458

DELTA G gas                 -56.7124               2.9103              0.4116
DELTA G solv                 56.1069               2.2933              0.3243

DELTA TOTAL                  -0.6054               1.9526              0.2761
```

Output file of ΔG_{bind} for inhibitor (3)-protein complex that resulted from MM-GBSA

File 5: MM-GBSA for ΔG_{bind} protein-inhibitor (3) complex

```
GENERALIZED BORN:

Complex:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2413.2870              21.5813              3.0521
EEL                      -19820.4055              30.7686              4.3513
EGB                       -4419.2772              22.1522              3.1328
ESURF                        83.6140               0.5858              0.0828

G gas                    -22233.6925              34.3084              4.8519
G solv                    -4335.6633              22.0179              3.1138

TOTAL                    -26569.3558              30.8426              4.3618

Receptor:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2371.3691              21.3473              3.0190
EEL                      -19879.6261              30.4154              4.3014
EGB                       -4435.8541              21.8024              3.0833
ESURF                        86.0346               0.5906              0.0835

G gas                    -22250.9952              34.4450              4.8713
G solv                    -4349.8194              21.6557              3.0626

TOTAL                    -26600.8147              31.0148              4.3861


Ligand:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                      -1.6425               0.7487              0.1059
EEL                          84.7296               1.3193              0.1866
EGB                         -21.0097               0.6344              0.0897
ESURF                         2.9199               0.0130              0.0018

G gas                        83.0871               1.5013              0.2123
G solv                      -18.0898               0.6318              0.0893

TOTAL                        64.9973               1.8088              0.2558


Differences (Complex - Receptor - Ligand):
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                     -40.2754               1.5773              0.2231
EEL                         -25.5090               2.5000              0.3536
EGB                          37.5866               2.4295              0.3436
ESURF                        -5.3405               0.0522              0.0074

DELTA G gas                 -65.7844               2.5661              0.3629
DELTA G solv                 32.2460               2.4181              0.3420

DELTA TOTAL                 -33.5384               1.4634              0.2070
```

Output file of ΔG_{bind} for inhibitor (3)-protein complex that resulted from MM-PBSA

File 6: MM-PBSA for ΔG_{bind} protein-inhibitor (3) complex

```
POISSON BOLTZMANN:

Complex:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                  -2413.2870              21.5813              3.0521
EEL                     -19820.4055              30.7686              4.3513
EPB                      -3993.3020              21.3042              3.0129
ENPOLAR                   2204.6542               4.0279              0.5696
EDISPER                  -1269.9511               3.1058              0.4392

G gas                   -22233.6925              34.3084              4.8519
G solv                   -3058.5989              22.0081              3.1124

TOTAL                   -25292.2914              30.6631              4.3364

Receptor:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                  -2371.3691              21.3473              3.0190
EEL                     -19879.6261              30.4154              4.3014
EPB                      -4025.4893              21.1439              2.9902
ENPOLAR                   2199.6821               4.0331              0.5704
EDISPER                  -1287.5893               3.0903              0.4370

G gas                   -22250.9952              34.4450              4.8713
G solv                   -3113.3965              21.6740              3.0652

TOTAL                   -25364.3917              31.0895              4.3967

Ligand:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                     -1.6425               0.7487              0.1059
EEL                         84.7296               1.3193              0.1866
EPB                        -19.4857               0.6965              0.0985
ENPOLAR                     32.7201               0.1509              0.0213
EDISPER                    -35.1051               0.1973              0.0279

G gas                       83.0871               1.5013              0.2123
G solv                     -21.8707               0.7350              0.1040

TOTAL                       61.2164               1.6242              0.2297

Differences (Complex - Receptor - Ligand):
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                    -40.2754               1.5773              0.2231
EEL                        -25.5090               2.5000              0.3536
EPB                         51.6731               3.3871              0.4790
ENPOLAR                    -27.7480               0.2755              0.0390
EDISPER                     52.7433               0.3431              0.0485

DELTA G gas                -65.7844               2.5661              0.3629
DELTA G solv                76.6684               3.2720              0.4627
```

Output file of ΔG_{bind} for inhibitor (4)-protein complex that resulted from MM-GBSA

File 7: MM-GBSA for ΔG_{bind} protein-inhibitor (4) complex

```
GENERALIZED BORN:

Complex:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                  -2353.3994              17.5785              2.4860
EEL                     -19658.0219              34.9942              4.9489
EGB                      -4319.8457              18.9849              2.6849
ESURF                       81.1668               0.6750              0.0955

G gas                   -22011.4213              37.1507              5.2539
G solv                   -4238.6789              18.7732              2.6549

TOTAL                   -26250.1002              33.7015              4.7661


Receptor:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                  -2308.7545              17.7632              2.5121
EEL                     -19710.5088              35.4969              5.0200
EGB                      -4319.0185              19.2179              2.7178
ESURF                       83.4388               0.6664              0.0942

G gas                   -22019.2633              37.8085              5.3469
G solv                   -4235.5797              18.9902              2.6856

TOTAL                   -26254.8430              33.7533              4.7734


Ligand:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                     -1.8345               0.6602              0.0934
EEL                         68.3767               1.3554              0.1917
EGB                        -29.0700               0.5917              0.0837
ESURF                        2.8460               0.0161              0.0023

G gas                       66.5422               1.4134              0.1999
G solv                     -26.2240               0.5889              0.0833

TOTAL                       40.3182               1.5645              0.2213


Differences (Complex - Receptor - Ligand):
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                    -42.8103               1.8014              0.2548
EEL                        -15.8898               3.5751              0.5056
EGB                         28.2428               3.3158              0.4689
ESURF                       -5.1180               0.0605              0.0086

DELTA G gas                -58.7002               3.1095              0.4397
DELTA G solv                23.1248               3.2962              0.4662

DELTA TOTAL                -35.5753               1.7954              0.2539
```

Output file of ΔG_{bind} for inhibitor (4)-protein complex that resulted from MM-PBSA

File 8: MM-PBSA for ΔG_{bind} protein-inhibitor (4) complex

```
POISSON BOLTZMANN:

Complex:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2353.3994              17.5785              2.4860
EEL                      -19658.0219              34.9942              4.9489
EPB                       -3927.1281              17.4236              2.4641
ENPOLAR                    2158.3678               3.8127              0.5392
EDISPER                   -1224.9080               3.1128              0.4402

G gas                    -22011.4213              37.1507              5.2539
G solv                    -2993.6683              17.1491              2.4252

TOTAL                    -25005.0896              31.3041              4.4271
Receptor:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2308.7545              17.7632              2.5121
EEL                      -19710.5088              35.4969              5.0200
EPB                       -3936.5249              17.6422              2.4950
ENPOLAR                    2151.7913               3.8115              0.5390
EDISPER                   -1240.8397               3.1114              0.4400

G gas                    -22019.2633              37.8085              5.3469
G solv                    -3025.5734              17.3662              2.4559

TOTAL                    -25044.8367              31.2005              4.4124

Ligand:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                      -1.8345               0.6602              0.0934
EEL                          68.3767               1.3554              0.1917
EPB                         -26.9996               0.6359              0.0899
ENPOLAR                      33.9443               0.1751              0.0248
EDISPER                     -38.3396               0.1965              0.0278

G gas                        66.5422               1.4134              0.1999
G solv                      -31.3950               0.7515              0.1063

TOTAL                        35.1472               1.3986              0.1978

Differences (Complex - Receptor - Ligand):
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                     -42.8103               1.8014              0.2548
EEL                         -15.8898               3.5751              0.5056
EPB                          36.3964               3.3776              0.4777
ENPOLAR                     -27.3677               0.2681              0.0379
EDISPER                      54.2713               0.3780              0.0535

DELTA G gas                 -58.7002               3.1095              0.4397
DELTA G solv                 63.3000               3.4647              0.4900
```

Output file of ΔG_{bind} for inhibitor (5)-protein complex that resulted from MM-GBSA

File 9: MM-GBSA for ΔG_{bind} protein-inhibitor (5) complex

```
GENERALIZED BORN:

Complex:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2334.6961              13.3627              1.8898
EEL                      -20111.3512              40.3633              5.7082
EGB                       -4065.6798              22.2102              3.1410
ESURF                        88.9084               0.4125              0.0583

G gas                    -22446.0473              41.1968              5.8261
G solv                    -3976.7714              22.1986              3.1394

TOTAL                    -26422.8187              31.1430              4.4043

Receptor:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2257.1401              12.9660              1.8337
EEL                      -20139.3788              39.9200              5.6455
EGB                       -4094.2921              22.2570              3.1476
ESURF                        94.0722               0.3976              0.0562

G gas                    -22396.5189              41.7082              5.8984
G solv                    -4000.2199              22.2518              3.1469

TOTAL                    -26396.7388              31.7675              4.4926

Ligand:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                      -1.1075               1.4022              0.1983
EEL                          83.6674               1.2794              0.1809
EGB                         -40.1660               0.7469              0.1056
ESURF                         4.4974               0.0174              0.0025

G gas                        82.5599               1.9406              0.2744
G solv                      -35.6687               0.7471              0.1056

TOTAL                        46.8913               1.9986              0.2826

Differences (Complex - Receptor - Ligand):
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                     -76.4485               3.2333              0.4573
EEL                         -55.6398               3.1700              0.4483
EGB                          68.7784               1.4713              0.2081
ESURF                        -9.6612               0.1114              0.0158

DELTA G gas                 -132.0884               3.3350              0.4716
DELTA G solv                  59.1172               1.4823              0.2096

DELTA TOTAL                  -72.9712               2.7883              0.3943
```

Output file of ΔG_{bind} for inhibitor (5)-protein complex that resulted from MM-PBSA

File 10: MM-PBSA for ΔG_{bind} protein-inhibitor (5) complex

```
POISSON BOLTZMANN:

Complex:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2334.6961              13.3627              1.8898
EEL                      -20111.3512              40.3633              5.7082
EPB                       -3542.1323              21.0261              2.9735
ENPOLAR                    2238.7802               1.8782              0.2656
EDISPER                   -1252.4585               1.7487              0.2473

G gas                    -22446.0473              41.1968              5.8261
G solv                    -2555.8106              20.6883              2.9258

TOTAL                    -25001.8578              32.5969              4.6099

Receptor:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                   -2257.1401              12.9660              1.8337
EEL                      -20139.3788              39.9200              5.6455
EPB                       -3573.4939              20.9491              2.9626
ENPOLAR                    2235.8833               1.8308              0.2589
EDISPER                   -1281.1244               1.7811              0.2519

G gas                    -22396.5189              41.7082              5.8984
G solv                    -2618.7350              20.6246              2.9168

TOTAL                    -25015.2538              33.3003              4.7094

Ligand:
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                      -1.1075               1.4022              0.1983
EEL                          83.6674               1.2794              0.1809
EPB                         -37.6917               0.6818              0.0964
ENPOLAR                      54.4222               0.1989              0.0281
EDISPER                     -53.3586               0.2648              0.0375

G gas                        82.5599               1.9406              0.2744
G solv                      -36.6281               0.7588              0.1073

TOTAL                        45.9319               1.9257              0.2723

Differences (Complex - Receptor - Ligand):
Energy Component            Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
VDWAALS                     -76.4485               3.2333              0.4573
EEL                         -55.6398               3.1700              0.4483
EPB                          69.0533               1.7267              0.2442
ENPOLAR                     -51.5253               0.2645              0.0374
EDISPER                      82.0245               0.6233              0.0881

DELTA G gas                 -132.0884               3.3350              0.4716
DELTA G solv                 99.5525               2.0608              0.2914
```

Output file of ΔS for inhibitor (1)-protein complex that resulted from Nmode

File 11: Entropic contribution of protein-inhibitor (1) complex

```
ENTROPY RESULTS (HARMONIC APPROXIMATION) CALCULATED WITH NMODE:

Complex:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               16.9773                0.0000              0.0000
Rotational                  17.5628                0.0068              0.0023
Vibrational               3217.1317                3.2815              1.0938
Total                     3251.6718                3.2827              1.0942

Receptor:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               16.9713                0.0000              0.0000
Rotational                  17.5549                0.0017              0.0006
Vibrational               3199.0826                3.2836              1.0945
Total                     3233.6089                3.2845              1.0948

Ligand:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               12.5107                0.0000              0.0000
Rotational                   9.6839                0.0000              0.0000
Vibrational                 11.0817                0.4783              0.1594
Total                       33.2764                0.4783              0.1594

Differences (Complex - Receptor - Ligand):
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational              -12.5047                0.0000              0.0000
Rotational                  -9.6760                0.0062              0.0021
Vibrational                  6.9673                3.2004              1.0668

DELTA S total=             -15.2135                3.2025              1.0675


-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
```

Output file of ΔS for inhibitor (2)-protein complex that resulted from Nmode

File 12: Entropic contribution of protein-inhibitor (2) complex

```
ENTROPY RESULTS (HARMONIC APPROXIMATION) CALCULATED WITH NMODE:

Complex:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               16.9773                0.0000              0.0000
Rotational                  17.5751                0.0023              0.0008
Vibrational               3221.8426                3.0359              1.0120
Total                     3256.3949                3.0350              1.0117

Receptor:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               16.9713                0.0000              0.0000
Rotational                  17.5723                0.0012              0.0004
Vibrational               3205.1449                2.6761              0.8920
Total                     3239.6885                2.6765              0.8922

Ligand:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               12.5148                0.0000              0.0000
Rotational                   9.6022                0.0000              0.0000
Vibrational                 13.3378                0.4639              0.1546
Total                       35.4548                0.4639              0.1546

Differences (Complex - Receptor - Ligand):
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational              -12.5089                0.0000              0.0000
Rotational                  -9.5994                0.0027              0.0009
Vibrational                  3.3599                2.6447              0.8816

DELTA S total=             -18.7484                2.6433              0.8811


-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
```

Output file of ΔS for inhibitor (3)-protein complex that resulted from Nmode

File 13: Entropic contribution of protein-inhibitor (3) complex

```
ENTROPY RESULTS (HARMONIC APPROXIMATION) CALCULATED WITH NMODE:

Complex:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               16.9734                0.0000              0.0000
Rotational                  17.5666                0.0027              0.0009
Vibrational               3205.8049                4.7201              1.5734
Total                     3240.3448                4.7214              1.5738

Receptor:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               16.9659                0.0000              0.0000
Rotational                  17.5490                0.0113              0.0038
Vibrational               3176.7016                9.3158              3.1053
Total                     3211.2167                9.3266              3.1089

Ligand:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               12.7072                0.0000              0.0000
Rotational                  10.2694                0.0022              0.0007
Vibrational                 18.7394                0.0090              0.0030
Total                       41.7160                0.0068              0.0023

Differences (Complex - Receptor - Ligand):
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational              -12.6997                0.0000              0.0000
Rotational                 -10.2518                0.0114              0.0038
Vibrational                 10.3639                7.2080              2.4027

DELTA S total=             -12.5878                7.2183              2.4061


-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
```

Output file of ΔS for inhibitor (4)-protein complex that resulted from Nmode

File 14: Entropic contribution of protein-inhibitor (4) complex

```
ENTROPY RESULTS (HARMONIC APPROXIMATION) CALCULATED WITH NMODE:

Complex:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               16.9561                0.0000              0.0000
Rotational                  17.5128                0.0017              0.0006
Vibrational               3132.5212                4.7875              1.5958
Total                     3166.9900                4.7872              1.5957

Receptor:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               16.9474                0.0000              0.0000
Rotational                  17.5062                0.0027              0.0009
Vibrational               3108.9020                7.1749              2.3916
Total                     3143.3558                7.1748              2.3916

Ligand:
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational               12.8252                0.0000              0.0000
Rotational                  10.4617                0.0001              0.0000
Vibrational                 20.9163                0.0024              0.0008
Total                       44.2035                0.0026              0.0009

Differences (Complex - Receptor - Ligand):
Entropy Term                Average              Std. Dev.   Std. Err. of Mean
-------------------------------------------------------------------------------
Translational              -12.8166                0.0000              0.0000
Rotational                 -10.4552                0.0026              0.0009
Vibrational                  2.7028                3.6238              1.2079

DELTA S total=             -20.5692                3.6237              1.2079


-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
```

Building a library file for SEP residue

The connectivity information in the pdb file was deleted. Using ***xLeap,*** atoms are bonded together manually. First, the SEP non-standard residue was loaded (Fig 1)

```
Xleap

SEP = loadpdb SEP.pdb
```

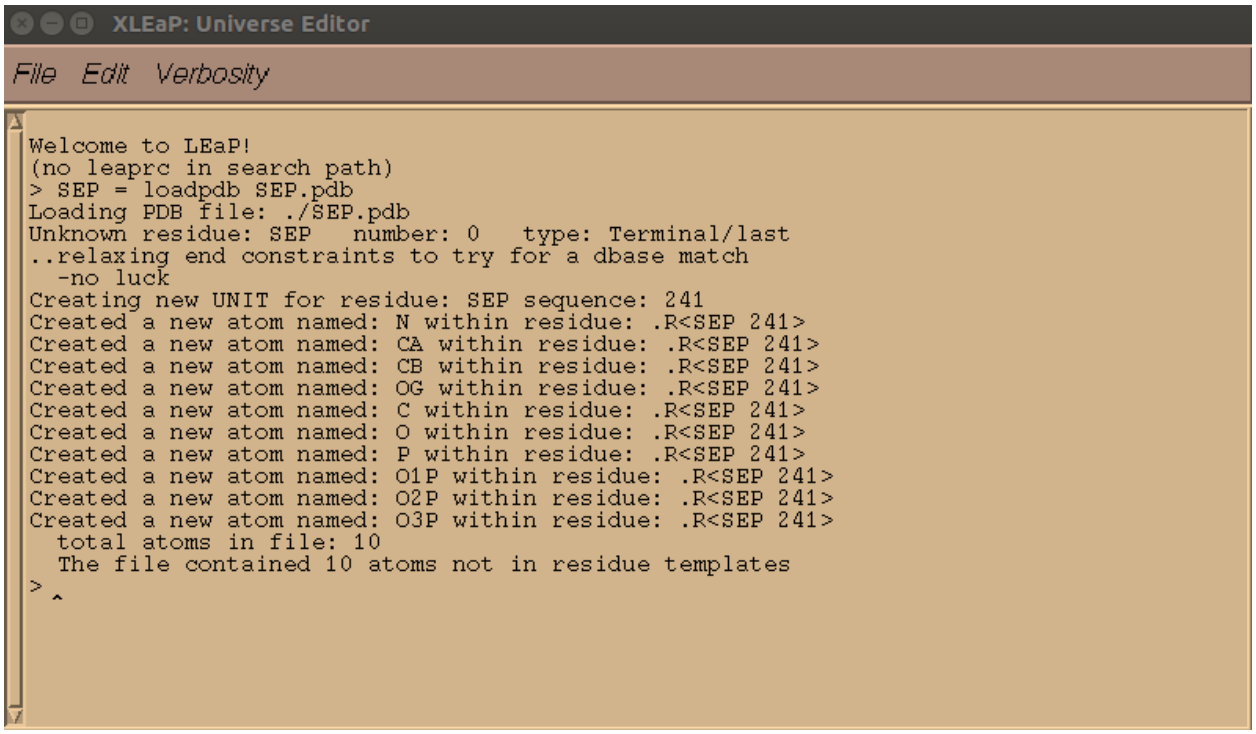


Figure 1: *XLeap* window shows loading atoms of SEP residue

All atoms of SEP residue were added to the new *UNIT*. It is important to check that ***xLeap*** "created" the correct atoms. The easiest way is to check for the total number of atoms in the file which should be 10.

The pdb file containing atom coordinates for SEP residue was loaded. ***XLeap*** does not have necessary SEP parameters and connectivity data, so this information was entered manually.

The SEP residue in ***xLeap*** was edited (Fig 2) by the command:

```
edit SEP
```

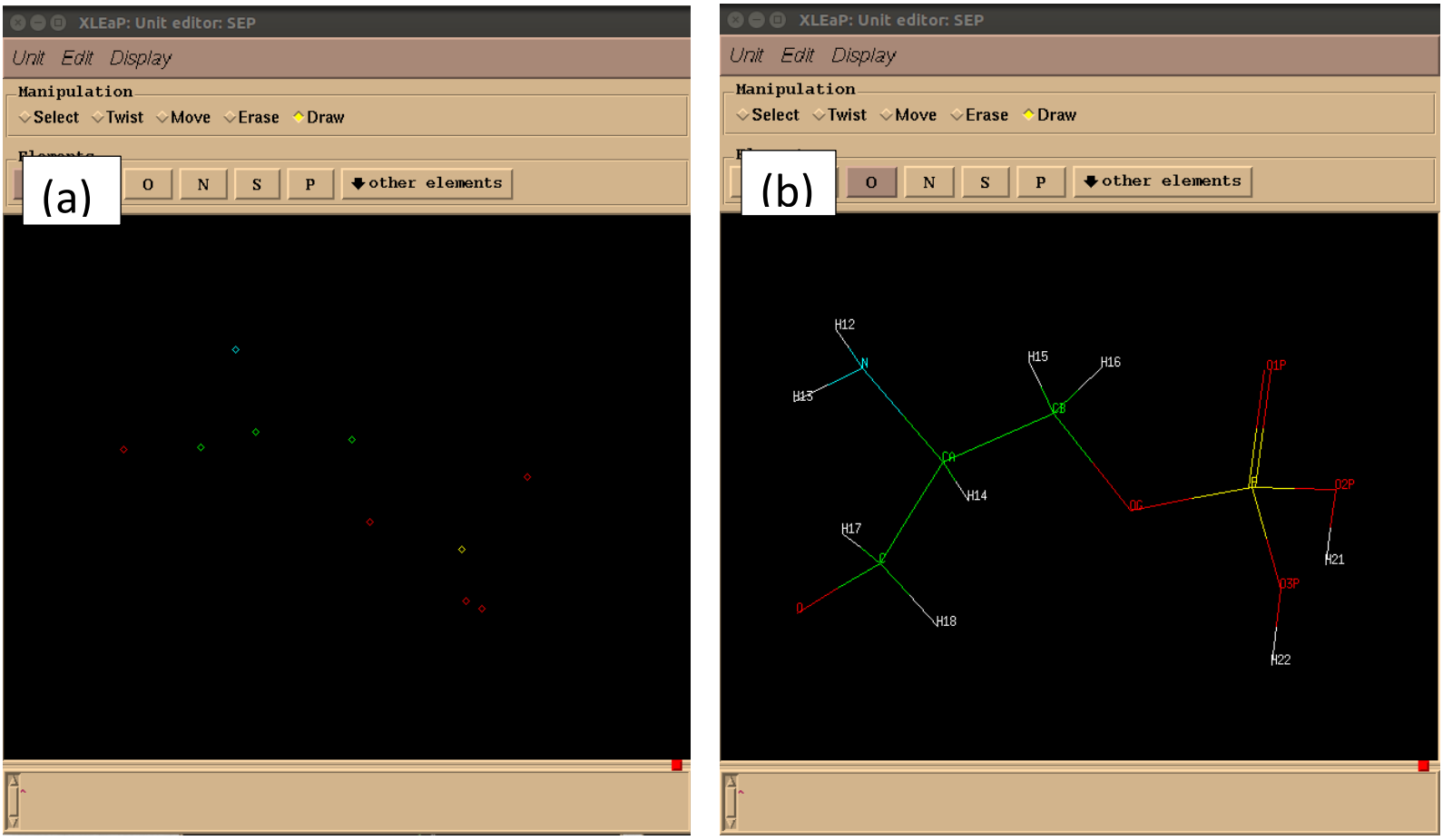


Figure 2: *XLeap* window shows (a) atoms of SEP residue before connecting between them, (b) SEP non-standard residue after connection

Parameters of each atom in SEP residue are given by ***parm99.dat*** file which lies ***in $AMBERHOME/dat/leap/parm*** table given in ***xLeap.*** The atom types of all SEP atoms were identified by blue table (Table 1).

Table 1: Blue table is given by *XLeap* shows all of atom types in SEP non-standard residue

XLEaP: Edit selected atoms: SEP

Table Operations

Table Editor: The table has no errors and 0 warnings.
Table Editor: Table has been 'saved' back to the program.

NAME	TYPE	CHARGE	ELEMENT	unused	PERT.name	PERT.type	DELTA.charge
N	NT	0.000000	N				0.000000
CA	CT	0.000000	C				0.000000
CB	CT	0.000000	C				0.000000
OG	OS	0.000000	O				0.000000
C	CT	0.000000	C				0.000000
O	OS	0.000000	O				0.000000
P	P	0.000000	P				0.000000
O1P	O2	0.000000	O				0.000000
O2P	OH	0.000000	O				0.000000
O3P	OH	0.000000	O				0.000000
H12	H	0.000000	H				0.000000
H13	H	0.000000	H				0.000000
H14	H1	0.000000	H				0.000000
H15	H1	0.000000	H				0.000000
H16	H1	0.000000	H				0.000000
H17	H1	0.000000	H				0.000000
H18	H1	0.000000	H				0.000000
H21	HO	0.000000	H				0.000000
H22	HO	0.000000	H				0.000000

The O2P and O3P atoms are oxygen atoms in hydroxyl group, so the type of this oxygen is OH in the blue table. But the O1P is oxygen atom in phosphate group, so the type of this atom is O2 according to PARM99.dat[17]

H12 and H13 are connected with nitrogen atom, so these atoms were assigned with an appropriate atom type which is H. HC and H1 were assigned for H in aliphatic bond to Carbon without electron withdrawing group and H in aliphatic bond to Carbon with one electron withdrawing group, respectively.[17]

Then library file of SEP was saved which will enable ***xLeap*** to recognize this residue in the future. This step is very essential to prevent the repetition of all of the previous steps each time. This was done by the following command.

```
Saveoff SEP SEP_leap.lib
Savepdb SEP SEP_leap.pdb
```

The missing bonds and angle parameters of SEP residue were identified by ***xLeap***. This was achieved by using the following commands

```
xleap -s -f $AMBERHOME/dat/leap/cmd/oldff/leaprc.ff99SB
loadoff SEP_leap.lib
check SEP
```

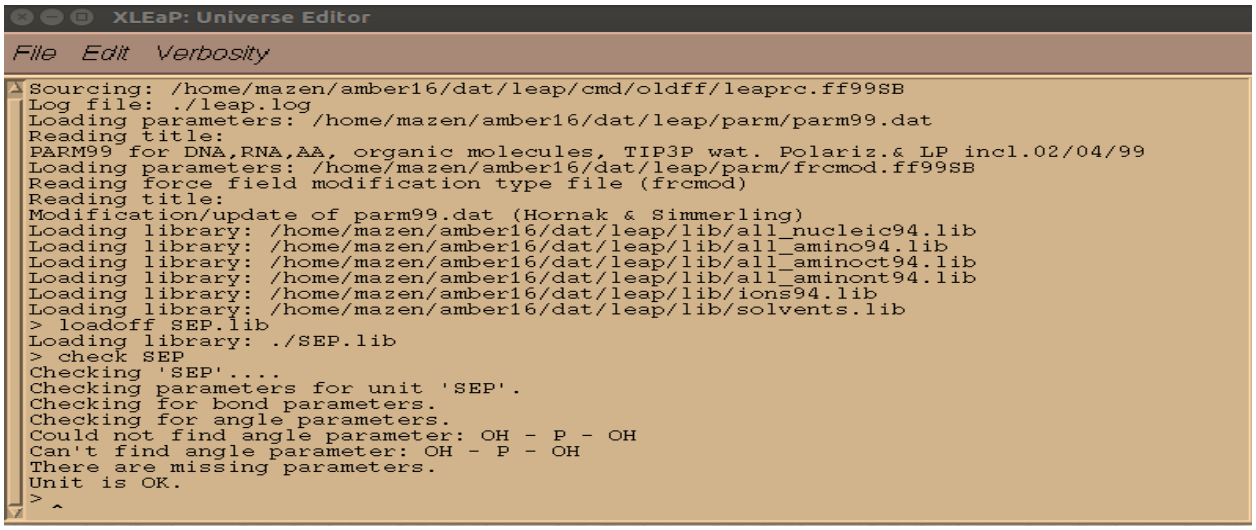


Figure 3: *XLeap* window shows the missing parameters of non-standard residue (SEP)

An *frcmod* file s required to provide all the bonds, angles and dihedral parameters that are not present in the standard FF99SB force field. The only missing parameter in SEP residue is the OH-P-OH angle parameter (Fig 3). So SEP_leap.frcmod file (see Appendix A) was created to define the missing OH-P-OH angle parameter. Then the *frcmod* file of SEP was loaded using this command

```
loadamberparams SEP_leap.frcmod
```

Finally, the SEP residue was successfully built using ***xLeap***, by check the residue (UNIT is OK) and by saving it as *prmtop* and *inpcrd* file (Fig 4).

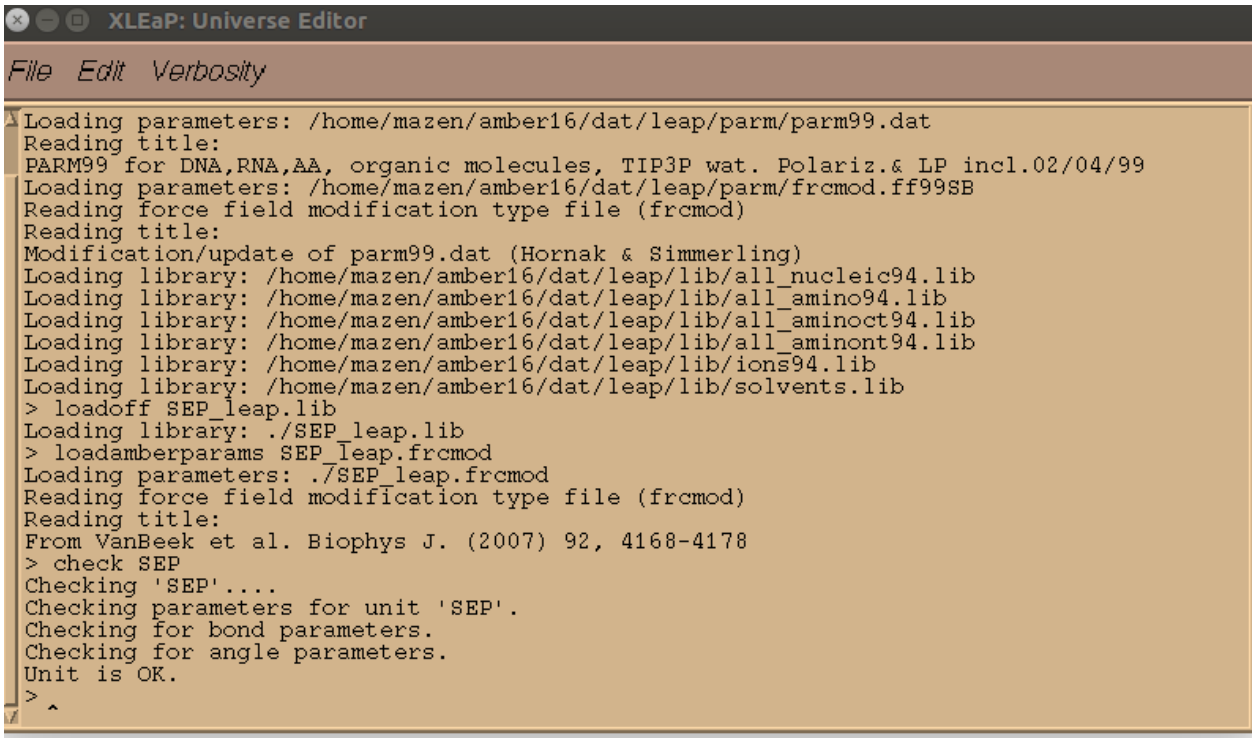


Figure 4: *XLeap* window shows that SEP residue succussefully built using *xLeap*

Creating *AMBER* input files

The *inpcrd* and *Prmtop* files are the coordinate files and molecular topology/parameter, respectively. These files are necessary for running molecular dynamics simulation of protein-ligand complexes using ***Sander***.

Antechamber is designed to be used with the "general AMBER force field (***GAFF***)", and was successfully used in the production of *frcmod* files and *mol2* files of inhibitors. ***GAFF*** force field covers most pharmaceutical molecules and is

compatible with *AMBER* force fields. ***GAFF*** is a complete force field and covers all the organic molecules that contain C, N, O, S, P, H, F, Cl, Br and I.[74]

The hydrogenated 6-methoxy-2-(1H-pyrazol-5-yl)-1H-benzimidazole (inhibitor 1) coordinates were done using ***Pymol***, then ***antechamber*** command was used to create the "*mol2*" file using the following command in terminal:

```
antechamber -i inhibitor1.pdb -fi pdb -o inhibitor1.mol2 -fo mol2 -
c bcc -s 2
```

This command line produced a number of files in CAPITALS. These files are used by antechamber and are not required here. These files are considered as intermediate files, but *mol2* file of inhibitor1 (see Appendix A) is the most important one because it reveals the definition of our inhibitor (1) residue, including all of the charges and atom types.

To specify any missing parameters (bonds, angles, dihedral angles) before we can create our *prmtop* and *inpcrd* files in ***Leap***, we run the *parmchk* command in terminal to test if all the parameters we require are available.

```
parmchk -i inhibitor1.mol2 -f mol2 -o inhibitor1.frcmod
```

Running this command produced a file called *inhibitor1.frcmod* (see Appendix A). This is a parameter file that can be loaded into ***xLeap*** in order to add missing parameters.

Subsequently, ***xLeap*** was used to form *prmtop* and *inpcrd* files using the following command was typed in terminal in order to open ***xLeap***:

```
xleap -s -f $AMBERHOME/dat/leap/cmd/leaprc.ff99SB
```

This command line starts ***xleap*** and loads the configuration files needed for *AMBER FF99SB* force field as shown in Figure 5.

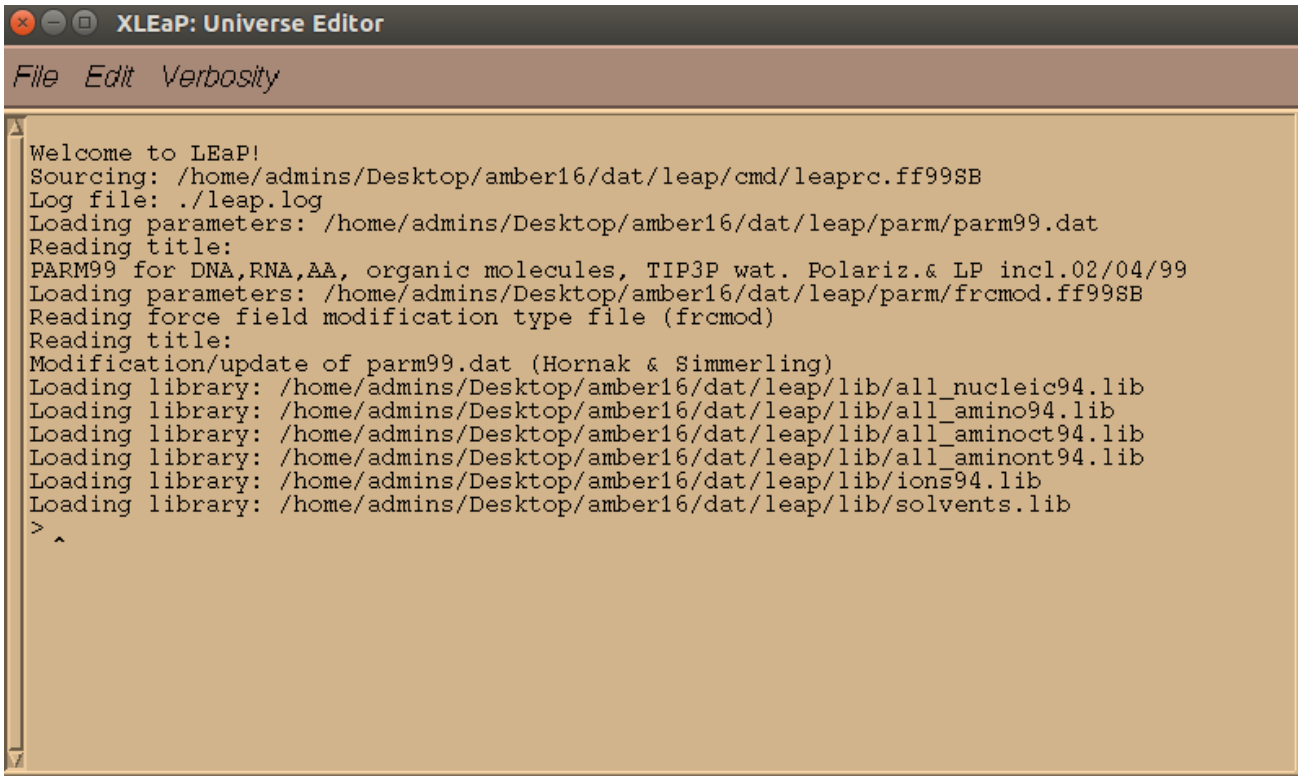


Figure 5: *XLeap* window shows FF99SB force field

To ensure that ***xLeap*** has the ***GAFF*** force field, it is loaded into ***xLeap*** by using command line:

```
Source leaprc.gaff
```

xLeap looks like this:

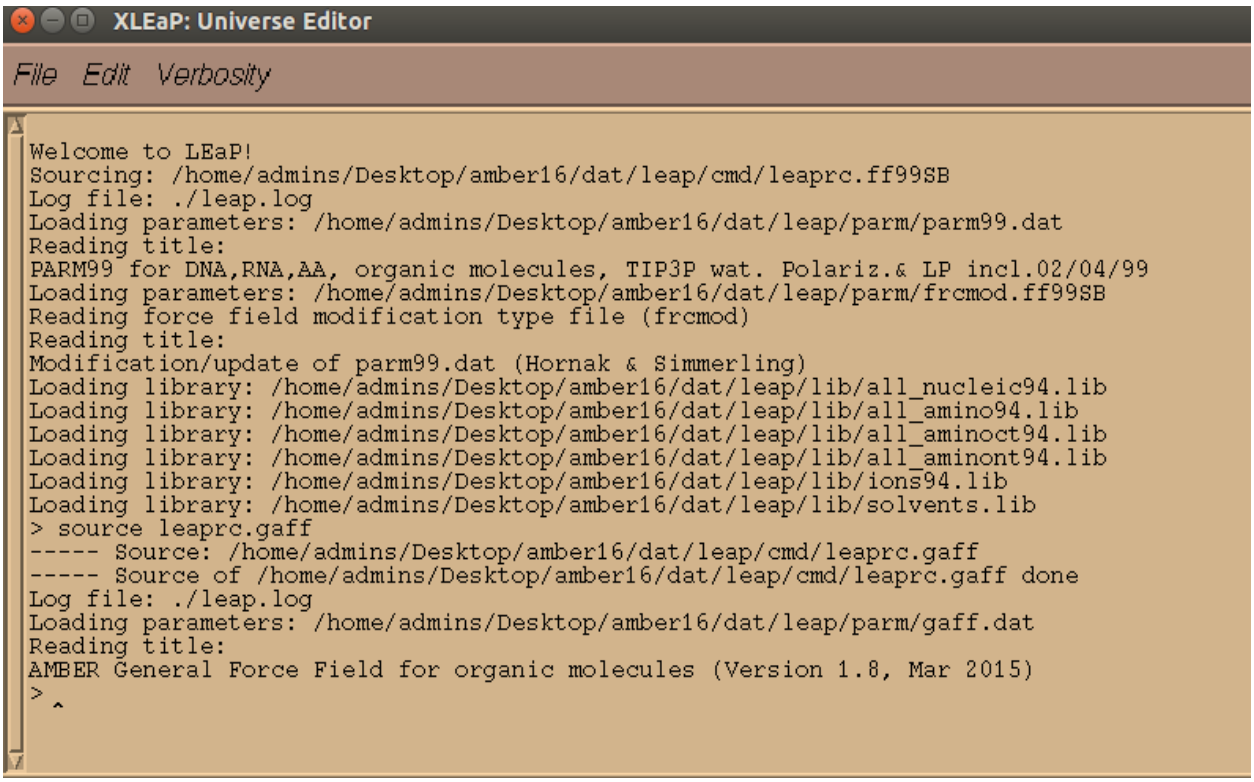


Figure 6: XLeap window shows preparing to load the protein-inhibitor (1) complex X-ray structure

Now inhibitor (1) unit (inhibitor1.mol2) is loaded:

```
62O = loadmol2 inhibitor1.mol2

loadamberparams inhibitor1.frcmod
```

The library file for inhibitor (1) was created, as well as the *prmtop and inpcrd* files using the command lines:

```
saveoff 62O inhibitor1.lib

saveamberparm 62O inhibitor1.prmtop inhibitor1.inpcrd
```

Inhibitor (1) can be seen (Fig 7 (a)) using edit command:

```
edit 62O
```

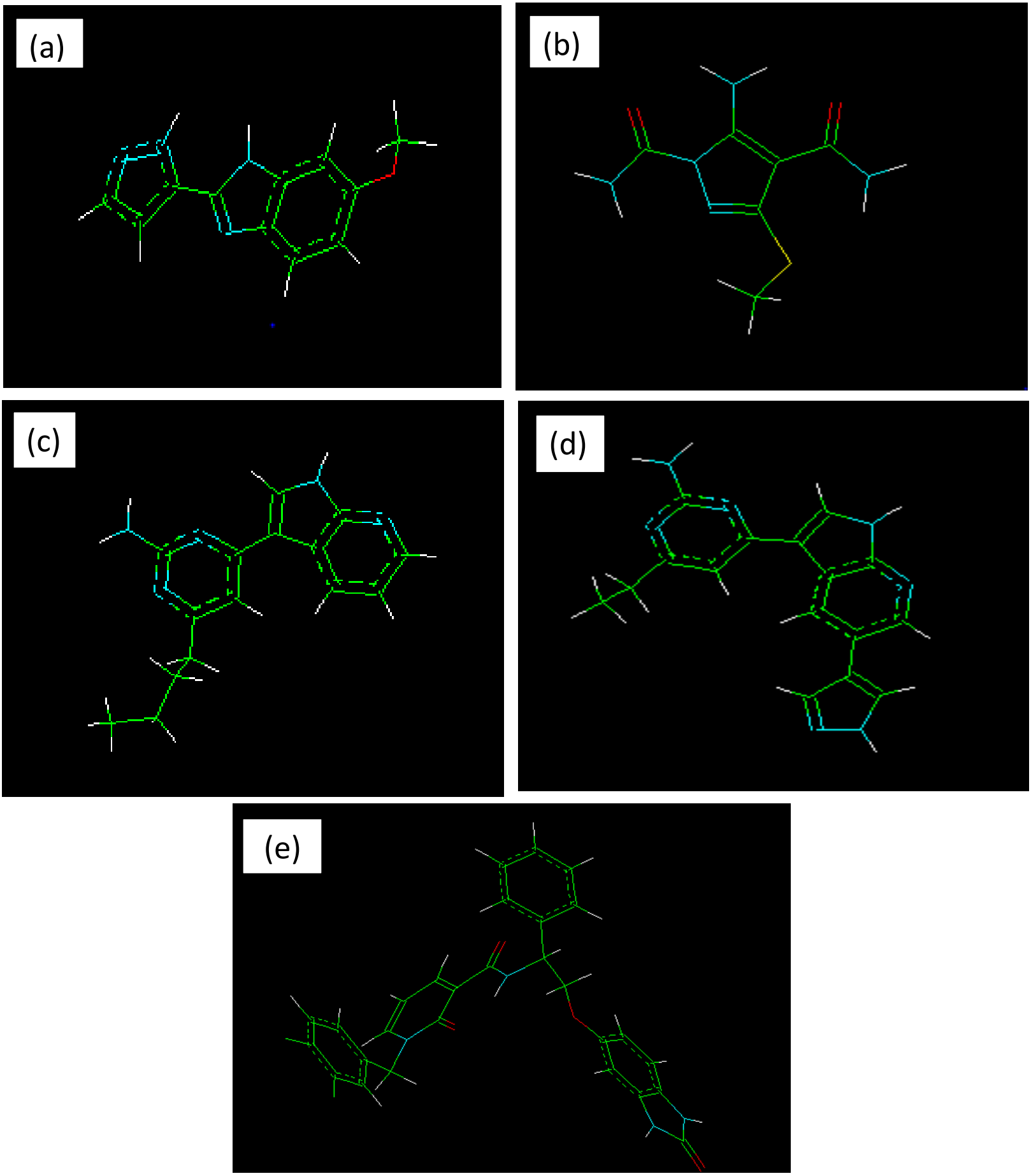


Figure 7: X*Leap* window showing the graphical representation of (a) inhibitor (1), (b) inhibitor (2), (c) inhibitor (3), (d) inhibitor (4), and (e) inhibitor (5), respectively

Now ***xLeap*** was ready to load the protein_inhibitor1.pdb file without having any problems. protein_inhibitor (1).pdb file was loaded into xleap after setting it as a new unit called "com" by writing the following commands:

```
Xleap -f $AMBERHOME/dat/leap/cmd/leaprc.ff99SB
Source leaprc.gaff
Loadoff SEP_leap.lib
Loadamberparams SEP_leap.frcmod
62O = loadmol2 62O.mol2
Loadamberparams 62O.frcmod
Com = loadpdb protein_inhibitor1_dry.pdb
Edit com
```

The ***xLeap*** window shows the graphical representation of protein_inhibitor1_dry.pdb (Figure 8).

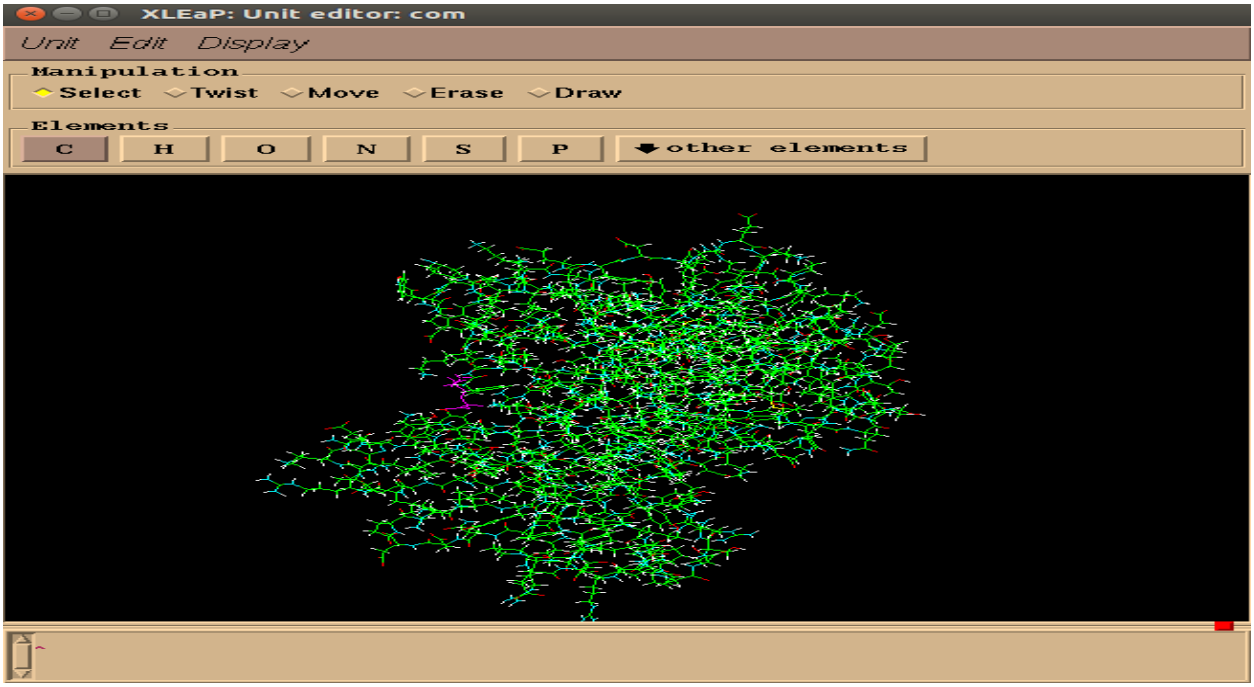


Figure 8: *Xleap* window showing the graphical representation of protein_inhibitor1_dry.pdb

```
set default PBRadii mbondi2

saveamberparm com protein_inhibitor1_dry.prmtop
protein_inhibitor1_dry.inpcrd

saveamberparm 62O inhibitor1.prmtop inhibitor1.inpcrd
```

$\Delta G^0_{nonpolar}$ was calculated using the default parameters γ = 0.00500 kcal/ Å^2 and β= 0.0000 kcal/mol this was achieved using the mbondi2 radius, because this method is affective in the calculation of non-polar solvation energy. The system was neutralized by adding counter ions:

```
charge com
```

Thus, two Cl- atoms were added to neutralize the protein_inhibitor1 complex. This task was done using the following order:

```
addions com Cl- 0
```

This command line causes a columbic potential on a grid of 1Å resolution and then puts the counter ions simultaneously at the points of lowest/greatest electrostatic potential (Figure 9).

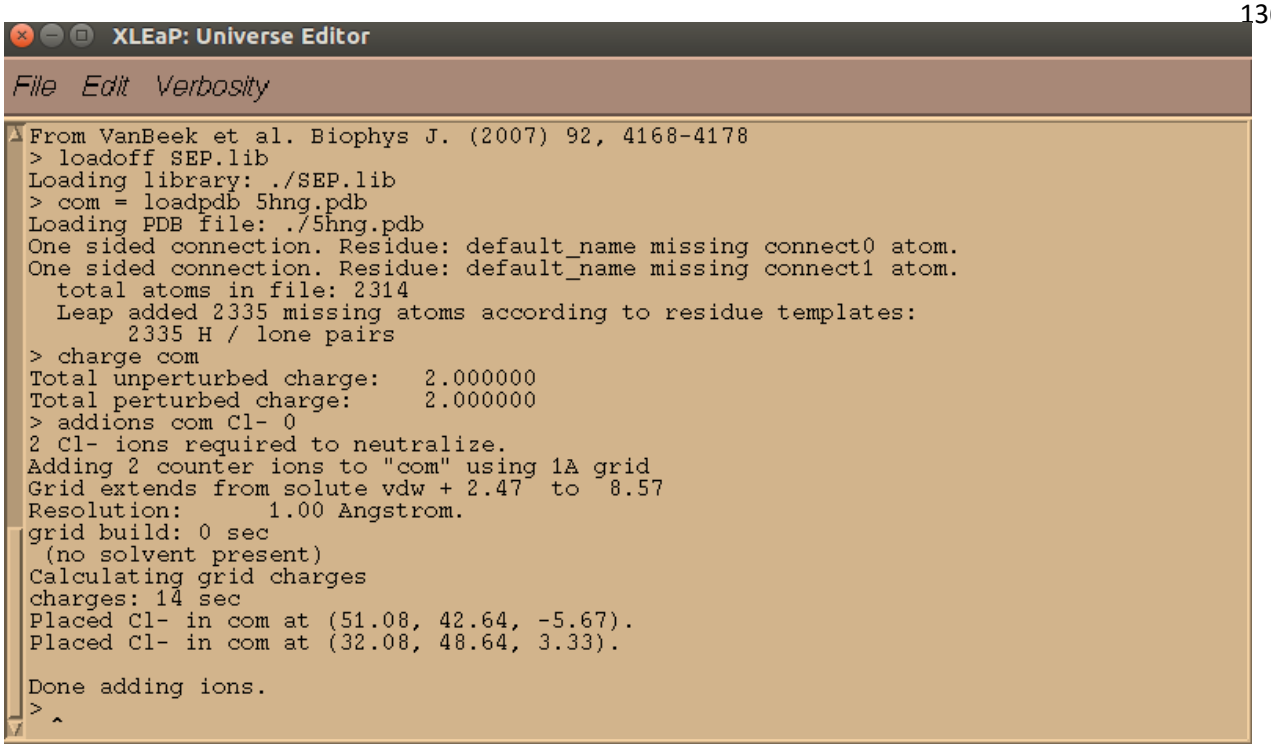


Figure 9: *XLeap* editor shows neutralization of the protein_inhibitor1_dry.pdb complex by addition of chlorine ions

Finally, the system was solvated using the following command to add a periodic rectangular box of TIP3P within a distance from the surface of the box to the closest atom of the solute was set to 10 Å in x, y, and z directions (Figure 2.10).[75]

```
solvatebox com TIP3PBOX 10.0
```

The *prmtop* and *inpcrd* files for the solvated system were saved using the Following commands:

```
saveamberparm com protein_inhibitor1_wat.prmtop
protein_inhibitor1_wat.inpcrd
```

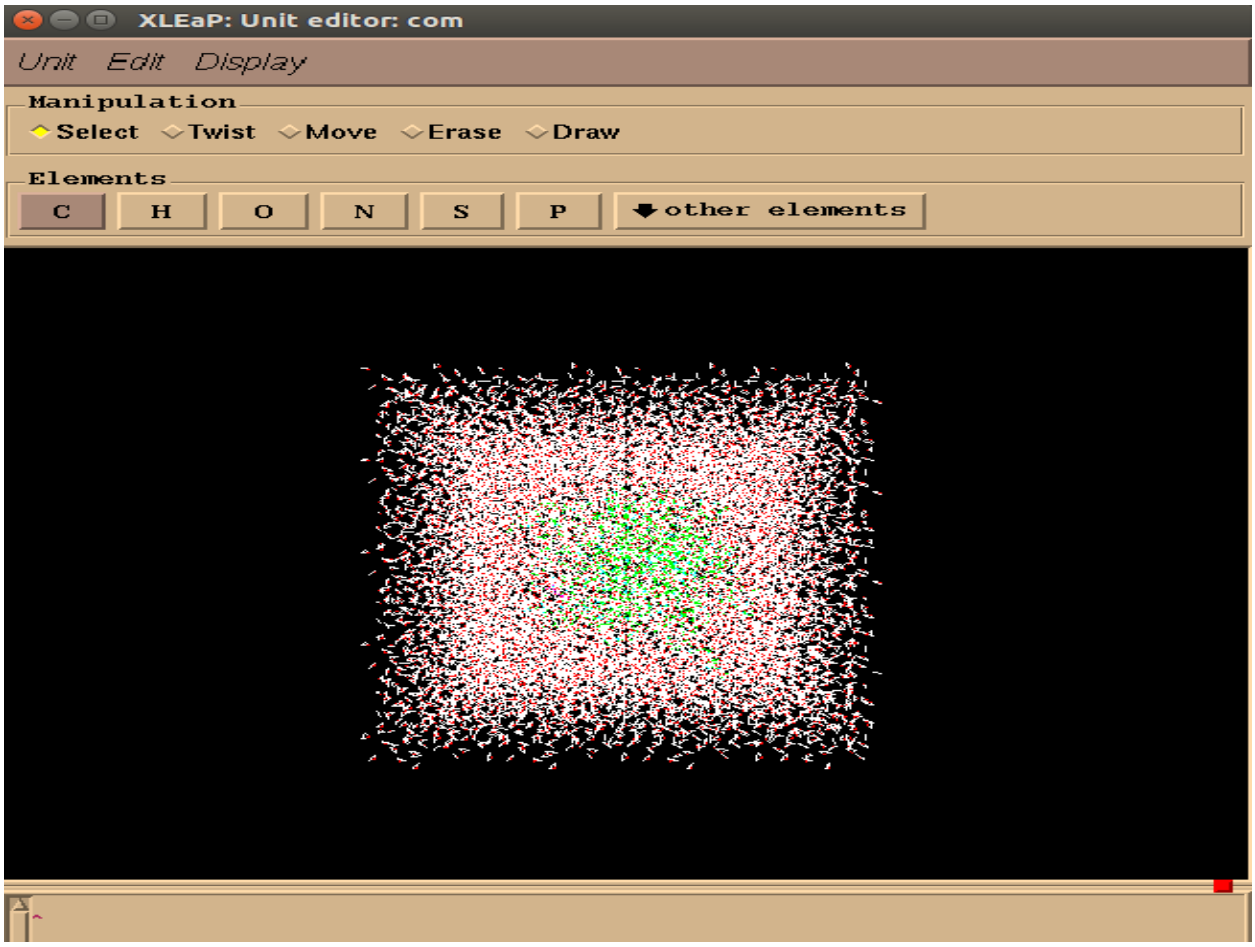


Figure 10: *Xleap* window showing the graphical representation of protein_inhibitor1_wat.pdb