University of Kentucky Doctoral Dissertations

Graduate School

2009

COMPUTATIONAL DESIGN OF 3-PHOSPHOINOSITIDE DEPENDENT KINASE-1 INHIBITORS AS POTENTIAL ANTI-CANCER AGENTS

Mohamed Diwan Mohideen AbdulHameed
University of Kentucky, mohamed.diwan@gmail.com

Click here to let us know how access to this document benefits you.

ABSTRACT OF DISSERTATION

Mohamed Diwan Mohideen AbdulHameed

The Graduate School

University of Kentucky

2009

COMPUTATIONAL DESIGN OF 3-PHOSPHOINOSITIDE DEPENDENT KINASE-1
INHIBITORS AS POTENTIAL ANTI-CANCER AGENTS

_____

ABSTRACT OF DISSERTATION
_____

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Pharmacy
at the University of Kentucky

By
Mohamed Diwan Mohideen AbdulHameed

Lexington, Kentucky

Director: Dr. Chang-Guo Zhan, Professor of Pharmaceutical Sciences

Lexington, Kentucky

2009

ABSTRACT OF DISSERTATION

COMPUTATIONAL DESIGN OF 3-PHOSPHOINOSITIDE DEPENDENT KINASE-1
INHIBITORS AS POTENTIAL ANTI-CANCER AGENTS

Computational drug design methods have great potential in drug discovery particularly in lead identification and lead optimization. 3-Phosphoinositide dependent kinase-1 (PDK1) is a protein kinase and a well validated anti-cancer target. Inhibitors of PDK1 have the potential to be developed as anti-cancer drugs. In this work, we have applied various novel computational drug design strategies to design and identify new PDK1 inhibitors with potential anti-cancer activity. We have pursued novel structure-based drug design strategies and identified a new binding mode for celecoxib and its derivatives binding with PDK1. This new binding mode provides a valuable basis for rational design of potent PDK1 inhibitors. In order to understand the structure-activity relationship of indolinone-based PDK1 inhibitors, we have carried out a combined molecular docking and three-dimensional quantitative structure-activity relationship (3D-QSAR) modeling study. The predictive ability of the developed 3D-QSAR models were validated using an external test set of compounds. An efficient strategy of the hierarchical virtual screening with increasing complexity was pursued to identify new hits against PDK1. Our approach uses a combination of ligand-based and structure-based virtual screening including shape-based filtering, rigid docking, and flexible docking. In addition, a more sophisticated molecular dynamics/molecular mechanics- Poisson-Boltzmann surface area (MD/MM-PBSA) analysis was used as the final filter in the virtual screening. Our screening strategy has led to the identification of a new PDK1 inhibitor. The anticancer activities of this compound have been confirmed by the anticancer activity assays of national cancer institute-developmental therapeutics program (NCI-DTP) using 60 cancer cell lines. The PDK1-inhibitor binding mode determined in this study may be valuable in future *de novo* drug design. The virtual screening approach tested and used in this study could also be applied to lead identification in other drug discovery efforts.

Mohamed Diwan Mohideen AbdulHameed
Student's Signature

Jul. 13, 2009
Date

COMPUTATIONAL DESIGN OF 3-PHOSPHOINOSITIDE DEPENDENT KINASE-1
INHIBITORS AS POTENTIAL ANTI-CANCER AGENTS

By
Mohamed Diwan Mohideen AbdulHameed

Dr. Chang-Guo Zhan
Director of Dissertation

Dr. Janice Buss
Director of Graduate Study

Jul. 13, 2009

RULES FOR THE USE OF DISSERTATION

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with permission of the author, and with the usual scholarly acknowledgements.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name   Date

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

DISSERTATION

Mohamed Diwan Mohideen AbdulHameed

The Graduate School

University of Kentucky

2009

COMPUTATIONAL DESIGN OF 3-PHOSPHOINOSITIDE DEPENDENT KINASE-1
INHIBITORS AS POTENTIAL ANTI-CANCER AGENTS

_____

DISSERTATION
_____

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Pharmacy
at the University of Kentucky

By
Mohamed Diwan Mohideen AbdulHameed

Lexington, Kentucky

Director: Dr. Chang-Guo Zhan, Professor of Pharmaceutical Sciences

Lexington, Kentucky

2009

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the guidance and support of several people. I would like to take the opportunity to acknowledge those who have helped me tremendously in this journey. First, I would like to express my sincere, heartfelt gratitude to my mentor and research advisor, Dr. Chang-Guo Zhan for his support, guidance, kindness and patience throughout my graduate studies. I am greatly indebted for his encouragement, advice and for setting a great lead to follow. His passion and enthusiasm for science and computational drug design is a source of great inspiration for me. I express my sincere thanks to my committee members, Dr. Hsin-Hsiung Tai, Dr. Kyung B. Kim and Dr. Peter Spielmann for their advice, time, and long-time support. I would also like to thank Dr. Xianglin Shi for being my external examiner.

I would like to express my heartfelt thanks to Dr. Adel Hamza for his help throughout my Ph.D study. His help in introducing me to technical details of simulations and countless hours of discussion has helped me a lot in my graduate work. I would also like to convey my thanks to my friend Mr. Jun Jun Liu. His suggestions and his help in scripting helped me a lot in my work. I would also like to thank Dr. Marius Vilkas for his support. Dr. Hamza, Mr. Liu, and Dr. Vilkas were a great team and their encouragement helped me a lot during my Ph.D work. I would like to thank Mr. Wenchao Yang for his help in experimental work. I am much thankful to all my lab members Ms. Yongmei Pan, Dr. Xiaoqin Huang, Dr. Daquan Gao, and Dr. Haiting Lu for all their support in the past 5 years. I would also like to thank our collaborator Dr. Wei Wang in University of New Mexico.

I would like to thank our DGS Dr. Janice Buss for her help, support and constant encouragement throughout my graduate study. I am extremely grateful to Ms. Catina Rossoll. Throughout my graduate study, she has been a great help to me. I acknowledge the Center for Computational Sciences (CCS) at University of Kentucky for providing enormous supercomputing time. I would like to thank all my friends in India and in Lexington for their love and support. I would also like to thank different cultural organizations in UKY and Lexington which helped in my transition to USA.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Computational drug design is a fast growing field and it has emerged as a major player to design new drugs in a rational manner. In this work, the application of computational drug design techniques including structure-based drug design, ligand-based three-dimensional quantitative structure-activity relationship (3D-QSAR) methods and combined virtual screening approaches were studied using a well known anti-cancer target 3-phosphoinositide dependent kinase-1 (PDK1). These computational drug design approaches have been successfully applied to both lead optimization and hit identification against PDK1. In this chapter the background information about computational drug design and PDK1 are presented.

1.1 Computer-aided drug design

Drug discovery is broadly defined by Burger as the process that includes identification and validation of a therapeutic target, development and validation of suitable assays, lead identification and characterization in vitro, lead optimization, formulation, pharmacological studies, pharmacokinetic and toxicity studies in animals, followed by three phases of clinical trials in humans.[1] This description of drug discovery itself underlines the fact that it is a very long and expensive process. Discovery and development of a single drug is estimated to take 12 to 24 years.[2] It is reported that the average cost to develop a new drug into market is more than US$800 million before 2004.[3] Historically, Paul Ehrlich postulated on the existence of chemoreceptors that can be exploited therapeutically and it marked the beginning of the modern drug therapy.[4] Initial focus was reported to be on the isolation and purification of active ingredients from natural products.[4] This evolved into the trial and error process of synthesis and in vivo screening of compounds leading to serendipitous discovery.[2] This was followed by rational approaches which were based on studying the underlying molecular mechanism involved and understanding the drug-receptor interactions.[4] Despite this systematic approach, the number of new

chemical entities (NCE) that are approved for market was reported to be declining in the recent years.[5] Only 18 NCEs were reported to be approved in the year 2005 and 2006.[6] There was a need for newer technologies to increase the efficiency of drug discovery process.[7] In this scenario, different experimental and computational approaches were developed to address the issues in drug discovery process. Some of the new approaches include combinatorial chemistry, high throughput screening and computational drug design.[8] Among these new technologies, computational drug discovery has been successful and helps to increase the efficiency of the drug discovery process.[7] Computational drug design is also commonly known as molecular modeling or computer-aided drug design (CADD). Cohen defines computational drug design as the approach "which uses a range of computerized techniques based on theoretical chemistry methods and experimental data that can be used to analyze molecules and molecular systems and to predict molecular and biological properties".[9] Simply this approach is the use of computers and computational methods to design and discover new drug molecules.[10] The increase in computing power, the availability of cheaper computers and development of new modeling software has increased the usefulness of this approach.[11] Rather than being a separate process, computational modeling is more useful when integrated with experimental studies.[7] CADD approaches are being implemented in all stages of drug discovery cycle namely target identification, lead identification, lead optimization, prediction of absorption, distribution, metabolism, elimination, and toxicity (ADMET) properties and design of compound libraries.[12] The main advantage of computational drug design is that it is faster and cheaper than other experimental approaches.[12] This approach is reported to play an active role in both lead identification and optimization. Computational drug design is reported to have played a major role in the successful development of marketed drugs like HIV protease inhibitor Viracept and the anti-influenza drug Relenza.[13] In these cases, computational approaches were used to predict the activity of the designed inhibitor prior to synthesis and testing.[14] Thus it helps to guide the experimental work and prioritize synthesis of the next compounds. Recently reported studies suggest that

computational drug design is also an important component for open-source drug discovery projects. [15 , 16] Different sources are being explored to carry out computational drug design in a cheaper way. A project called the 'screen saver project' made use of the free computer time (screen saver time) of users throughout the world and successfully carried out computational screening against many therapeutic targets and identified new hits.[17] There are two broad strategies in computational drug design. They are

1. Ligand-based drug design
2. Structure-based (or Receptor-based) drug design

Ligand based drug design (LBDD) which is also known as indirect drug design make use of the ligand molecule or a series of ligand molecules as the starting point and explore their properties to identify better lead molecules.[7,8] Based on the study of known ligands a hypothetical receptor active site can be proposed.[18] The structure based drug design which is also known as direct drug design make use of either a solved X-ray crystal/NMR structure or modeled structure of the target protein or other macromolecule.[7,19] This target structure is used as the starting point.[19] The binding mode of ligand is then studied to design new ligands with increased affinity to the target.[19]
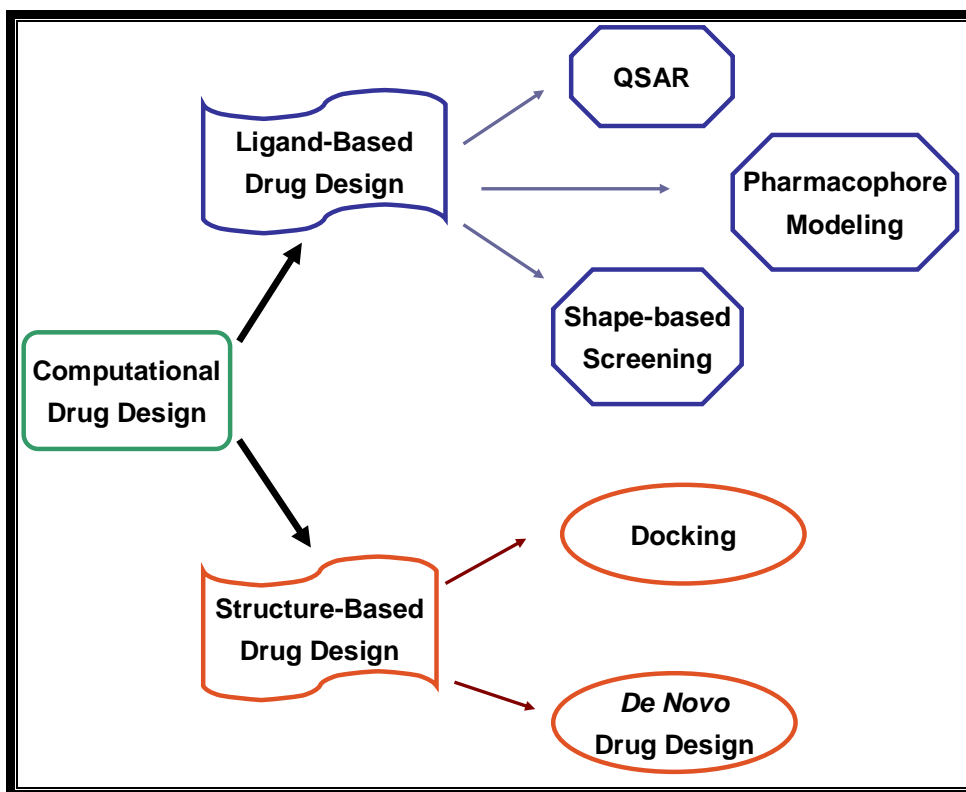
Figure 1.1 Flowchart of computational drug design strategies

1.1.1 Ligand-based drug design

The structures of protein or other macromolecules are not available for many of the potential drug targets. In such cases, ligand-based drug design strategies can be utilized.[7] These strategies are based on the concept that biological activity depends on the ligand's structure, molecular constitution and charge distribution.[20] When a series of ligands with their biological activity are available, then attempt can be made to correlate their structure with their biological activities.[21] Quantitative structure-activity relationship (QSAR) is the "technique that quantifies the relationship between structure and biological data".[22] It involves the study of descriptors to mathematically correlate the chemical structure with biological activity.[23] The descriptors are parameters which are used to represent the structural features in numerical values.[24] The different types of descriptors commonly used in QSAR models include lipophilicity parameter like LogP, electronic descriptors like Hammet's constants and dipole moments, geometrical descriptors like Taft's steric parameter and Verloop's STERIMOL parameters, and finally topological descriptors

like molecular connectivity index.[20,25] Three essential prerequisites for developing a QSAR model are "1) descriptors of the chemical's structure 2) measures of the chemical's activity and 3) statistical techniques to quantify the relationship".[20] The most attractive feature of the QSAR model is that the activity of new molecules to be synthesized in a series could be predicted.[21] QSAR methods have some common assumptions. They are 1) the compounds in the series being studied interact with the same biological target through non-covalent interactions; 2) structurally similar compounds interact in the same pocket and are similarly oriented in the active site.[7] The dynamics of the receptor system are usually neglected in ligand-based studies.[7] The first QSAR model was reported by Hansch *et al.*[26,27] It was developed by relating the biological activity to various physico-chemical parameters using suitable statistical regression analysis. The type of QSAR approach which uses the physicochemical properties as molecular descriptors is commonly known as the Hansch approach or linear free energy relationships (LFER) approach.[28] Free-Wilson approach is another type of QSAR method in which indicator variables that denote the presence or absence of particular groups were used as descriptors instead of the physicochemical properties.[28] In addition to predicting the activity of new molecule, the QSAR model can also provide information on the mechanism of action of the compounds.[22] It also helps to understand the different aspects of chemical-biological interaction.[29] Some of the successful application of 2D-QSAR has been reported in the literature.[30,31] Some of the limitations of the 2D-QSAR methods include 1) the lack of physicochemical parameters for either the whole molecule or its substituents; 2) the absence of 3D molecular properties which are expected to better explain the receptor-ligand interactions.[32] The 3D-QSAR tried to address the deficits of 2D-QSAR methods. 3D-QSAR models are "quantitative models that relate the biological activity of small molecules with their properties calculated in 3D space".[7] Some of the well known 3D-QSAR methods are Comparative Molecular Field Analysis (CoMFA), Comparative Molecular Similarity Indices Analysis (CoMSIA), Molecular Field Analysis (MFA) and Receptor Surface Analysis (RSA).[33,34,35]

Some of the general steps involved in 3D-QSAR are: identification of bioactive conformation which is usually used as the template, alignment of compounds to the template, computation of molecular properties, correlation analysis of these properties with activity, and visualization of regions around the molecule which are either favorable or unfavorable for activity.[32] Alignment of different compounds using similar groups or common atoms is one of the crucial steps in many 3D-QSAR methods.[36] The major advantages of 3D-QSAR is the ability to predict the activity of new molecules which is not used in the development of the model as well as the presence of contour maps showing favorable and non-favorable areas around the molecules in 3D-space.[37]

Pharmacophore modeling and shape based screening are the other ligand-based drug design methods. Pharmacophore can be defined as the three dimensional representation of essential chemical features of a compound needed for its biological activity.[38] A 3D pharmacophore gives "the relationship between the groups or chemical features, by defining distance ranges between groups, angles between groups or planes, and exclusion spheres".[39] Based on the structure-activity relationship of diverse ligands the common pharmacophore can also be identified.[40,41] Once a pharmacophore is developed, it can be used in 3D database searching to retrieve novel compounds that would match the pharmacophore, without having the topological features of known active compounds.[42] Thus it helps to identify new class of molecules. Some of the most commonly used programs for pharmacophore development are Catalyst, Phase and Ligand scout.[38] More recently the shape of the known ligand was successfully used in virtual screening process to carry out scaffold hopping or to find new inhibitors from database searching.[43]

1.1.2 Structure-based drug design

Structure-based drug design (SBDD) methods are used when the detailed structural information of the macromolecular target is available.[39] Docking studies and de novo drug design are two different methods used in structure-based drug design.[8] SBDD is now considered as an integrated part of drug discovery cycle.[44]

The main theme of SBDD is to find ligands complementary to the receptor active site. Hence the structure of the macromolecular target is essential.[39] If the experimentally solved structure of the target is not available, computational modeling methods like homology modeling or *ab initio* modeling can be used.[45,46] The modeled structures can then be used to study the ligand-receptor interactions. The docking process helps to understand the structure-activity relationship of known ligands based on their molecular level interactions with the active site residues.[47] Visual analysis of docked results allows us make chemical modifications in a rational manner. The two important aspects of docking algorithm are 1) to find the correct binding mode and 2) to estimate the binding energy and rank order a given series of ligands.[48] Virtual screening is the process by which whole databases of molecules were screened against target of interest to identify new hits.[13] The identified compounds can then be either synthesized or purchased from commercial sources and tested for activity.[13] This computational screening strategy helps to refine millions of virtual compounds *in silico* and reduces the amount of actual experimental tests needed.[49] Different docking programs currently available include FlexX, Fred, Dock, AutoDock and Glide.[39,50] The docking programs are limited by their ability to consider protein dynamics, solvation and entropic effect of ligand binding.[51,52] *De novo* design involves design of totally new inhibitors using only the active site as the starting point.[8] In this case, fragments which are complementary to the different regions of active sites were identified and then joined using suitable linker groups.[8] The problem with this strategy is that some computationally designed compounds could not be synthesized.

1.2 3-Phosphoinositide dependent kinase-1 (PDK1) as a test case

1.2.1 Cancer

Cancer is defined as a group of diseases characterized by uncontrolled growth and spread of abnormal cells.[53] It is a life-threatening disease condition worldwide. A total of 7.6 million people were estimated to have died due to cancer worldwide in the year 2008 alone.[54] Others have estimated that the global cancer burden will

increase to 27 million new cases by 2050.[55] In USA, cancer is the second most common cause of death.[56] According to the American cancer society, nearly 1,437,180 new cancer cases are expected to be diagnosed in 2008.[53] It is estimated that cancer death rate will be 1500 people per day in USA alone.[53] Cancer can develop in any individual at any time in their lifetime. But the risk of being diagnosed with cancer is reported to be increasing with age.[57] Nearly 80% of all cancers are diagnosed in patients age 55 and older.[57] About 12% of childhood deaths in USA are due to cancer.[55] The overall cost of medical treatment of cancer in 2007 is reported to be $89 billion in USA alone.[53] More than 100 different types of cancers were reported so far.[58] Among the different types of cancers, lung cancer is reported to be the major type of cancer occurring worldwide.[59] In the year 2000 itself 1.1million deaths were reported to be due to lung cancer.[59] The Prostate, lung and colorectal cancers are the three most commonly diagnosed cancers in men living in developed countries.[55] Breast, colorectal and lung cancer are the three most commonly occurring cancer in women living in developed countries.[55] In developing countries the most commonly diagnosed cancers are lung, stomach and liver in men, and breast, cervix and stomach in women.[55] Leukemia and lymphoma are reported to be the most commonly occurring cancer in children in most countries.[60] The leading sites producing cancer mortality for males and females in the year 2007 were given in Figure 1.2 and Figure 1.3
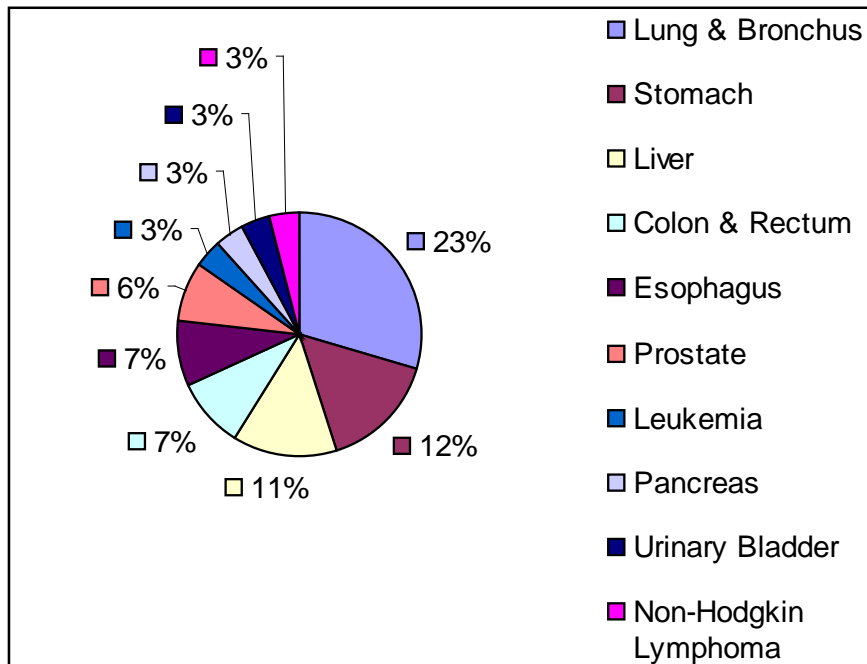
Figure 1.2 Leading sites producing cancer mortality reported worldwide for males. Graph created using the data reported in ref. 55.
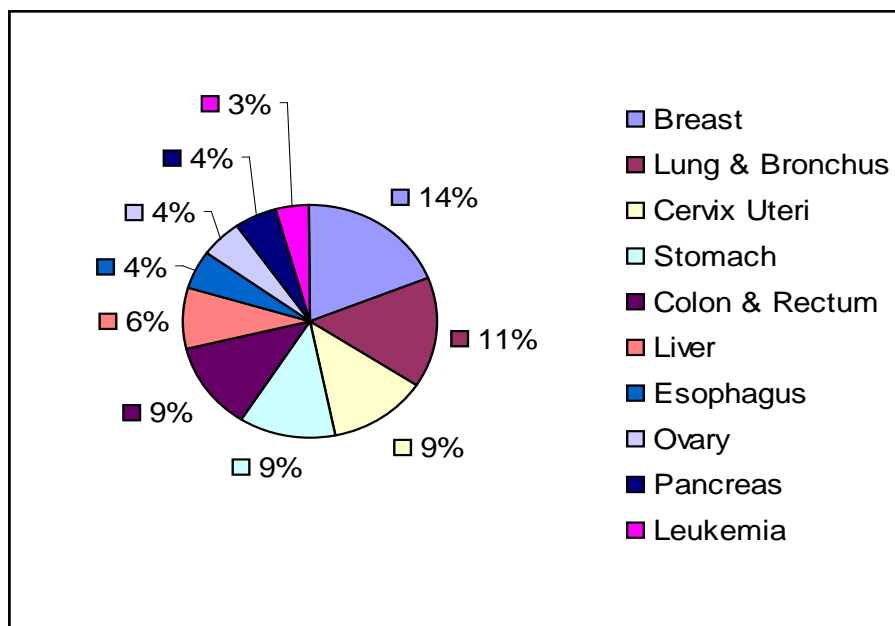
Figure 1.3 Leading sites producing cancer mortality reported worldwide for females. Graph created using the data reported in ref. 55.

The mechanism by which cancer is produced in an organism is reported to be multifactorial, multistage and complex process.[61,62] The current dominant view regarding the molecular basis of human cancer is that it is due to the "accumulation of multiple mutations within genes of a single cell that drives the neoplastic

9

transformation, ultimately leading to tumorigenesis".[63]

The molecular mechanisms involved in carcinogenesis and the pattern of spread of cells from the primary site is reported to differ in different types of cancers.[58] In spite of these differences, cancer cells were reported to exhibit some characteristic features. Hanahan *et al.* have defined six hallmarks that are present in most of the cancers.[64] The six hallmarks of cancer are: "1. Acquiring capability for self-sufficiency in growth signals 2. Insensitivity to anti-growth signals, 3. Evasion of apoptotic cell death, 4. Acquiring capability for limitless replicative potential, 5. Angiogenesis or the ability to form new blood vessels, and 6. Tissue invasion and metastasis".[64]

The conventional chemotherapy is the main treatment option available to the cancer patients.[65] Surgery and radiation therapy are two other most commonly used treatment options for localized tumors.[65] The drugs which are active against a cancer of one tissue are reported to be mostly not effective against cancer of other tissues.[66] Most of the chemotherapeutic drugs non-specifically target almost all dividing cells and causes many toxic side-effects.[67] They have very low therapeutic index. The side-effects associated with cancer treatment as well as cost of treatment makes it a dangerous disease. Selectively targeting cancer cells without inhibiting the normal cells is an area of great interest in anti-cancer drug design.[66] Very few cancers can be completely cured.[66] In most of the cases the goal of the treatment is to achieve remissions and increase the survival rate of the patients.[68] The rise of resistance to the currently available drugs treatment is another major hurdle in cancer therapy.[69,70]

For many types of cancer we still need better treatment options.[71] There is also a wide variation in the global survival rate for different types of cancer. For example, there is a considerable improvement in the five year survival rate in the case of breast, colorectal and prostate cancer in developed nations.[72] But for the same types of cancer the survival rate is lower in developing or underdeveloped countries.[72] The five year survival rate for lung cancer patients in U.S.A is 15% only.[73] The 5 year survival rate for pancreatic cancer is less than 5%.[74] Once the main line treatment fails, currently there are no available drugs for some cancers like pancreatic cancer

and the five year survival rate is low.[75] So there is an urgent need to discover new compounds that could be developed as anti-cancer drugs. The newer approach in cancer treatment is to target the altered cell survival pathways which are over expressed in cancer cells.[76] This approach is based on the assumption that targeting over-expressed pathways will be more selective in destroying the cancer cells and will have fewer effects on normal cells.[76] Protein kinases which are important constituents of cellular signal transduction pathways are frequently altered or overexpressed in cancer cells.[77,78] This has led to increased focus on kinases for targeted anti-cancer drug discovery efforts.[79,80]

1.2.2 Protein Kinase

Protein kinases are critical components of cellular signal transduction cascades.[81] There are reported to be over 500 protein kinases in the human genome and they are considered as the second largest group of drug targets.[82,83] Kinases catalyze the transfer of the terminal (or gamma) phosphate group of adenosine triphosphate (ATP) to the specific hydroxyl group of serine, threonine, or tyrosine residues in a protein substrate.[84] This phosphorylation acts as a key event in signal transduction and ultimately helps to transfer the signal from extracellular to intracellular environment.[84,85] The kinase mediated signal transduction plays a key role in a number of important cellular processes like cell growth, cell differentiation, maintaining cytoskeletal integrity and apoptosis.[86] Hence aberrant kinase signaling either from an activating mutation or over expression was reported to play a role in diseases like cancer, inflammation, diabetes neurodegeneration and psoriasis.[87] In cancer cells, the kinase signaling pathways are often altered, resulting in uncontrolled growth and increased capability to invade surrounding tissue.[88] Agents targeting altered kinase pathways are currently being developed as the next generation anti-cancer agents.[89,90] Among different kinase signaling pathways, altered or constitutively activated phosphoinositide 3-kinase (PI3K) pathway is reported in many human cancers.[91,92,93] Such aberrant PI3K pathway is implicated in tumor development and progression and also in tumor's response to cancer

treatment.[94] 3-Phosphoinositide-dependent protein kinase-1 (PDK1) is a pivotal kinase for the PI3K pathway[95] and is an attractive target for developing anticancer therapeutics.[96]

1.2.3 3-Phosphoinositide dependent kinase-1 (PDK1)

PDK1 is a 63 kDa Serine/Threonine kinase and is ubiquitously expressed in human tissues.[97] It is a 556 amino acid containing enzyme.[98] PDK1 possesses an N-terminal kinase domain and C-terminal pleckstrin homology (PH) domain.[98] PH domain is involved in the interaction with the phosphatidyl-inositol 3,4,5 triphosphate (PtdIns(3,4,5)P$_3$) in the membrane, and kinase domain is involved in activation of other downstream kinases.[99] PDK1 is noted as the master regulator of c<u>A</u>MP-dependent, c<u>G</u>MP-dependent, protein kinase <u>C</u> (AGC) kinase family.[100] It is known to activate at least 23 downstream AGC kinases.[101] These include protein kinase B (PKB, also known as Akt),[102,103] p70 ribosomal S6 kinase (S6K), [104] serum- and glucocorticoid-induced protein kinase (SGK)[105]and protein kinase C (PKC) isoforms.[106,107]

X-ray crystal structures are available for PDK1 binding with its inhibitors in the protein data bank (PDB).[108,109,110,111] Like all other protein kinases, the catalytic core of PDK1 has a bilobal domain, with predominantly α-helical C-terminal lobe and an N-terminal lobe consisting mainly of β-sheets.[108] The adenosine triphosphate (ATP) binding region is located in the hinge region between the N- and C-terminal lobes.[108] As in other kinases, the ATP binding site in PDK1 can be divided into adenine region, sugar region, and phosphate region.[112] As PDK1 itself belongs to the same AGC family as its substrates, it needs to be phosphorylated at the activation loop or T-loop.[100] PDK1 is reported to be constitutively active and has an intrinsic ability to phosphorylate its own T-loop at Ser241 residue.[113] This autophosphorylation is found to be mediated by intermolecular reaction.[114] The activation loop is a common feature in kinase superfamily.[115] PDK1 kinase domain has hydrophobic motif (HM) pocket or PDK1 interacting fragment (PIF) pocket.[108] AGC kinases were reported to be activated by phosphorylation in two regions

namely T-loop region and hydrophobic motif region.[97] PDK1 is different from other AGC kinases in that it does not have a separate hydrophobic motif.[100] Studies have shown that PDK1 is the major T-loop kinase of other AGC kinases.[116]

PDK1 function is reported to be regulated by substrate conformation and sub-cellular location.[95] The various kinases involved in PDK1 signaling pathway is given in Figure 1.5. Activation of upstream PI3K leads to the synthesis of phosphatidyl-inositol 3,4,5 triphosphate [PtdIns(3,4,5)P$_3$]/ phosphatidyl-inositol 3,4 diphosphate [PtdIns(3,4)P$_2$] at the plasma membrane.[97] This results in PKB's and PDK1's PH domain's interaction with PtdIns(3,4,5)P3/PtdIns(3,4)P2 and thus causes translocation to membrane and co-localization.[99] The interaction of the PH domain with PtdIns(3,4,5)P$_3$ relieves autoinhibition of PKB and PDK1 phosphorylates PKB on Thr308.[95] Substrates other than PKB was reported to interact with PIF pocket of PDK1 using their phosphorylated hydrophobic motif and then subsequently gets activated by PDK1.[117] Thus PDK1 acts as the sensor of substrate conformation and is regulated by substrate conformation.
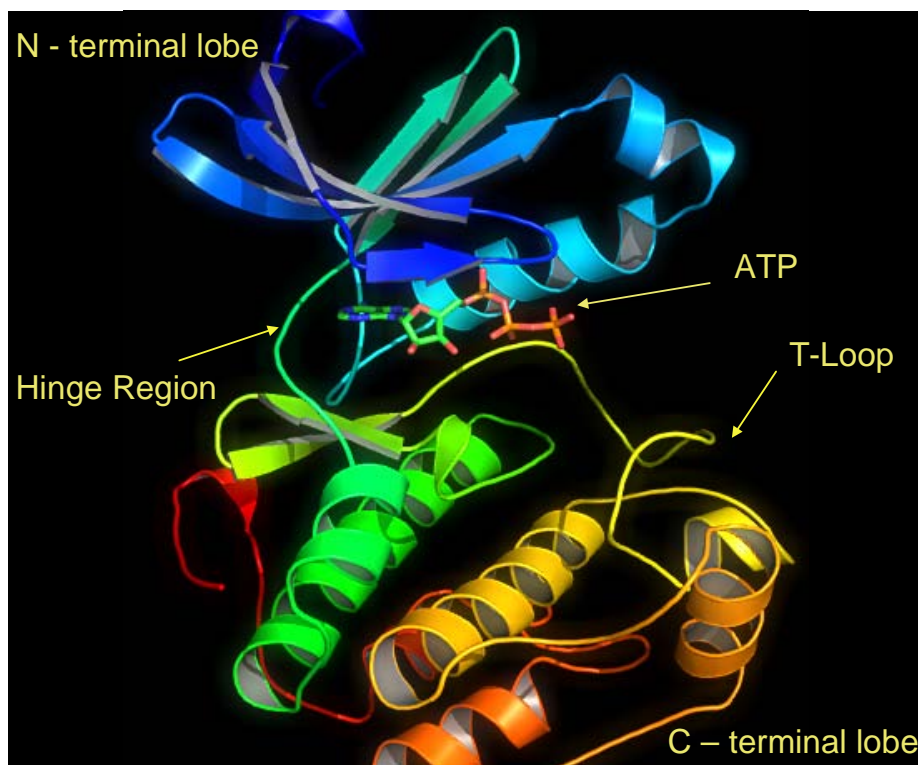
Figure 1.4 Structure of PDK1 kinase domain with ATP molecule (PDB ID: 2BIY)[101]
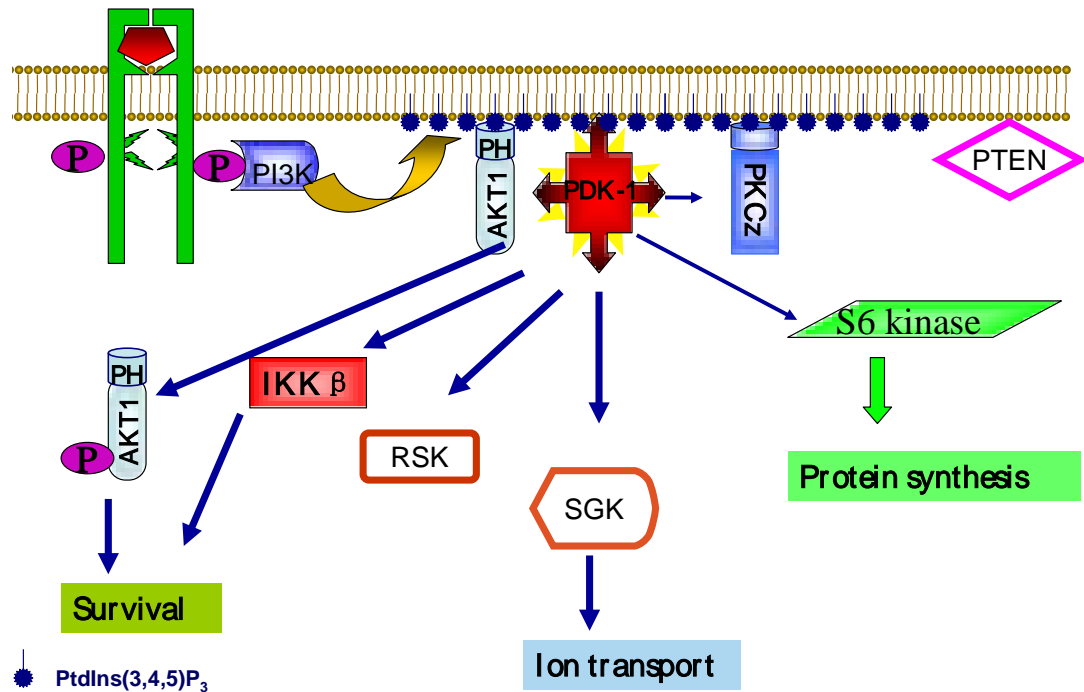
Figure 1.5 PDK1 signaling pathway

1.2.4 PDK1 as anti-cancer target

The tumor suppressor, Phosphatase and Tensin homologue deleted on chromosome Ten (PTEN), acts as the phosphatidylinositol 3-phosphatase[118] and thus down regulates the PDK1-mediated growth and signaling pathway. Mutations in PTEN resulting in elevated levels of PtdIns(3,4,5)P$_3$ was reported in many human cancers and in such cases inhibition of PDK1 is expected to mimic the tumor suppressing activity of PTEN.[119] Moreover, most of the downstream proteins of PDK1 are also implicated in cancers.[95,120,121,122,123] As PDK1 is the master regulator of such AGC kinases, it is in a unique position to control all these proteins.

Accumulating pharmacologic and genetic evidences supports the potential role of PDK1 as anticancer target.[124,125] It has been reported that over-expression of PDK1 in mammary cells resulted in their transformation *in vitro* and tumor formation *in vivo*.[125] Elevated levels of PDK1 phosphorylation was also reported in metastasized breast tumors.[126] Studies also showed that PDK1 phosphorylation is associated with downstream kinase activation and together correlated with invasive breast tumor.[126] PDK1 is over-expressed in pre-malignant and low to high grade

ovarian carcinomas.[127] PDK1/AKT pathway is also reported to be activated in Rhabdomyosarcoma (RMS).[128] Targeting PDK1 with anti-sense oligonucleotides has showed a marked reduction of cell proliferation and survival and also an increased rate of apoptosis than that observed in PI3K or PKB inhibition.[124] PDK1 was also reported as a potential target for sensitizing breast cancer cells to chemotherapeutic agents.[129] Knock down of PDK1 was recently reported to enhance the anti-tumor effect of EGFR inhibitor.[130] These studies also show that PDK1 inhibitors will help to improve the clinical response to EGFR inhibitors. PDK1 is also reported to play a role in the motility of cancer cells.[131] Lack of PDK1 is reported to cause inhibition of cell proliferation in mouse embryonic fibroblasts (MEFs).[132] A recent study shows that PDK1 mediates cell survival through another distinct IκB kinase-β (IKKB)/NFkB pathway in addition to AKT pathway.[133] PDK1-hypomorphic mice which express only 10% of normal levels of PDK1 were reported to be viable and fertile.[134] This finding shows that inhibition of PDK1 can be achieved without severe toxicity. A more recent study using PDK1-hypomorphic mice has shown that the reduced PDK1 expression in PTEN$^{+/-}$ mice markedly protected the animal from a wide range of tumors.[135] Thus PDK1 has become a well validated anticancer target. ATP competitive PDK1 inhibitors will compete with ATP molecule to bind in the PDK1 active site and prevents the transfer of phosphate group from ATP to downstream substrate proteins (Figure 1.6). This will result in the blockade of PDK1 mediated signal transduction. Development of PDK1 inhibitor could lead to development of better treatment options for cancer.
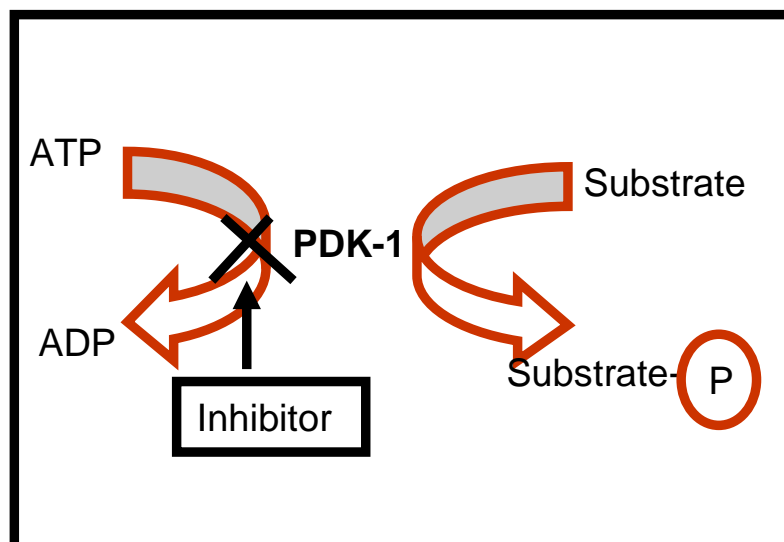
Figure 1.6 Schematic diagram of PDK1 inhibitor action. PDK1 inhibitor block the transfer of phosphate group from ATP molecule to substrate proteins

1.3 Overview of present work

Computational drug design approaches were reported to aid in the drug design process. This approach helps to increase the efficiency of the drug discovery process as well as reduce the experimental work done. We were interested to explore the different aspects of computational drug design. As PDK1 is a well validated anti-cancer target we selected it as our target molecule. The availability of X-ray crystal structure of PDK1 allowed us to explore both the structure-based and ligand-based drug design strategies.

Identification of correct binding mode is very essential for structure-based lead optimization. Chapter 2 details our study of the binding mode of PDK1 with known inhibitors including celecoxib and its derivatives using computational techniques like molecular docking, molecular dynamics (MD) simulation, and molecular mechanics-Poisson-Boltzmann surface area method (MM-PBSA). For celecoxib and its derivatives, we had identified a new binding mode which serves as a better starting point for structure-based drug design and further lead optimization of this series.

Ligand-based drug design strategies can be explored to develop predictive

3D-QSAR models as well as to understand the SAR for a given series. In Chapter 3, our ligand-based study of PDK1 inhibitors belonging to the series of novel indolinone derivatives was discussed. We have developed predictive 3D-QSAR models by exploring ligand-based and receptor-based 3D-QSAR model development strategies. Our models were internally and externally well validated. It could be used in the further lead optimization of this series.

Chapter 4 details our work on virtual screening studies. We have developed a novel hierarchical virtual screening strategy which combines ligand-based and structure-based virtual screening. We have identified new hits in this process. One of the compounds was found to have anti-cancer activity against many cancer cell lines tested at national cancer institute (NCI) screening.

CHAPTER 2

2. IDENTIFICATION OF NEW BINDING MODE AND ITS APPLICATION IN LEAD OPTIMIZATION

Computational drug design strategies are utilized in both lead optimization and lead identification processes. Lead optimization is the process by which a known inhibitor (also called as the lead compound) of a particular target is used as the starting point and further more potent inhibitors were developed. Computational lead optimization strategies can adopt structure-based, ligand-based or combined approaches. As discussed in the introduction, PDK1 is a well known anti-cancer target and there is a need to develop new inhibitors for PDK1. This will help towards the development of new anti-cancer drugs. In this chapter, we discuss the application of structure-based lead optimization strategy against PDK1. Celecoxib and its derivatives were used as the lead compounds and different sophisticated computational approaches were utilized to understand the binding mode of these inhibitors. The obtained binding mode serves as the starting point for further rational design of new inhibitors. Part of the contents in chapter 2 was adapted from published article, AbdulHameed *et al. J. Phys. Chem. B.* **2006**, 110, 26365-26374.[136]

2.1 Celecoxib and its derivatives as lead molecules in PDK1 inhibitor design

Celecoxib is a selective cyclooxygenase-2 (COX-2) inhibitor and used as a non-steroidal anti-inflammatory drug (NSAID).[137] It is a well-known FDA-approved drug in the market since 1998.[137] Celecoxib (Commercial Name: Celebrex; Company: Pfizer) is the first selective COX-2 inhibitor to be approved in the market.[138] It is approved for the treatment of osteoarthritis and rheumatoid arthritis.[139] Initial randomized clinical trials showed that the celecoxib produced better anti-inflammatory activity than placebo.[138] Clinical trials also showed that the celecoxib was equivalent in efficacy as the conventional NSAIDs and produced less gastrointestinal side-effects than traditional NSAIDs.[138] The gastrointestinal ulcers

and bleeding are some of the severe side effects of traditional NSAIDs.[139] The treatment of conditions like arthritis requires taking medications for a long time. Hence side-effects like ulcers severely affect the patients, particularly the elderly patients. Statistics shows that in United States alone, a total of ~16,000 patients with arthritis died due to GI toxicity of NSAIDs during the year 1997.[140] Nearly 200,000 to 400,000 hospitalizations were due to the GI complications of NSAID therapy. [141] The cost of treating the GI complications due to NSAID therapy is around $0.8 to 1.6 billion/year in United States alone.[141] Hence the discovery of selective COX-2 inhibitors was considered as a significant breakthrough in the treatment of inflammatory conditions. The success of this class of drugs can be understood by the fact that celecoxib is remembered as the fastest-selling drug in history.[142] Despite the unprecedented success of celecoxib in the market, stomach and intestinal ulcers still occur with the use of celecoxib although the incidence is less than with other NSAIDs in short-term studies.[143] Further clinical trials showed that the increased gastrointestinal protection of celecoxib is lost, when the patient is simultaneously taking low-dose aspirin.[144] Moreover the therapeutic uses of selective COX-2 inhibitors were also limited by the occurrence of cardiovascular side-effects.[145] Rofecoxib (brand name: Vioxx) is also a selective COX-2 inhibitor but was withdrawn from the market due its cardiovascular effects.[145] A boxed warning has been added to the celecoxib label highlighting the potential of cardiovascular risk associated with the use of this drug. [146] Researchers have identified that the cardiovascular side-effect is not related to a particular drug but a class effect that arises from the inhibition of cyclooxygenase pathway.[147] Cyclooxygenase is a key enzyme in the prostaglandin (PG) bio-synthesis pathway. By inhibiting COX-2, the functions of all downstream PG synthases are blocked, including prostacyclin synthase (PGIS) for the conversion of prostaglandin H_2 (PGH_2) to prostaglandin I_2 (PGI_2).[139] PGI_2 is reported to have a vasodilatory role.[148] Blocking the production of PGI_2 has been reported to play a role in cardiovascular side effects.[148] In addition to the anti-inflammatory effect, celecoxib is in the spotlight for its anti-cancer activity. Epidemiological studies have shown that the use of celecoxib and other NSAIDs

19

were associated with a reduced risk of colon cancer.[149] The use of celecoxib for cancer prevention and treatment has become a very hot topic in the field of cancer research.[150,151,152,153,154,155,156,157] Celecoxib was reported to cause significant reduction in the number of colorectal polyps.[158] It is the only NSAID that has been approved for the familial adenomatous polyposis treatment.[159] Overexpression of COX-2 is reported in variety of malignancies.[160] Inhibition of COX-2 by selective inhibitors may account for the anti-carcinogenic activity of these compounds.[137] But there is a growing body of evidence to show the COX-2 independent anti-cancer activity of NSAIDs. For example, studies have shown that celecoxib has apoptotic effect which is independent of COX-2 inhibitory activity,[161] implying that the well-known anticancer activity of celecoxib is not due to the inhibition of COX-2. Different COX-2 inhibitors have the same ability to inhibit COX-2 but have different apoptosis-inducing activity.[162] Celecoxib produces rapid induction of apoptosis.[162] Other potent COX-2 inhibitors like rofecoxib and DuP697 is reported to have very low apoptosis inducing activity.[162] This shows that the apoptosis activity may be independent of COX-2 inhibitory activity. The cardio-vascular side effect is considered as a class effect whereas the beneficial anti-cancer activity is not considered as a class effect of coxibs.[159] Experiments on the effect of selective COX-2 inhibitors in COX-2 overexpressing (HCA-7) and COX-2 negative (HCT-116) human colon cancer cells showed that only celecoxib has anti-proliferative effect.[159] This study showed that related coxibs like valdecoxib, etoricoxib and rofecoxib has very weak or no anti-proliferative effects in vitro and in vivo.[159] Celecoxib is reported to produce similar apoptosis effect in BxPC-3 pancreatic carcinoma cell line which does not express COX-2 compared to MIA-PaCA-2 cell line in which COX-2 is highly expressed.[163] Studies have shown that anti-sense mediated depletion of COX-2 did not produce apoptosis in prostate cancer cells.[161] It was also identified that the sensitivity of COX-2 inhibitor induced apoptosis is independent of COX-2 expression.[161] Further, a systematic structure-activity relationship (SAR) study was performed to identify the functional groups essential for apoptotic activity of celecoxib and similar compounds.[162] In the

work by Zhu *et al.* the molecule was divided into 3 regions and the structural requirements for apoptosis inducing activity and COX-2 inhibition was explored.[162] Through this detailed SAR analysis it was proved that the structural requirement for apoptosis induction is different from COX-2 inhibitory activity. Another study has shown that celecoxib inhibits PDK1 kinase activity.[164] Zhu *et al.*[165] demonstrated that a series of celecoxib derivatives, through certain modification of the aryl group and other groups of celecoxib, are more potent PDK1 inhibitors that have the desired anticancer activity, but these celecoxib derivatives do not inhibit COX-2. Without inhibiting COX-2, these compounds are not expected to have the common side effects of the COX-2 inhibitors. Further structural modifications of celecoxib could eventually lead to more potent PDK1 inhibitors as potentially more efficient anticancer drugs that do not have the side effects of celecoxib. According to Zhu *et al.*,[165] celecoxib has an IC_{50} value of 48 μM, whereas a series of celecoxib derivatives have IC_{50} values at the low micromolar, for inhibiting PDK1 kinase activity.

To further develop a novel derivative of celecoxib with a lower IC_{50} value for PDK1, one first needs to understand how celecoxib and its known derivatives bind with PDK1. X-ray crystal structures are available for PDK1 binding with other PDK1 inhibitors in the active site.[108, 109, 110, 111] To begin a structure-based study, first we have to understand the nature of the active site. As mentioned in chapter 1, the catalytic core of PDK1 has a bilobal domain, with predominantly α-helical C-terminal lobe and an N-terminal lobe consisting mainly of β-sheets.[108] The adenosine triphosphate (ATP) binding region is located in the hinge region between the N- and C-terminal lobes.[108] Based on the interaction of ATP, the ATP binding site in kinase is usually divided into adenine region, sugar region, and phosphate region.[112] In PDK1, as in other kinases, adenine region consists of hydrophobic residues and the adenine moiety also forms the conserved hydrogen bonds with the backbone of hinge region.[108] The sugar region is hydrophilic and consists of acidic residues. The phosphate region is a solvent-exposed area.[166] The ATP binding site of PDK1 is shown in Figure 2.1
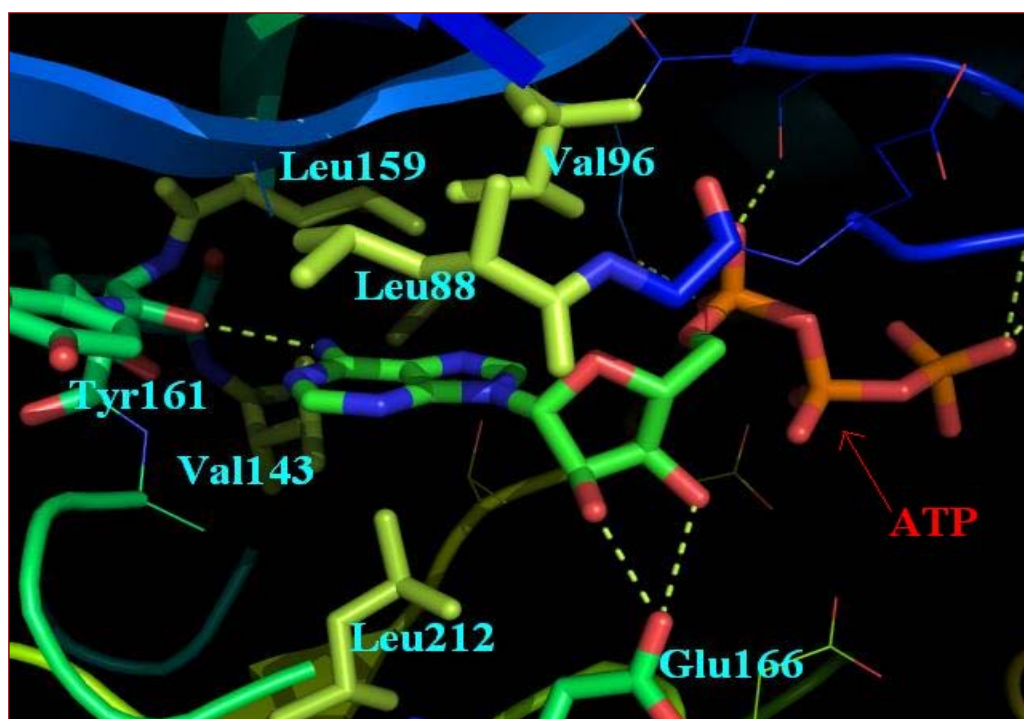
Figure 2.1 The active site of PDK1 kinase domain with bound ATP molecule

By using the available X-ray crystal structures, Zhu *et al.*[165] further proposed a microscopic binding mode for PDK1 binding with celecoxib and its derivatives based on some molecular docking tests with the compounds. Obviously, it is crucial to know whether the proposed microscopic binding mode is reasonable or not for future rational drug design targeting PDK1. A reasonable microscopic binding mode for PDK1 binding with celecoxib and its derivatives could guide rational design of significantly more potent PDK1 inhibitors. As discussed in the chapter 1, molecular docking methods are now being routinely used to predict protein-ligand binding modes.[167] Some important shortcomings of the currently used docking programs are associated with the difficulty of reliably accounting for the solvation/desolvation effects on and entropic contributions to the protein-ligand binding.[167] In addition, it is also difficult to account for the effects of protein dynamics on the microscopic binding during the simple docking process. Nevertheless, the protein-ligand binding structures obtained from molecular docking can be considered as reasonable starting structures for molecular dynamics (MD) simulations in water that can more

reasonably account for the solvation effects and the dynamics of the protein-ligand binding. Hence, in further considering the limit of traditional molecular docking methods, we re-examined all of the possible microscopic binding modes for PDK1 binding with celecoxib and its derivatives (Figure 2.2), by using more sophisticated computational modeling techniques, including molecular dynamics (MD) simulations and molecular mechanics-Poisson-Boltzmann surface area (MM-PBSA) calculations,[168] in addition to the molecular docking. The MD simulations allow us to obtain a dynamically stable protein-ligand binding mode associated with a stable MD trajectory. Further, the MD simulations are followed by the MM-PBSA calculations accounting for the contributions of the gas phase interactions, bulk solvent, and entropy to the binding free energies. When a ligand has multiple possible modes of binding with a protein, the relative binding free energies obtained from the MM-PBSA calculations can be used to determine the most favorable binding mode. The MD simulations followed by MM-PBSA calculations have been used to accurately predict the protein-protein and protein-ligand binding free energies for other systems.[168,169]
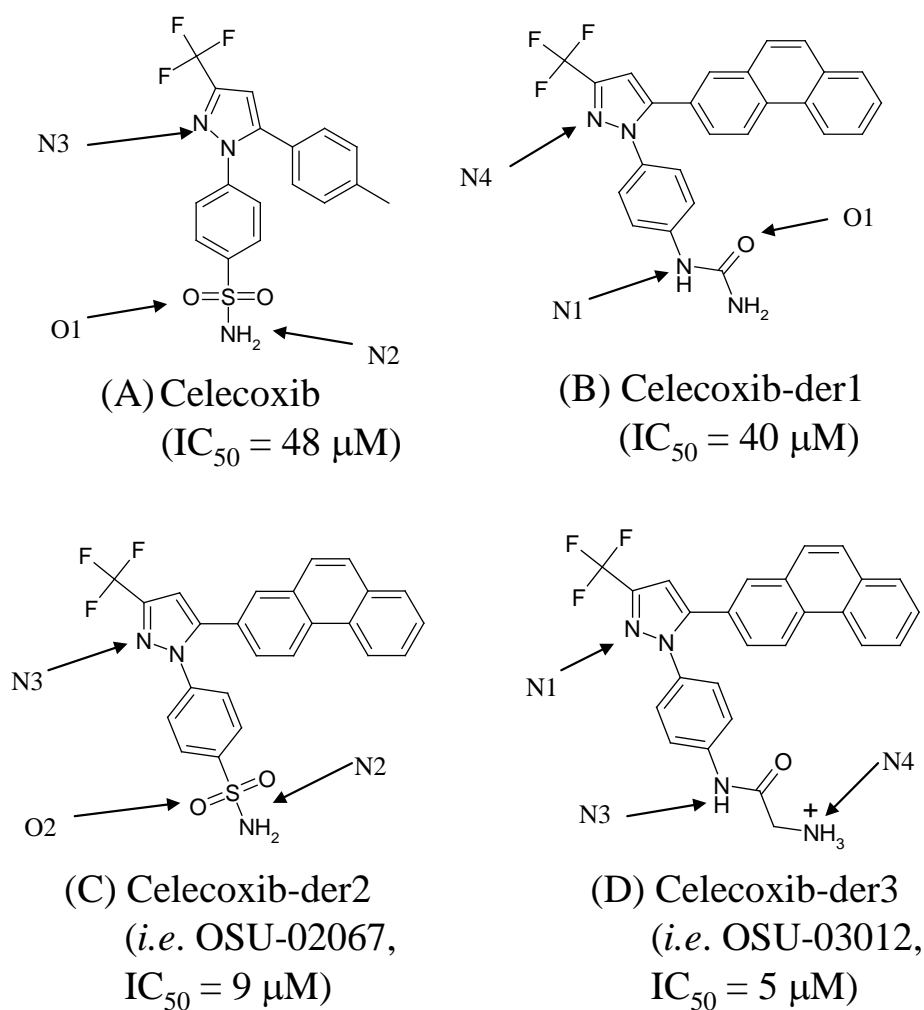
Figure 2.2 Molecular structures of celecoxib and its representative derivatives.

In this study, molecular docking followed by MD simulations and MM-PBSA binding free energy calculations has allowed us to examine various possible microscopic binding modes for a given ligand with PDK1 and to theoretically determine the most favorable binding mode, *i.e.* the one with the lowest binding free energy. Thus, we have identified a new, more favorable binding mode for PDK1 binding with celecoxib and its derivatives. Based on the new binding mode, one can better understand the available SAR data and the calculated binding free energies were all in good agreement with available experimental activity data, providing valuable new insights for future rational design of more potent PDK1 inhibitors.

(A) Staurosporine (IC_{50} = 6 nM) (B) BX-320 (IC_{50} = 39 nM)

(C) LY-333531 (IC_{50} = 0.75 μM) (D) BIM-1 (IC_{50} = 9 μM)
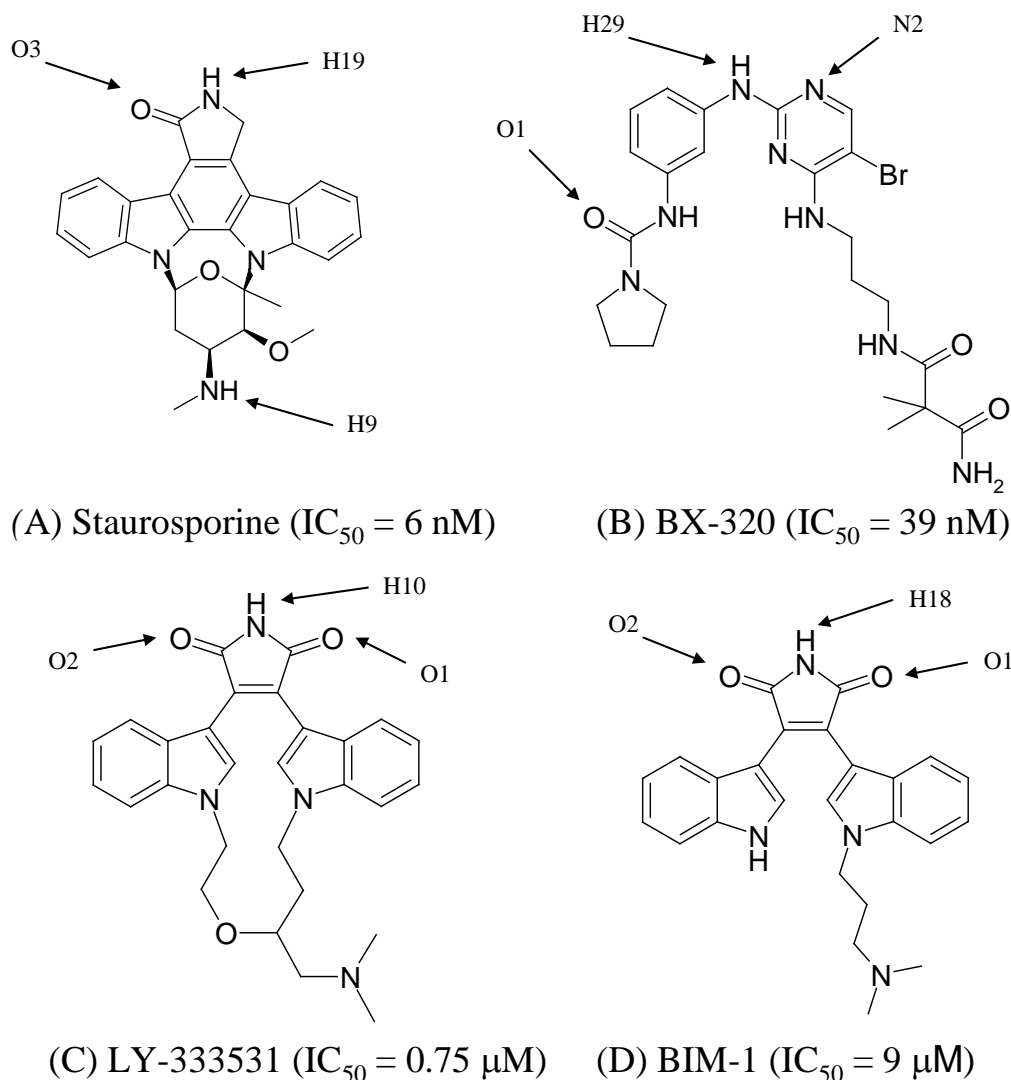
Figure 2.3 Molecular structures of staurosporine, BX-320, LY-333531, and BIM-1.

2.2 Computational Methods

2.2.1 Molecular docking

A few X-ray crystal structures are available for PDK1 binding with inhibitors in the protein data bank (PDB).[170] The starting protein structure used for our molecular docking is the X-ray crystal structure of ATP-bound PDK1 (pdb code: 2BIY),[101] as this structure is more complete as compared with other X-ray crystal structures of PDK1. The missing side chain atoms of residues Gln73, Arg75, Glu153, Lys228, Arg238, Lys304, Glu343, Glu348, and Lys357 were modeled using Sybyl 7.0 program.[171]

Molecular docking was carried out by using two different docking programs in order to obtain different initial binding structures for MD simulations. One is FlexX module of the Tripos software.[172] The FlexX uses incremental docking algorithm called pose clustering.[172] In this program, the ligand molecule is split up into fragments and the core fragment is first placed according to given scoring function[173] The molecule is built up using the tree search, which performs the conformational search of the ligand and calculates its binding score.[172] Only the best conformation was built upon and the others are discarded. As discussed in the introduction, the important steps in docking include generation of conformations, placement in the active site and scoring the poses. FlexX uses MIMUMBA approach for the generation of conformations and spherical interaction surfaces to model the molecular interactions.[48] The interaction type, interacting group and interacting geometry were parameterized to account for interactions such as hydrogen bonds, salt bridges, metal coordination, Π-Π interaction and hydrophobic interaction.[48] The "FlexX score" is used as the scoring function. FlexX score contains "a hydrogen bonding term, a penalty for protein-ligand overlap, a pairwise hydrophobic potential and additional terms for hydrophobic contacts".[48] The receptor description file for our docking with the FlexX module was built upon using the PDK1 structure[110] and the residues within 5 Å around the ligand were defined as the active site. The docking was performed for four representative ligands, *i.e.* celecoxib and its three representative derivatives depicted in Figure 2.2, with different IC_{50} values for PDK1. The molecular structures of these ligands were drawn and energy-minimized by using the Sybyl 7.0 program with the Tripos force field.[174] For docking with each ligand, 30 poses were saved for further examination before the best possible poses were chosen based on the docking scores.

The other program used in our molecular docking is DOCK,[175,176,177] which uses incremental construction algorithm for flexible docking.[178] We also used the DOCK program because a previously reported binding mode[165] could not be obtained from the molecular docking using the FlexX for some protein-ligand binding systems. The DOCK program was used only in these special cases. The

ligand orientations were scored through the use of a force field-based energy scoring function and the top-scored binding structure was selected. Both the FlexX and DOCK programs use incremental docking algorithm for flexible docking of ligands. The major difference is that whereas the FlexX program uses an empirical scoring function based on the work of Bohm[179] and Klebe,[180] the DOCK program uses a grid-based energy scoring function.[181] In the DOCK approach, the binding site is filled with spheres and the center of each sphere was used as a matching point to guide the orientation of ligand in the active site.[182] DOCK uses Anchor and Grow algorithm for incremental construction of the ligands in the active site.[182] The files of protein and ligands required for using the DOCK 5.2 program were prepared by using Sybyl 7.0. The active site spheres were prepared using SPHGEN program.[183] The spheres were selected such that they cover the entire ligand-binding region. These spheres served to orient ligands in the active site.[183] Grid calculations were carried out by using 0.3 Å grid spacing.

2.2.2 Molecular dynamics

The possible protein-ligand binding complexes obtained from the molecular docking with the ligands depicted in Figure 2.2 were used as the initial structures for MD simulations. Besides, we also examined other four ligands that have X-ray crystal structures available for their binding with PDK1, *i.e.* Staurosporine (pdb code: 1OKY),[110] BX-320 (pdb code: 1Z5M),[111] LY-333531 (pdb code: 1UU8),[109] and BIM-1 (pdb code: 1UU3),[109] and carried out MD simulations on these four protein-ligand complexes. The residues #230 to #241 were missing in these structures. To complete the protein structures for these four complexes, each of these four incomplete X-ray crystal structures were superimposed with the complete protein structure constructed from the structure in 2BIY. Thus we obtained the complete protein structure binding with Staurosporine, BX-320, LY-333531, and BIM-1. These complete complex structures were also used as the initial structures for MD simulations.

The MD simulations were performed using Sander module of Amber 8

program.[184] The partial atomic charges for the ligand atoms were calculated using the RESP protocol[185] after electrostatic potential calculations at Hartree-Fock (HF) level with 6-31G* basis set using Gaussian 03 program.[186] Each PDK1-ligand binding complex was neutralized by adding counterions (three Cl^- ions) and was solvated in a rectangular box of TIP3P water molecules[187] with a minimum solute wall distance of 10 Å. The solvated systems were carefully equilibrated and fully energy-minimized. These systems were gradually heated from T = 10 K to T = 298.15 K in 35 ps before a production MD simulation run for 1 ns, or longer, until a stable MD trajectory was obtained for each of the simulated systems. The time step used for the MD simulations was 2 fs. Periodic boundary conditions in the NPT ensemble at T = 298.15 K with Berendsen temperature coupling[188] and P = 1 atm with isotropic molecule-based scaling[185] were applied. The SHAKE algorithm[189] was used to fix all covalent bonds containing hydrogen atoms. The non-bonded pair list was updated every 10 steps. The particle mesh Ewald (PME) method[190] was used to treat long-range electrostatic interactions. Restrain was placed on the C-alpha backbone atoms during the MD run. A residue-based cutoff of 12 Å was utilized for the non-covalent interactions. The time-dependent geometric parameters were carefully examined to make sure that we obtained a stable MD trajectory for each simulated protein-ligand binding system. The coordinates of the simulated system were collected every 1 ps during the simulation. The equally distributed 100 snapshots of the simulated structure within the stable MD trajectory were used to perform the MM-PBSA calculations.

2.2.3 Binding free energy calculation

The binding free energies were calculated by using the molecular mechanics-Poisson-Boltzmann surface area (MM-PBSA) free energy calculation method.[168] In the MM-PBSA method, the free energy of the receptor/protein-inhibitor binding, ΔG_{bind}, is obtained from the difference between the free energies of the receptor/protein-ligand complex (G_{cpx}) and the unbound receptor/protein (G_{rec}) and ligand (G_{lig}) as following:

$$\Delta G_{bind} = G_{cpx} - G_{rec} - G_{lig} \tag{2-1}$$

The binding free energy (ΔG_{bind}) was evaluated as a sum of the changes in the molecular mechanical (MM) gas-phase binding energy (ΔE_{MM}), solvation free energy (ΔG_{solv}), and entropic contribution (-TΔS):

$$\Delta G_{bind} = \Delta E_{bind} - T\Delta S \tag{2-2}$$

$$\Delta E_{bind} = \Delta E_{MM} + \Delta G_{sol} \tag{2-3}$$

$$\Delta E_{MM} = \Delta E_{ele} + \Delta E_{vdw} \tag{2-4}$$

$$\Delta G_{sol} = \Delta G_{PB} + \Delta G_{np} \tag{2-5}$$

$$\Delta G_{np} = \gamma * SASA + \beta \tag{2-6}$$

The MM binding energies were calculated with the ANAL module of Amber8 program. Electrostatic solvation free energy was calculated by the finite-difference solution to the Poisson–Boltzmann equation (ΔG_{PB}) as implemented in the Delphi program.[191,192] Used in the solvation calculation was a grid spacing of 0.5 Å, 80%-filled grid box, an exterior dielectric constant of 80, and an interior dielectric constant of 1. Parse radius[193] and Amber charge[194] were used in the Delphi calculations. The radius used for the solvent probe is 1.4 Å. The MSMS program[195] was used to calculate the SASA for the estimation of the non-polar solvation energy (ΔG_{np}) using Eq.(2-6) with the default parameters, $i.e.$ $\gamma = 0.00542$ kcal/Å^2 and $\beta = 0.92$ kcal/mol. Further, the entropic contribution, -TΔS, to the binding free energy was also calculated at T = 298.15 K by using the NMODE module of Amber8 program which is based on a combination of the standard classical statistical formulas[196] and normal mode analysis.[197]

The final ΔE_{MM} and ΔG_{sol} values for each protein-ligand binding mode were taken as the averages of the respective ΔE_{MM} and ΔG_{sol} values calculated for the 100 snapshots. For these energetic calculations, the ligand and the receptor were studied using the snapshots from the complex simulation (without performing further energy minimization); the water molecules and counter ions were stripped away before MM-PBSA analysis. To evaluate –TΔS, a full energy minimization was first performed on each species (receptor, ligand, or complex) prior to the normal mode analysis. As the full energy minimization and normal mode analysis are much more

time-consuming than the single-point energy calculations, we only evaluated the –TΔS values for 10 of the above 100 snapshots (*i.e.* one from each 10 of the above 100 snapshots). The final -TΔS value for each protein-ligand binding mode was taken as the average of the -TΔS values calculated for the 10 snapshots.

2.3 Validation study and identification of new binding mode

2.3.1 Binding with staurosporine, BX-320, LY-333531, and BIM-1.

X-ray crystal structures have been reported for PDK1 binding with some PDK1 inhibitors, including staurosporine, BX-320, LY-333531, and BIM-1 depicted in Figure 2.3. Through MD simulations and MM-PBSA calculations on PDK1 binding with these four inhibitors whose binding modes have been known in the X-ray crystal structures, we tested whether the computational protocol used in this study, *i.e.* MD simulations followed by MM-PBSA calculations, can reasonably predict the PDK1-ligand binding free energy for a given binding mode.

The MD simulation for each of the PDK1-ligand binding structures was quickly stabilized, because there was no any significant structural change during the simulation (see Figure 2.5 for the MD trajectories). The MD-simulated PDK1-ligand binding structures (Figure 2.4) were all essentially the same as the corresponding X-ray crystal structures. We did not note any significant difference between the MD-simulated structures and the corresponding X-ray crystal structures. The energetic results obtained from the MM-PBSA calculations for these binding structures are summarized in Table 2.1, in comparison with available experimental data. As seen in Table 2.1, the calculated binding free energies (ΔG_{bind}) were all in good agreement with the corresponding experimental ΔG_{bind} values (derived from the experimental IC_{50} values reported in literature); the average absolute deviation is ~1.2 kcal/mol. More importantly, the qualitative order of the calculated ΔG_{bind} values was completely consistent with that of the experimental ΔG_{bind} values. The reasonable agreement between the calculated and experimental binding free energies showed that the computational protocol was adequate for predicting PDK1 binding with its ligands.
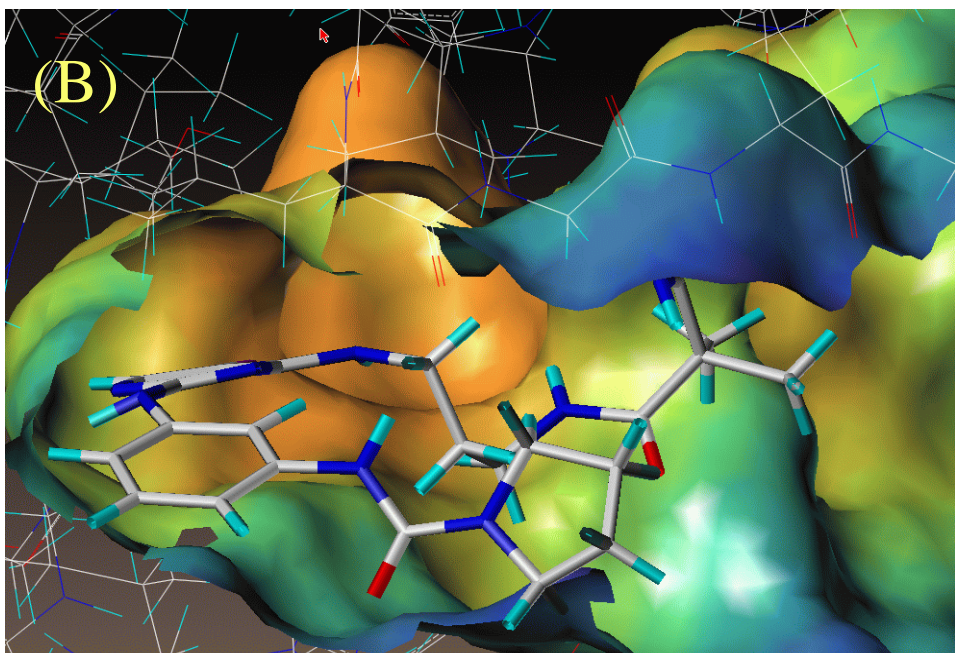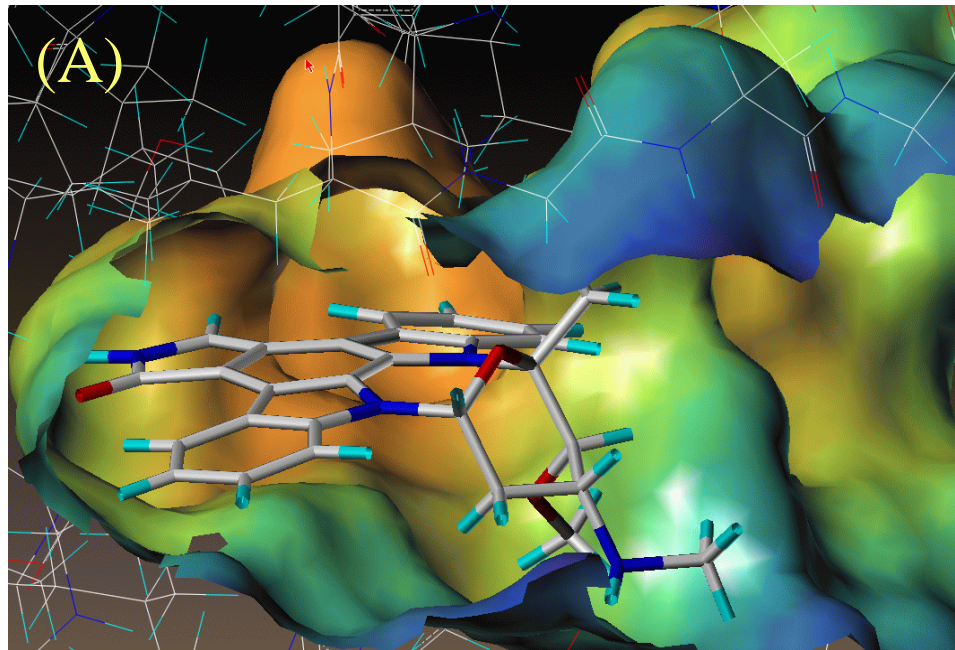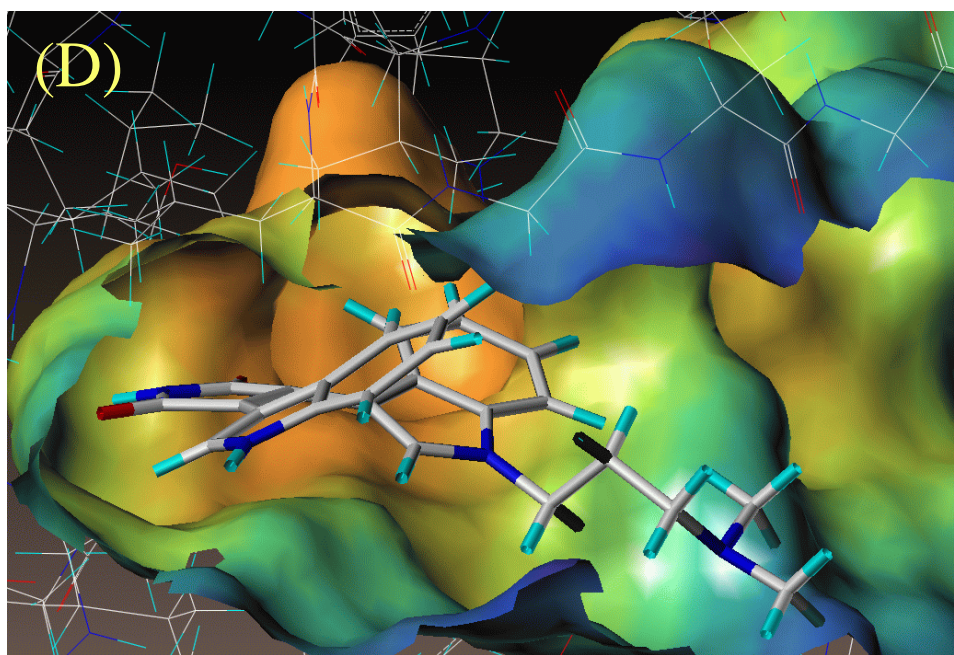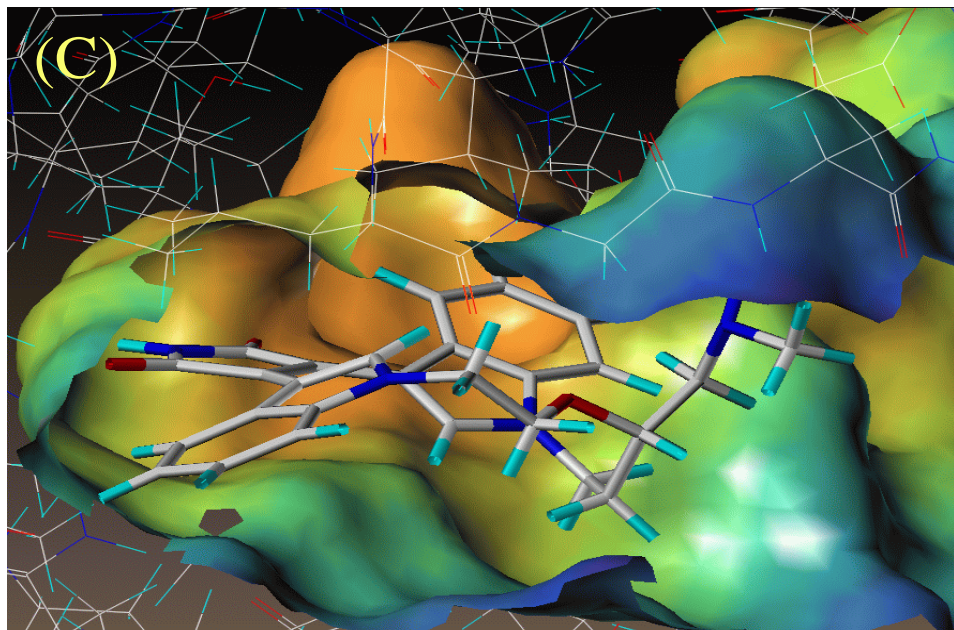
Figure 2.4 MD-simulated structures of PDK1 binding with (A) staurosporine, (B) BX-320, (C) LY-333531, and (D) BIM-1.

(A) Staurosporine with PDK1

(B) BX-320 with PDK1

(C) LY333531 with PDK1
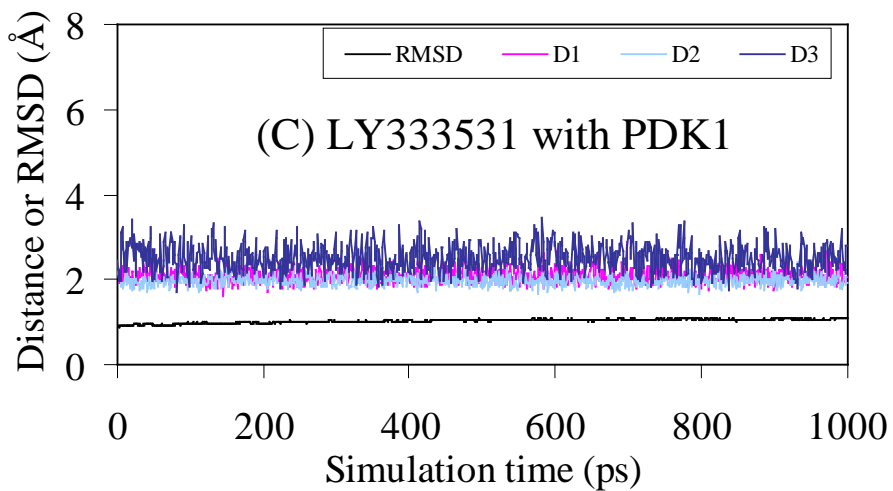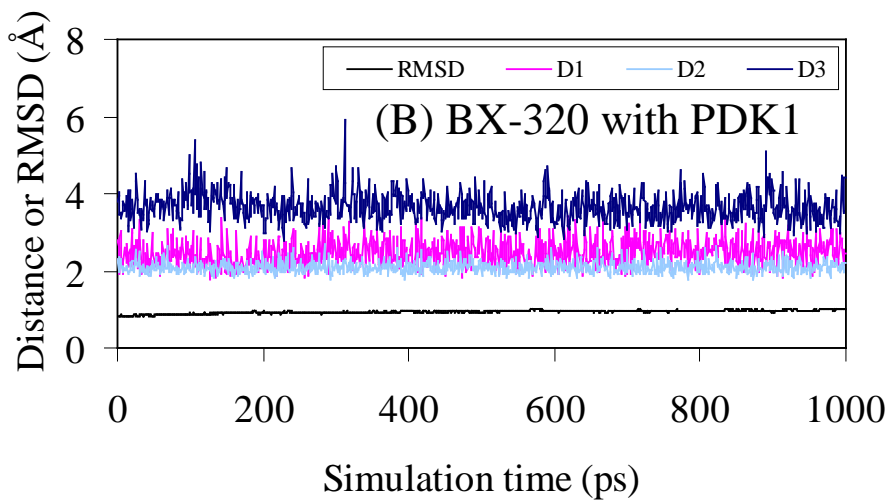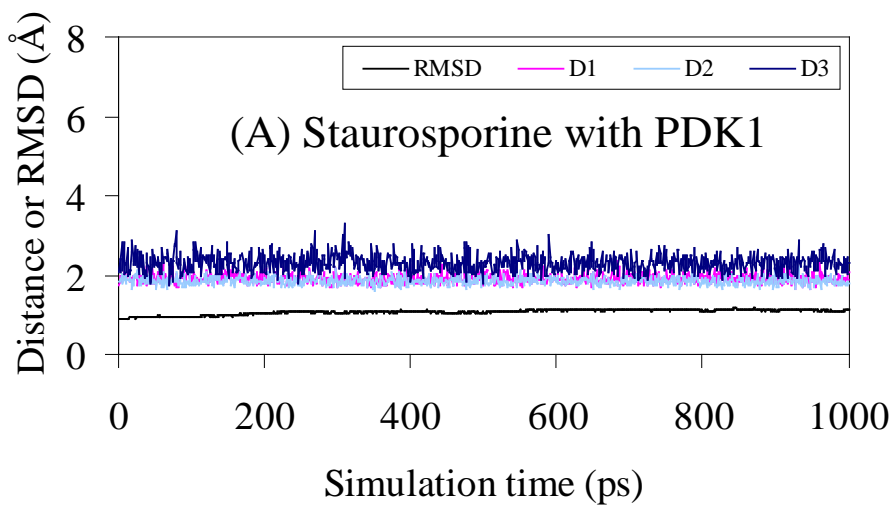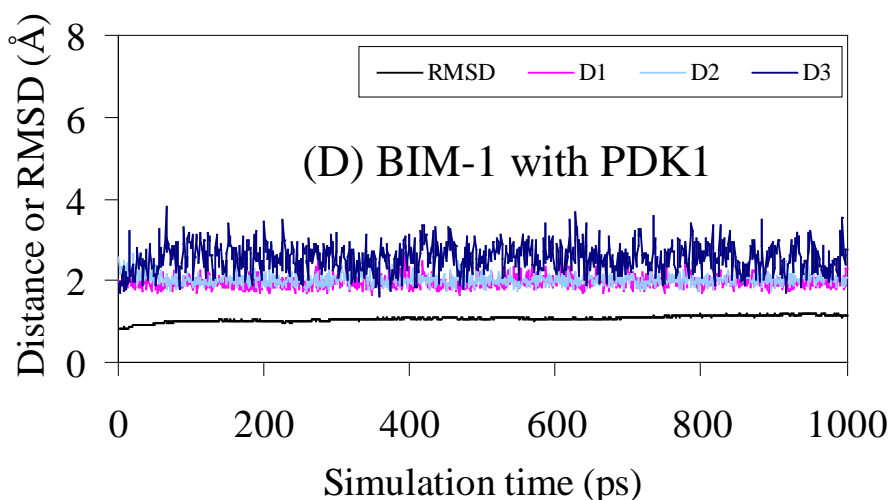
Figure 2.5 Plots of MD-simulated internuclear distances *versus* simulation time for PDK1 binding with staurosporine, BX-320, LY-333531, and BIM-1 (see Figure 2.3 for the structures and the positions of atoms referred here). (A) staurosporine: D1 refers to the distance between O3 atom of Staurosporine and the NH hydrogen of Ala162 backbone, D2 the distance between H19 atom of staurosporine and the carbonyl oxygen of Ser160 backbone, and D3 the distance between H9 atom of staurosporine and the carbonyl oxygen of Glu209 backbone. (B) BX-320: D1 refers to the distance between H29 atom of BX-320 and the carbonyl oxygen of Ala162 backbone, D2 the distance between N2 atom of BX-320 and the NH hydrogen of Ala162 backbone, and D3 the distance between O1 atom of BX-320 and the NH nitrogen atom of Glu166 backbone. (C) LY-333531: D1 refers to the distance between O2 atom of LY-333531 and the NH hydrogen of Ala162 backbone, D2 the distance between H10 atom of LY-333531 and the carbonyl oxygen of Ser160 backbone, and D3 the distance between O1 atom of LY-333531 and the hydroxyl hydrogen of Thr222 side chain. (D) BIM-1: D1 refers to the distance between O2 atom of BIM-1 and the NH hydrogen of Ala162 backbone, D2 the distance between H18 atom of BIM-1 and the oxygen atom of Ser160 backbone carbonyl group, and D3 the distance between O1 atom of BIM-1 and the hydroxyl hydrogen of Thr222 side chain.

34

Table 2.1 Binding free energies (kcal/mol) calculated at T = 298.15 K and P = 1 atm for PDK1 binding with representative inhibitors in comparison with the corresponding experimental data.[136]

Inhibitor	Calc.[a]				Expt.[b]
	ΔE_{MM}	ΔG_{sol}	$-T\Delta S$	ΔG_{bind}	ΔG_{bind}
Staurosporine	-93.8 (1.4)	60.3 (1.3)	21.7 (2.5)	-11.8 (1.9)	-11.2
BX-320	-127.5 (2.1)	97.3 (2.2)	18.7 (1.9)	-11.5 (2.1)	-10.1
LY-333531	-107.7 (1.4)	77.4 (1.6)	20.1 (1.9)	-10.2 (1.8)	-8.4
BIM-1	-82.8 (1.4)	59.0 (1.8)	16.1 (2.4)	-7.7 (2.1)	-6.9

[a] The MM-PBSA calculations were performed on 100 snapshots along a stable MD trajectory for each PDK1-inhibitor binding complex. The results given in the table are the average values calculated for the 100 snapshots.

[b] Experimental binding free energies were calculated from the experimental IC_{50} values reported in refs.109, 110 and 111 *via* $\Delta G_{bind} = RT \times \ln K_d \approx RT \times \ln IC_{50}$.

2.3.2 Binding with celecoxib and its derivatives

To understand how PDK1 binds with celecoxib and its derivatives, four representative compounds (depicted in Figure 2.2) were considered whose experimental IC_{50} values for inhibiting PDK1 kinase activity range from 48 μM to 5 μM. The molecular docking revealed two quite different possible binding modes for each of these compounds binding with PDK1 and the scores are given in Table 2.2. For all the compounds FlexX gave a new binding mode. DOCK results showed only the mode which was similar to that proposed by Zhu *et al.*[165] Further MD simulations led to a stable MD trajectory, as seen in Figure 2.10 to 2.13, for each of the binding modes (Figures 2.6 to 2.9). So, we could not simply judge which binding mode is more reasonable without further evaluating the relative binding free energies.
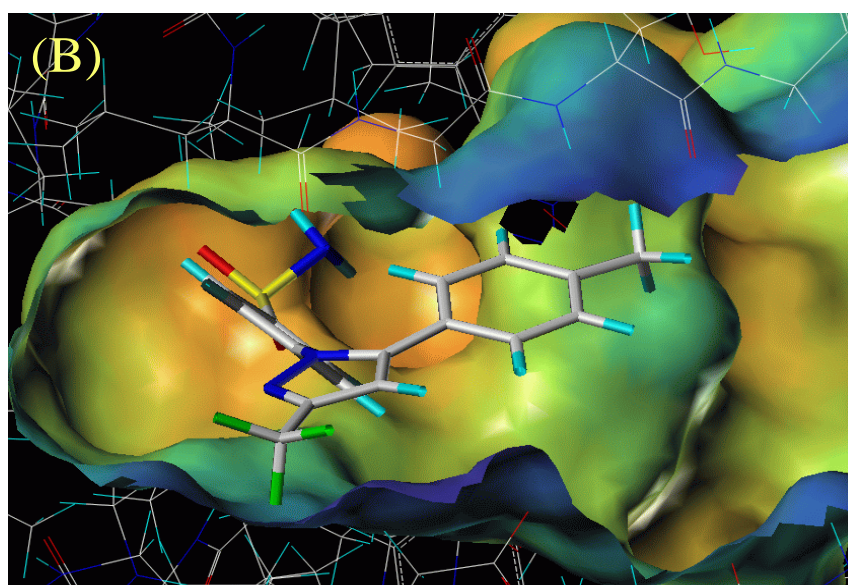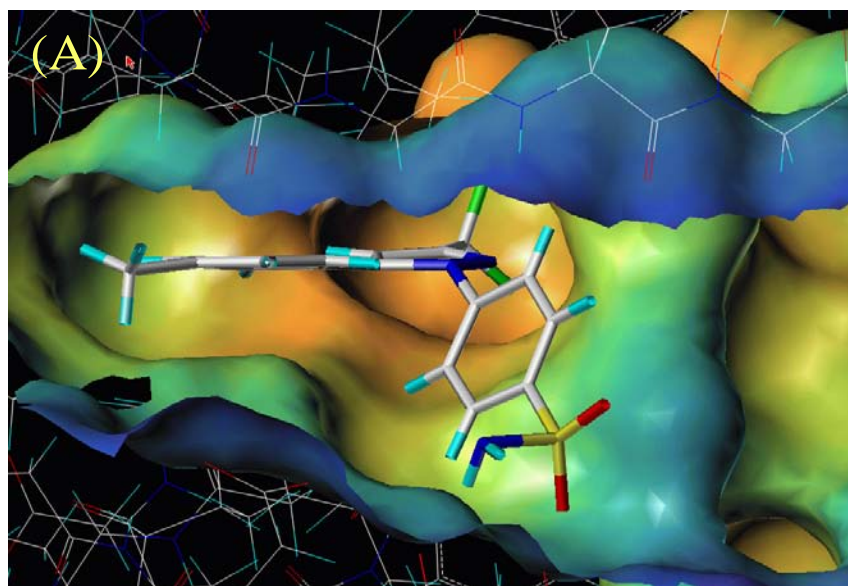
Figure 2.6 MD-simulated structures of PDK1 binding with celecoxib: (A) new binding mode determined in our work (ref. 136); (B) the binding mode similar to that proposed by Zhu *et al*. (ref.165).
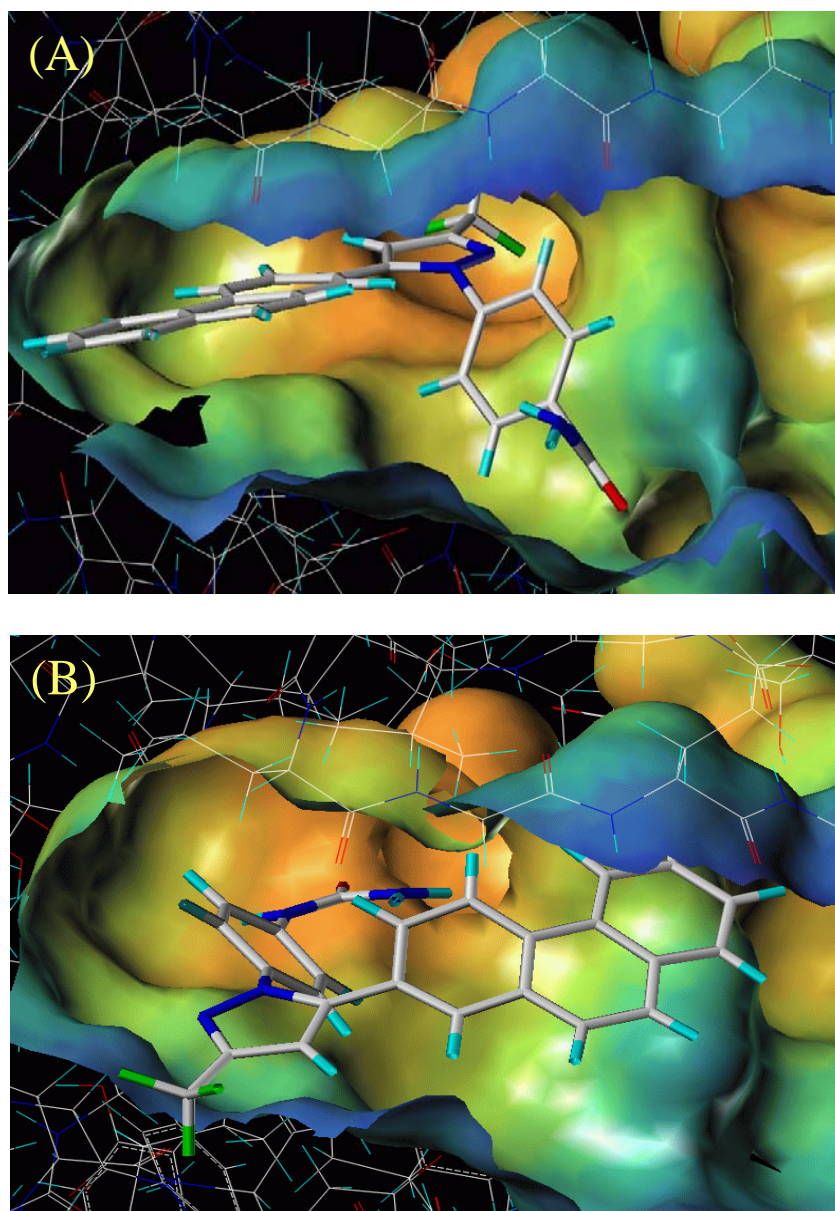
Figure 2.7 MD-simulated structures of PDK1 binding with celecoxib-der1: (A) new binding mode determined in our work (ref. 136); (B) the binding mode similar to that proposed by Zhu *et al*. (ref. 165).
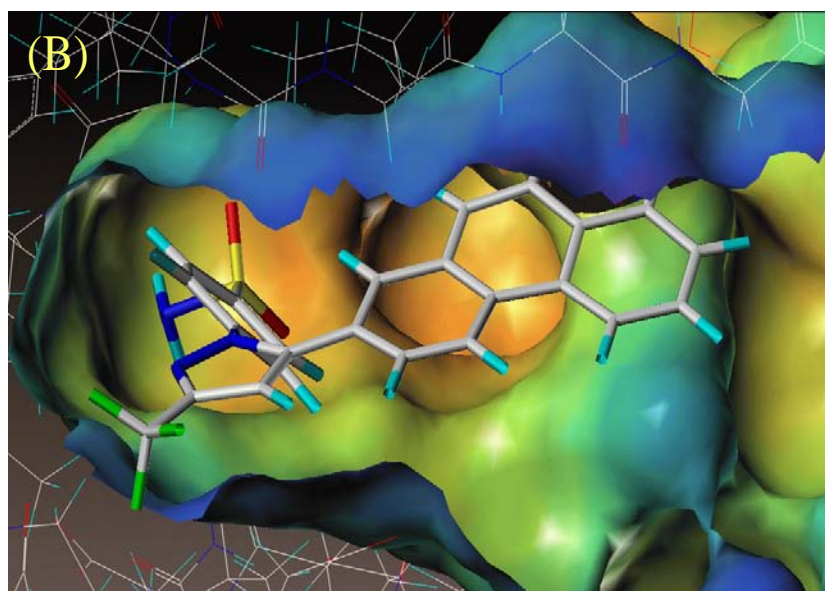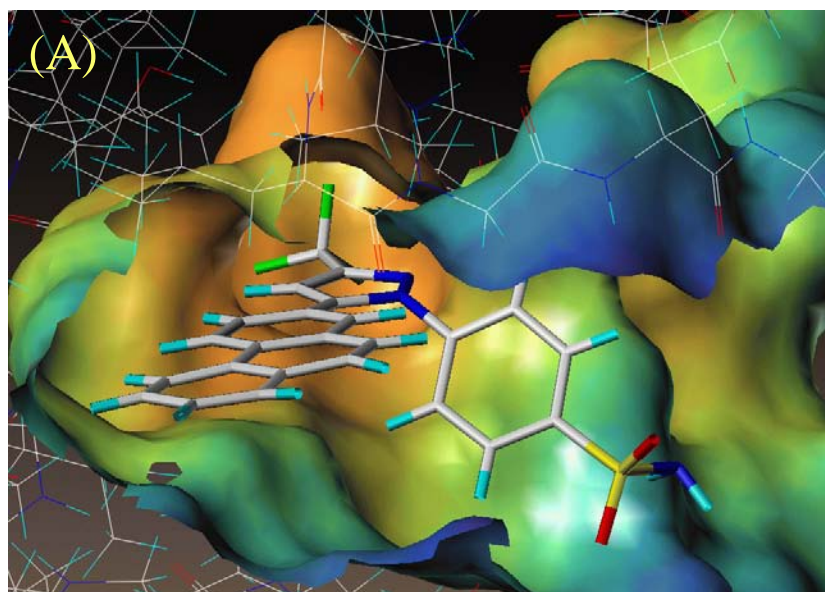
Figure 2.8 MD-simulated structures of PDK1 binding with celecoxib-der2: (A) new binding mode determined in our work (ref. 136); (B) the binding mode similar to that proposed by Zhu *et al*. (ref. 165).
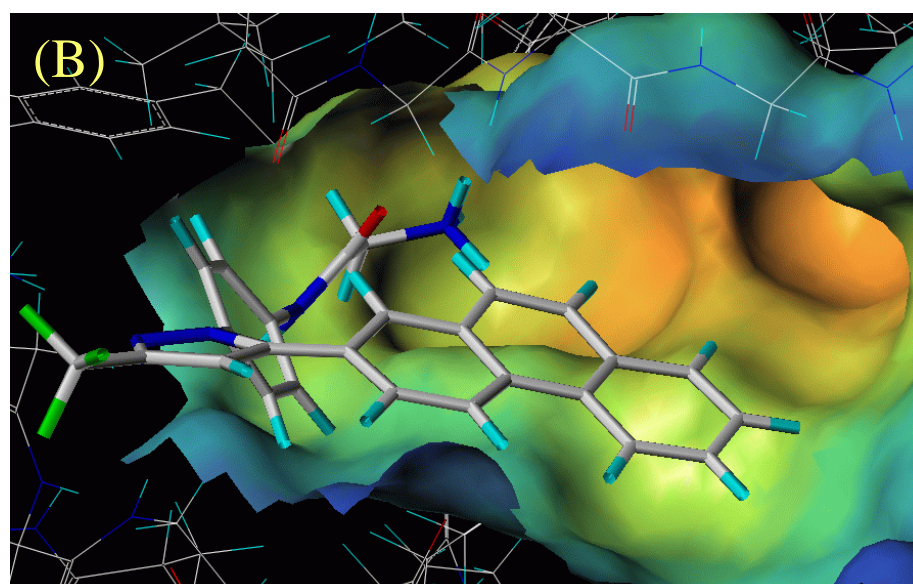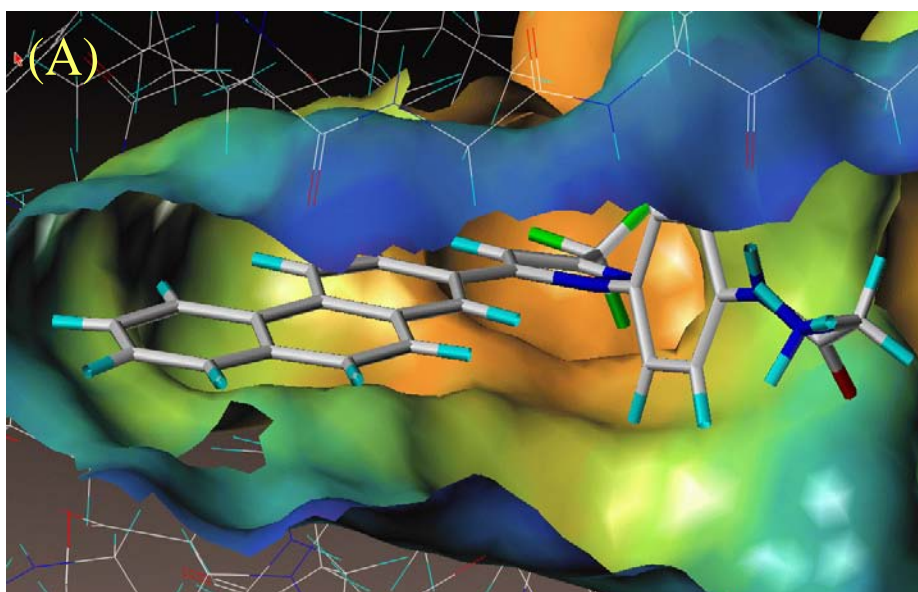
Figure 2.9 MD-simulated structures of PDK1 binding with celecoxib-der3: (A) new binding mode determined in our work (ref.136); (B) the binding mode similar to that proposed by Zhu *et al*. (ref. 165).
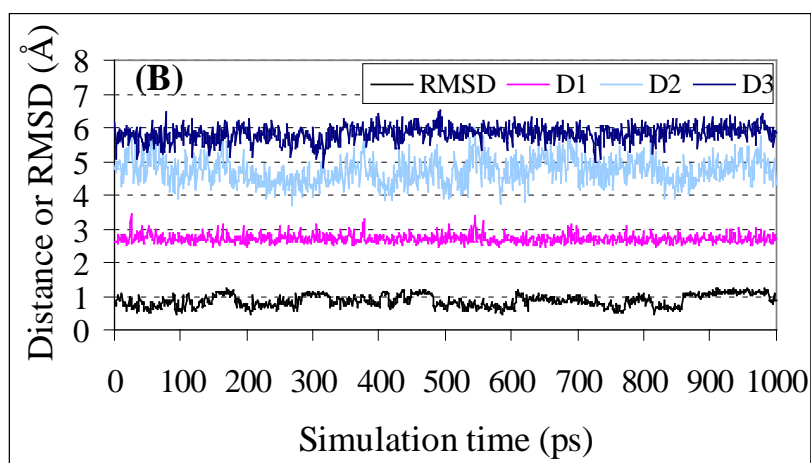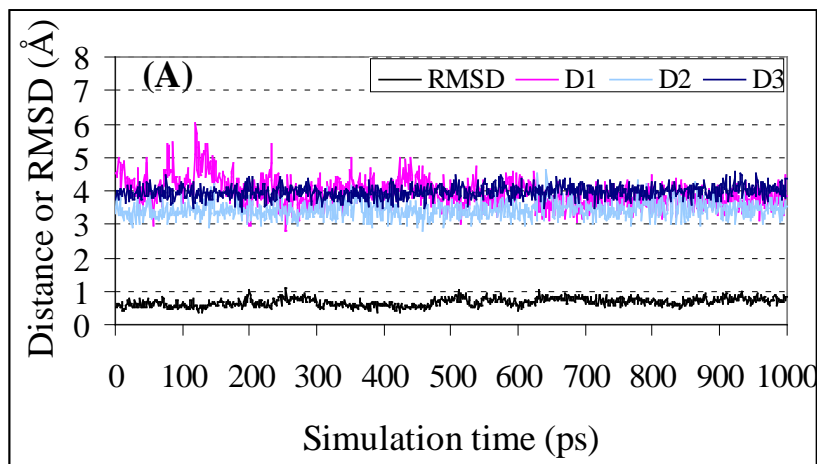
Figure 2.10 Plots of MD-simulated internuclear distances and RMSD for atomic positions of the ligand *versus* simulation time for PDK1 binding with celecoxib-der1 (see Figure 2.2 for the structure). (A) The new binding mode determined in our work for the first time (ref. 136): D1 refers to the distance between N1 atom of celecoxib-der1 and one of the two oxygen atoms of Asp223 side chain carboxyl group, D2 the distance between N4 atom of celecoxib-der1 and the nitrogen atom of Lys111 side chain, and D3 the distance between N4 atom of celecoxib-der1 and the nitrogen atom of Asp223 backbone. (B) The binding mode similar to that proposed by Zhu *et al*. (ref. 165): D1 refers to the distance between O1 atom of celecoxib-der1 and the hydroxyl oxygen of Thr222 side chain, D2 the distance between N1 atom of celecoxib-der1 and the carbonyl oxygen of Ser160 backbone, and D3 the distance between N4 atom of celecoxib-der1 and the nitrogen atom of Glu166 backbone.

40
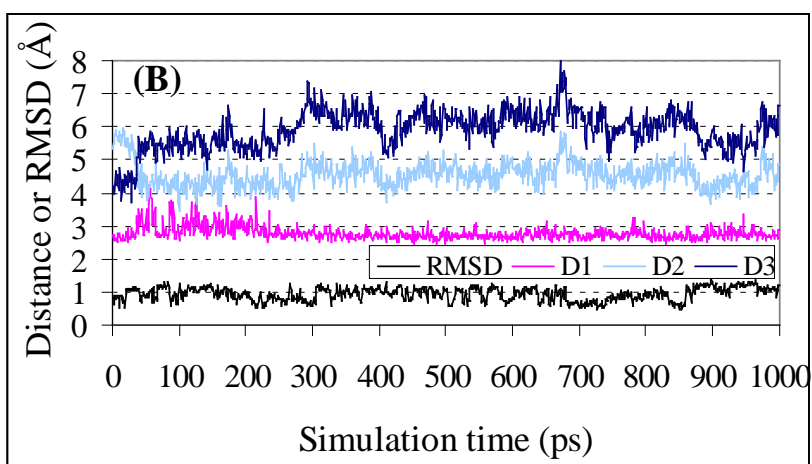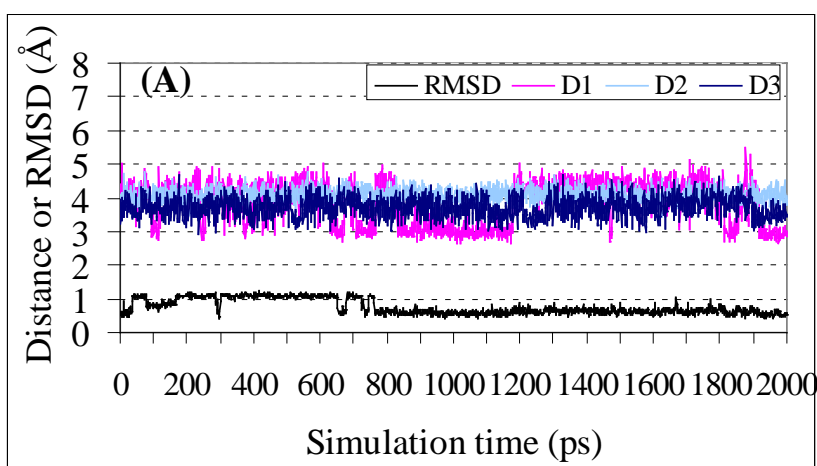
Figure 2.11 Plots of MD-simulated internuclear distances and RMSD for atomic positions of the ligand *versus* simulation time for PDK1 binding with celecoxib (see Figure 2.2 for the structure). (A) The new binding mode determined in our work for the first time (ref. 136): D1 refers to the distance between N2 atom of celecoxib and the carbonyl oxygen of Glu209 backbone, D2 the distance between N3 atom of celecoxib and the NH nitrogen of Asp223 backbone, and D3 the distance between N3 atom of celecoxib and the nitrogen atom of Lys111 side chain. (B) The binding mode similar to that proposed by Zhu *et al*. (ref. 165): D1 refers to the distance between O1 atom of celecoxib and the hydroxyl oxygen of Thr222 side chain, D2 the distance between N3 atom of celecoxib and the NH nitrogen of Glu166 backbone, and D3 the distance between N3 atom of celecoxib and the carbonyl oxygen of Ala162 backbone.

41
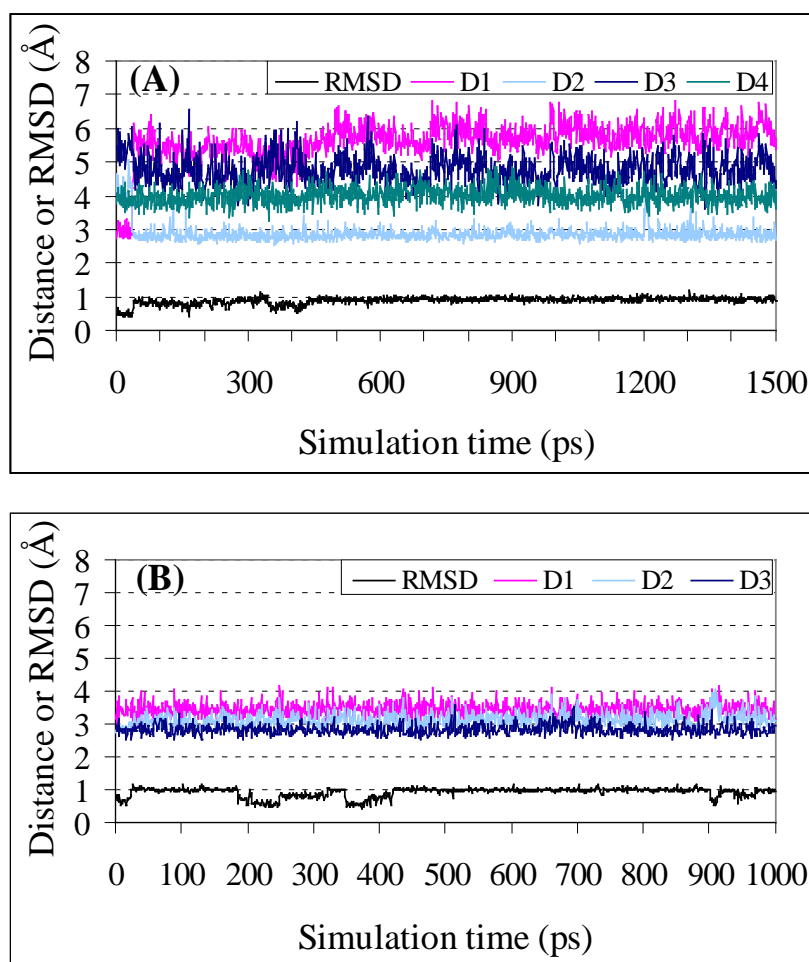
Figure 2.12 Plots of MD-simulated internuclear distances and RMSD for atomic positions of the ligand *versus* simulation time for PDK1 binding with celecoxib-der2 (see Figure 2.2 for the structure). (A) The new binding mode determined in our work for the first time (ref. 136): D1 refers to the distance between N2 atom of celecoxib-der2 and the carbonyl oxygen of Glu209 backbone, D2 the distance between N2 atom of celecoxib-der2 and one of the two oxygen atoms of Asp223 side chain carboxyl group, D3 the distance between O2 atom of celecoxib-der2 and the nitrogen atom of Asn210 backbone, and D4 the distance between N3 atom of celecoxib-der2 and the nitrogen atom of Asp223 backbone. (B) The binding mode similar to that proposed by Zhu *et al.* (ref. 165): D1 refers to the distance between N2 atom of celecoxib-der2 and the nitrogen atom of Ala162 backbone, D2 the distance between N2 atom of celecoxib-der2 and the carbonyl oxygen of Ser160 backbone, and D3 the distance between O2 atom of celecoxib-der2 and the hydroxyl oxygen of Thr222 side chain.
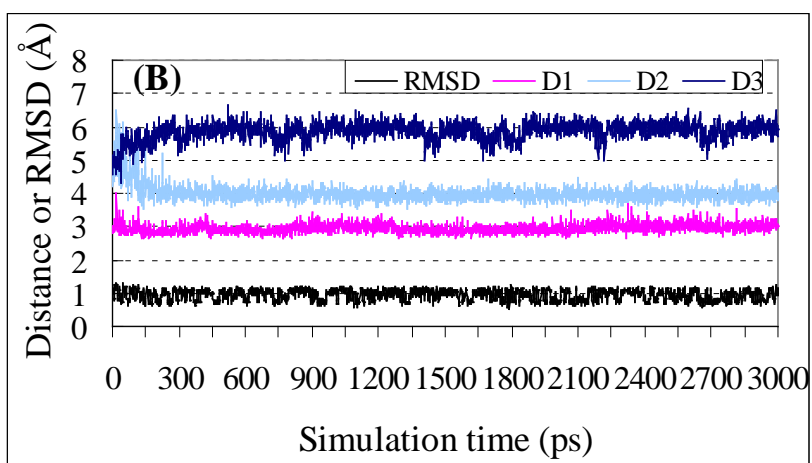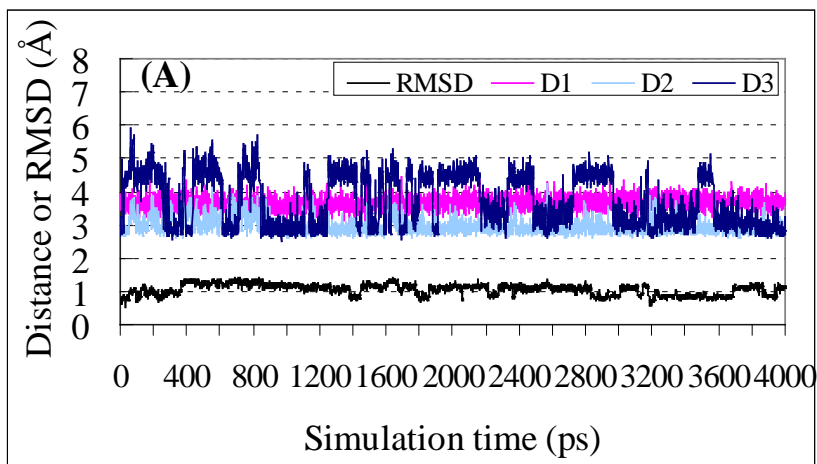
Figure 2.13 Plots of MD-simulated internuclear distances and RMSD for atomic positions of the ligand *versus* simulation time for PDK1 binding with celecoxib-der3 (see Figure 2.1 for the structure). (A) The new binding mode determined in our work for the first time (ref. 136): D1 refers to the distance between N1 atom of celecoxib-der3 and the nitrogen atom of Asp223 backbone, D2 the distance between N3 atom of celecoxib-der3 and the carbonyl oxygen of Glu90 backbone, and D3 the distance between N4 atom of celecoxib-der3 and the carbonyl oxygen atom of Glu90 backbone. (B) The binding mode similar to that proposed by Zhu *et al*. (ref. 165): D1 refers to the distance between N3 atom of celecoxib-der3 and the carbonyl oxygen of Ala162 backbone, D2 the distance between N4 atom of celecoxib-der3 and the nitrogen atom of Ala162 backbone, and D3 the distance between N1 atom of celecoxib-der3 and the nitrogen atom of Glu166 backbone.

Based on the molecular docking followed by the MD simulations, one of the two possible binding modes (see mode "B" depicted in Figures 2.6 to 2.9) is similar to that proposed by Zhu *et al.*[165] For a remarkable feature of this binding mode, the 4-methyl-phenyl group of celecoxib (or the phenanthrene ring of the celecoxib derivatives) was positioned in the ATP sugar/phosphate binding region formed by Gly89, Glu90, Gly91, and Ser94 residues. This hydrophilic environment looks more appropriate to accommodate a hydrophilic functional group of a ligand. The only hydrophobic residue nearby the 4-methyl-phenyl group of celecoxib is Val96 at a distance >3.5 Å. Glu166 is close to both the pyrazole ring and 4-methyl-phenyl of celecoxib. The other binding mode (see mode "A" depicted in Figure 2.6) found in our docking and MD simulations for the first time is totally different. In this newly found binding mode, the 4-methyl-phenyl moiety of celecoxib stays in the adenine pocket and is sandwiched between the hydrophobic residues Val96 and Leu212, the sulfonamide moiety extends into the sugar/phosphate pocket and forms a hydrogen bond with the carbonyl oxygen of Glu209 backbone, and a nitrogen on the pyrazole ring forms a hydrogen bond with the protonated amine group on the Lys111 side chain. This binding mode allows the CF_3 group to stay in a small hydrophobic pocket defined by residues Leu159, Phe157, Leu113, Met134, and Val143. The conserved residues Lys111 and Glu130 are nearby this small pocket. Further, all the celecoxib derivatives examined in this study bind with PDK1 in a similar way as celecoxib, as seen in Figures 2.6 to 2.9. For example, when the 4-methyl-phenyl group of celecoxib is replaced by a phenanthrene ring, celecoxib becomes celecoxib-der2 (see Figure 2.2) and the phenanthrene ring of celecoxib-der2 also stays in the adenine pocket although the phenanthrene ring is surrounded by more hydrophobic residues (including Leu88 in addition to Val96 and Leu212). An interesting difference is that the terminal sulfonamide group of this derivative forms a hydrogen bond with Asp223, instead of Glu209, for the binding with celecoxib. The celecoxib-der1 and celecoxib-der3 retain the same hydrophobic interactions between the phenanthrene ring and the adenine pocket as that found in celecoxib-der2 binding with PDK1. These derivatives show flexibility in their

interactions with the sugar/phosphate pocket. The amino group of celecoxib-der3 forms a hydrogen bond with the carbonyl oxygen of Glu90 backbone, whereas celecoxib-der1 does not have such a strong hydrogen bond with PDK1.

MM-PBSA calculations were performed to calculate the binding free energies for all of the MD-simulated binding structures and the calculated binding free energies are summarized in Table 2.2. We note that the binding free energy for a given binding mode is determined by all of the contributions from the protein-ligand interactions in the gas phase, solvation/desolvation effects, and entropy change. In comparison between the two different modes of PDK1 binding with a given ligand (*i.e.* celecoxib or its derivative), one binding mode might be much more favorable than the other in the gas phase, but the solvation/desolvation effects could make a big difference in changing the relative binding free energies. As seen in Table 2.2, for each of these compounds binding with PDK1, the binding free energy calculated for our newly found binding mode is always significantly lower than that calculated for the other binding mode (similar to that proposed by Zhu *et al.*[165]), which suggests that our newly found binding mode is significantly more favorable. According to the new binding mode, the binding free energies calculated for PDK1 binding with all of these compounds are reasonably close to the corresponding experimental binding free energies (derived from the experimental IC_{50} values reported by Zhu *et al.*[165]); the average absolute deviation is ~0.9 kcal/mol. This agreement further supports that the new binding mode found for PDK1 binding with celecoxib and its derivatives in this study is more reasonable.

Table 2.2 Binding free energies (kcal/mol) calculated at T = 298.15 K and P = 1 atm for PDK1 binding with celecoxib and its derivatives (der1, der2, and der3) in comparison with the corresponding experimental data.[a,136]

Compound	Binding mode	Docking scoring [d]	MM-PBSA[e]				Expt.[f]
			ΔE_{MM}	ΔG_{sol}	$-T\Delta S$	ΔG_{bind}	ΔG_{bind}
Celecoxib	This work[b]	-3.1	-56.5 (1.7)	40.1 (1.6)	11.3 (0.2)	-5.1 (1.1)	-5.9
	Zhu *et al.*[c]	(-43.8)	-59.1 (1.4)	48.5 (1.6)	11.0 (1.9)	0.4 (1.6)	
Celecoxib-der1	This work[b]	-4.9	-71.2 (2.0)	54.8 (2.0)	11.9 (1.2)	-4.5 (1.6)	-6.0
	Zhu *et al.*[c]	(-47.9)	-61.0 (3.1)	50.9 (1.5)	13.2 (1.3)	3.1 (1.9)	
Celecoxib-der2 (*i.e.* OSU-02067)	This work[b]	-5.6	-64.6 (1.7)	45.1 (1.4)	12.1 (2.0)	-7.3 (1.7)	-6.9
	Zhu *et al.*[c]	(-49.4)	-71.5 (1.7)	55.6 (1.5)	11.8 (1.5)	-3.7 (1.4)	
Celecoxib-der3 (*i.e.* OSU-03012)	This work[b]	-6.5	-96.8 (2.9)	78.0 (2.8)	12.6 (2.0)	-6.2 (2.0)	-7.2
	Zhu *et al.*[c]	-4.4	-48.2 (1.9)	45.9 (2.5)	12.1 (0.8)	9.8 (1.4)	

[a] All energies are in kcal/mol.

[b] The new binding mode determined in our study (ref. 136).

[c] The binding mode reported by Zhu *et al*. (ref. 165).

[d] Scores of the molecular docking using the FlexX program. The values in parentheses are scores of the molecular docking using the DOCK program; the DOCK program was used for these binding structures because the molecular docking using the FlexX program did not reveal these binding structures. We note that the DOCK score is the gas phase interaction energy which does not account for the solvation. Thus, the two types of docking scores cannot be compared.

[e] The MM-PBSA calculations were performed on 100 snapshots along a stable MD trajectory for each PDK1-inhibitor binding mode. The results given in the table are the average values calculated for the 100 snapshots. The values in the parentheses are the root-mean-square fluctuations of the calculated energies.

[f] Experimental binding free energies are calculated from IC_{50} (ref 165) *via* $\Delta G_{bind}=$

$RT \ln K_d \approx RT \ln IC_{50.}$

2.3.3 Significance of new binding mode

Further, the new, more favorable binding mode found in this study can help to better understand the substituent effects of the celecoxib derivatives on the previously reported biological activity. [165,162] For example, some experimental trends can be understood easily now when we know that the 4-methyl-phenyl group of celecoxib stays in the hydrophobic adenine pocket. When the 4-methyl-phenyl group of celecoxib is replaced by a phenanthrene group, celecoxib becomes celecoxib-der2 and the IC_{50} value changes from 48 μM to 9 μM.[165] The activity increase from celecoxib to celecoxib-der2 is attributed to the increase of the size of the hydrophobic functional group (from the 4-methyl-phenyl to phenanthrene) interacting with the hydrophobic adenine pocket of PDK1. So, the hydrophobic interaction for PDK1 binding with celecoxib-der2 should be stronger than that with celecoxib. The additional interaction of celecoxib-der2 with Leu88 residue is shown in Figure 2.14. When 4-methyl-phenyl group of celecoxib is replaced by 4-amino-phenyl group, the hydrophobic interaction is expected to decrease and, therefore, its binding affinity with PDK1 is expected to decrease, which is consistent with the experimental observation of the decrease in the apoptotic activity.[162] The molecular structures of celecoxib-der1 and celecoxib-der3 differ only by a -CH_2- group, but show a significant difference in the IC_{50} value (40 μM for celecoxib-der1 *versus* 5 μM for celecoxib-der3), as seen in Figure 2.2. According to our newly found binding mode, the amino group of celecoxib-der3 forms a hydrogen bond with the carbonyl oxygen of Glu90 backbone, whereas the similar hydrogen bonding does not exist in celecoxib-der1 binding with PDK1 (Figure 2.15). The difference in the hydrogen bonding reveals that celecoxib-der3 binds with PDK1 stronger than celecoxib-der1 with PDK1, which explains why the IC_{50} value of celecoxib-der3 is significantly lower than that of celecoxib-der1. Previously reported SAR studies by *Zhu et al.*[162] identified that CF_3 group, tolyl group and sulfonamide group of celecoxib are essential for its apoptosis activity. Removal of CF_3 group was reported

to produce a loss of apoptosis activity of celecoxib. In our newly proposed binding mode, as seen in Figure 2.16, the CF_3 group occupies a small hydrophobic pocket defined by residues Leu159, Phe157, Leu113, Met134, and Val143. The tolyl group occupies the hydrophobic adenine pocket and the sulfonamide group has interaction in the sugar binding region. We also analyzed whether our newly proposed binding mode can explain the difference in apoptotic activity between celecoxib and rofecoxib. In the case of Rofecoxib, the furanone ring will have unfavorable interaction with the hydrophobic residues (Figure 2.16). Due to the absence of sulfonamide group, there is a loss of hydrogen bonding interaction which is observed in celecoxib in sugar pocket. This explains the loss of apoptotic activity for rofecoxib.

Another interesting aspect in our newly proposed binding mode is the absence of hydrogen bonding interaction with the hinge region. To the best of our knowledge, we have identified for the first time a new binding mode of kinase inhibitor without hinge hydrogen bonding interaction. A recent review by Liu *et al.* in 2006 reports that almost all kinase inhibitors has a hydrogen bond interaction with the hinge region.[198] Our newly proposed binding mode will serve as a good starting point for the design of new class of PDK1 inhibitors without hinge hydrogen bonding region. Further experimental tests are needed to analyze whether binding without hinge hydrogen bond interactions confers selectivity for the inhibitors. More recently, after publication of our result, an X-ray crystal structure of cyclin-dependent kinase-2 (CDK2) with inhibitor was deposited in PDB (PDB ID: 2UZD).[199] Richardson *et al.* shows that the new class inhibitors interact with CDK2 without hinge hydrogen bond interaction (Figure 2.17).[199] This X-ray crystal structure further serves as an experimental proof for design of inhibitors without hinge hydrogen bond interaction and it is consistent with our newly proposed binding mode for celecoxib and its derivatives.

Our newly proposed binding mode serves as the starting point for further optimization of this series. *De novo* drug design of new inhibitors can be carried out based on the study of the binding mode and the environment around the inhibitors.

Based on our initial study, we propose that the celecoxib and its derivatives could be modified in two regions as shown in Figure 2.18. Additional interactions could be introduced through substitution of suitable groups either in hinge region or in sugar binding region or both. In kinase inhibitor design, a small hydrophobic pocket deep inside the active site has been exploited for achieving potency and selectivity.[79] "Gate-keeper" is the term used to denote the residues which are present at the entrance of the hydrophobic pocket in kinases. A small gate-keeper residue like threonine has been reported to allow bulky groups to enter the hydrophobic pocket whereas larger gate-keeper residues like phenylalanine restrict access to this pocket.[200] In the case of PDK1, Leu159 is present as the gate-keeper residue. In our newly found binding mode, the CF_3 group of celecoxib and its derivatives occupies the hydrophobic pocket. We propose that further studies involving isosteric replacement of CF_3 group with methyl or iodide groups will help to understand the significance of interaction of this group. Studies using X-ray crystallography or point mutations would also help to further confirm the new binding mode experimentally.
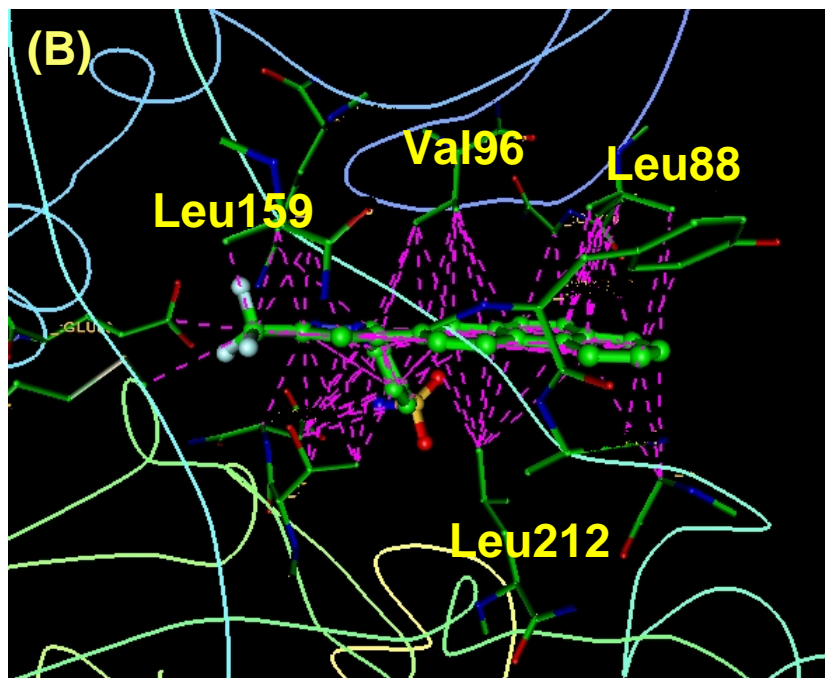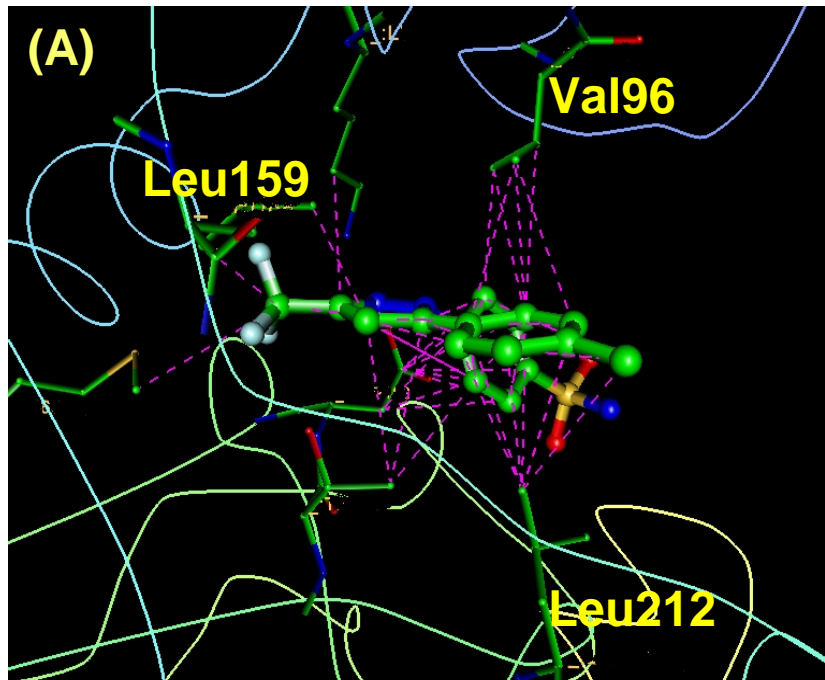
Figure 2.14 PDK1 binding with celecoxib and celecoxib-der2 in the newly identified binding mode. PDK1 binding with (A) celecoxib and (B) celecoxib-der2. The violet dotted lines show hydrophobic contacts.
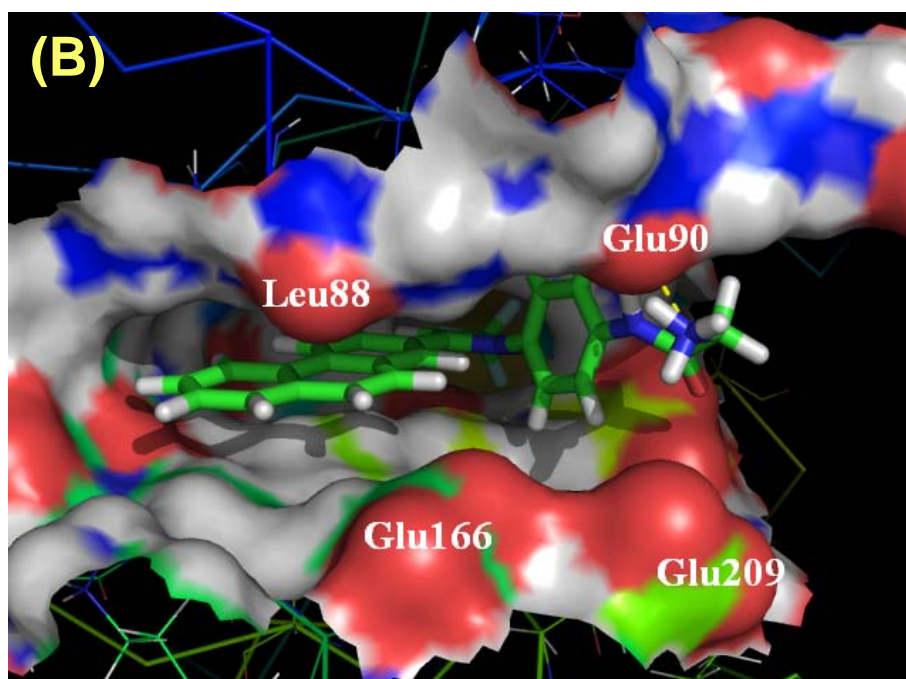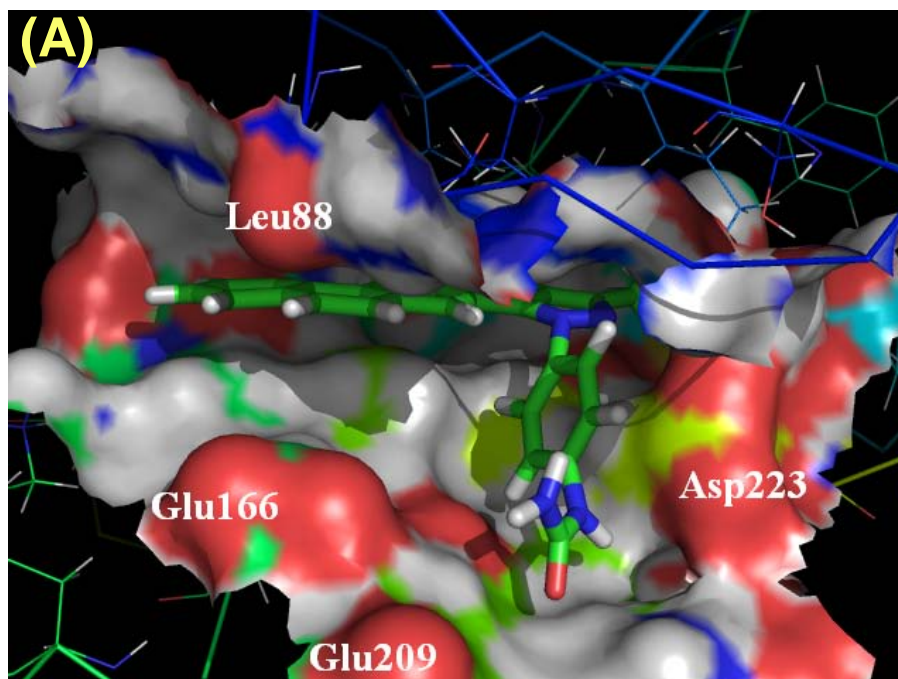
Figure 2.15 PDK1 binding with inhibitors in the newly identified binding mode. PDK1 binding with (A) celecoxib-der1 and (B) celecoxib-der3. The violet dotted lines show hydrophobic contacts.
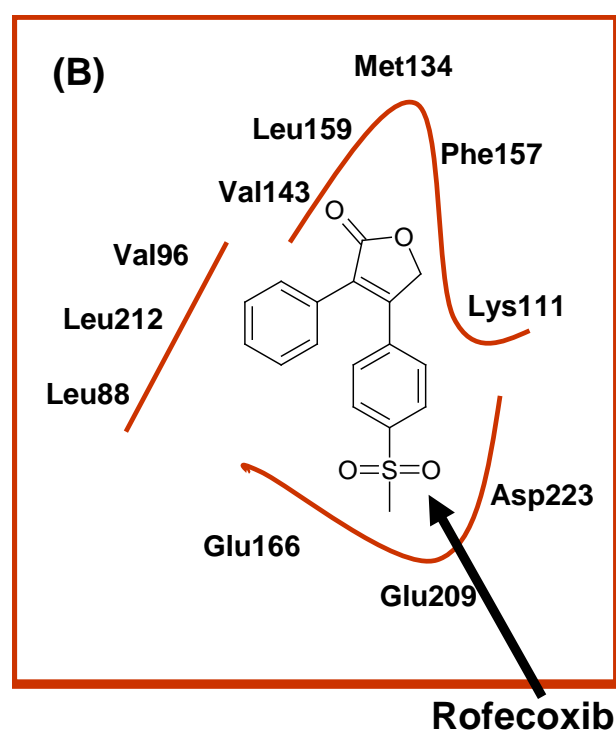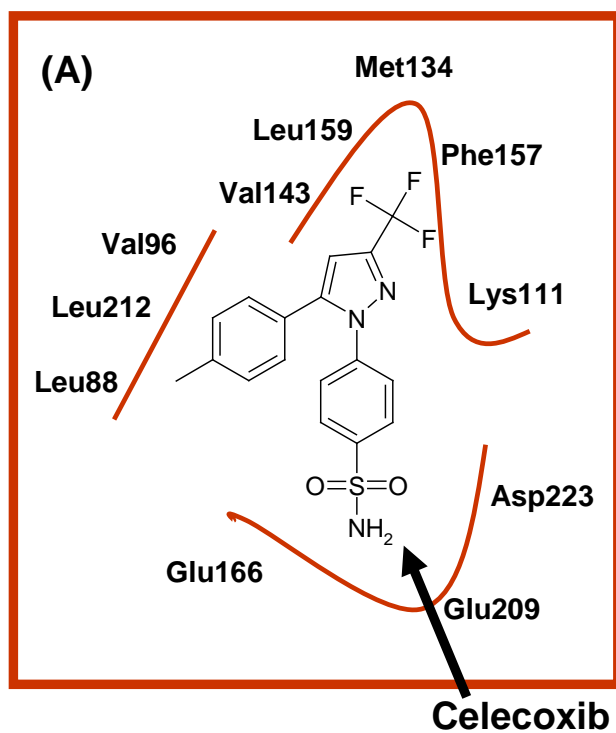
Figure 2.16 Schematic diagram showing PDK1 binding with inhibitors in the newly identified binding mode. PDK1 binding with (A) celecoxib and (B) rofecoxib.
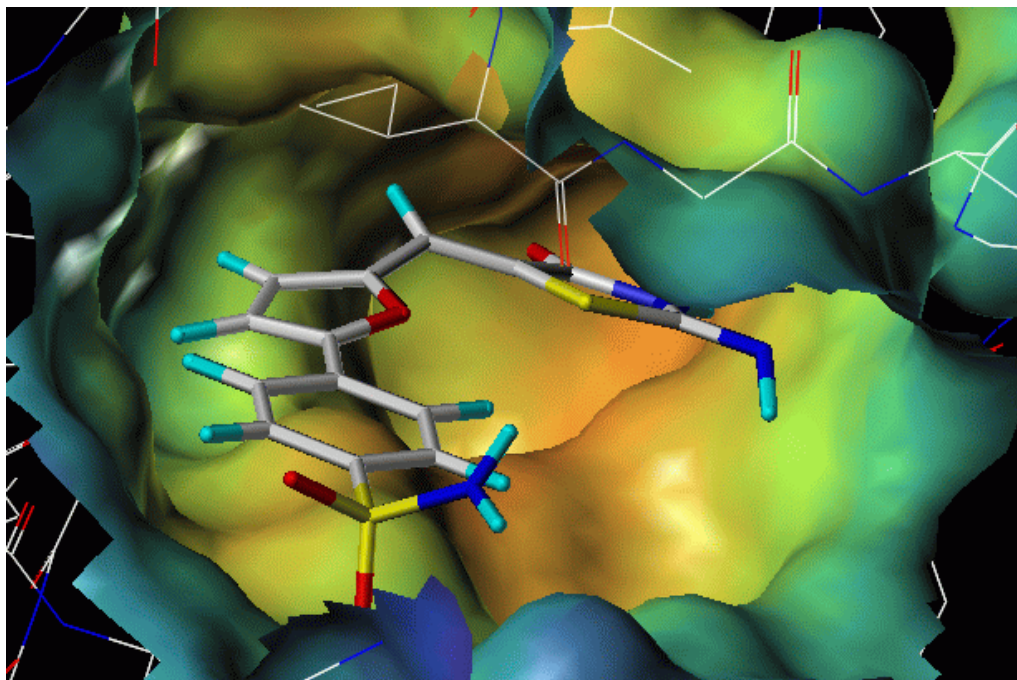
Figure 2.17 CDK2 binding with inhibitor without hinge hydrogen bond interaction (PDB ID: 2UZD).[199]

2.4 Summary of main computational insights

Extensive molecular docking and combined molecular dynamics (MD) simulations and molecular mechanics-Poisson-Boltzmann surface area (MM-PBSA) binding free energy calculations have demonstrated a new, more favorable microscopic binding mode concerning how celecoxib and its derivatives bind with 3-Phosphoinositide-dependent protein kinase-1 (PDK1). The new binding mode was remarkably different from that proposed previously based on simple molecular docking tests. For example, the 4-methyl-phenyl moiety of celecoxib stays in the hydrophobic adenine pocket and is sandwiched between the hydrophobic residues Val96 and Leu212 according to the new binding mode, whereas the 4-methyl-phenyl moiety is positioned in the ATP sugar/phosphate binding region (which is hydrophilic) according to the previously proposed binding mode. For celecoxib and all of the derivatives binding with PDK1, the binding free energies determined for the new binding mode by the combined MD simulations and MM-PBSA calculations were always significantly lower than those determined for the

previously proposed binding mode. So the new binding was energetically more favorable. Thus we have identified a totally new binding mode for the first time. This binding mode can be used for further *de novo* design of novel PDK1 inhibitors.
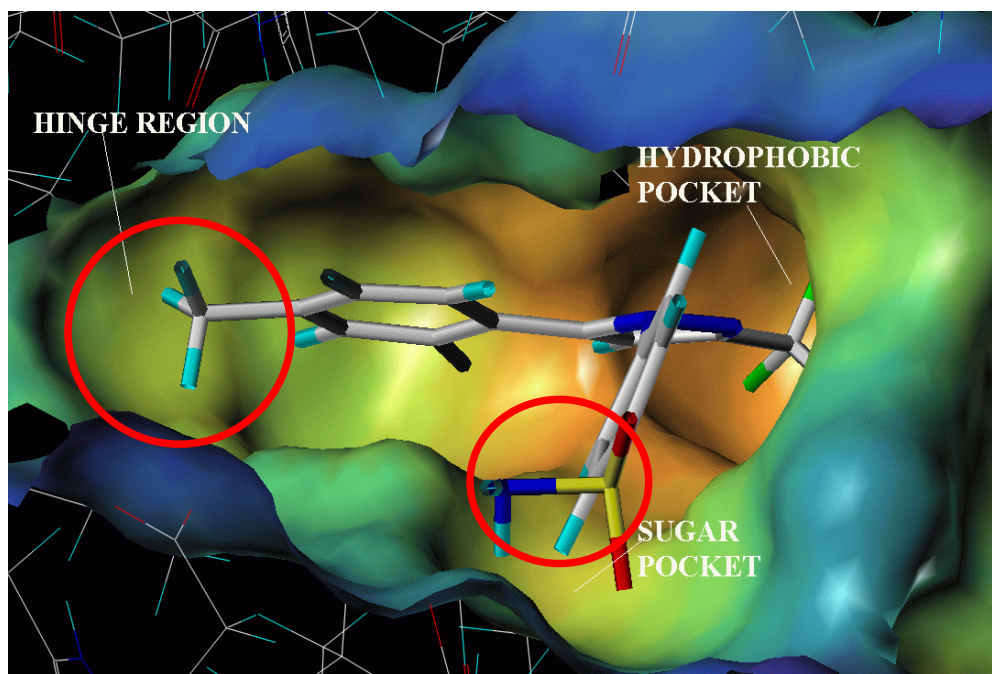
Figure 2.18 PDK1 binding with celecoxib in the newly identified binding mode. Potential region for *de novo* design are marked in red circle.

For all of the representative PDK1 inhibitors examined in this study, based on the most favorable binding modes, the calculated binding free energies were all in good agreement with the corresponding experimental activity data. All of the computational results reported in this study strongly suggest that (1) the new microscopic binding mode determined in this work is reliable; (2) the determined new, more favorable binding mode can help to better understand the substituent effects of the celecoxib derivatives on the previously reported biological activity data; and (3) the computational protocol tested in this study, *i.e.* the molecular docking followed by the combined MD simulations and MM-PBSA calculations, is reliable and accurate for predicting protein-ligand binding structures and binding free energies, whereas only performing simple molecular docking could lead to a

wrong binding mode and thus mislead the drug design. These new insights will be valuable not only for future rational design of novel, more potent PDK1-specific inhibitors as promising anticancer therapeutics, but also for rational design of drugs targeting other proteins.

CHAPTER 3

3. LIGAND-BASED DRUG DESIGN: DEVELOPMENT OF PREDICTIVE 3D-QSAR MODELS

Computational lead optimization strategies can adopt structure-based, ligand-based or combined approaches. PDK1 is a well known anti-cancer target and application of such computational drug design approaches will help towards the development of potent PDK1 inhibitors and ultimately lead to development of anti-cancer agents. In the previous chapter, we discussed the application of computational structure-based lead optimization strategy against PDK1. We used celecoxib and its derivatives as the lead compounds and identified the binding mode of these inhibitors which serves as the starting point for further rational design of new inhibitors. In this chapter, we discuss the application of ligand-based lead optimization strategy and combined approaches against PDK1. Part of the contents in chapter 3 was adapted from published article, AbdulHameed *et al. J. Chem. Inf. Mol. Mod.* **2008**, *48*, 1760-1772.[201]

3.1 Ligand-based modeling study on indolinone derivatives as PDK1 inhibitors

Computational approaches like structure-based and ligand-based ones have been found to be valuable in further optimization and development of novel inhibitors. Ligand-based three-dimensional quantitative structure-activity relationship (3D-QSAR) approaches, including comparative molecular field analysis (CoMFA)[202] and comparative molecular similarity indices analysis (CoMSIA),[203] were reported to be effective for understanding the structure-activity relationships.[37] The 3D-QSAR modeling is useful to predict the activity of new molecules to be synthesized.[37] 3D-QSAR methods serve as an important complement to the structure-based methods.[37] CoMFA and CoMSIA are two 3D-QSAR methods that have been successfully employed in drug design.[204,205] These methods were useful in the lead optimization and also in understanding the drug-target

interaction.[206,207,208] In CoMFA, the biological activity of molecules is correlated with their steric and electrostatic interaction energies.[209] The steric and electrostatic interaction energies are calculated using Lennard-Jones potential and Coulombic potential respectively.[202] In CoMSIA, similarity indices are calculated at regularly placed grid points for the aligned molecules.[203] CoMSIA includes additional molecular descriptors like hydrophobic fields and hydrogen bond donor and acceptor fields.[205] Both 3D-QSAR methods give contour maps as output that can be used to get some general insights into the topological features of the binding site.[208]

Kinase inhibitors, such as Imatinib, Erlotinib and Sunitinib, have recently been approved for clinical use in the market as anti-cancer agents.[210] The success of these new drugs, has given a new impetus for developing better anti-cancer agents by targeting kinases. Since PDK1 is a well validated anti-cancer target, ligand-based drug design strategies could be used in lead optimization of PDK1 inhibitors. More recently, a new series of indolinone derivatives were reported as PDK1 inhibitors.[211,212] This new series of compounds were developed using an initial hit identified from a high throughput screening (HTS) study.[211, 212] Indolinones are well known kinase inhibitor scaffold but a novel scaffold with respect to PDK1. Developing predictive 3D-QSAR will be useful to predict the activity of new compounds belonging to indolinone series as PDK1 inhibitors.

In this present work, the 3D-QSAR study on PDK1 inhibitors using CoMFA and CoMSIA methods was carried out. We also carried out molecular docking of the PDK1 inhibitors and further used the docked inhibitor structures to develop a separate set of docking-based (or receptor-based) 3D-QSAR models. Such a study allows us to compare the contour maps from two different strategies. Since the X-ray crystal structure of PDK1 binding with inhibitors are available, this study allows us to validate the potential and usefulness of the contour plots from ligand-based methods. In this study, the 3D-QSAR models obtained from both the ligand-based and receptor-based methods were all found to be statistically valid. The models were capable of predicting the activity of the test-set molecules. The contour plots obtained from the 3D-QSAR models correlate well with the detailed

interactions between the ligands and active site residues in the docked PDK1-inhibitor binding structures. The developed computational models are expected to help the lead optimization and future rational drug design and discovery efforts.

3.2. Computational Methods

3.2.1 Data sets and alignment

All compounds used in the present study were reported recently by Islam and co-workers[211,212] as inhibitors of PDK1. Of the 70 compounds reported, 56 compounds were used as a training set and the remaining 14 compounds were used as a test set, based on a random selection. The compounds in the test set have a range of biological activity values similar to that of the training set. The IC_{50} values were converted into pIC_{50} (*i.e.* $-logIC_{50}$) values. The pIC_{50} values are preferred due to convention, as the negative logarithms give larger values to most active compounds.[213] The pIC_{50} values of the compounds studied cover an interval of more than 3 log units. The structures of the compounds and their pIC_{50} values are given in Table 3.1.

Identification of the bioactive conformation and molecular alignment of compounds are two important steps in a 3D-QSAR study.[214] The X-ray crystal structures of this class of inhibitors bound with PDK1 are available from the protein data bank (PDB).[170] The bound conformation of compound **53** (PDB ID: 2PE2)[212] had a relatively better resolution (2.13Å) and was used as a template for alignment molecular structures. The 3D structures of all compounds were built using the SYBYL software (Tripos, Inc.).[171] The geometries of all compounds were optimized by using semiempirical PM3 method. The optimized geometries were used to perform single-point *ab initio* calculations at the HF/6-31G* level in order to determine the electrostatic potential (ESP)-fitted atomic charges, *i.e.* the ESP charges, that fit to the electrostatic potential at points selected according to the Merz-Singh-Kollman scheme.[215,216] Two alignment methods were used. One method, denoted by Alignment-I, is a substructure-based alignment. In this method,

all compounds in the dataset were aligned to a common substructure (Figure 3.1) using the 'align database' command in the SYBYL software. The other method, denoted by Alignment–II, is the atom/centroid-based alignment. The selected atoms and centroids of the molecules were used for alignment using the root-mean-squares (RMS) fit method. The determined substructure and the reference atoms/centroids used for the alignment were shown in Figure 3.1.

Table 3.1 Molecular structures of compounds used in the training and test sets and their PDK1 inhibitory activity [a]
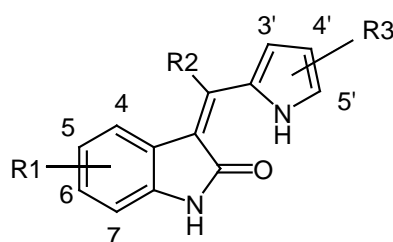
Compd.	R1	R2	R3	pIC50
1	H	H	H	5.74
2	4-Me	H	H	6.29
3[b]	5-Me	H	H	5.57
4	7-Me	H	H	4.58
5	4-OH	H	H	6.00
6[b]	5-OH	H	H	6.47
7	6-OH	H	H	6.05
8	7-OH	H	H	4.58
9	5-OH	Me	H	7.10
10[b]	5-OH	H	3'-Me	6.55
11	5-OH	H	4'-Me	6.55
12	5-OH	H	5'-Me	6.17
13	5-OH	H	5'-Et	5.64
14	5-OH	H	3',5'-Me	5.96
15	5-OMe	H	H	6.24
16	5-SO$_2$NH$_2$	H	H	6.54
17	5-CO$_2$Me	H	H	6.17
18	5-CO$_2$H	H	H	6.59
19	5-CONH$_2$	H	H	6.92
20	5-Tetrazole	H	H	6.70
21[b]	5-NH$_2$	H	H	6.24
22[b]	5-CN	H	H	6.00
23	5-CH$_2$NH$_2$	H	H	6.00

Table 3.1 (Continued)

Compd.	R1	R2	R3	pIC50
24	5-NHSO$_2$Me	H	H	6.28
25[b]	5-NHCOMe	H	H	7.26
26	5-NHCONH$_2$	H	H	7.74
27	5-SO$_2$NH$_2$	Me	H	7.17
28[b]	5-SO$_2$NH$_2$	Et	H	7.85
29	5-SO$_2$NH$_2$	Ph	H	7.54
30	5-SO$_2$NH$_2$	CO$_2$Et	H	7.47
31	5-SO$_2$NH$_2$	CO$_2$H	H	5.38
32	5-SO$_2$NH$_2$	CONH$_2$	H	5.41
33	5-SO$_2$NH$_2$	CONHEt	H	5.21
34	5-SO$_2$NH$_2$	CONEt$_2$	H	5.27
35	5-NHCONH$_2$	Me	H	8.30
36	5-NHCONH$_2$	Et	H	8.52
37	5-NHCONH$_2$	Ph-3-NH$_2$	H	8.05
38[b]	5-NHCONH$_2$	4-Pyridine	H	8.00
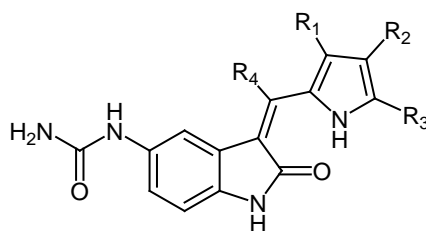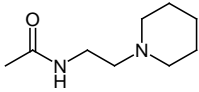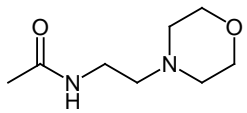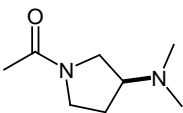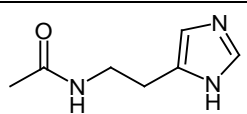
Compd.	R1	R2	R3	R4	pIC50
39	H	COO-	H	H	8.05
40	CH$_3$	COO-	H	H	8.05
41	CH$_3$	(CH$_2$)$_2$COO-	H	H	7.41
42	CH$_3$	(CH$_2$)$_2$COO-	CH3	H	7.24
43	H	CONH(CH$_2$)$_2$N(CH$_3$)$_2$	H	H	7.57
44	H		H	H	7.96
45	H		H	H	7.54
46[b]	H		H	H	7.51
47	H		H	H	7.85

60

Table 3.1 (Continued)
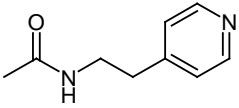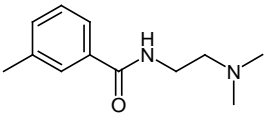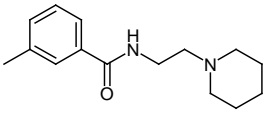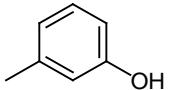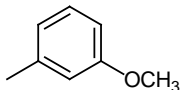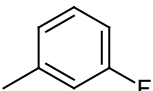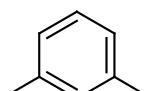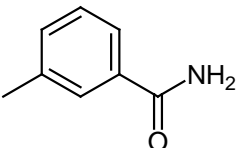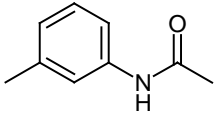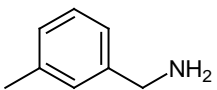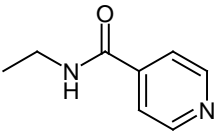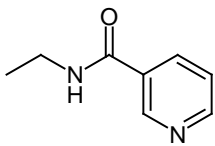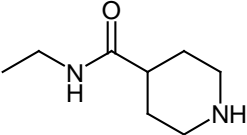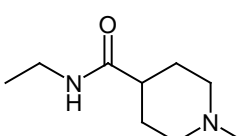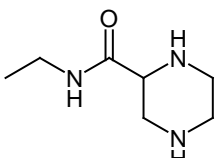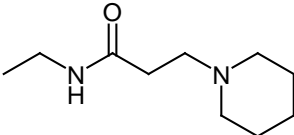
Compd.	R1	R2	R3	R4	pIC50
48	H	(structure: CH$_3$C(O)NH-CH$_2$CH$_2$-4-pyridyl)	H	H	7.77
49	H	-Phenyl	H	H	7.68
50	H	3-pyridyl	H	H	8.10
51	H	-Phenyl-3-carboxylic acid	H	H	7.35
52	H	(structure: 3-methylbenzamide N-CH$_2$CH$_2$-N(CH$_3$)$_2$)	H	H	8.40
53	H	(structure: 3-methylbenzamide N-CH$_2$CH$_2$-piperidinyl)	H	H	8.40
54[b]	H	(structure: 3-methylphenol)	H	H	8.22
55	H	(structure: 3-methyl-OCH$_3$ phenyl)	H	H	7.72
56[b]	H	(structure: 3-methylfluorobenzene)	H	H	7.29
57	H	(structure: 3,5-dimethylphenyl)	H	H	7.43
58	H	(structure: 3-methylbenzamide NH$_2$)	H	H	8.30
59[b]	H	(structure: 3-methyl-N-acetylaniline)	H	H	8.10
60	H	(structure: 3-methylbenzylamine CH$_2$NH$_2$)	H	H	8.52
61	H	-CH$_2$NH$_2$	H	CH$_3$	8.00
62[b]	H	-CHNHCOCH$_3$	H	CH$_3$	8.10
63	H	(structure: N-ethyl isonicotinamide)	H	CH$_3$	8.30

Table 3.1 (Continued)

Compd.	R1	R2	R3	R4	pIC50
64	H		H	CH$_3$	8.00
65	H		H	CH$_3$	8.22
66	H		H	CH$_3$	8.40
67[b]	H		H	CH$_3$	8.30
68	H	-CH$_2$NHCOCH$_2$NH$_2$	H	CH$_3$	8.30
69	H	-CH$_2$NHCOCH$_2$OH	H	CH$_3$	8.00
70	H		H	CH$_3$	8.52

[a] pIC$_{50}$ values calculated from the IC$_{50}$ data in refs. **211** and **212**. [b] Compounds used in the test set were based on random selection.
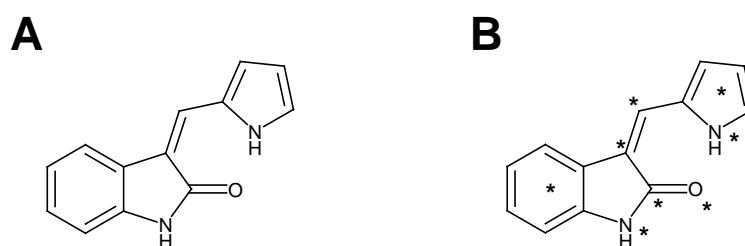
Figure 3.1. The substructure and atoms/centroids used in alignments (A) The substructure used for the 'Common substructure-based alignment' (Alignment-I); (B) The atoms/centroids used in 'Atom/Centroid based alignment' is shown in asterisk (Alignment-II).
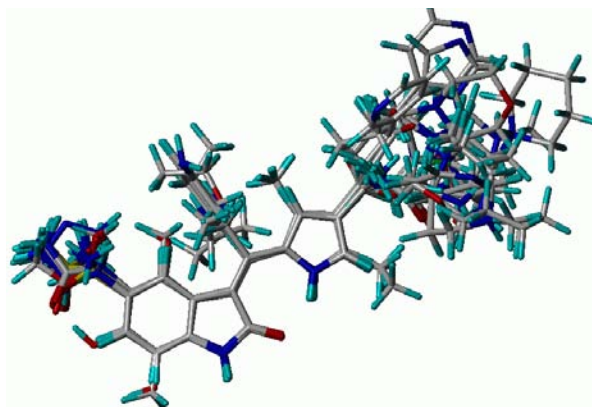
Figure 3.2 Superposition of all compounds in the training and test sets based on the common substructure based alignment (Alignment-I).

3.2.2 3D-QSAR analysis

The CoMFA and CoMSIA models were generated by using the SYBYL software with the default parameters. A regularly placed grid of 2.0 Å was created around the aligned molecules. In CoMFA, the steric and electrostatic fields were calculated at each intersection lattice point of the grid. A sp^3 carbon atom with charge +1.00 was used as a probe atom. The steric and electrostatic fields were truncated at +30.00 kcal/mol.

In CoMSIA, the steric, electrostatic, hydrophobic, hydrogen bond donor and acceptor descriptors were calculated at each lattice intersection of a regularly placed grid of 2.0 Å. A probe atom with radius 1.0 Å, charge +1.0, and hydrophobicity of +1.0 was used to calculate the respective fields. The attenuation factor α was set to 0.3. CoMSIA similarity indices ($\mathbf{A}_F$) for molecule j with atom i at grid point q are calculated by Eq.(3-1):

$$\mathrm{A}_{\mathrm{F},k}{}^{q}(j) = -\sum \omega_{\mathrm{probe},k}\,\omega_{ik}\,e^{-\alpha r_{iq}{}^{2}} \tag{3-1}$$

where k represents the steric, electrostatic, hydrophobic, H-bond donor, or H-bond acceptor descriptor. A Gaussian-type distance-dependence was used between the grid point q and each atom i of the molecule.

The partial least-square (PLS)[217] analysis was used to derive the 3D-QSAR models. Sample-distance PLS (SAMPLS) algorithm [218] was used for the leave-one-out (LOO) cross-validation. The optimum number of the components

identified in the cross-validation was used in the final analysis. To further validate the model, 100 runs of bootstrap analyses were performed. The models were also rigorously analyzed by performing cross-validation using 5 groups and 2 groups in the training set. Since only the number of groups and the number of validation times can be controlled in the SYBYL 'cross-validation' process, the process was repeated for 25 times. The models were also evaluated for their ability to predict the activity of molecules in the test set. The predictive r^2 (denoted by r^2_{pred}) for molecules in the test set was calculated by using the Eq.(3-2):

$$r^2_{pred} = (SD - PRESS) / SD \qquad (3\text{-}2)$$

where SD is the sum of the squared deviations of the individual biological activity values for the test-set molecules from the mean activity value of the test-set molecules, and PRESS is the sum of the squared deviations of the predicted activity values from the actual activity values of the test-set molecules.

3.2.3. Molecular docking

Molecular docking was carried out to understand the detailed binding modes of PDK1 binding with this series of inhibitors and develop receptor-based 3D-QSAR models. X-ray crystal structures are available for PDK1 binding with three different compounds in this series. The PDB IDs for X-ray crystal structures of PDK1 bound with compounds **9**, **35**, and **53** are 2PE0, 2PE1, 2PE2, respectively.[211,212] The X-ray crystal structure (PDB ID: 2PE2) of PDK1 was used in molecular docking. Used in the docking were the optimized geometries and the calculated ESP charges for all of the compounds. Initial docking test runs were carried out on compounds **9, 35**, and **53** using four different docking approaches and the ability to reproduce the X-ray crystal structures were analyzed. Of the four docking approaches used in this study, three are associated with the use of the popularly used docking programs, *i.e.* DOCK,[178] AutoDock,[219] and FlexX.[172] The fourth is an approach combining FRED docking[220,221] with energy minimization calculations on the docked binding structures, denoted by FRED-EM for convenience. Based on our initial docking test

runs, we finally selected FlexX and FRED-EM for molecular docking with all 70 ligands. It should be noted that the different docking programs have different default choices/parameters and the above results are limited to this series of molecules and should not be taken as the mutual comparison of software performance in general.

DOCK. The DOCK 6.1 program uses the anchor and grow strategy for flexible ligand docking.[178] The protein was considered to be rigid and the ligand molecules were flexible. The ligand orientations were scored through the use of a force field-based energy scoring function and the top-scored binding structure was selected. The active site spheres were prepared using SPHGEN program.[183] The spheres were selected such that they cover the entire ligand-binding region. These spheres served to orient ligands in the active site.[183] A box was created to enclose the spheres and the energetic grid was created by the GRID module of DOCK program. Grid calculations were carried out by using 0.3 Å grid spacing. Both the anchor and the ligand were minimized. 500 minimization iterations and a clash overlap value of 0.5 were used for the docking runs.

AutoDock. The AutoDock 4.0 program performs automated and flexible ligand docking.[219] The AutoDock Tools (ADT) was used in the preparation of protein and ligands. The molecular docking was performed using the Lamarkian genetic algorithm (LGA). The size of the grid, in which both the enzyme and the ligand were embedded, was set to be 60 Å × 60 Å × 60 Å along the x, y, and z directions of Cartesian coordinate system. This size of grid is large enough to cover all the protein atoms near the docking site. The default parameters in the AutoDock were used in this study. For docking with each ligand, the top-10 docked poses were compared to the corresponding X-ray crystal structure.

FlexX. We also analyzed docking using FlexX module of SYBYL.[172] In FlexX method, the protein is kept rigid and the ligand flexibility is explored. In this program, the ligand molecule is split up into fragments and is built up by performing conformational search of the ligand in the active site.[173] Only the best conformation was built upon and the others are discarded. The active site was defined as residues within 6.5Å around the bound ligand. Based on the crystal structure the region with

residues Leu88, Val96, Leu212, Ser160, and Ala162 were defined as the core sub-pocket. The ligands were docked using the multiple-ligand docking option of the FlexX. The top-30 docked orientations were generated for each ligand and the best docked structure was selected.

FRED docking followed by energy minimization (FRED-EM). We first generated and energy-minimized various molecular orientations and conformations of each ligand by using Omega (Open Eye Scientific Software) and Amber9 programs,[222] producing ~600 different minimum-energy conformations. Omega sampling is capable of selecting a ligand conformation similar to that of the targeted X-ray crystal structure by using an appropriate option (the default) including a low-energy cutoff to discard high-energy conformations, a low RMSD value below which two conformations are considered to be similar, and a maximum of 500 to 1000 output conformations.[223] We checked to make sure that the sampling was sufficient enough to include at least one conformer of the scaffold (3-[(Pyrrol-2-yl)-meth-(Z)-ylidene]-1,3-dihydro-indol-2-one) similar to the one found in the X-ray complex. FRED (OpenEye Scientific Software) docking calculations were carried out using protein structures with all hydrogen atoms and with the binding site definitions provided by FRED Receptor program (Open Eye Scientific Software).

FRED docking roughly consisted of two steps, *i.e.* shape fitting and optimization. During the shape fitting, the ligand was placed into a 0.5 Å-resolution grid box encompassing all active-site atoms (including hydrogen atoms) using a smooth Gaussian potential.[220] A series of two optimization filters were then processed, consisting of rigid-body optimization and optimization of the ligand pose in the dihedral angle space. In the optimization step, four scoring functions are available: Gaussian shape scoring,[220] ChemScore,[224] PLP,[225] and ScreenScore.[226] Preliminary docking trials led us to select ChemScore for the optimization filters. In separate docking runs, the conformational poses of each ligand that passed the shape-fitting and optimization filters were submitted to the energy minimization using AMBER9 program. During the energy minimization in vacuum, the

non-bonded cutoff and the dielectric constant were set to group-based (20 Å cutoff distance) and distance-dependent ($\varepsilon=4r$), respectively.[227] The pose with the lowest interaction energy was selected as the best binding mode.

3.3. Combined receptor-based and ligand-based 3D-QSAR models and their evaluation

3.3.1. 3D-QSAR models

The results obtained from the CoMFA and CoMSIA models using a training set of 56 compounds are summarized in Table 3.2. The structural alignment of the compounds is one of the important steps in the development of a 3D-QSAR model.[202] Hence we tested two different alignment rules namely the substructure-based alignment (Alignment-I) and atom/centroid-based alignment (Alignment-II). For all of the 3D-QSAR models, the leave-one-out (LOO) cross-validation was performed first to identify the cross-validated correlation coefficient (q^2) values. Then the number of components identified in the LOO cross-validation process was used in the final non-cross-validated PLS run. The developed 3D-QSAR models were analyzed in terms of a number of statistical parameters, namely the q^2, non-cross-validated correlation coefficient (r^2), standard error estimate (SEE), and F-statistic values. All models developed have a good q^2 value of >0.6. A q^2 value of greater than 0.5 is usually considered significant.[228] In CoMFA analyses, the substructure-based alignment (Alignment-I) led to the models with larger q^2 and r^2 values than those obtained from Alignment-II. In the case of CoMSIA, different combinations of the descriptors were also studied. The CoMSIA model with steric, electrostatic, hydrophobic, donor, and acceptor descriptors was associated with larger q^2 and r^2 values. It has been reported that, in addition to the q^2 value, r^2_{pred} should also be used to choose the predictive QSAR models.[229] Hence the CoMSIA model (denoted by CoMSIA-1d) from alignment I was selected for final analysis based on the larger r^2_{pred} value and used to predict the activity of the compounds (see Table 3.2).

The CoMFA-1a model (see Table 3.2) has a q^2 value of 0.737 and an r^2 value

of 0.907. The model was developed with 5 components. It has an F value of 97.053 and an SEE value of 0.354. The CoMSIA-1d model has a q^2 value of 0.824 and an r^2 value of 0.991. It was developed with 9 components and has an F value of 589.262 and an SEE value of 0.112. The high r^2, q^2, and F values along with the low SEE values suggest that the models are reasonable and should have a good predictive ability. The results also reveal that the electrostatic and hydrophobic and hydrogen bond donor features play an important role in determining the biological activity of these inhibitors. In CoMSIA, the steric feature is found to have the least contribution to the activity. Based on the CoMFA-1a model, the contributions from the steric and electrostatic fields were found to be around ~46% and ~54%, respectively. According to the CoMSIA-1d model, the steric field has ~8% contribution, the electrostatic field has ~23% contribution, hydrophobic field has ~23% contribution, donor feature has 24%, and acceptor feature has ~22% contribution.

The 3D-QSAR models were further validated using an external test set of 14 compounds. CoMFA-1a and CoMSIA-1d models all gave good predictions of both the training- and test-set compounds (see Supporting Information). Both the CoMFA-1a and CoMSIA-1d models have the larger r^2_{pred} values, *i.e.* 0.812 and 0.883, respectively. The predicted activity of the compounds and their residuals (deviations) are given in Table 3.5, and the plots obtained were depicted in Figure 3.2. In both models (CoMFA-1a and CoMSIA-1d) the deviations of the predicted pIC_{50} values from the corresponding experimental pIC_{50} values are always smaller than 1 log unit.

We also performed additional cross-validation analyses using more groups in the training set. The CoMFA-1a and CoMSIA-1d models were further analyzed by additional rigorous statistical cross-validation using five and two groups in the training set. Each cross-validation process was repeated for 25 times and the results are tabulated in Table 3.3. In the case of cross-validation using two groups, the training set is divided into two groups and 50% of the compounds are dropped out in the training process and the activities of those compounds are predicted. In the case of cross-validation using 5 groups, 20% of the compounds are left out in the training

process and the activity values of those compounds are predicted. As shown in Table 3.3, the average q^2 values obtained in this way was only slightly lower than the q^2 values obtained with the LOO method. These results suggest that the high q^2 values were not obtained by chance correlation and the obtained CoMFA-1a and CoMSIA-1d models are stable and valid. To further obtain statistical confidence limits, 100-runs bootstrap analyses were also carried out. The bootstrapping results are considered as good indicators concerning whether there exist possible chance correlations.[230] Our analyses gave a bootstrap r^2 value of 0.942 and 0.994 for CoMFA-1a and CoMSIA-1d, respectively. The high values of bootstrap r^2 further confirm the robustness of our CoMFA-1a and CoMSIA-1d models.

Table 3.2 Summary of the results obtained from the CoMFA and CoMSIA analyses[a,201]

PLS statistic	Alignment-I				Alignment-II			
	CoMFA 1a	CoMSIA 1b	CoMSIA 1c	CoMSIA 1d	CoMFA 1e	CoMSIA 1f	CoMSIA 1g	CoMSIA 1h
	SE	SE	SEH	SEHDA	SE	SE	SEH	SEHDA
q^2	0.737	0.712	0.772	0.824	0.716	0.732	0.782	0.837
SEP	0.595	0.623	0.548	0.507	0.619	0.611	0.536	0.493
r^2	0.907	0.900	0.921	0.991	0.882	0.897	0.919	0.993
SEE	0.354	0.366	0.323	0.112	0.398	0.372	0.327	0.099
F value	97.053	90.362	148.011	589.262	74.774	86.976	144.231	685.762
NOC	5	5	4	9	5	5	4	10
r^2_{pred}	0.812	0.834	0.911	0.883	0.743	0.819	0.909	0.877
r^2_{bs} [b]	0.942	n/a	n/a	0.994	n/a	n/a	n/a	n/a
Fraction								
Steric	0.456	0.266	0.122	0.088	0.459	0.281	0.128	0.093
Electrostatic	0.544	0.734	0.476	0.230	0.541	0.719	0.460	0.227
Hydrophobic	n/a	n/a	0.402	0.230	n/a	n/a	0.412	0.233
Donor	n/a	n/a	n/a	0.235	n/a	n/a	n/a	0.223
Acceptor	n/a	n/a	n/a	0.217	n/a	n/a	n/a	0.224

[a] CoMFA and CoMSIA analyses were preformed using the training set of 56 compounds; Alignment-I refers to substructure-based alignment; Alignment-II refers to atom/centroid-based alignment; S=steric, E=electrostatic, H=hydrophobic,

D=donor and A=acceptor. [b]100 runs of bootstrap analysis.
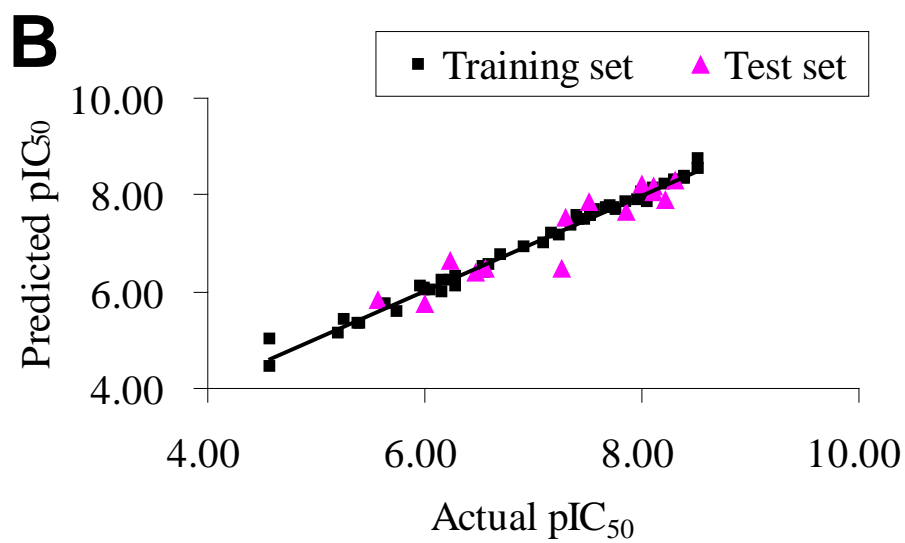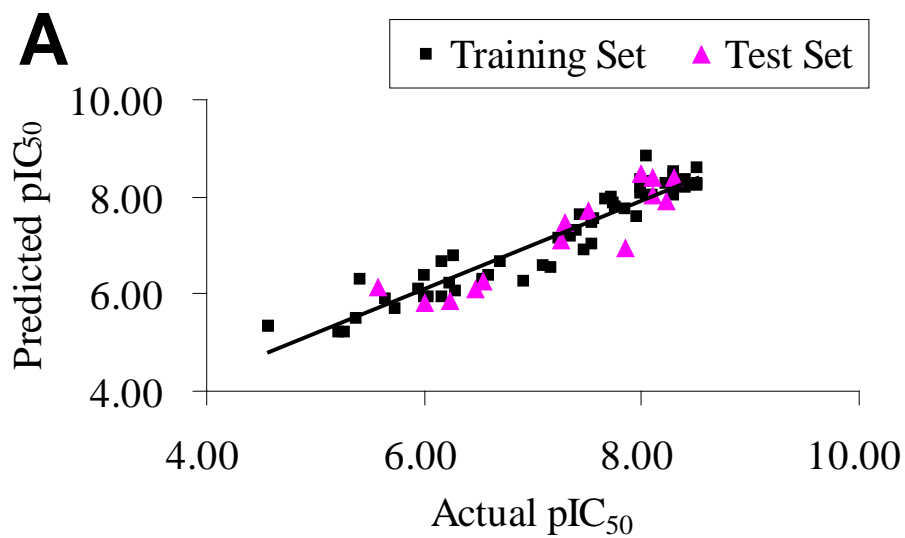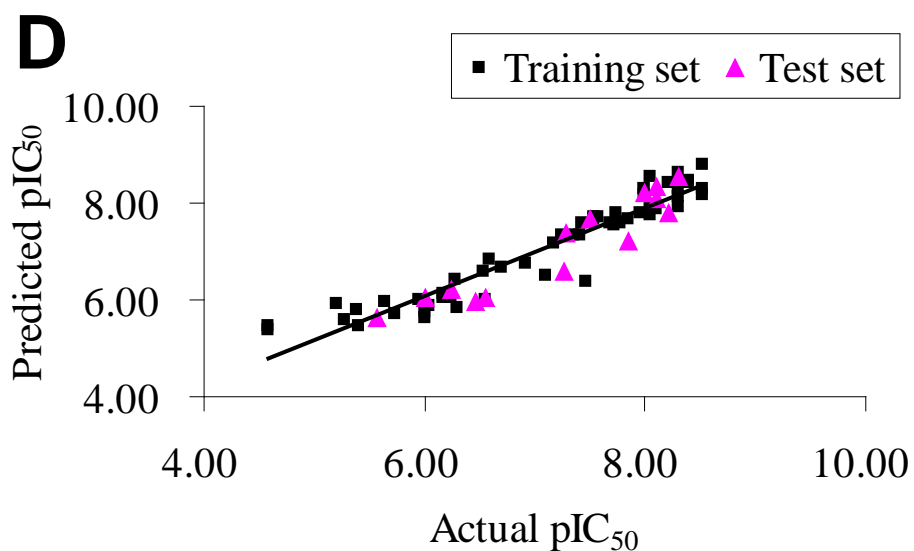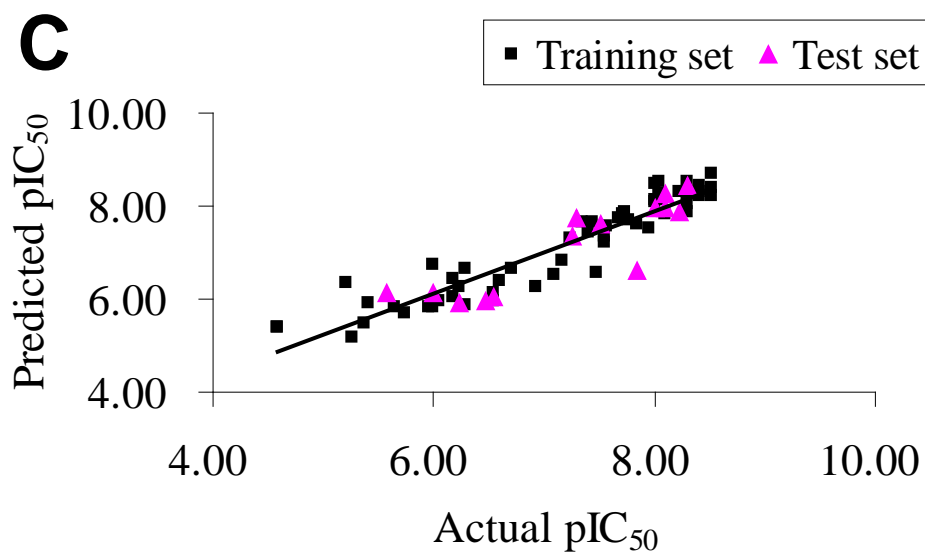
Figure 3.3 Plots of the predicted pIC_{50} values *versus* the actual pIC_{50} values using the training set of 56 compounds and the test set of 14 compounds.[201] (A) CoMFA-1a model using substructure-based alignment; (B) CoMSIA-1d model using substructure-based alignment; (C) CoMFA-2a model using the docking-based alignment; (D) CoMSIA-2b model using docking-based alignment.

Table 3.3 Result of the cross-validation analyses using five and two groups[201]

	$q^{2\,a}$				
	Using five groups			Using two groups	
	CoMFA-1a	CoMSIA-1d		CoMFA-1a	CoMSIA-1d
mean	0.725	0.794		0.67	0.730
high	0.779	0.863		0.781	0.821
low	0.631	0.699		0.562	0.637

[a] Cross-validated q^2 values obtained from using five and two groups with the optimum number of components on average of 25 runs.

3.3.2 3D-QSAR contour maps

One of the attractive features of the CoMFA and CoMSIA modeling is the visualization of the results as 3D-coefficient contour plots. The contour maps were generated as scalar products of coefficients and standard deviation, associated with each CoMFA or CoMSIA column. The maps generated depict regions having scaled coefficients greater than 80% (favored) or less than 20% (disfavored). In the case of CoMFA, the green contour shows favorable steric interaction and the yellow contours show the region where steric group is not favored. The red contour shows favorable electronegative region and the blue contour shows the region where electropositive region is favored. These contour maps (as depicted in Figures 3.4 and 3.5) give us some general insight into the nature of the receptor-ligand binding regions.

CoMFA contour maps One of the most active compounds in the series (compound **60**) is shown superimposed with the CoMFA contour maps in Figure 3.4. The yellow contour region near #7 position of indolinone ring shows that, substituents at this position have unfavorable steric interactions. This is consistent with the reported experimental results. Compounds **4** and **8** have a methyl and hydroxyl group, respectively, at the #7 position and both compounds have a lower activity. The yellow contours are present below the plane of the indolinone group. This shows that there will be an unfavorable steric interaction if bulky substitutions were introduced in that region. This is in agreement with the reported X-ray crystal structure of PDK1 bound with compound **53** (PDB ID: 2PE2).[212] The indolinone

group is sandwiched between hydrophobic residues. If bulky groups are introduced below the plane of indolinone group, they will have an unfavorable steric interaction with Leu212.
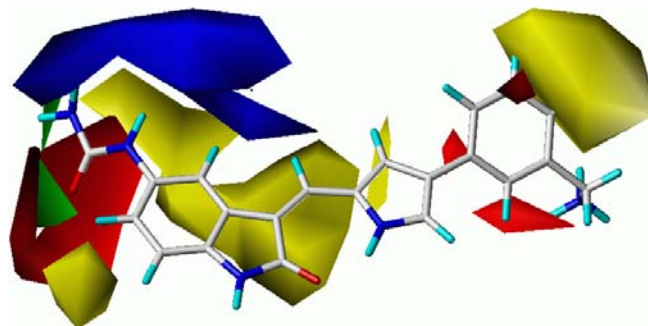
Figure 3.4 CoMFA steric and electrostatic contour maps around compound **60**; Green isopleths enclose areas where steric interaction is favored. Yellow contours are areas where the steric interaction is disfavored. Blue region represents the area where electropositive group is favorable for the binding. Red region refers to the area where an electronegative group is favorable for the binding.

The green region near the urea group shows a favorable steric interaction in this position. The red contour near the carbonyl of urea group shows that electronegative groups are favored in this region. This explains the lower activity of compound **23**. Compound **23** has ($-CH_2NH_2$) substituent at #5 position of indolinone group and has an IC_{50} value of 1000 nM. Similarly, red contour region (favorable electronegative group) is found in 4′ substituents at the pyrrole group. Compounds **39** to **48** have carbonyl group at this position and have an IC_{50} value in low nanomolar range. Blue contour region shows that electropositive groups are favored in this region. The $-NH_2$ of urea group is present near this region and explains the increased activity of compound **26** ($IC_{50} = 18$ nM) compared to compound **25** ($IC_{50} = 55$ nM).

CoMSIA contour maps In the case of CoMSIA, we get additional insight from the hydrophobic, acceptor, and donor features. The CoMSIA steric and electrostatic contour maps were similar to the ones obtained from the CoMFA model. Figure 3.5a shows compound **60** superimposed on the CoMSIA steric, electrostatic, and hydrophobic contour plot.

Figure 3.5. CoMSIA contour maps around compound **60.** (A) CoMSIA steric, electrostatic, and hydrophobic contour maps around compound **60**. Green and Yellow contours represent regions with favorable and unfavorable steric interactions. Blue and red contours represent regions which favor electropositive and electronegative groups respectively. Cyan isopleths enclose areas where hydrophobic groups could enhance the activity. White contours represent area where hydrophobic groups are disfavored. (B) CoMSIA acceptor and donor contour maps around compound **60**. Orange isopleths indicate regions where acceptor group is

favored and red contour indicates region where acceptor groups are not favored. Magenta and purple contours represent favorable and unfavorable hydrogen bond donor regions, respectively.

The yellow contour near #7 position of indolinone ring shows an unfavorable steric interaction at this region. This is also observed in the CoMFA model and explains the lower activity of compounds **4** and **7**. In addition, the CoMSIA model shows that #5' position of pyrrole ring has an unfavorable steric interaction. This agrees with the reported experimental results. Compounds **12** to **14** and **42** have a methyl or ethyl substituent at this position and have a lower activity. Green contour is present between #4 position of indolinone group and the methylidenyl ($>C=CH-$) group at the #3 position. This shows that, steric interaction is favored at this region. This agrees with the reported experimental results. Compounds **27** to **29** have bulky substituents in this region and have a higher activity. Similarly, when we compare compound **2** (IC_{50} = 510 nM) with **5** (IC_{50} = 1000 nM), we find that 4-methyl group (compound **2**) is slightly favored than 4-OH group (compound **5**) in this region, which is consistent with the relative IC_{50} values. The red contour near the carbonyl of urea group at #5 position of indolinone ring shows that electronegative groups are favored in this region. The blue contours occur immediately adjacent to it. It shows the favorable electropositive region. Another favorable electronegative region is present near the pyrrole ring. The favorable hydrophobic region (cyan contour) is present near the methylidenyl ($>C=CH-$) group at #3 position of indolinone ring (Figure 3.5a). This is consistent with the increased activity observed for compounds **27** to **29**. In Figure 3.5a, white contour showing an unfavorable hydrophobic interaction region is present near #5 position of indolinone group. This explains the lower activity of compound **3** (IC_{50} = 2700 nM) compared to compound **6** (IC_{50} = 340 nM). The hydrogen bond donor and acceptor contour superimposed on compound **60** is shown in Figure 3.5b. The orange contour shows a favorable hydrogen bond acceptor region and the red contour shows regions in which hydrogen bond acceptor is not favored. The magenta color shows the favorable hydrogen bond donor contour and the purple color shows regions in which a

hydrogen bond donor group is not favored. From the contour plot we can see that the carbonyl oxygen of urea is present in the favorable hydrogen bond acceptor group region. This also agrees with the favorable electronegative contour observed in Figure 3.5a. The favorable hydrogen bond donor contour is present near the $-NH_2$ of urea group. This is in agreement with the reported X-ray crystal structure of PDK1 bound with compound **53** (PDB ID: 2PE2).[212] In the crystal structure, the carboxylate side chain of Glu130 is present near $-NH_2$ of urea group and could have a favorable interaction. There are unfavorable hydrogen bond donor and acceptor contours (Figure 3.5b, red contour and magenta contour) above the plane of the methylidenyl ($>C=CH-$) group (at #3 position of indolinone ring). Taking into account the hydrophobic contour from Figure 3.5a, we find that, this region has only a favorable hydrophobic interaction. Donor and acceptor groups are not preferred in this region. This explains the lower activity of compounds **31** to **34**, as they have either donor or acceptor group in this region. There is an additional unfavorable hydrogen bond donor contour (purple) near the $>NH$ of indolinone group. The X-ray crystal structure (PDB ID: 2PE2) shows that, the -NH- of indolinone group has a hydrogen bonding interaction with the backbone carbonyl oxygen of Ser160. The unfavorable hydrogen bond donor contour is present in the place where carbonyl group of serine backbone is present. So this purple contour represents that the -NH- of indolinone ring is the optimal or ideal group at this position. If a hydrogen bond donor substitution is made at #1 position of indolinone ring (instead of -NH- on the ring), there will be clash with the backbone atom and such a hydrogen bond donor will not be favorable for the activity.

3.3.3 Binding structures and docking-based 3D-QSAR models

In addition to the ligand-based 3D-QSAR, we have also performed molecular docking for all of the 70 inhibitors to understand the nature of interactions of these compounds with PDK1. We also carried out molecular docking and receptor-based 3D-QSAR modeling, *i.e.* using the docked poses of the 70 compounds in the PDK1 active site. The contour plots from the receptor-based 3D-QSAR models directly

relate the favorable and unfavorable contours to the corresponding detailed protein-ligand interactions in the active site. X-ray crystal structures are available for PDK1 binding with three different compounds (**9**, **35**, and **53**) in this series. The PDB IDs for X-ray crystal structures of PDK1 binding with compounds **9**, **35**, and **53** are 2PE0, 2PE1, and 2PE2, respectively.[211,212] As the first step, the docking reliability was tested using the PDK1 structure in 2PE2 for docking with compounds **9**, **35**, and **53**. For molecular docking with all ligands, we used the optimized geometries and calculated ESP charges. Three commonly used docking programs (DOCK, AutoDock, and FlexX) and the aforementioned FRED-EM approach were used for these three inhibitors. We found that only FlexX and FRED-EM were able to reproduce the X-ray crystal structures for all of the three ligands. The AutoDock was unable to reproduce the pose of compound **35** in PDK1. All of the 10 AutoDock poses for compound **35** were significantly different from that observed in the X-ray crystal structure. In the case of DOCK, compound **53** failed to dock into the active site.

In light of the above docking tests, the FlexX and FRED-EM were finally selected for carrying out molecular docking for all of the 70 inhibitors. In the case of FlexX-based docking, 69 compounds were docked in a similar pose. One exception is compound **37**, in which the -NH- of pyrrole ring is flipped by ~180° compared to the X-ray crystal structure. So, compound **37** was omitted from the training set of the 3D-QSAR modeling using the molecular structures obtained from the FlexX docking. The FRED-EM approach was able to dock all of the 70 compounds into the active site with a similar pose. The docked poses of all 70 compounds using this method were shown in Figure 3.6. The docked poses serve as a very good starting point for carrying out 3D-QSAR modeling. As discussed earlier, the alignment of compound structures plays a key role in developing successful 3D-QSAR models. Hence the docked poses of the ligands were used to develop receptor-based 3D-QSAR models.

In the development of 3D-QSAR models based on FlexX docking and FRED-EM docking, we found that the models developed with FRED-EM docking

(CoMFA-2a and CoMFA-2b) had larger q^2, r^2, and r^2_{pred} values. Hence CoMFA-2a and CoMSIA-2b were used in our further analysis below. The results are summarized in Table 3.4.
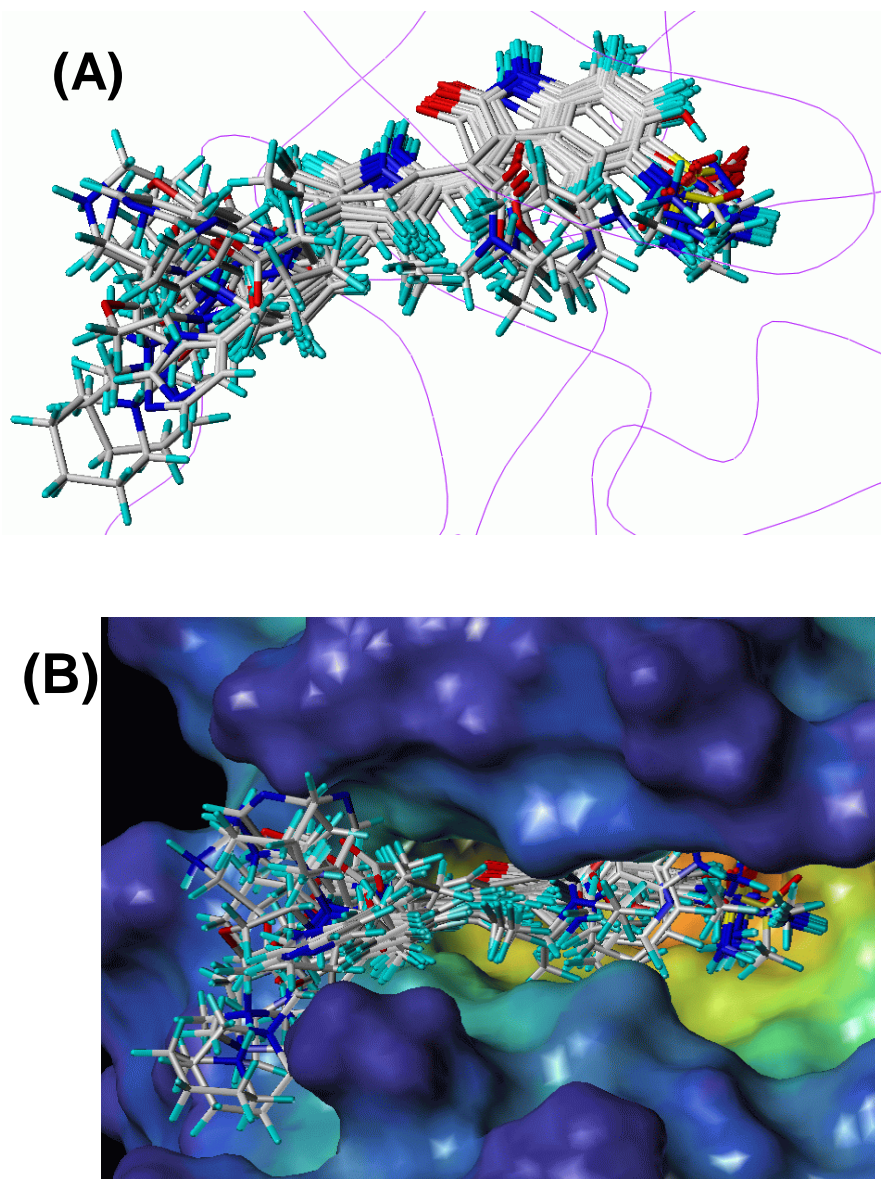
Figure 3.6 Docked poses of all 70 inhibitors in the active site of PDK1. (A) The trace view of PDK1 backbone is shown in magenta color. (B) Different view of the 70 docked molecules; PDK1 is shown in Connolly surface and the surface is colored according to cavity depth.

Table 3.4 Summary of the results obtained from the docking-based CoMFA and CoMSIA analyses[201]

| PLS statistic | FRED-EM[a] | | FlexX[b] | |
| | COMFA-2a | COMSIA-2b | CoMFA-3a | CoMSIA-3b |
	SE	SEHDA	SE	SEHDA
q^2	0.729	0.79	0.662	0.738
SEP	0.598	0.521	0.663	0.615
r^2	0.884	0.909	0.854	0.992
SEE	0.392	0.343	0.437	0.109
F value	96.878	173.423	99.093	698.659
NOC	4	3	3	8
r^2_{pred}	0.736	0.840	0.726	0.806
r^2_{bs}[c]	0.911	0.926	n/a	n/a
Fraction				
Steric	0.428	0.079	0.459	0.086
Electrostatic	0.572	0.283	0.541	0.248
Hydrophobic	n/a	0.245	n/a	0.229
Donor	n/a	0.237	n/a	0.271
Acceptor	n/a	0.156	n/a	0.167

[a] CoMFA and CoMSIA analyses were preformed using the training set of 56 compounds and the alignment is based on the docked poses from FRED-EM method. [b] CoMFA and CoMSIA analyses were preformed using the training set of 55 compounds and the alignment is based on the FlexX docked poses; S=steric, E=electrostatic, H=hydrophobic, D=donor and A=acceptor. [c] 100 runs of bootstrap analysis.

CoMFA. The CoMFA-2a model developed with the docked poses has a q^2 value of 0.729 and r^2 value of 0.884. The validity of the model was tested with the external test set of 14 compounds. CoMFA-2a model gives satisfactory activity (pIC_{50}) predictions for both the training and test sets. It is also interesting to note that the contour plots obtained from the docking-based model correlate well with the detailed interactions between the compounds and the active site residues (Figure 3.7b). The predicted activity of the compounds and their residuals are provided in Table 3.6 and the plots obtained are depicted in Figure 3.3c. The deviations of the predicted pIC_{50} values are greater than 1 log unit only for two compounds (**28** and **33**). The residuals between actual and predicted values for compounds **28** and **33** are

-1.25 and -1.12, respectively. A yellow contour present near #7 position of the indolinone ring represents an unfavorable steric interaction in this region. This is consistent with the corresponding protein-ligand binding structure obtained from molecular docking. A comparison between the docked structure and the contour plots reveals that the yellow contour is present in the region of Leu159 and Val143. Hence, bulky substitutions at #7 position of indolinone ring will have an unfavorable steric interaction. This also explains the lower activity of compounds **4** and **8**. Another unfavorable steric interaction region (yellow contour) is present near the pyrrole and phenyl ring. Our docked model shows that a bulky substituent at this position will have an unfavorable steric interaction with the backbone of Lys163 and side chain of Tyr161. The red contour represents a favorable electronegative region and is present near the carbonyl of urea group. According to the docked structure, this favorable electronegative region is surrounded by Lys111, Thr222, and backbone -NH- group of Asp223. This explains the increased activity expected from the introduction of electronegative groups in this region. The red contour around the phenyl ring shows a favorable electronegative region. This contour is present near the backbone -NH- of Gly165 residue. Thus it is expected to have a favorable interaction with electronegative groups like carbonyl. The amino group on 3″ position of phenyl ring is present near the backbone carbonyl group of Asn164. This explains the increased activity of compounds with electropositive groups in this region. Blue contour is present in the region around urea group. Glu130 and Ser94 are present in this region. This reveals that the activity of the compound can be increased further by extending the substituents and introducing suitable electropositive groups to interact with these residues.

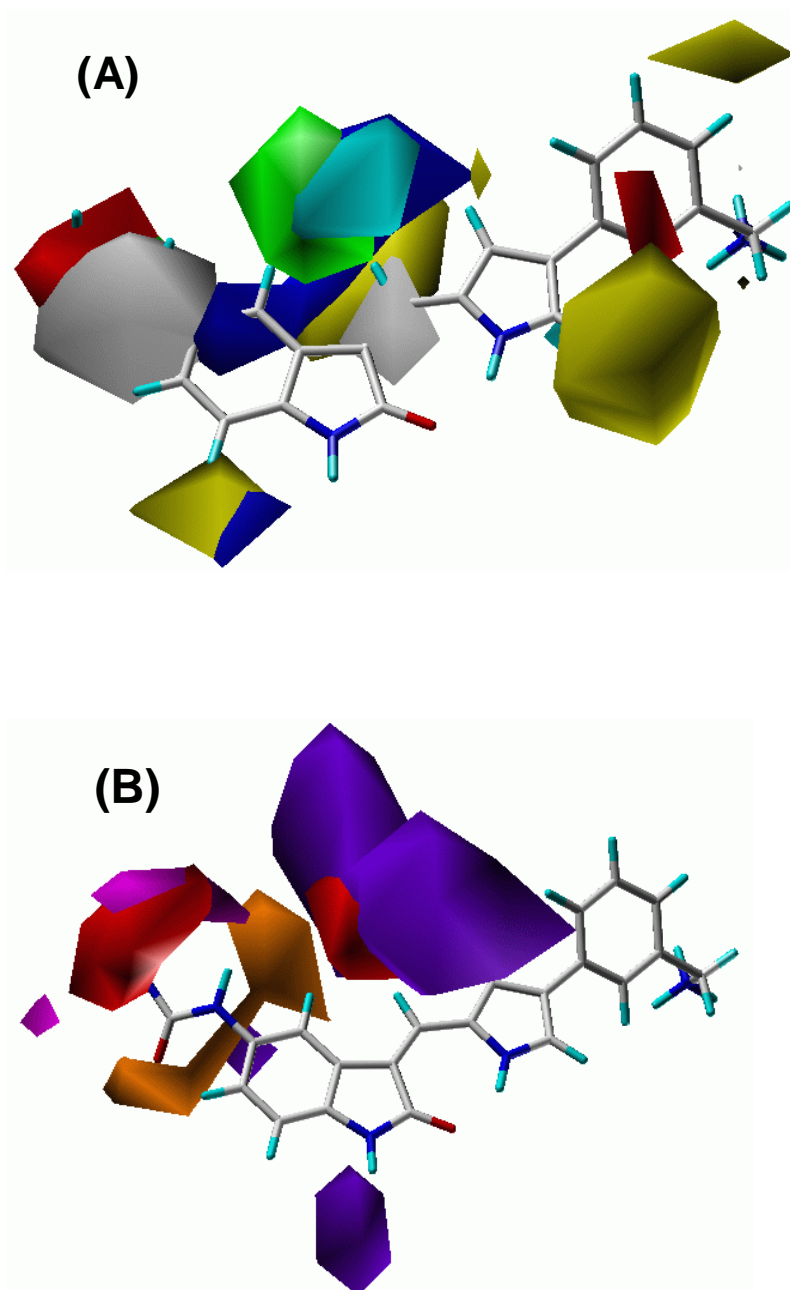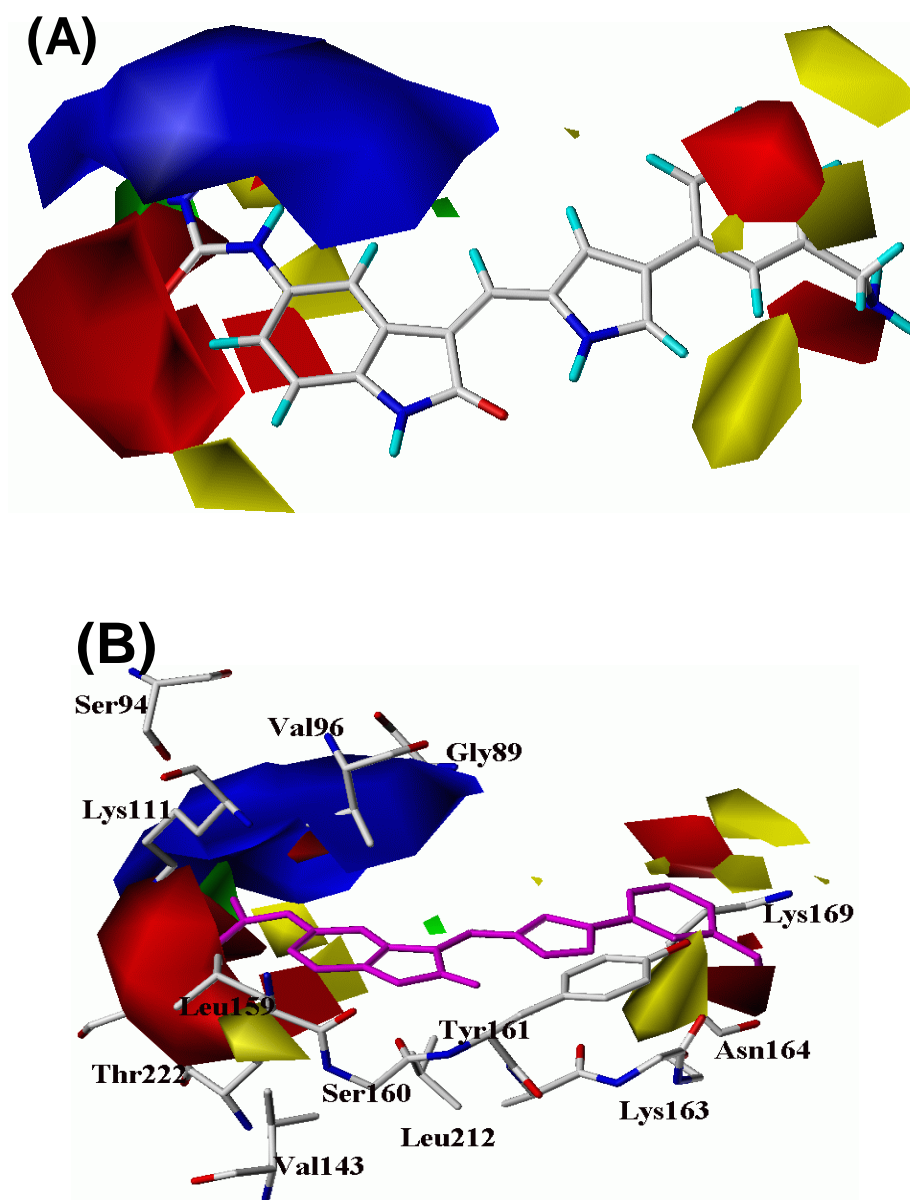**Figure 3.7** CoMFA contour maps around compound **60** (A) CoMFA steric and electrostatic contour maps around compound **60**; Green isopleths enclose areas where a steric bulk could enhance the activity. Yellow contours are areas where the steric interaction is disfavored. Blue region represents the area where a positive charge is favorable for the binding. Red region refers to the area where a negative charge is favorable for the binding. (B) The CoMFA contour plots were shown superimposed with PDK1 active site residues. Compound **60** is shown in magenta color. Hydrogen atoms and some side-chains were not shown for better clarity of this figure.
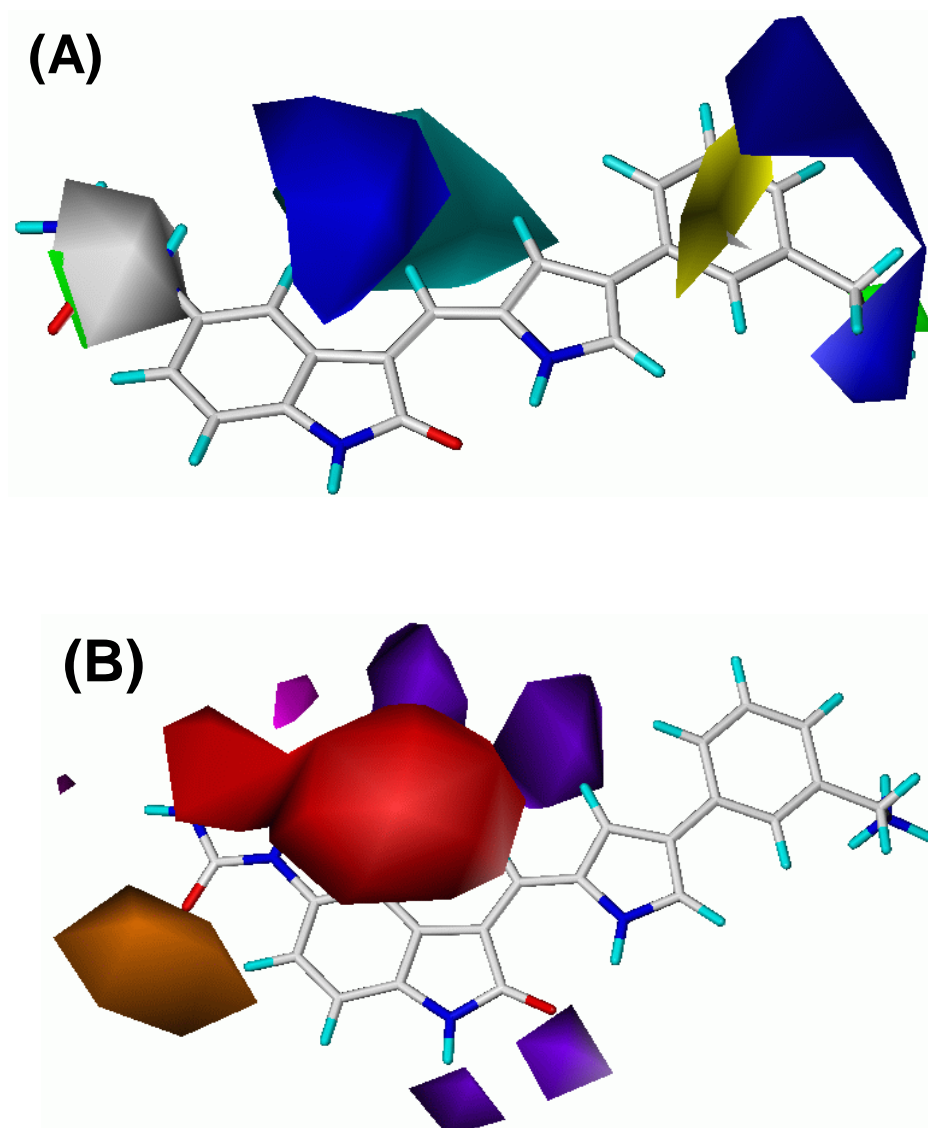
Figure 3.8 CoMSIA contour maps around compound **60** (A) CoMSIA steric, electrostatic, and hydrophobic contour maps around compound **60**. Green isopleths indicates region where there is favorable steric interaction. Yellow contours are areas where the steric interaction is disfavored. Blue region represents the area where a positive charge is favorable for the binding. Red region refers to the area where a negative charge is favorable for the binding. Cyan isopleths enclose areas where hydrophobic groups could enhance the activity. White region represents area where hydrophobic groups are disfavored. (B) CoMSIA acceptor and donor contour maps around compound **60**. Orange isopleths indicate regions where acceptor group is favored and red regions indicate where acceptor groups are not favored. Magenta

region indicates areas where hydrogen bond donor groups are favored. The purple region indicated areas where hydrogen bond donors are disfavored.

CoMSIA. The CoMSIA model developed with the docked poses has a q^2 value of 0.79 and r^2 value of 0.909. The validity of the model was tested with the external test set of 14 compounds. CoMFA-2b model gives satisfactory activity (pIC_{50}) predictions for both the training and test sets. The predicted activity of the compounds and their residuals are provided in Table 3.6 and the plots obtained were depicted in Figure 3.3d. The deviation of the predicted pIC_{50} value from the corresponding experimental value was greater than 1 log unit only for compound **30**. The residual (deviations) between actual and predicted values for compound **30** is 1.09. In Figure 3.8a, the green contour along with white contour around the urea group shows that hydrophobic substituents are not favored at this region but the urea group has a favorable steric fit into the pocket surrounding this group. This is in agreement with the crystal structure as this group is surrounded by Thr222, Lys111, and Asp 223. Compound **26** with urea derivative has a higher activity (IC_{50} = 18 nM) than compounds **16** (IC_{50} = 290 nM) and **18** (IC_{50} = 260 nM) with 5-sulfonamide and 5-carboxylate groups, respectively. The cyan contour shows a favorable hydrophobic interaction with substituents in the methylidenyl (>C=CH-) group (at #3 position of indolinone ring). This is in agreement with the experimental observation that compounds **36** to **38** with a hydrophobic susbtituent in this region have an increased activity. The blue contour shows a favorable electropositive region near the amino group on the phenyl ring. The amino group of this molecule is present near the backbone carbonyl group of Asn164. This explains the increased activity of compounds with electropositive groups in this region. The orange contour shows a favorable acceptor region which corresponds to the interaction with Lys111. The red contour in Figure 3.7b corresponds to the unfavorable acceptor groups. Hydrophobic residues like Val96 and Leu88 are present in this region. This explains the unfavorable hydrogen bond donor (purple) and acceptor (red) contour in this region. The unfavorable hydrogen bond donor contour near the 2-carbonyl group of indolinone ring corresponds to the -NH- backbone of Ala162. This shows that

H-bond donor is not favored in this region.

3.4. Significance and limitation of the combined 3D-QSAR modeling and molecular docking study

Indolinone is one of the ten most frequently encountered sub-structures in kinase inhibitors.[231] Reasonable 3D-QSAR models will be of great help in further rational design of new ligands of indolinone series as PDK1 inhibitors. Through a combined use of 3D-QSAR modeling and molecular docking, we have developed ligand-based and docking-based 3D-QSAR models for the PDK1 inhibitors. The high r^2 and q^2 values of the 3D-QSAR models suggest that PDK-1 binds with all of the examined inhibitors in a similar binding mode. Combining molecular docking with 3D-QSAR modeling offers a more interesting, integrated approach and allows us to utilize structural information of the protein for 3D-QSAR modeling. The advantage of docking-based model is that we can directly superimpose the contour plots into the protein active site. Such superimposition will also allow us to check the correlation between the contour plots and the corresponding receptor residues present near them. Since the docked pose gives the bioactive conformation of the ligands, this method helps to overcome the error which may arise by using an incorrect conformation of the ligand. The results obtained from molecular docking and those from 3D-QSAR modeling can complement and validate each other. All of the structural insights obtained from molecular docking and 3D-QSAR contour maps are consistent with the available experimental activity data, suggesting that the microscopic enzyme-inhibitor binding structures obtained from the molecular docking are reasonable and that the developed 3D-QSAR models are reliable.

Another interesting aspect is the nature of contour plots from two different approaches namely ligand-based and receptor-based QSAR modeling. In the former, no receptor information is used, whereas in the latter docked poses were used as the starting alignment to carry out QSAR analysis. The contour plots from both approaches are very similar. As discussed above, the contour maps obtained from both models were consistent with the active site. These results show one of the

potential applications of ligand-based study in drug design. For many important therapeutic targets, the experimentally solved X-ray structure is not available. But there are known ligands against such targets which were developed through traditional medicinal chemistry approach. In such cases, ligand-based approaches like 3D-QSAR can be used to get an insight into the nature of the active site. We have previously shown one such application in understanding the nature of active site of microsomal prostaglandin E synthase-1 and its interaction with the inhibitors.[206] In our present study, we have also found that the contour plots from ligand-based alignment which does not include any receptor information are consistent with the receptor active site.

One possible disadvantage of the docking-based 3D-QSAR method is that the small variation in docking pose could give rise to "noise" in the 3D-QSAR model. It is reported that the slight variation in the docked poses could create variations in the ligand field and will weaken the ability of the partial least-squares (PLS) to detect the real "signal".[232] The general limitation of 3D-QSAR modeling is that it cannot be used if the ligands have different microscopic binding modes with the enzyme. In such cases, the combined molecular docking and 3D-QSAR modeling approach cannot be used. Whenever 3D-QSAR modeling can be used for a series of compounds, developing both ligand-based and docking-based 3D-QSAR models could help us to overcome some possible error which may occur if we use one particular method.

Table 3.5 The actual and predicted inhibitory activity values (pIC_{50}) and the residuals of the training- and test-set molecules.[201]

Compound	Actual pIC_{50}	CoMFA-1a Predicted pIC_{50}	Residual	CoMSIA-1d Predicted pIC_{50}	Residual
1	5.74	5.69	0.05	5.58	0.16
2	6.29	6.04	0.26	6.12	0.17
4	4.58	5.33	-0.75	5.00	-0.42
5	6.00	5.93	0.07	6.04	-0.04
7	6.05	5.93	0.12	6.03	0.02
8	4.58	5.32	-0.74	4.45	0.13
9	7.10	6.56	0.54	7.00	0.10
11	6.55	6.29	0.26	6.41	0.14
12	6.17	5.94	0.23	5.99	0.18
13	5.64	5.91	-0.27	5.76	-0.12
14	5.96	6.09	-0.13	6.09	-0.13
15	6.24	6.22	0.02	6.23	0.01
16	6.54	6.24	0.30	6.52	0.02
17	6.17	6.65	-0.48	6.21	-0.04
18	6.59	6.36	0.23	6.57	0.02
19	6.92	6.26	0.66	6.91	0.01
20	6.70	6.65	0.06	6.74	-0.04
23	6.00	6.37	-0.37	6.08	-0.08
24	6.28	6.77	-0.49	6.30	-0.02
26	7.74	7.86	-0.12	7.74	0.00
27	7.17	6.52	0.65	7.18	-0.01
29	7.54	7.02	0.52	7.57	-0.03
30	7.47	6.91	0.56	7.49	-0.02
31	5.38	5.47	-0.09	5.33	0.05
32	5.41	6.29	-0.88	5.34	0.07
33	5.21	5.21	0.00	5.13	0.08
34	5.27	5.20	0.07	5.43	-0.16
35	8.30	8.09	0.21	8.31	-0.01
36	8.52	8.22	0.30	8.74	-0.22
37	8.05	8.82	-0.77	7.84	0.21
39	8.05	8.04	0.01	7.96	0.09
40	8.05	8.29	-0.24	8.01	0.04
41	7.41	7.31	0.10	7.57	-0.16
42	7.24	7.16	0.08	7.16	0.08
43	7.57	7.52	0.05	7.70	-0.13
44	7.96	7.60	0.36	7.89	0.07
45	7.54	7.44	0.10	7.57	-0.03
47	7.85	7.75	0.10	7.84	0.01

Table 3.5 (Continued)

Compound	Actual pIC$_{50}$	Predicted pIC$_{50}$	Residual	Predicted pIC$_{50}$	Residual
48	7.77	7.78	-0.01	7.68	0.09
49	7.68	7.93	-0.25	7.71	-0.03
50	8.10	8.03	0.07	8.12	-0.02
51	7.35	7.18	0.17	7.38	-0.03
52	8.40	8.21	0.19	8.34	0.06
53	8.40	8.21	0.19	8.36	0.04
55	7.72	7.99	-0.27	7.76	-0.04
57	7.43	7.61	-0.18	7.49	-0.06
58	8.30	8.02	0.28	8.30	0.00
60	8.52	8.28	0.24	8.52	0.00
61	8.00	8.18	-0.18	8.02	-0.02
63	8.30	8.50	-0.20	8.27	0.03
64	8.00	8.06	-0.06	8.06	-0.06
65	8.22	8.28	-0.06	8.22	0.00
66	8.40	8.36	0.04	8.38	0.02
68	8.30	8.43	-0.13	8.30	0.01
69	8.00	8.35	-0.35	7.96	0.04
70	8.52	8.58	-0.06	8.54	-0.02
Test set					
3	5.57	6.12	-0.55	5.83	-0.26
6	6.47	6.09	0.38	6.38	0.09
10	6.55	6.24	0.31	6.47	0.08
21	6.24	5.86	0.38	6.63	-0.39
22	6.00	5.83	0.17	5.75	0.25
25	7.26	7.12	0.14	6.48	0.78
28	7.85	6.94	0.91	7.65	0.20
38	8.00	8.48	-0.48	8.22	-0.22
46	7.51	7.71	-0.20	7.84	-0.33
54	8.22	7.91	0.31	7.88	0.34
56	7.29	7.45	-0.16	7.53	-0.24
59	8.10	8.01	0.09	8.05	0.05
62	8.10	8.39	-0.29	8.16	-0.06
67	8.30	8.41	-0.10	8.29	0.01

Table 3.6 The actual and predicted inhibitory activity values (pIC$_{50}$) and the residuals of the training- and test-set molecules using docking based 3D-QSAR.[201]

		CoMFA-2a		CoMSIA-2b	
Compound	Actual pIC$_{50}$	Predicted pIC$_{50}$	Residual	Predicted pIC$_{50}$	Residual
1	5.74	5.71	0.03	5.69	0.05

Table 3.6 (Continued)

Compound	Actual pIC$_{50}$	Predicted pIC$_{50}$	Residual	Predicted pIC$_{50}$	Residual
2	6.29	5.86	0.43	5.85	0.44
4	4.58	5.38	-0.80	5.47	-0.89
5	6.00	5.83	0.17	5.64	0.36
7	6.05	5.96	0.09	5.86	0.19
8	4.58	5.39	-0.81	5.36	-0.78
9	7.10	6.52	0.59	6.50	0.60
11	6.55	6.14	0.41	6.00	0.55
12	6.17	6.03	0.14	6.06	0.11
13	5.64	5.83	-0.19	5.95	-0.31
14	5.96	5.84	0.12	6.00	-0.04
15	6.24	6.27	-0.03	6.04	0.20
16	6.54	5.95	0.60	6.57	-0.03
17	6.17	6.46	-0.29	6.13	0.04
18	6.59	6.38	0.21	6.85	-0.26
19	6.92	6.26	0.66	6.74	0.18
20	6.70	6.64	0.06	6.65	0.05
23	6.00	6.73	-0.73	5.74	0.26
24	6.28	6.64	-0.36	6.43	-0.15
26	7.74	7.86	-0.12	7.78	-0.04
27	7.17	6.82	0.35	7.18	0.00
29	7.54	7.20	0.34	7.72	-0.18
30	7.47	6.56	0.91	6.38	1.09
31	5.38	5.48	-0.10	5.79	-0.41
32	5.41	5.93	-0.52	5.45	-0.04
33	5.21	6.33	-1.12	5.90	-0.69
34	5.27	5.16	0.11	5.57	-0.30
35	8.30	7.89	0.41	7.92	0.38
36	8.52	8.20	0.32	8.19	0.33
37	8.05	8.51	-0.46	8.54	-0.49
39	8.05	8.24	-0.19	7.89	0.16
40	8.05	8.14	-0.09	7.76	0.29
41	7.41	7.42	-0.01	7.34	0.07
42	7.24	7.30	-0.05	7.35	-0.11
43	7.57	7.56	0.01	7.72	-0.15
44	7.96	7.54	0.42	7.80	0.16
45	7.54	7.37	0.18	7.66	-0.12
47	7.85	7.63	0.22	7.68	0.17
48	7.77	7.71	0.06	7.58	0.19
49	7.68	7.73	-0.05	7.56	0.12
50	8.10	7.81	0.29	7.87	0.23
51	7.35	7.66	-0.31	7.33	0.02

Table 3.6 (Continued)

Compound	Actual pIC_{50}	Predicted pIC_{50}	Residual	Predicted pIC_{50}	Residual
52	8.40	8.30	0.10	8.42	-0.02
53	8.40	8.44	-0.04	8.45	-0.05
55	7.72	7.83	-0.11	7.55	0.17
57	7.43	7.67	-0.24	7.57	-0.14
58	8.30	8.04	0.26	8.12	0.18
60	8.52	8.41	0.12	8.28	0.24
61	8.00	8.48	-0.48	8.30	-0.30
63	8.30	8.30	0.00	8.31	-0.01
64	8.00	8.14	-0.14	8.17	-0.17
65	8.22	8.32	-0.10	8.42	-0.20
66	8.40	8.22	0.18	8.45	-0.05
68	8.30	8.50	-0.20	8.64	-0.34
69	8.00	8.08	-0.08	8.31	-0.31
70	8.52	8.71	-0.19	8.79	-0.27
Test set					
3	5.57	6.14	-0.57	5.64	-0.07
6	6.47	5.96	0.51	5.96	0.51
10	6.55	6.03	0.52	6.06	0.49
21	6.24	5.93	0.31	6.19	0.05
22	6.00	6.12	-0.12	6.05	-0.05
25	7.26	7.35	-0.09	6.57	0.69
28	7.85	6.61	1.25	7.21	0.64
38	8.00	7.95	0.05	8.20	-0.20
46	7.51	7.61	-0.10	7.68	-0.17
54	8.22	7.89	0.34	7.79	0.43
56	7.29	7.75	-0.46	7.39	-0.10
59	8.10	7.96	0.14	8.08	0.02
62	8.10	8.26	-0.16	8.33	-0.23
67	8.30	8.44	-0.14	8.53	-0.23

3.5 Summary of computational insights

The combined 3D-QSAR modeling and molecular docking resulted in valuable insights into PDK1 binding with 70 inhibitors and their structure-activity correlation. In this present study we first developed a 3D-QSAR model using ligand-based alignment. We compared different alignment techniques and found that the substructure based alignment method gives the best ligand-based 3D-QSAR model. Use of higher-level optimization/charge calculation methods were reported to be

able to reproduce the experimental structure.[233] Hence in our present study, we utilized PM3 for geometry optimization and the HF 6-31G* method for the charge calculations. The drawback that could be expected for ligand-based 3D-QSAR method is the difference between the template used for alignment and the real bio-active conformation of the molecule. So we also performed receptor-based 3D-QSAR. We first performed a control run and checked the ability of different docking programs to reproduce the three X-ray crystal structures of the ligands binding with PDK1. We selected FRED-EM as the docking method to dock all the ligands. Thus we have used different strategies to develop robust ligand-based and receptor-based 3D-QSAR models. The models developed using both methods were statistically valid with better q^2, r^2 values. Finally the models were also able to predict the external test set of compounds which were not used in QSAR model development. This shows the usefulness of the models in predicting the activity of new compounds. Molecular docking revealed the detailed structures of PDK1 binding with the compounds. The interactions identified from the CoMFA and CoMSIA 3D-contour maps correlate well with the specific interactions between the inhibitors and the amino acid residues identified in the docked binding structures. The fact that both methods gave similar results could point towards the general usefulness of ligand-based approaches in cases where the target structure is not available. Moreover, the combined modeling using two different approaches is a better approach because it helps to overcome the error which may occur if we use one particular method. The 3D contour maps obtained from the CoMFA and CoMSIA models in combination with the detailed PDK1-inhibitor binding structures obtained from molecular docking helped to better interpret the structure-activity relationship of these PDK1 inhibitors and provide valuable insights into rational drug design for further optimization of the biological activity of the PDK1 inhibitors.

CHAPTER 4

4. HIT IDENTIFICATION: APPLICATION OF HIERARCHICAL VIRTUAL SCREENING

Computational drug design approaches are used in both lead optimization and hit/lead identification. As discussed before, PDK1 is a well known anti-cancer target and application of such computational drug design approaches will help towards the development of potent PDK1 inhibitors and ultimately lead to development of new anti-cancer agents. In the previous chapters, we discussed the application of computational lead optimization strategies against PDK1. We have used structure-based lead optimization strategy for celecoxib and its derivatives. We have used ligand-based and combined approaches in lead optimization of indolinone derivatives. In both cases, previously known inhibitors were used as a starting point to understand and design new PDK1 inhibitors. We were also interested in identifying new compounds as PDK1 inhibitors. Computational virtual screening is a very useful tool for hit identification. In this chapter, we discuss the application of a hierarchical virtual screening strategy to identify hits against PDK1. Part of the contents in chapter 4 was adapted from submitted article, AbdulHameed *et al. J. Comp. Aided Mol. Des.* **2009.**

4.1 Computational virtual screening approaches

Virtual screening methods are used for both lead identification and optimization in various drug discovery projects.[234] This method involves screening a large database of millions of compounds using computational approaches and selection of smaller number of compounds for final experimental testing.[235] Virtual screening is expected to reduce the amount of experimental work and has the potential to reduce the cost and time of the initial drug discovery phase. This makes virtual screening a very popular approach. A recent study comparing virtual screening against experimental high-throughput screening (HTS) for the same target

shows that virtual screening performs better than HTS.[236] A number of recent studies have successfully identified new hits through virtual screening.[237,238] Inspite of its success, the current virtual screening approaches have their own limitations including consideration of the flexibility of active site residues during docking and accurately accounting for the solvation/desolvation effects of protein-ligand interaction.[239] Many strategies have been proposed to improve the virtual screening approaches.[240,241] Screening strategies which use fast and quick filters in the initial steps and more complex filters in later stages were reported to be better than simple screening using one method.[242] The development of commercially available compound database (ZINC) has further helped virtual screening applications by reducing the time spent in preparing the databases.[243] Recent research papers shows that the performance of ligand-based virtual screening approach is faster and perform better than simple docking methods.[234] Absorption, Distribution, Metabolism, and Excretion (ADME) properties were also reported to play a major role in the development of a hit into a potential drug molecule.[242] Application of appropriate ADME filters in the early stage of virtual screening will help in identifying a more suitable hit molecule. The computational binding free energy calculation methods like linear interaction energy (LIE) and molecular mechanics/Poisson-Boltzmann surface area (MM-PBSA) were reported to be useful for the rapid estimation of the binding free energy of an inhibitor with the macromolecular target.[242] The MM-PBSA method has been successfully used for calculating the binding free energies of protein-protein and protein-inhibitor interactions.[168] There are very few reports about the use of MM-PBSA method as a part of virtual screening strategy to identify new lead molecules. One recent report has shown the ability of MM-PBSA method to retain the known ligand when the method is used as a part of virtual screening strategy.[242] We are interested to analyze the hierarchical virtual screening approach which uses a combination ligand-based and structure-based approach along with the more sophisticated MM-PBSA method as a final filter.

  As PDK1 is a well validated and a promising anti-cancer target, we chose it

as the target for our study. Identifying new hits as PDK1 inhibitor will help towards the development of better treatment options for cancers. We have pursued an efficient strategy of the hierarchical virtual screening with increasing complexity to identify new hits against PDK1. We used ligand-based method ROCS [244,43] and Rigid docking FRED[245] as the initial filters followed by flexible docking using FlexX [172] and finally molecular dynamics (MD) and MM-PBSA methods were used. Applying hierarchical (step-wise) virtual screening strategy has helped us to identify a nanomolar PDK1 inhibitor which is active in cell based anti-cancer screening.

4.2 Simulation methods

4.2.1 Preparation of database and protein

In this work we have used the ligand-based screening, rigid docking, flexible docking, and molecular dynamics (MD) simulations followed by the MM-PBSA calculations as screening filters in virtual screening. The flow chart of steps used in our virtual screening strategy is given in Figure 4.1. The first step is the selection of compound database and preparing protein structure for virtual screening. A subset of Zinc 7 database[243] containing 688,086 compounds was used in this study. Zinc database represents an annotated database with curated molecules suitable for virtual screening.[243] The subset is made up of compounds from two major commercial compound suppliers IBScreen (570,038 compounds) and Sigma Aldrich (115,588 compounds). We also included a subset of 2,460 FDA-approved drugs in our screening dataset.

We chose to screen compounds from these two major vendors as it would be convenient to purchase the compounds obtained as hits. We also included a natural product database (2,900 compounds) built in our lab. This database was created using resources like Comprehensive herbal medicinal information system for cancer (CHMIS-C) database[246] and other natural product literatures.[247] Thus a total of 690,986 compounds were used for this virtual screening study.
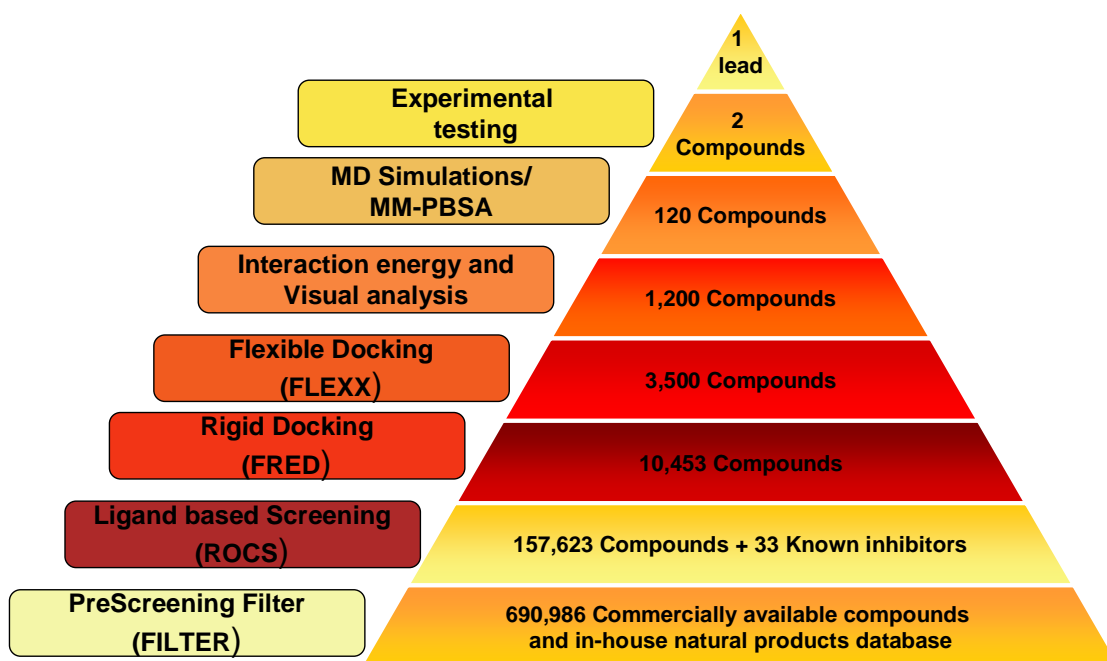
Figure 4.1 Flowchart of the virtual screening strategy used in this study.

As discussed previously few X-ray crystal structures in the protein data bank (PDB) are available for studying PDK1 binding with potential inhibitors. The starting protein structure used in this study is the X-ray crystal structure of adenosine-tri-phosphate (ATP) bound PDK1 (pdb code: 2BIY)[101]. The missing side chain atoms of residues Gln73, Arg75, Glu153, Lys228, Arg238, Lys304, Glu343, Glu348, and Lys357 were modeled using Sybyl 7.0 program. We decided to use a known inhibitor, *i.e.* BX-320,[111] as the bound ligand. The obtained complete protein structure binding with BX-320 was used to carry out an MD simulation. The final stable, MD-simulated structure was used as the initial structure for our virtual screening study.

4.2.2 Pre-screening filter

The entire database was filtered with FILTER v.1.1.1 (OpenEye scientific software, www.eyesopen.com) to eliminate inappropriate or undesirable compounds.[248] The Filter is a molecular screening tool that uses a combination of physical property calculations and functional group knowledge to assess libraries and ultimately remove non-lead-like compounds.[241] Physical properties, atomic or

functional group content, and molecular graph topology are the three categories used by this program to allow or remove a given molecule in the database. The default lead-like filter available in this program was used with some minor variations. The main parameters that we used involve: molecular weight (minimal value = 150 Da, Maximum value = 440 Da, and rings (min=0, max=3), rotatable bonds (min=0, max=10), allowed elements (H, C, N, O, F, S, Cl, and Br), hydrogen bond donor (max=6), hydrogen bond acceptor (max=10). We filtered out molecules with XlogP greater than 4.0, which violates more than one Lipinski rule of five or if it is a known aggregator. Predicted aggregator and GSK_VEBER calculation which is present in the default script was not used in our study. The resulting library contained 157,623 compounds in mol2 format. The Filter input file used in this study is given in Appendix A.

4.2.3 Control data set

In order to check the nature of the run and to evaluate the process, we added a control dataset of 33 known PDK-1 inhibitors. 17 compounds were collected from various scientific reports [109, 110, 111,211, 212 ,249] and 16 compounds from patent applications [250,251,252,253,254]. Thus a total of 33 known inhibitors were used as the control set to quantify the performance of the virtual screen run. To evaluate the performance of the virtual screening run, hit rate and enrichment factor suggested by Wang et $al.$[242] were calculated by using Eqs.(1) and (2):

$$\text{Hit Rate} = \frac{\text{Number of known inhibitors that passed the filter}}{\text{Total number of known inhibitor in the database}} \tag{4-1}$$

$$\text{Enrichment Factor} = \text{Hit Rate} \times \frac{\text{Number of compounds in the database}}{\text{Number of compounds that passed the filter}} \tag{4-2}$$

Enrichment factor is a commonly used metric in virtual screening run.[255] The higher the enrichment factor, the better the performance of a virtual screening run is. The enrichment factor indicates that the final database obtained after the screening is enriched with the active compounds compared to the initial database where it is randomly scattered.[255]

4.2.4 Conformer Generation

We first generated multiple conformations of each ligand in the database by using OMEGA.[256] OMEGA is a systematic conformer generation tool. The algorithm implemented in OMEGA has two main components: 1) model building and 2) torsion driving. It dissects the molecule into fragments and generates fragment library. The molecules were then reassembled from the fragments. Molecular assembly is accomplished by simple vector alignment. By reassembling the fragments, it regenerates many possible combinations and submits each conformer to a simplified energy evaluation. Then all conformers below an energy threshold were compared and those within a certain RMS distance are clustered into a single representation. Atom typing, energy calculations, and geometry optimization in OMEGA were performed using the Merck Molecular Force Field (MMFF). The maximum allowed conformations per compound was set to 400 and the energy window (the value used to discard high-energy conformations) was set to 10 kcal/mol. The default values of OMEGA program were used for other parameters.

4.2.5 Filter-1 (Ligand-based virtual screening)

Shape-based screening was used as the first filter in our virtual screening. The program ROCS was used.[244,43] It was used to find the similarity between the molecules based on their shape. The basic idea behind ROCS is that two objects cannot have the same shape if their volumes are not the same. This method tries to find and quantify the maximal overlap of the volume of two molecules. The goal of this method is to find molecules that can adopt shapes extraordinarily similar to the query without necessarily having similar atom types and bonding patterns.[43] So molecules are optimally aligned and matches are based on the volume overlap. Similarity is quantitated by shape Tanimoto.   This shape Tanimoto has quantity 1.0 if two shapes are identical, and 0.0 if completely different. In the present work ROCS shape searching with chemical complementarity is used. The chemical

complementarities searches used the ImplicitMillsDean chemical forcefield, (also called color forcefield) which defines six chemical types: hydrogen-bond donors, hydrogen-bond acceptors, hydrophobes, anions, cations, and rings. Both shape Tanimoto and color forcefield were used in the ranking of ligands by ROCS.

From the 157,623 compounds, a multi-conformational database was generated using the OMEGA and screened with ROCS. We utilized the co-crystallized structures of 3 PDK1 inhibitors from PDB databank as the ligand query and selected the top ~2.5% (4000) hits for each query. The three ligand queries have three different scaffolds including amino-pyrimidine, bisindolylmaleimide, and indolinone and were retrieved from PDB IDs: 1Z5M, 1UU3, and 1PEO respectively.[109,111,211]

4.2.6 Filter-2 (Rigid docking)

FRED (OpenEye Scientific Software) docking calculations were carried out using protein structures with all hydrogen atoms and with the binding site definitions provided by FRED Receptor program (OpenEye Scientific Software).[245] FRED docking roughly consisted of two steps, *i.e.* shape fitting and optimization. During the shape fitting, the ligand was placed into a 0.5 Å-resolution grid box encompassing all active-site atoms (including hydrogen atoms) using a smooth Gaussian potential.[220]

FRED carries out exhaustive docking of multi-conformer ligand and generation of poses. The pose ensemble was then filtered to reject poses that do not have sufficient shape complimentarity to the protein's active site followed by rejection of poses that do not have at least one heavy atom in adenine pocket near Leu212. The top-ranked poses are optimized by solid body optimization and refined with MMFF forcefield. Preliminary docking trials led us to select OEchemscore for the optimization filters.

4.2.7 Filter-3 (Flexible docking)

We also analyzed docking using FlexX module of SYBYL. The active site was defined as residues within 6.5 Å around the bound ligand. Based on the crystal

structure, the region with residues Leu88, Val96, Leu212, Ala162, and Lys111 were defined as the core sub-pocket. The ligands were docked using the multiple-ligand docking option of the FlexX.

4.2.8 Molecular dynamics in vacuum

We further refined the PDK1-ligand binding structures with the energy minimization followed by a 20-ps MD simulation at $T = 298.15$ K for each binding structure. The energy minimization was performed by using the steepest descent algorithm first until the maximum energy derivative was smaller than 4 kcal/mol/Å and then using the conjugated gradient algorithm until the maximum energy derivative was smaller than 0.001 kcal mol^{-1} Å^{-1}. During the energy minimization and MD simulation, only the ligand and residue side chains in the binding pocket were kept free to move. The nonbonded interaction cutoff and the dielectric constant were set up to group-based (20-Å cutoff distance) and distance-dependent ($\varepsilon = 4r$)[227,257] to mimic the solvent environment, respectively. The MD simulation was performed by using Amber8 program with a time step of 1 fs.

4.2.9 Molecular dynamics in water

The general procedure for carrying out the MD simulations in water is essentially the same as that used in our earlier computational studies in chapter 2. Briefly, the MD simulations were performed using Sander module of Amber8 program.[184] The partial atomic charges for the ligand atoms were calculated using the RESP protocol[185] after electrostatic potential calculations at Hartree-Fock (HF) level with 6-31G* basis set using Gaussian03 program.[186] The PDK1-ligand binding complex was neutralized by adding appropriate counter ions and was solvated in a rectangular box of TIP3P water molecules[187] with a minimum solute-wall distance of 10 Å. The solvated systems were carefully equilibrated and fully energy-minimized. These systems were gradually heated from T = 10 K to T = 298.15 K in 35 ps before a production MD simulation run. The time step used for the MD simulations was 2 fs. Periodic boundary conditions in the NPT ensemble at

T = 298.15 K with Berendsen temperature coupling[188] and P = 1 atm with isotropic molecule-based scaling were applied. The SHAKE algorithm was used to fix all covalent bonds containing hydrogen atoms.[189] The particle mesh Ewald (PME) method[190] was used to treat long-range electrostatic interactions. Restrain was placed on the C-alpha backbone atoms during the MD run. A residue-based cutoff of 12 Å was utilized to the non-covalent interactions. Production MD was then carried out for 1 nanosecond (ns) or more with 2 fs time step. The time-dependent geometric parameters were carefully examined to make sure that we obtained a stable MD trajectory for each simulated protein-ligand binding system. The coordinates of the simulated system were collected every 1 ps during the simulation. 100 snapshots of the simulated structure within the stable MD trajectory were used to perform the MM-PBSA calculations.

4.2.10 Binding free energy calculation

The binding free energies were calculated by using the molecular mechanics-Poisson-Boltzmann surface area (MM-PBSA) free energy calculation method.[168] The MM binding energies were calculated with the Sander module of Amber8 program. Electrostatic solvation free energy was calculated by the finite-difference solution to the Poisson–Boltzmann equation (ΔG_{PB}) as implemented in the Delphi program.[192] The radius used for the solvent probe is 1.4 Å. The MSMS program[195] was used to calculate the SASA for the estimation of the non-polar solvation energy (ΔG_{np}) using with the default parameters, $i.e.$ $\gamma = 0.00542$ kcal/Å^2 and $\beta = 0.92$ kcal/mol. Further, the entropic contribution, -TΔS, to the binding free energy was also calculated at T = 298.15 K by using the NMODE module of Amber8 program which is based on a combination of the standard classical statistical formulas and normal mode analysis.[196,197] The final binding free energy calculated for each protein-ligand binding mode was taken as the average of the ΔG_{bind} values calculated for the 100 snapshots.

4.2.11 PDK1 kinase assay

The *in vitro* kinase assays were performed using the Invitrogen PDK1 assay kit (P2884) according to vendor's instructions. This assay was based on the ability of recombinant PDK1, in the presence of DMSO or the inhibitor, to phosphorylate its substrate peptide (P2925). These phosphopeptides generated during the kinase reaction of PDK1 competes with the fluorescein-labeled phosphopeptides (called as tracer) for binding to anti-phosphothreonine peptide-specific antibodies. This binding is then quantified using fluorescence polarization (FP) technique. FP value was measured using TECAN GENIOS PRO microplate reader.

4.2.12 Cell-based assay

The most active compounds in the kinase assay were submitted to the developmental therapeutics program (DTP) at national cancer institute (NCI).[258, 259] The compounds are screened against 60 human cancer cell lines. In the NCI screen, 60 human cancer cell lines were treated 48 h with 10-fold dilutions of compounds at a minimum of five concentrations (0.01 to 100 μM). The screening was done using a sulforhodamine B protein assay to estimate the cell viability or growth. Using seven absorbance measurements [time zero, (Tz), Control growth, (C), and test growth in the presence of drug at five different concentration levels (Ti)], the percentage growth was calculated at each of the drug concentration levels. The percentage growth was calculated as: [(Ti-Tz)/(C-Tz)]* 100.[258]

4.3 Hierarchical virtual screening analysis and identification of new hits

We have used rapid and lower level methods at the beginning and turned to the more accurate and quantitative methods at the end. In this prospective analysis, we have utilized a subset of 688,086 commercially available inhibitors from Zinc database. In addition to screening commercially available compounds, we included another promising strategy in lead identification, namely the virtual screening of natural products. Natural products are a source of diverse chemical compounds. So we also included an in-house database of 2,900 natural products. Thus a total of

690,986 compounds were subjected to the virtual screening study.

4.3.1 Pre-screening filter & control dataset

We applied the physicochemical property filters (ADME/Tox, Lipinski's rule of five) as the first step (pre-screening filter). The overall reasoning is that it is better to focus on identifying new leads with good potencies as well as good ADMET and pharmacokinetic properties. Previous reports have suggested that it would be more beneficial if we analyze the datasets for drug-like property at the beginning of the virtual screen run.[242] False positives are one of the major problems in virtual screen run. Removing the non-drug like molecule will also help to improve the screening efficinecy and time. We used the lead-like filter with the FILTER software from OpenEye. It was used to remove compounds with reactive functional groups, compounds that violate more than one Lipinski rule of five, exceeding a maximum 2-dimensional polar surface area *etc*. Thus we initially eliminated inappropriate or undesirable compounds from the dataset. Application of this pre-screening filter to the initial dataset of 690,986 compounds gave a final dataset of 157,623 compounds.

The pre-filtered database of 157,623 compounds was used as the starting point for this virtual screening study. Our aim was to perform a prospective analysis to identify new PDK1 inhibitors. In order to check the nature of the run and to evaluate the process, we added a control dataset of 33 known PDK-1 inhibitors reported in literature and patents.

4.3.2 Ligand-based virtual screening

Recent reports have shown the success ligand-based virtual screening method such as ROCS in identifying the known inhibitors and in new lead identification.[43] Because of its speed and good performance in previously reported virtual screening studies,[260,261] ROCS was chosen as the filter for the first ligand-based virtual screening filter. We used three different scaffolds as the query molecule in ROCS. The reason for using three different scaffolds simultaneously is that the choice of the ligand molecule will have a critical effect on the results from ligand-based screening

and the use of diverse ligands but in their native state will help us to overcome the bias of using any one ligand as the query molecule. Moreover it has been reported that the better enrichment can be obtained by using more than one query simultaneously.[262] By selecting the top 4,000 (~2.5%) most similar compounds to the query ligands in this filter we got a total of 12,000 compounds. Removing the repeated structures gave us a total of 10,431 final compounds. This filter was able to retain 28 of the 33 known inhibitors. The hit rate for this step was 85% and the enrichment factor was 12.84. The enrichment factor of ~13 means that we have a 13-fold higher chance to find a real inhibitor in this filtered dataset than from the initial database.

4.3.3 Rigid & Flexible docking

The results from ROCS were subjected to rigid docking using the FRED software from OpenEye. This step is more sophisticated than Filter-1 as it introduces the active site shape in the screening. Since it performs rigid docking, conformations have to be generated again for the hits from ROCS. OMEGA was used to generate the conformations using the default values. On average, 30 conformations were generated per molecule. FRED docking involves exhaustive docking of multi-conformer ligand and generation of poses. OEChemscore was used in the selection of hits. A total of 3,500 compounds were obtained by using the score cutoff of -32.6. 22 known inhibitors were found in this filtered set. The hit rate for this step is 76%. The hit rate for the first two filters together was 66% and the enrichment factor was 29.7

Next we carried out flexible docking using FlexX. In this method, the ligand was treated as flexible and the receptor was held rigid. The FlexX uses anchor and grow strategy to account for flexible ligand docking. Selecting the top 1,200 compounds from FlexX docking gave us a final set of 1,200 compounds. 19 known inhibitors were found in this final set. The hit rate for this step is 86%. The hit rate for the three filters together was 58% and the enrichment factor was 76.2. This means that we have ~76-fold higher chances to find a real inhibitor in this final

dataset (obtained after three filters) than from whole database.

4.3.4 Refinement through interaction energy calculation and visual analysis

To further refine the result, the docked ligands were refined by MD simulations in vacuum. The interaction energies (vdw and electrostatic) of PDK1-ligand complexes were used along with visual analysis and refinement. It should be noted that the active site of PDK1 as well as all known kinases can be divided into 3 regions, namely the adenine pocket, sugar pocket, and phosphate pocket.[112] In order to refine the hits, the compounds that did not fit into the adenine pocket were removed first. We also checked whether there are common chemical features present in the known PDK1 inhibitors, as this could help us in further refining the hits. Superposition of the available PDK1-inhibitor X-ray crystal structures showed that all the inhibitors had a common hydrogen bond acceptor (in the same position) which interacts with the backbone NH group of Ala162 in the hinge region. Some inhibitors have an additional hydrogen bond interaction with the backbone carbonyl group of Ala162. Based on these observations, we removed the compounds that have very low binding affinity (*i.e.* high interaction energy) and that do not have at least one interaction with either i) the backbone -NH group of Ala162 or ii) with the backbone carbonyl group of Ala162 in the hinge region or iii) with Lys111 residue. 120 hits were selected from this process and were further refined using molecular dynamics simulations and MM-PBSA analysis discussed below.

4.3.5 Molecular dynamics simulations and Binding free energy calculations

Our final filter made use of molecular dynamics simulations and binding free energy calculations. Explicit water molecules were used in the molecular dynamics simulation. The ligands and the side chains of the protein were all flexible during this solvated molecular dynamics run for ~1.0 ns. The docked molecules that have stable MD trajectories were selected for further analysis. The binding free energy was calculated using the MM-PBSA method. The compounds having MM-PBSA scores better than -7.0 kcal/mol was chosen as the final hit set. From this hit list, two

compounds were cherry picked (based on the commercial availability or ease of synthesis) for initial experimental testing. The compound chosen for wet experimental study, is ethyl-4-[[2-[-(4-ethoxycarbonylphenyl)amino]-6-methylpyrimidin-4yl]amino] benzoate (**1**) which can be purchased from Analogix, Inc. (Burlington, WI)**.** The other compound is (**2**) which had to be synthesized in house, as it belongs to the natural product database and it was not commercially available.

4.3.6 PDK1 assay

Compound **1** was purchased from Analogix, Inc. (Burlington, WI). Compound **2** belongs to natural product database and was not commercially available. It was synthesized by our colloborator Dr. Wei Wang at University of New Mexico. Mr. Wenchao Yang in our lab carried out the wet experimental tests. Wet experimental tests on the inhibitory activities of the identified compounds were carried out using fluorescence polarization assays. We carried out the PDK1 assays along with a standard reference (Staurosporine) in order to make sure that the inhibitory activity data obtained for the two compounds are comparable to the previously reported inhibitory activity of known compound. Our experimental tests revealed that both compounds **1** and **2** can significantly inhibit PDK1. According to the activity data, we obtained $IC_{50} = $ ~200 nM for compound **1** and $IC_{50} = $ ~20 μM for compound **2**. The natural product compound **2** is norathyirol, the aglycone of mangiferin. Thus by starting with a database of 690,986 molecules, we were able to obtain two new inhibitors of PDK1, by combining sequential virtual screening with experimental testing.

4.3.7 Cell-based assay

For a successful hit it is also important for the compound to be active in cell-based assay. So the activity of the identified PDK1 inhibitors was analyzed further in the cell-based assays. The compound was submitted to the NCI DTP for the 60 cancer cell lines screening. The new PDK1 inhibitors was found to

significantly inhibit many cancer cell lines. When tested at 10 micromolar concentration, norathyriol was found to have a significantly lower inhibitory activity than compound **1**. This is qualitatively consistent with our PDK1 inhibition data. The inhibition data of the new compound against 4 selected cell lines was shown in Figure 4.5. Our new lead compound **1** also effectively inhibit many cancer cell lines, such as multiple myeloma cell line RPMI8226, non non-small cell lung cancer cell line HOP92, colon cancer cell line KM12, CNS cancer cells SF268, Melanoma cell LOX IMVI, Ovarian cancer cells OVCAR-3, Renal cancer cells RXF 393, Prostate cancer cell line PC-3, and Breast cancer cell line HS578T. The growth inhibitory activity of compound **1** against various cell lines are given in Table 4-2.

4.3.8 Binding mode of new hits

Understanding the binding mode of this new PDK1 inhibitor will be of great help in further lead optimization studies. Figures 4.3 and 4.5 shows the binding mode of the hits with PDK1. The compound **1** has a hydrogen bond interaction with the backbone carbonyl group of Ala162 in the hinge region. The pyrimidine ring is present in the hydrophobic adenine pocket of PDK1. It is surrounded by residues Leu212 and Leu88. There is another hydrogen bond interaction between the backbone carbonyl group of Leu88 and the –NH group of the inhibitor. The ester group is present in the sugar region in the pocket does not interact with any specific residue. Based on the PDK1-inhibitor binding modes, the identified PDK1 inhibitor may be used as new starting point in future *de novo* drug design and discovery efforts.
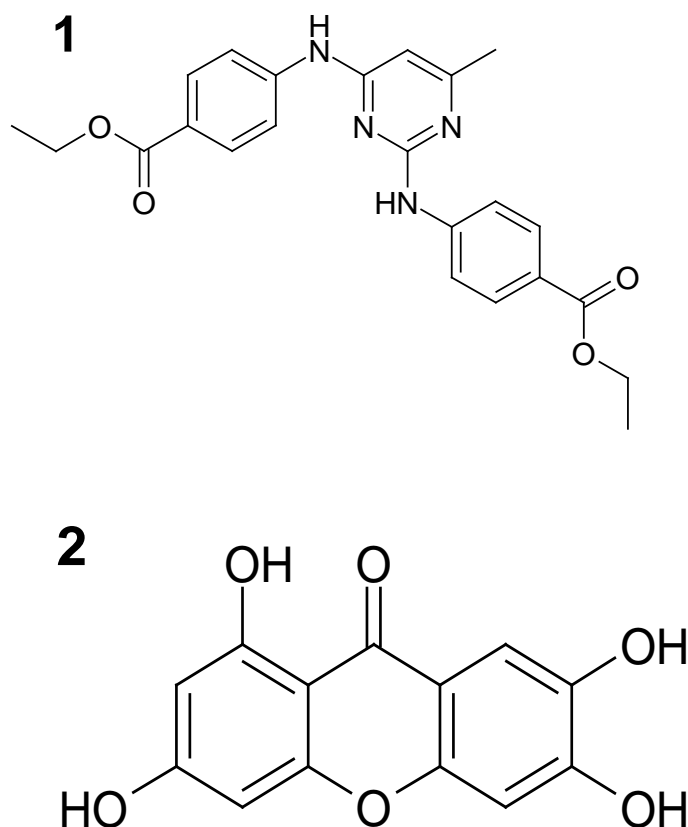
Figure 4.2 Molecular structures of new hits identified through virtual screening.

Table 4.1 Binding free energies (kcal/mol) calculated at T = 298.15 K and P = 1 atm for PDK1 binding with representative inhibitors in comparison with the corresponding experimental data.

Inhibitor	Calc.[a]				Expt.
	ΔE_{MM}	ΔG_{sol}	$-T\Delta S$	ΔG_{bind}	ΔG_{bind}
Celecoxib[b]	-56.5	40.1	11.3	-5.1	-5.9
Compound **1**	-59.6	35.5	12.5	-11.5	-9.1
Compound **2**	-81.3	59.8	13.3	-8.2	-6.4

[a] The MM-PBSA calculations were performed on 100 snapshots along a stable MD trajectory for each PDK1-inhibitor binding complex. The results given in the table are the average values calculated for the 100 snapshots. [b] The data for celecoxib was reported in reference [136].
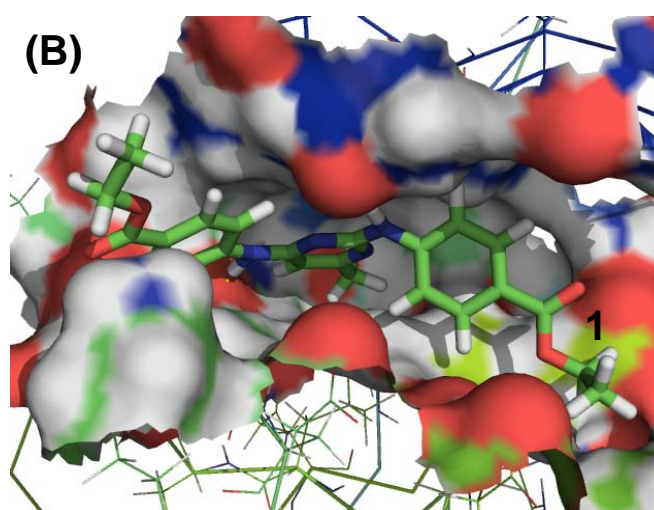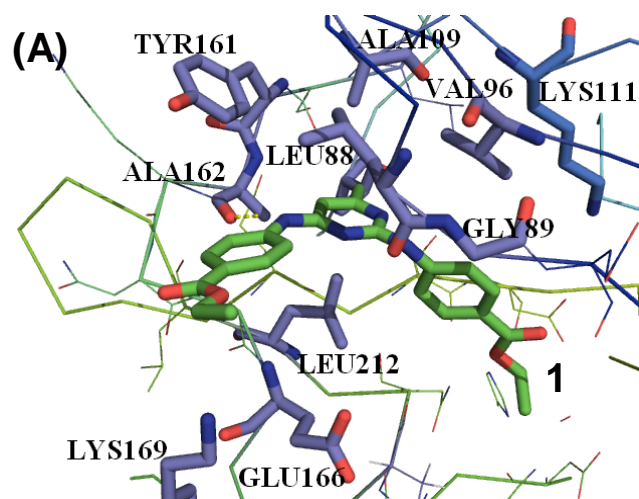
Figure 4.3 Binding of compound **1** in the PDK1 active site. Compound **1** is shown in sticks with carbon atoms in green color.
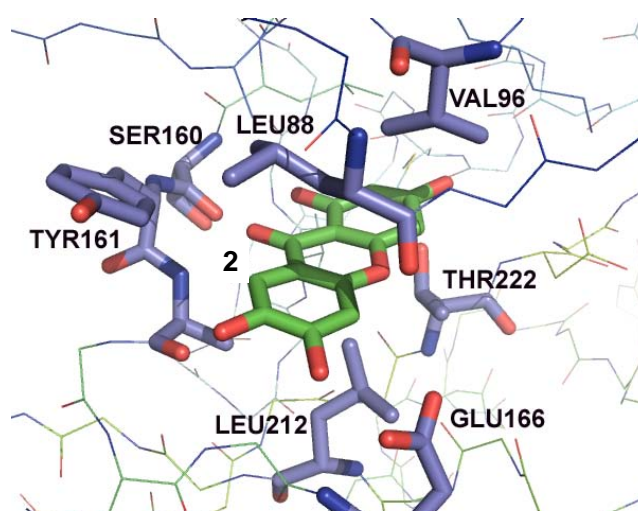
Figure 4.4 Binding of compound **2** in the PDK1 active site. Compound **2** is shown in sticks with carbon atoms in green color.

(A)

Leukemia: RPMI8226
GI_{50}: 3.54 µ M

(B)

Non-small cell
Lung cancer: HOP-92
GI_{50}: 3.25 µ M

(C)

CNS cancer: SF-295
GI_{50}: 8.64 µ M

(D)

Prostate cancer: PC-3
GI_{50}: 6.02 µ M

Log_{10} of inhibitor concentration (Molar)
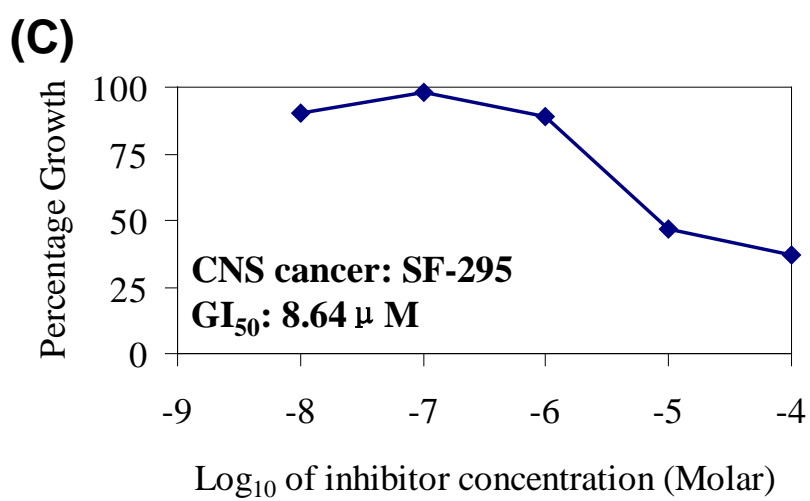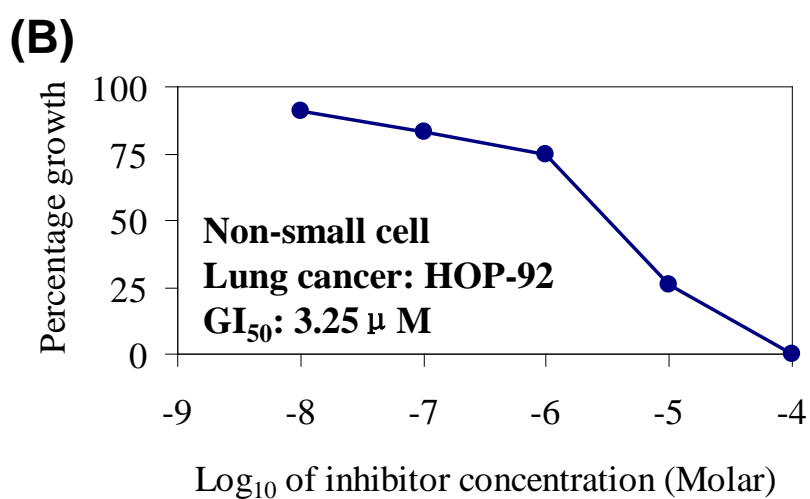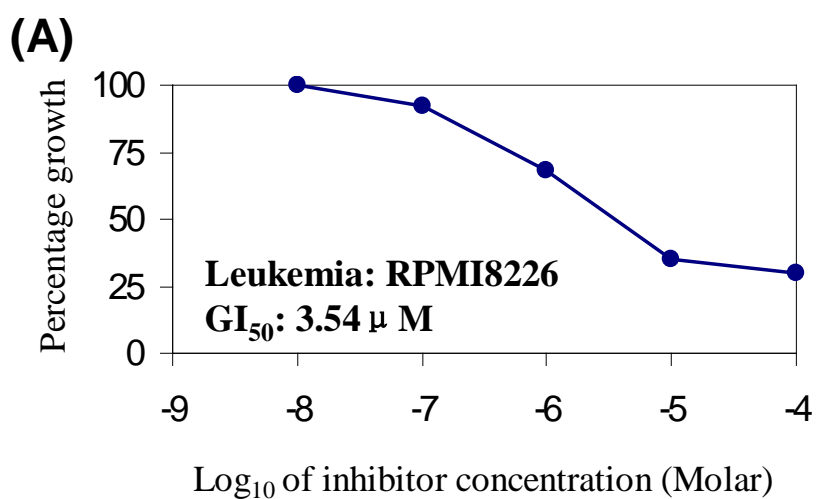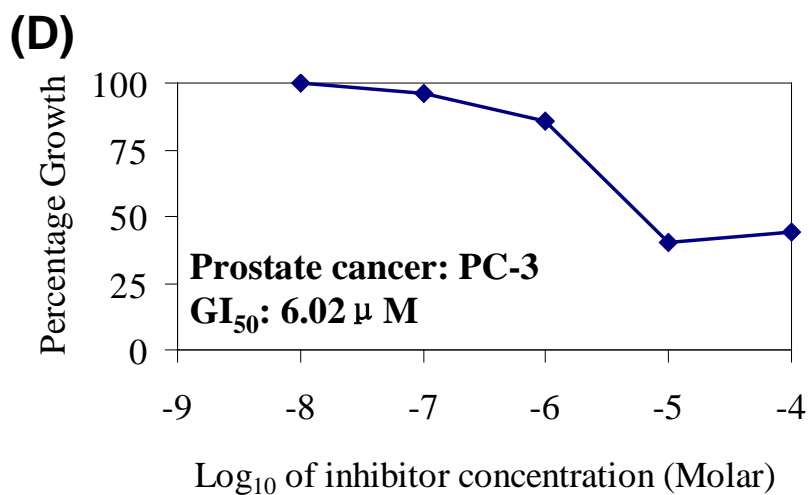
Figure 4.5 Growth inhibitory effect of compound **1** against representative cancer cell lines

(A)

(B)
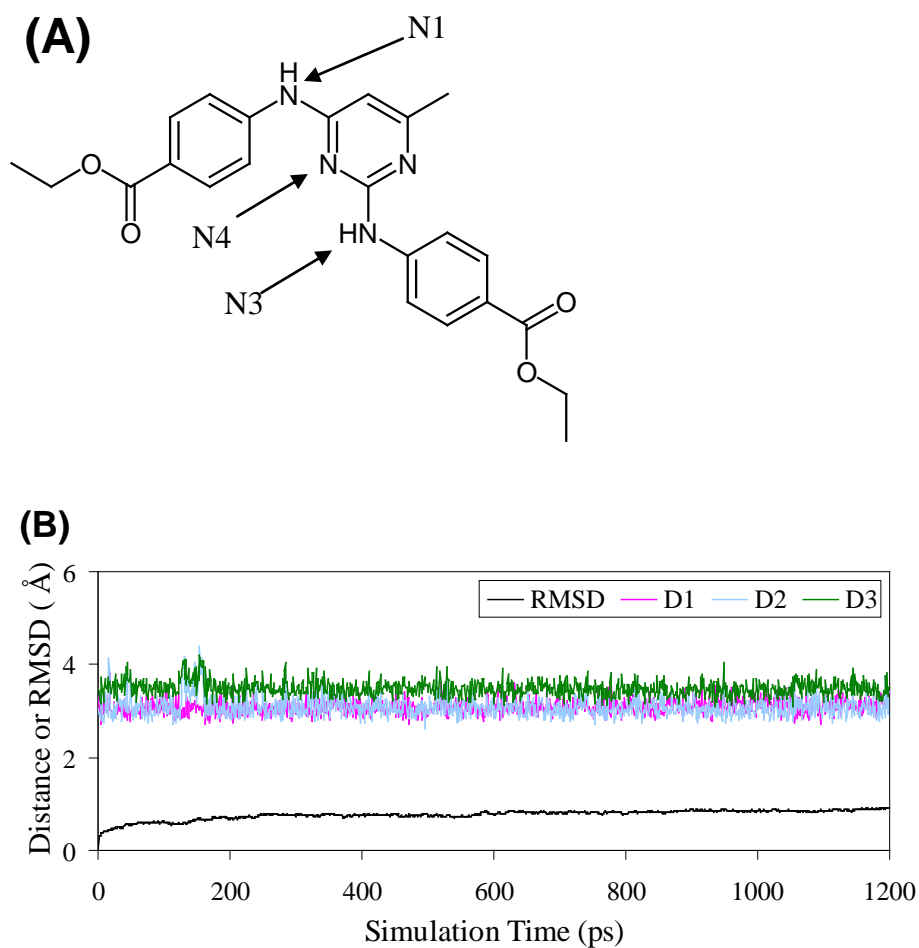
Figure 4.6 Plots of MD-simulated internuclear distances *versus* simulation time for PDK1 binding with compound **1**. D1 refers to the distance between N1 atom of compound **1** and the carbonyl oxygen of Ala162 backbone, D2 the distance between

N3 atom of compound **1** and the carbonyl oxygen of Leu88 backbone, and D3 the distance between N4 atom of compound **1** and the carbonyl oxygen of Leu88 backbone.

(A)

(B)
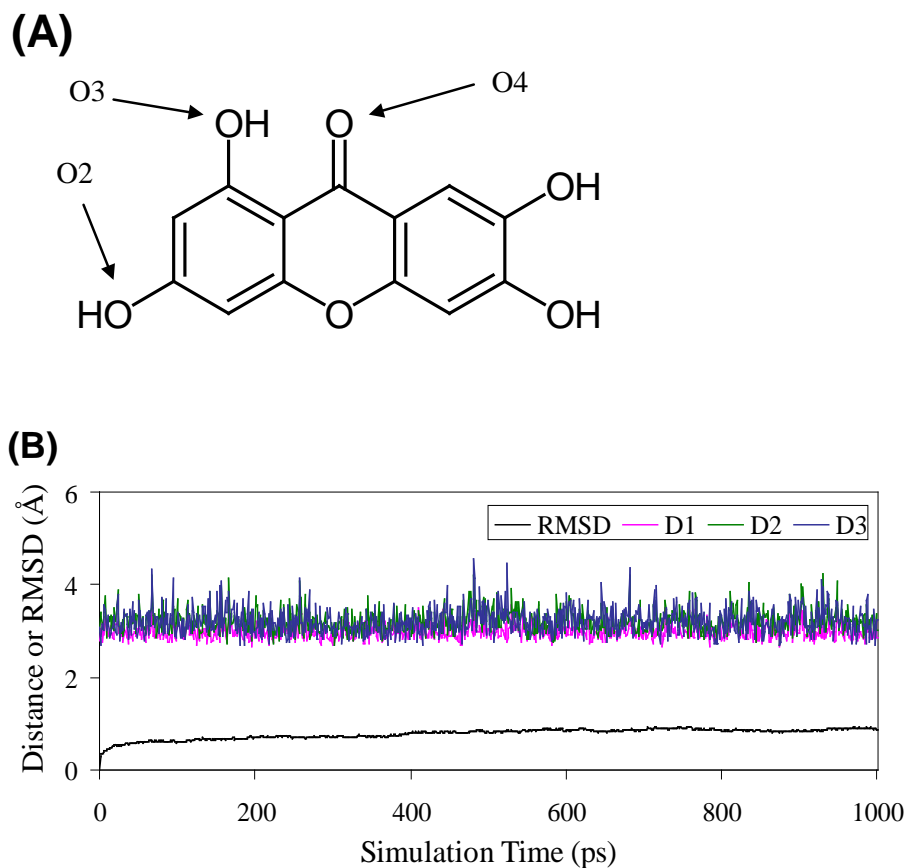
Figure **4.7** Plots of MD-simulated internuclear distances *versus* simulation time for PDK1 binding with compound **2**. D1 refers to the distance between O4 atom of compound **2** and the NH nitrogen of Ala162 backbone, D2 the distance between O3 atom of compound **2** and the carbonyl oxygen of Ser160 backbone, and D3 the distance between O2 atom of compound **2** and the hydroxyl oxygen of Thr222 sidechain.

Table 4.2 Growth inhibitory activity of compound **1**[a] against various cell lines in NCI human cancer cell line panel

No.	Cell Line (Panel name)	GI_{50} (μM)
1.	RPMI-8226 (Leukemia)	3.54
2.	HOP-92 (Non-Small cell Lung Cancer)	3.25
3.	KM12 (Colon Cancer)	3.62
4.	SF-268 (CNS cancer)	3.80
5.	SF-295 (CNS Cancer)	8.64
6.	U251 (CNS Cancer)	9.88
7.	LOX IMVI (Melanoma)	7.24
8.	OVCAR-3 (Ovarian Cancer)	2.02
9.	RXF 393 (Renal Cancer)	5.05
10.	PC-3 (Prostate Cancer)	6.02
11.	DU-145 (Prostate Cancer)	5.39

[a] Data obtained from the NCI-DTP's in-vitro human cancer cell line screen[259]

4.4 Key Findings

A hierarchical virtual screening strategy has been tested and employed to identify new lead compounds against PDK1. We utilized the fast filtering, such as ligand-based screening in the beginning, and carried out the more sophisticated MD/MM-PBSA analysis as a final filter in the virtual screening. The virtual screening of 690,986 compounds eventually led to the selection of five compounds for wet experiments. Among the five compounds selected for wet experiments, two compounds were found to significantly inhibit PDK1. We also used a control dataset to evaluate the virtual screening run. The hit rate and the enrichment factor for the first three filtering steps were found to be 58% and 76.2, respectively. Thus, using virtual screening stratgey combined with the MD simulations and MM-PBSA binding energy analysis, we were able to identify new hits for PDK1 with reduced the amount of experimental efforts. The compounds identified in this study serves as a good starting point for future rational drug design of PDK1 inhibitors. Our virtual screening approach may also be applied in discovery of new lead molecules for other drug targets.

CHAPTER 5

5. SUMMARY AND FUTURE DIRECTIONS

Computational drug design approach has a great potential in accelerating the drug discovery process. Structure-based drug design and ligand-based drug design are the two broad classes of computational drug design. PDK1 is a well validated anti-cancer target and developing inhibitors for PDK1 has the potential to be developed as the anti-cancer therapeutics. In this work, we have showed the potential of computational approaches in drug discovery using PDK1 as our target. We have used both structure-based and ligand-based drug design strategies as well as combined approaches for lead optimization and hit identification against PDK1.

In our structure-based lead optimization, we used celecoxib and its derivatives as our lead molecules. For any structure-based drug design study it is crucial to know whether the proposed binding mode is reasonable or not for future rational drug design. Understanding the molecular basis of interaction is essential to carry out structure-guided lead optimization. One cannot expect to have success in a computational drug design starting with a wrong binding mode. In order to find the reasonable microscopic binding mode for PDK1 binding with celecoxib and its derivatives, we carried out molecular docking, followed by molecular dynamics simulations, and found a totally new binding mode. Since different docking programs gave different results, we used more sophisticated computational techniques like molecular dynamics simulations and MM-PBSA binding free energy calculations. For all of the representative PDK1 inhibitors examined in this study, based on the most favorable binding modes, the calculated binding free energies were all in good agreement with the corresponding experimental activity data. The determined new, more favorable binding mode can explain the SAR for this series of inhibitors.

Another significant contribution is that we have showed for the first time that the novel binding mode without the hydrogen bond interaction in the hinge region is possible for a kinase inhibitor. It will be interesting to analyze in future studies

whether absence of such interaction contributes to the selectivity of the inhibitors against different kinases. Further exploring this type of inhibitors for other kinases has the potential to find new class of inhibitors which will help to overcome crowded intellectual property rights in this field. Moreover we have also shown that only performing simple molecular docking could lead to a wrong binding mode and thus mislead the rational drug design. The computational protocol tested in this study, *i.e*. the molecular docking followed by the combined MD simulations and MM-PBSA calculations, is reliable and accurate for predicting protein-ligand binding structures and binding free energies. This protocol could be used in other drug discovery efforts.

We have also carried out ligand-based lead optimization using indolinone derivatives as our lead molecule. We have developed both ligand-based and receptor-based 3D-QSAR models. This analysis led to the development of satisfactory 3D-QSAR models for predicting the biological activity of new compounds. Our models were able to predict the activity of external test set of compounds which were not used in model development. All of the results obtained from the 3D-QSAR analysis and molecular docking are consistent with the experimental activity data. Since Indolinone is one of the ten most frequently encountered sub-structures in kinase inhibitors, our 3D-QSAR models have the potential to predict and guide the rational design of new ligands of indolinone series as PDK1 inhibitors. We have also showed that the usefulness of ligand-based models. The contour plots obtained from the ligand based models is consistent with the receptor active site. There are many therapeutic targets which do not have the experimentally determined structures. Our study shows that the ligand-based approaches like CoMFA have the potential to give insight into the nature of the active site of the targets in such cases.

We have used hierarchical virtual screening of increasing complexity to identify new hits against PDK1. We have used the ligand-based screening, rigid docking, flexible docking, and molecular dynamics (MD) simulations followed by the MM-PBSA calculations as screening filters in virtual screening. Using this

strategy we have identified a promising new hit for the development of a novel anticancer therapeutic. The inhibitor identified in this study serves as a good starting point for future rational drug design of PDK1 inhibitors. Our virtual screening approach may also be applied to discovery of new lead molecules for other drug targets.

De novo drug design can be carried out starting from the new hit identified from virtual screening. The computational design can be followed by chemical synthesis and *in vitro* enzyme inhibition assays for the inhibitory activity and selectivity. The inhibitors can be tested for their anticancer activities *in vitro*. Further extensive retrospective validation of the shape-based virtual screening approaches and hierarchical virtual screening approaches can be done. Such large-scale validation using already known inhibitors will help to prove the applicability of this approach in a wide variety of target and also document potential advantages and limitations of this approach.

The insights obtained from this research work can also be applied in new drug discovery projects against other therapeutic targets. The different scenarios for new projects include: 1) projects where the experimentally solved structure of target is available; 2) experimentally solved structure of target is not available but structure of related protein (which can serve as template) is available; 3) structures of target and template are not available; 4) experimentally solved structure of target is not available but ligands of target available. In the first case, when the X-ray crystal structure of the target is available, we may carry out structure-based drug design by using the same procedures as we used in this project. Specifically, we may carry out virtual screening and *de novo* design for discovery of new inhibitors. If the experimentally solved structure is not available but the X-ray structure of a related protein (template) is available, then we can carry out homology modeling of the target structure using the template structure. The modeled protein structure will serve as the starting point for structure-based drug design. Then we may follow the same strategies used in this present work, namely virtual screening and *de novo* design, using the modeled target structure. In the third scenario, *i.e.* when both the target structure and template

structure are all unavailable, it might be possible to perform *ab initio* prediction of the target protein structure, as our lab did for anti-inflammatory target mPGES-1.[46] Site-directed mutagenesis experiments may be performed to validate the 3D structure resulted from the *ab initio* protein structure prediction. If the 3D structure is consistent with the experimental studies, then it is reasonable to use the 3D structure for further computational drug design as described above. Finally, if we have to work on a new project without any experimentally solved structure of the target but if there are known ligands of the target, then we may utilize ligand-based drug design strategies as we have done in this work (chapter 3). Ligand-based 3D-QSAR methods like CoMFA and CoMSIA can be utilized to understand the structure-activity relationship and to aid in the design of new inhibitors of this series. The contour plots can also be used to get insights into the nature of the target active site as we have shown in our studies. [201,206]

REFERENCES

[1] Wolff, M. F. Ed. *Burger's medicinal chemistry and drug discovery;* Vol. 1; John Wiley & Sons, Inc: Somerset,NJ 1995.

[2] Lombardino, J. G.; LoweIII, J. A. The role of medicinal chemist in drug discovery-then and now. *Nat. Rev. Drug Discov.* **2004**, *3*, 853-862.

[3] Rawlins, M. D. Cutting the cost of drug development. *Nat. Rev. Drug Discov*. **2004**, *3*, 360-364.

[4] Drews, J. Drug Discovery: A historical perspective. *Science* **2000**, *287*, 1960-1964.

[5] Muller, G. Medicinal Chemistry of target family- directed master keys. *Drug Discov. Today* **2003**, *8*, 681-691.

[6] Owens, J. 2006 drug approvals: finding the niche. *Nat. Rev. Drug Discov*. **2007**, *6*, 99-101.

[7] Ooms, F. Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. *Curr. Med. Chem.* **2000**, *7*, 141-158.

[8] Veselovsky, A. V.; Ivanov, A. S. Strategy of computer-aided drug design. *Curr. drug targets infect. disord.* **2003**, *3*, 33-40.

[9] Cohen, N. C. The molecular modeling perspective in drug design. In *Guidebook on molecular modeling in drug design*; Cohen, N. C. Ed.; Academic Press: San Diego, 1996.

[10] Zhong, S.; Macias, A. T.; Mackerrell Jr, A. D. Computational identification of inhibitors of protein-protein interactions. *Curr. Top. Med. Chem.* **2007**, *7*, 63-82.

[11] Wong, C. F.; McCammon, J. A. Protein flexibility and computer-aided drug design. *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 31-45.

[12] Young, D. C. *Computational drug design. A guide for computational and medicinal chemists*; John Wiley & Sons, Inc: Hoboken, NJ. 2009.

[13] Shoichet, B. K. Virtual screening of chemical libraries. *Nature* **2004**, *432*, 862-865.

[14] Reddy, M. R.; Erion, M. D. Computer aided drug design strategies used in

discovery of fructose 1,6 biphosphatase inhibitors. *Curr. Pharm. Des.* **2005**, *11*, 283-294.

15 Orti, L.; Carbajo, R. J.; Pieper, U.; Eswar, N.; Maurer, S. M.; Rai, A. K.; Taylor, G.; Todd, M. H.; Pineda-Lucena, A.; Sali, A.; Marti-Renom, M. A. A kernal for open source drug discovery in tropical diseases. *PLoS Negl. Trop. Dis.* **2009**, *3*, e418.

16 Delano, W. L. The case for open source software in drug discovery. *Drug Discov. Today* **2005**, *10*, 213-217.

17 Richards, W. D. Virtual screening using grid computing. The screen saver project. *Nat. Rev. Drug Discov.* **2002**, *1*, 551-555.

18 Costanzi, S.; Tikhonova, I. G.; Harden, T. K.; Jacobson, K. A. Ligand and structure based methodologies for the prediction of the activity of G protein-coupled receptor ligands. *J. Comput. Aided Mol. Des.* **2008**. [Epub ahead of print].

19 Congreve, M.; Murray, C. W.; Blundell, T. L. Structural biology and drug discovery. *Drug Discov. Today* **2005**, *10*, 895-907.

20 Nendza, M. *Structure-activity relationships in environmental sciences*; Chapman & Hall. London, 1998.

21 Price, N. R.; Watkins, R. W. Quantitative structure-activity relationships (QSAR) in predicting the environmental safety of pesticides. *Pestic. outlook* **2003**, *14*, 127-129.

22 Gund, T. Molecular modeling of small molecules. In *Guidebook on molecular modeling in drug design*; Cohen, N. C. Ed.; Academic Press: San Diego, 1996.

23 Hansch, C. The physicochemical approach to drug design and discovery (QSAR). *Drug Develop. Res.* **1981**, *1*, 267-309.

24 Estrada, E.; Uriarte, E. Recent advances on the role of topological indices in drug discovery research. *Curr. Med. Chem.* **2001**, *8*, 1573-1588.

25 Selassie, C. D.; Mekapati, S. B. ; Verma, R. P. QSAR : Then and now. *Curr. Med. Chem.* **2002**, *2*, 1357-1379.

26 Hansch, C.; Maloney, P. P.; Fujita, T.; Muir, R. M. Correlation of the biological activity of phenoxyacetic acids with Hammett substituent constants and partition

coefficients. *Nature*. **1962**, *194*, 178-180.

[27] Hansch, C.; Fujita, T. ρ-σ-Π Analysis. A method for the correlation of biological activity and chemical structure. *J. Am. Chem. Soc*. **1964**, *86*, 1616-1626.

[28] Martin, Y. C. A practitioner's perspective of the role of Quatitative structure-activity analysis in medicinal chemistry. *J. Med. Chem.* **1981**, *24*, 229-237.

[29] Garg, R.; Kurup, A.; Mekapati, S. B.; Hansch, C. Cyclooxygenase inhibitors: A comparative QSAR study. *Chem. Rev.* **2003**, *103*, 703-731.

[30] Boyd, D. B. Success of computer-assisted molecular design, In *Reviews in computational chemistry.* Lipkowitz, K. B.; Boyd, D. B. Eds.; VCH, New York, 1990; pp. 355-371.

[31] Kubinyi, H. Quantitative structure-activity relationships (QSAR) and molecular modelling in cancer research. *J. Cancer Res. Clin. Oncol.* **1990**, *116*, 529-537.

[32] Akamatsu, M. Current state and perspectives of 3D-QSAR. *Curr. Top. Med. Chem.* **2002**, *2*, 1381-1394.

[33] Doweyko, A. M. 3D-QSAR illusions. *J. Comput. Aided Mol. Des.* **2004**, *18*, 587-596.

[34] Hahn, M. Receptor Surface Models. 1. Definition and Construction. *J. Med. Chem.* **1995**, *38*, 2080-2090.

[35] Hahn, M.; Rogers, D. Receptor Surface Models. 2. Application to Quantitative structure activity relationship studies. *J. Med. Chem.* **1995**, *38*, 2091-2102.

[36] Bohl, C. E.; Chang, C.; Mohler, M. L.; Chen, J.; Miller, D. D.; Swann, P. W.; Dalton, J. T. A ligand based approach to identify quantitative structure activity relationship for the androgen receptor. *J. Med. Chem.* **2004**, *47*, 3765-3776.

[37] Martin, Y. C. 3D QSAR: Current state, scope, and limitations. *Perspec. Drug. Discov. Design.* **1998**, *12*, 3-23.

[38] Wolber, G.; Seidel, T.; Bendix, F.; Langer, T. Molecule-pharmacophore superpositioning and pattern mathcing in computational drug design. *Drug Discov. Today* **2008**, *13*, 23-29.

[39] Joseph-McCarthy, D. Computational approaches to structure-based ligand design.

Pharmacol. Ther. **1999**, *84*, 179-191.

[40] Mayer, D.; Naylor, C. B.; Motoc, I.; Marshall, G. R. A unique geometry of the active site of angiotensin-converting enzyme consistent with structure-activity studies. *J. Comput. Aided Mol. Des.* **1987**, *1*, 3-16.

[41] Roche, O.; Nettekoven, M.; Vifian, W.; Sarmiento,R. M. R. Refinement of histamine H3 ligands pharmacophore model leads to a new class of potent and selective naphthalene inverse agonists. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4377-4379.

[42] Kurogi, Y.; Guner, O. F. Pharmacophore modeling and three-dimensional database searching for drug design using catalyst. *Curr. Med. Chem.* **2001**, *8*, 1035-1055.

[43] Rush TS, Grant JA, Mosyak L, Nicholls, A. A shape based 3D scaffold hopping method and its application to a bacterial protein-protein interaction. *J. Med. Chem.* **2005,** *48*, 1489-1495.

[44] Scapin, G. Structural biology and drug discovery. *Curr. Pharm. Des.* **2006**, *12*, 2087-2097.

[45] Fischer, P. M. Computer chemistry approaches to drug discovery in signal transduction. *Biotechnol. J.* **2008**, *3*, 452-470.

[46] Huang, X.; Yan, Weili, Y.; Gao, D.; Tong, M.; Tai, H-H.; Zhan, C-G. Structural and functional characterization of human microsomal prostaglandin E synthase-1 by computational modeling and site directed mutagenesis. *Bioorg. Med. Chem.* **2006**, *14*, 3553-3562.

[47] Itai, A.; Mizutani, M. Y.; Nishibata, Y.; Tomioka, N. Computer-assited new lead design. In *Guidebook on molecular modeling in drug design*; Cohen, N. C., Ed.; Academic Press: San Diego, 1996.

[48] Gastreich, M.; Lemmen, C.; Briem, H.; Rarey, M. Addressing the virtual screening challenge: the flex* approach. In *Virtual screening in drug discovery*. Alvarez, J. Shoichet, B. Eds.; CRC Press, 2005.

[49] Klebe, G.; Virtual ligand screening: strategies, perspectives and limitations. *Drug*

Discov. Today **2006**, *11*, 580-594.

50 Kapetanovic, I. M. Computer-aided drug discovery and development (CADD): *In silico* chemico biological approach. *Chem. Biol. Interact.* **2008**, *171*, 165-176.

51 Kitchen, D. B.; Decorenz, H.; Furr, J. R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **2004**, *3*, 935-949.

52 Teague, S. J. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discov.* **2003**, *2*, 527-541.

53 American Cancer Society. *Cancer Facts & Figurs 2008*. Atlanta; American caner society; **2008**.

54 Hirpara, K. V.; Aggarwal, P.; Mukherjee, A. J.; Joshi, N.; Burman, A. C. Quercetin and its derivatives: synthesis, pharmacological uses with special emphasis on anti-tumor properties and prodrug with enhanced bio-availability. *Anticancer Agents Med. Chem.* **2009**, *9*, 138-161.

55 Garcia, M.; Jemal, A.; Ward, E. M.; Center, M. M.; Hao, Y.; Siegel, R. L.; Thun, M. J. Global cancer facts & figures 2007. Atlanta, GA: American Cancer Society, 2007.

56 Collins I.; Workman, P. New approaches to molecular cancer therapeutics. *Nature Chem. Biol.* **2006**, 2, 689-700.

57 Panno, J. *Cancer: The role of genes, lifestyle and environment.* Facts on file, Inc. New York. 2005.

58 Pecorino, L. *Molecular biology of cancer. Mechanisms, targets and therapeutics.* Oxford university press Inc. 2008.

59 Parkin, D. M.; Bray, F.; Ferlay, J.; Pisani, P. Estimating the world cancer burden: Globocan 2000. *Int. J. Canc.* **2001**, *94*, 153-156.

60 Ross, J. A.; olshan, A. F. Pediatric cancer in the United States: The children's oncology group epidemiology research program. *Cancer Epidemiol. Biomarkers Prev.* **2004**, *13*, 1552-1554.

61 Carbone, M.; Pass, H. I. Multistep and multifactorial carcinogenesis: when does a

contributing factor become a carcinogen? *Semin. Cancer Biol.* **2004**, *14*, 399-405.

[62] Clavel, J. Progress in the epidemiological understanding of gene-environment interactions in major diseases: cancer. *C R Biol.* **2007**, *330*, 306-317.

[63] Coleman, W. B.; Tsongalis, G. J. Eds.; *The molecular basis of human cancer.* Humana Press. New Jersey. 2002.

[64] Hanahan, D.; Weinberg, R. A. The hallmarks of cancer. *Cell*, **2000**, *100*, 57-70.

[65] Gibbs, J. B. Mechanism-based target identification and drug discovery in cancer research. *Science* **2000**, *287*, 1969-1973.

[66] Remers, W. A. Antineoplastic agents. In *Wilson and Gisvold's textbook of organic medicinal and pharmaceutical chemistry*; Block, J. H.; Beale Jr, J. M. Eds.; Lippincott Williams & Wilkins. Philadelphia. 2004.

[67] Cao, Y. Molecular mechanisms and therapeutic development of angiogenesis inhibitors. *Adv. Cancer Res.* **2008**, *100*, 113-131.

[68] Bennett, P. N., Brown, M. J., *Clinical Pharmacology;* Churchill Livingstone: 2008

[69] Raguz, S.; Yague, E. Resistance to chemotherapy: new treatments and novel insights into an old problem. *Br. J. Cancer.* **2008**, *99*, 387-391.

[70] Szakacs, G.; Paterson, J.K.; Ludwig, J. A.; Booth-Genthe, C.; Gottesman, M.M. Targeting multidrug resistance in cancer. *Nat. Rev. Drug. Disc.* **2006**, *5*, 219-234.

[71] Jain, R. K.; Tomaso, E. D.; Duda, D. G.; Loeffler, J. S.; Sorensen, A. G.; Batchelor, T. T. Angiogenesis in brain tumor. *Nat. Rev. Neurosci.* **2007**, *8*, 610-622.

[72] Coleman, M. P.; Quaresma, M.; Berrino, F.; Lutz, J. M.; Angelis, R.; Capocaccia, R.; baili, P.; Rachet, B.; Gatta, G.; hakulinen, T.; Micheli, A.; Sant, M.; Weir, H. K.; Elwood, J. M.; Tsukuma, H.; Koifman, S.; E Silva, G. A.; Franscisci, S.; Santaguilani, M.; Verdecchia, A.; Storm, H. H.; Young, J. L.; Concord working group. Cancer survival rate in five continents: a worldwide population-based study (CONCORD). *Lancet Oncol.* **2008**, *9*, 730-756.

[73] Luu, M.; Sabo, E.; De la Monte, S. M.; Greaves, W.; Wang, J.; Tavares, R.; Simao, L.; Wands, J. R.; Resnick, M. B.; Wang, L. Prognostic value of aspartyl (asparaginyl)

–β-hydroxylase humbug expression in non-small cell lung carcinoma. *Hum. Pathol.* **2009**, *40*, 639-644.

[74] Yokoyama, Y.; Nimura, Y.; Nagino, M. Advances in the treatment of pancreatic cancer: limitations of surgery and evaluation of new therapeutic strategies. *Surg. Today.* **2009**, *39*, 466-475.

[75] Li, D.; Xie, K.; Wolff, R.; Abbruzzesse, J. L. Pancreatic Cancer. *Lancet* **2004**, *363*, 1049-1057.

[76] Downward, J. Targeting RAS signaling pathways in cancer therapy. *Nat. Rev.Cancer* **2003**, *1*, 11-22.

[77] McGovern, S. L.; Shoichet, B. K. Kinase inhibitors: Not just for kinases anymore. *J. Med. Chem.* **2003**, *46*, 1478-1483.

[78] Cohen, P. Protein kinases – the major drug targets of the twenty-first century? *Nat. Rev. Drug Discov.* **2002**, *1*, 309-315.

[79] Noble, M. E.; Endicott, J. A.; Johnson, L. N. Protein kinase inhibitors: Insight into drug design from structure. *Science*, **2004**, *303*, 1800-1805

[80] Liao, J.J.L. Molecular recognition of protein kianse binding pocket for design of potent and selective kinase inhibitors. *J. Med. Chem.* **2007**, *50*, 409-424.

[81] Fabian, M.A.; Biggs III, W.H.; Treiber, D.K.; Atteridge, C.E.; Azimioara, M.D.; Benedetti, M.G.; Carter, T.A.; Ciceri, P.; Edeen, P.T.; Floyd, M.; Ford, J.M.; Galvin, M.; Gerlach, J.L.; Gortzfeld, R.M.; Herrgard, S.; Insko, D.E.; Insko, M.A.; Lai, A.G.; Lelias, J-M.; Mehta, S.A.; Milanov, Z.V.; Velasco, A.M.; Wodicka, L.M.; Patel, H.K.; Zarrinkar, P.P.; Lockhart, D.J. A small molecule-kinase interaction map for clinical kinase inhibitors. *Nature Biotech.* **2005**, *23*, 329-336.

[82] Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science*, **2002**, *298*, 1912-1934.

[83] Akritopoulou-Zanze, I.; Hajduk, P. J. Kinase-targeted libraries: The design and synthesis of novel, potent and selective kinase inhibitors. *Drug Discov. Today* **2009**, *14*, 291-297.

[84] Cherry, M.; Williams, D. H. Recent kinase and kinase inhibitor X-ray structures:

Mechanisms of inhibition and selectivity insights. *Curr. Med. Chem.* **2004**, *11*, 663-673.

[85] Halazy, S. Signal transduction: An exciting field of investigation for small molecule drug discovery. *Molecules* **2003**, *8*, 349-358.

[86] Gill, A. New lead generation strategies for protein kinase inhibitors-fragment based screening approaches. *Mini Rev. Med. Chem.* **2004**, *4*, 301-311.

[87] Bogoyevitch, M. A.; Fairlie, D. P. A new paradigm for protein kinase inhibition: blocking phosphorylation without directly targeting ATP binding. *Drug. Discov. Today* **2007**, *12*, 622-633.

[88] Bianco, R.; Melisi, D.; Ciardiello, F.; Tortora, G. Key cancer cell signal transduction pathways as therapeutic targets. *Eur. J. Cancer* **2006**, *42*, 290-294.

[89] Dancey, J.; Sausville, E. A. Issues and progress with protein kinase inhibitors for cancer treatment. *Nat. Rev. Drug Discov.* **2003**, *2*, 296-313.

[90] Zhang, J.; Yang, P. L.; Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer* **2009**, *9*, 28-39.

[91] Luo, J.; Manning, B.D.; Cantley, L.C. Targeting the PI3K-Akt pathway in human cancer: Rationale and promise. *Cancer Cell* **2003**, *4*, 257-262.

[92] Hennessy, B.T.; Smith, D.L.; Ram, P.T.; Lu, Y.; Mills, G.B. Exploiting the PI3K/Akt pathway for cancer drug discovery. *Nature Rev. Drug Discov.* **2005**, *4*, 988-1004.

[93] Bader, A.G.; Kang, S.; Zhao, L.; Vogt, P.K. Oncogenic PI3K deregulates transcription and translation. *Nature Rev. Cancer* **2005**, *5*, 921-929.

[94] Vara, J.A.F.; Casado, E.; Castro, J.D.; Cejas, P.; Belda-Iniesta, C.; Gonzalez-Baron, M. PI3K/Akt signalling pathway and cancer. *Cancer Treat. Rev.* **2004**, *30*, 193-204.

[95] Toker, A.; Newton. A.C. Cellular signaling: pivoting around PDK-1. *Cell.* **2000**, *103*, 185-188.

[96] Fujita, N.; Tsuruo, T. Survival-signaling pathway as a promising target for cancer chemotherapy. *Cancer Chemother. Pharmacol.* **2003**, *52*, S24-S28.

[97] Vanhaesebroeck, B.; Alessi, D.R. The PI3K-PDK1 connection: more than just a road to PKB. *Biochem. J.* **2000**, *346*, 561-576.

[98] Alessi, D.R.; Deak, M.; Casamayor, A.; Caudwell, F.B.; Morrice, N.; Norman, D.G.; Gaffney, P.; Reese, C.B.; MacDougall, C.N.; Harbison, D.; Ashworth, A.; Bownes, M. 3-Phosphoinositide-dependent protein kinase-1 (PDK1): structural and functional homology with the Drosophila DSTPK61 kinase. *Curr. Biol.* **1997**, *7*, 776-789.

[99] Currie, R.A.; Walker, K.S.; Gray, A.; Deak, M.; Casamayor, A.; Downes, C.P.; Cohen, P.; Alessi, D.R.; Lucocq, J. Role of phosphatidylinositol 3,4,5-triphosphate in regulating the activity and localization of 3-phosphoinositide dependent protein kinase-1. *Biochem. J.* **1999**, *337*, 575-583.

[100] Mora, A.; Komander, D.; van Aalten, D.M.F.; Alessi, D.R. PDK1, the master regulator of AGC kinase signal transduction. *Semin. Cell. Dev. Biol.* **2004**, *15*, 161-170.

[101] Komander, D.; Kular, G.; Deak, M.; Alessi, D.R.; VanAalten, D.M.F. Role of T-loop phosphorylation in PDK1 activation, stability, and substrate binding. *J. Biol. Chem.* **2005**, *280*, 18797-18802.

[102] Alessi, D.R.; James, S.R.; Downes, C.P.; Holmes, A.B.; Gaffney, P.R.J.; Reese, C.B.; Cohen, P. Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase Bα. *Curr. Biol.* **1997**, *7*, 261-269.

[103] Brazil, D.P.; Hemmings, B.A.; Ten years of protein kinase B signaling: a hard Akt to follow. *Trends Biochem. Sci.* **2001**, *26*, 657-664.

[104] Pullen, N.; Dennis, P.B.; Andjelkovic, M.; Dufner, A.; Kozma, S.C.; Hemmings, B.A.; Thomas, G. Phosphorylation and activation of p70s6k by PDK1. *Science* **1998**, *279*, 707-710.

[105] Perrotti, N.; He, R.A.; Phillips, S.A.; Haft, C.R.; Taylor, S.I. Activation of serum- and glucocorticoid-induced protein kinase (sgk) by cyclic AMP and insulin. *J. Biol. Chem.* **2001**, *276*, 9406-9412.

[106] Newton, A.C.; Regulation of the ABC kinases by phosphorylation: protein

kinase C as a paradigm. *Biochem. J.* **2003**, *370*, 361-371.

[107] Belham, C.; Wu, S.; Avruch, J. Intracellular signaling: PDK1 – a kinase at the hub of things. *Current Biology* **1999**, *9*, R93-R96.

[108] Biondi, R.M.; Komander, D.; Thomas, C.C.; Lizcano, J.M.; Deak, M.; Alessi, D.R.; VanAalten, D.M.F. High resolution crystal structure of the human PDK1 catalytic domain defines the regulatory phosphopeptide docking site. *EMBO J.* **2002**, *21*, 4219-4228.

[109] Komander, D.; Kular, G.S.; Schuttelkopf, A.W.; Prakash, K.R.C.; Bain, J.; Elliott, M.; Garrido-franco, M.; Kozikowski, A.P.; Alessi, D.R.; VanAalten, D.M.F. Interaction of LY333531 and other bisindolyl malemide inhibitors with PDK1. *Structure* **2004**, *12*, 215-226.

[110] Komander, D.; Kular, G.S.; Bain. J.; Elliott, M.; Alessi, D.R.; Van Aalten, D.M.F. Structural basis for UCN-01 (7-hydroxy staurosporine) specifically and PDK1 (3-phosphoinositide dependent protein kinase-1) inhibition. *Biochem. J.* **2003**, *375*, 255-262.

[111] Feldman, R.I.; Wu, J.M.; Polokoff, M.A.; Kochanny, M.J.; Dinter, H.; Zhu, D.; Biroc, S.L.; Alicke, B.; Bryant, J.; Yuan, S.; Buckman, B.O.; Lentz, D.; Ferrer, M.; Whitlow, M.; Alder, M.; Finster, S.; Chang, Z.; Arnaiz, D.O. Novel small molecule inhibitors of 3-phosphoinositide dependent kinase-1. *J. Biol. Chem.* **2005**, *280*, 19867-19874.

[112] Vulpetti, A.; Bosotti, R. Sequence and structural analysis of kinase ATP pocket residues. *IL Farmaco* **2004**, *59*, 759-765.

[113] Casamayor, A.; Morrice, N.; Alessi, D. R. Phosphorylation of Ser 241 is essential for the activity of PDK1 identification of five sites of phosphorylation in vivo. *Biochem. J.* **1999**, *342*, 287-292.

[114] Wick, M. J.; Ramos, F. J.; Chen, H.; Quon, M. J.; Dong, L. Q.; Liu, F. Mouse 3'-phosphoinositide-dependent protein kianse-1 undergoes dimerization and trans-phosphorylation in the activation loop. *J. Biol. Chem.* **2003**, *278*, 42913-42919.

[115] Ballif, B.A.; Shimamura, A.; Pae, E.; Blenis, J. Disruption of 3-Phosphoinositide-dependent kinase 1 (PDK1) signaling by the anti-tumorigenic and anti-proliferative agent N-α-tosyl-L-phenylalanyl chloromethyl ketone. *J. Biol. Chem.* **2001**, *276*, 12466-12475.

[116] Dong, L. Q.; Liu, F. PDK2: the missing piece in the receptor tyrosine kinase signaling pathway puzzle. *Am. J. Physiol. Endocrinol. Metab.* **2005**, *289*, E187-E196.

[117] Biondi, R. M. Phosphoinositide-dependent protein kinase 1, a sensor of protein conformation. *Trends Biochem Sci.* **2004**, *29*, 136-142.

[118] Leslie, N.R.; Downes, C.P. PTEN: the downside of PI3-kinase signaling. *Cell. Sign.* **2002**, *14*, 285-295.

[119] Gao, X.; Yo, P.; Harris, T.K. Improved yields for baculovirus-mediated expression of human His6-PDK1 and His-PKBβ/AKT2 and characterization of phosphor-specific isoforms for design of inhibitors that stabilize inactive conformations. *Protein Expr. Purif.* **2005**, *43*, 44-56.

[120] Kim, D.; Chung, J. Akt: Versatile mediator of cell survival and beyond. *J. Biochem. Mol. Biol*. **2002**, *35*, 106-115.

[121] Sahoo, S.; Brickley, D.R.; Kocherginsky, M.; Conzen, S.D. Coordinate expression of PI3-Kinase downstream effectors serum and glucocorticoid-induced kinase (SGK-1) and Akt-1 in human breast cancer. *Eur. J. Canc*. **2005**, *41*, 2754-2759.

[122] Gao, N.; Flynn, D.C.; Zhang, Z.; Zhong, X.S.; Walker, V.; Liu, K.J.; Shi, X.; Jiang, B.H. G1 cell cycle progression and the expression of G1 cyclins are regulated by PI3K/Akt/mTOR/p70S6K1 signaling in human ovarian cancer cells. *Am. J. Physiol. Cell. Physiol.* **2004**, *287*, C281-C291.

[123] Smith, J.A.; Poteet-Smith, C.E.; Xu, Y.; Errington, T.M.; Hecht, S.M.; Lannignan, D.A. Identification of the first specific inhibitor of p90 ribosomal S6 kinase (RSK) reveals an unexpected role for RSK in cancer cell proliferation. *Cancer. Res*. **2005**, *65*, 1027-1034.

[124] Flynn, P.; Wong, M.; Zavar, M.; Dean, N.M.; Stokoe, D. Inhibition of PDK-1 activity causes a reduction in cell proliferation and survival. *Curr. Biol.* **2000,** *10,* 1439-1442.

[125] Zeng. X.; Xu, H.; Glazer. R.I. Transformation of mammary epithelial cells by 3-Phosphoinositide dependent protein kinase-1 (PDK-1) is associated with the induction of protein kinase Cα. *Canc. Res*. **2002**, *62*, 3538-3543.

[126] Lin, H-J.; Hsieh, F-C.; Song, H.; Lin, J. Elevated phosphorylation and activation of PDK-1/AKT pathway in human breast cancer. *British J. cancer* **2005**, *93*, 1372-1381.

[127] Ahmed, N.; Riley, C.; Quinn, M. A. An immunohistochemical perspective of PPARβ and one of its putative targets PDK1 in normal ovaries, benign and malignant ovarian tumors. *Br. J. Cancer* **2008**, *98*, 1415-1424.

[128] Cen, L.; Hsieh, F-C.; Lin, H-J.; Chen, C-S.; Qualman, S. J.; Lin, J. PDK-1/AKT pathway as a novel therapeutic target in rhabdomyosarcoma cells using OSU-03012 compound. *Br. J. Cancer.* **2007**, *97*, 785-791.

[129] Liang, K.; Lu, Y.; Li, X. ; Zeng, X.; Glazer, R.I.; Mills, G.B.; Fan, Z. Differential roles of phosphoinositide-dependent protein kinase-1 and Akt1 expression and phosphorylation in breast cancer cell resistance to plaxitaxel, doxorubicin and gemcitabine. *Mol. Pharmacol.* **2006**, *70*, 1045-1052.

[130] Zhang, Q.; Thomas, S.M.; Lui, V.W.Y.; Xi, S.; Siegfried, J.M.; Fan, H.; Smithgall, T.E.; Mills, G.B.; Grandis, J.R. Phosphorylation of TNF-α converting enzyme by gastrin releasing peptide induces amphiregulin release and EGF receptor activation. *Proc. Natl. Acad. Sci. USA.* **2006**, *103*, 6901-6906.

[131] Pinner, S.; Sahai, E. PDK1 regulates cancer cell motility by antagonising inhibition of ROCK1 by RhoE. *Nature Cell Biol.* **2008**, *2*, 127-137.

[132] Nakamura, K.; Sakaue, H.; Nishizawa, A.; Matsuki, Y.; Gomi, H.; Watanabe, E.; Hiramatsu, R.; Tamamori-Adachi, M.; Kitajima, S.; Noda, T.; Ogawa, W.; Kasuga, M. PDK1 regulates cell proliferation and cell cycle progression through control of cyclin D1 and P27^{Kip1} expression. *J. Biol. Chem.* **2008**, *25*, 17702-17711.

[133] Tanaka, H.; Fujita, N.; Tsuruo, T. 3-Phosphoinositide dependent protein kinase-1 mediated IκB kinase β (IKKB) phosphorylation activates NF-κB signaling. *J. Biol. Chem.* **2005**, *280*, 40965-40973.

[134] Lawlor, M.A.; Mora, A.; Ashby, P.R.; Williams, M.R.; Murray-trait, V.; Malone, L.; Prescott, A.R.; Lucocq, J.M.; Alessi, D.R. Essential role of PDK1 in regulating cell size and development in mice. *EMBO J.* **2002**, *21*, 3728-3738.

[135] Bayascas, J.R.; Leslie, N.R.; Parsons, R.; Fleming, S., Alessi, D.R. Hypomorphic mutation of PDK1 suppresses tumorigenesis in PTEN+/- mice. *Curr. Biol.* **2005**, *15*, 1839-1846.

[136] AbdulHameed, M. D.M.; Hamza, A.; Zhan, C-G. Microscopic modes and free energies of 3-phosphoinositide dependent kinaes-1 (PDK1) binding with celecoxib and other inhibitors. *J. Phys. Chem. B.* **2006**, *110*, 26365-26374.

[137] Grosch, S.; Maier, T. J.; Schiffmann, S.; Geisslinger, G. Cyclooxygenase-2 (COX-2) independent anticarcinogenic effects of selective COX-2 inhibitors. *J. Natl. Cancer Inst.* **2006**, *98*, 736-747.

[138] Clemett, D.; Goa, K. L. Celecoxib. A review of its use in osteoarthritis, Rheumatoid arthritis and Acute pain. *Drugs.* **2000**, *59*, 957-980.

[139] Everts, B.; Wahrborg, P.; Hedner, T. COX-2 specific inhibitors- the emergence of a new class of analgesic and anti-inflammatory drugs. *Clin. Rheumatol.* **2000**, *19*, 331-343.

[140] Dajani, E. Z.; Islam, K. Cardiovascular and gastrointestinal toxicity of selective cyclo-oxygenase-2 inhibitors in man. J. Physiol. Pharmacol. 2008, 59, 117-133.

[141] Kalpan-Machilis, B.; Klostermeyer, B. S. The Cyclooxygenase-2 inhibitors: Safety and Effectiveness. *Ann. Pharmacother.* **1999**, *33*, 979-988.

[142] Lehmann, F. S.; Beglinger, C. Impact of COX-2 inhibitors in common clinical practice gastroenterologist's perspective. *Curr. Top. Med. Chem.* **2005**, *5*, 449-464.

[143] http://www.medicinenet.com/celecoxib/article.htm.

[144] Shi, S.; Klotz, U. Clinical use and pharmacological properties of selective COX-2 inhibitors. *Eur. J. Clin. Pharmacol.* **2008**, *64*, 233-252

[145] Davies, N. M.; Jamali, F. COX-2 selective inhibitors cardiac toxicity: getting to the heart of the matter. *J. Pharm. Pharmaceut. Sci.* **2004**, *7*, 332-336.

[146] Mitchell, A. J.; Warner, T. D. COX isoforms in the cardiovascular system: understanding the activities of non-steroidal anti-inflammatory drugs. *Nat. Rev. Drug Discov.* **2006**, *5*, 75-86.

[147] Grosser, T.; Fries, S.; FitzGerald, G. A. Biological basis of the cardiovascular consequences of COX-2 inhibition: therapeutic challenges and opportunities. *J. Clin. Invest.* **2006**, *116*, 4-15.

[148] Cheng, Y.; Wang, M.; Yu, Y.; Lawson, J.; Funk, C. D.; FitzGerald, G. A. Cyclooxygenases, microsomal prostaglandin E synthase-1, and cardiovascular function. *J. Clin. Invest.* **2006**, *116*, 1391-1399.

[149] Reddy, B. S.; Rao, C. V. Novel approaches for colon cancer prevention by cyclooxygenase-2 inhibitors. *J. Environ. Pathol. Toxicol. Oncol.* **2002**, *21*, 155-164.

[150] Suganuma, M.; Kurusu, M.; Suzuki, K.; Tasaki, E.; Fujiki, H. Green tea polyphenol stimulates cancer preventive effects of celecoxib in human lung cancer cells by upregulation of GADD153 gene. *Int. J. Cancer* **2006**, *119*, 33-40.

[151] Reckamp, K.L.; Krysan, K.; Morrow, J.D.; Milne, G.L.; Newman, R.A.; Tucker, C.; Elashoff, R.M.; Dubinett, S.M.; Figlin, R.A. A phase I trial to determine the optimal biological dose of celecoxib when combined with erlotinib in advanced non-small cell lung cancer. *Clin. Cancer Res.* **2006**, *12,* 3381-3388.

[152] Ferrandina, G.; Ranelletti, F.O.; Legge, F.; Salutari, V.; Martinelli, E.; Fattorossi, A.; Lorusso, D.; Zannoni, G.; Vellone, V.; Pagila, A.; Scambia, G. Celecoxib up-regulates the expression of the ζ chain of T cell receptor complex in tumor-infiltrating lymphocytes in human cervical cancer. *Clin. Cancer Res.* **2006**, *12*, 2055-2060.

[153] Zhang, G-S.; Liu, D-S.; Dai, C-W.; Li, R-J. Antitumor effects of celecoxib on K562 leukemia cells are mediated by cell-cycle arrest, caspase-3 activation, and downregulation of COX-2 expression and are synergistic with hydroxyurea or Imatinib. *Am. J. Hematology* **2006**, *81*, 242-255.

154 Hahn, T.; Alvarez, I.; Kobie, J.J.; Ramanathapuram, L.; Dial, S.; Fulton, A.; Besselsen, D.; Walker, E.; Akporiaye, E.T.; Short-term dietary administration of celecoxib enhances the efficacy of tumor lysate-pulsed dendritic cell vaccines in treating murine breast cancer. *Int. J. Cancer* **2006**, *118*, 2220-2231.

155 Klenke, F.M.; Gebhard, M-M.; Ewerbeck, V.; Abdollahi, A.; Huber, P.E.; Sckell. A. The selective Cox-2 inhibitor celecoxib suppresses angiogenesis and growth of secondary bone tumors: An intravital microscopy study in mice. *BMC Cancer* **2006**, *6*, 9.

156 Pruthi, R.S.; Derksen, J.E.; Moore, D.; Carson, C.C.; Grigson, G.; Watkins, C.; Wallen, E. PhaseII trial of celecoxib in prostrate-specific antigen recurrent prostate cancer after definitive radiation therapy or radical prostatectomy. *Clin. Cancer Res*. **2006**, *12*, 2172-2177.

157 Harris, R.E.; Beebe-Donk, J.; Alshafie, G.A. Reduction in the risk of human breast cancer by selective cyclooxygenase-2 (COX-2) inhibitors. *BMC cancer* **2006**, *6*, 27.

158 Steinbach, G.; Lynch, P. M.; Phillips, R. K.; Wallace, M. H.; Hawk, E.; Gordon, G. B.; Wakabayashi, N.; Saunders, B.; Shen, Y.; Fujimura, T.; Su, L. K.; Levin, B. The effect of celecoxib, a cyclo-oxygenase inhibitor, in familial adenomatous polyposis. *N. Engl. J. Med.* **2000**, *342*, 1946-1952.

159 Schiffmann, S.; Maier, T. J.; Wobst, I.; Janssen, A.; Corban-William, H.; Angioni, C.; Geisslinger, G.; Grosch, S. The anti-proliferative potency of celecoxib is not a class effect of coxibs. *Biochem. Pharmacol.* **2008**, *76*, 179-187.

160 Gallego, G. A.; Prado, S. D.; Fonseca, P. J.; Campelo, R. G.; Espinosa, J. C.; Aparicio, L. M. A. Cyclooxygenase-2 (COX-2): a molecular target in prostate cancer. *Clin. Transl. Oncol.* **2007**, *9*, 694-702.

161 Song, X.Q.; Lin, H.P.; Johnson, A.J.; Tseng, P.H.; Yang, Y.T.; Kulp, S.K.; Chen, C.-S. Cyclooxygenase-2, player or spectator in cyclooxygenase-2 inhibitor-induced apoptosis in prostate cancer cells. *J. Nat. Cancer Ins.*. **2002**, *94*, 585-591.

162 Zhu, J.; Song, X.; Lin, H-P.; Young, D.C.; Yan, S.; Marquez, V.E.; Chen, C-S.

154 Hahn, T.; Alvarez, I.; Kobie, J.J.; Ramanathapuram, L.; Dial, S.; Fulton, A.; Besselsen, D.; Walker, E.; Akporiaye, E.T.; Short-term dietary administration of celecoxib enhances the efficacy of tumor lysate-pulsed dendritic cell vaccines in treating murine breast cancer. *Int. J. Cancer* **2006**, *118*, 2220-2231.

155 Klenke, F.M.; Gebhard, M-M.; Ewerbeck, V.; Abdollahi, A.; Huber, P.E.; Sckell. A. The selective Cox-2 inhibitor celecoxib suppresses angiogenesis and growth of secondary bone tumors: An intravital microscopy study in mice. *BMC Cancer* **2006**, *6*, 9.

156 Pruthi, R.S.; Derksen, J.E.; Moore, D.; Carson, C.C.; Grigson, G.; Watkins, C.; Wallen, E. PhaseII trial of celecoxib in prostrate-specific antigen recurrent prostate cancer after definitive radiation therapy or radical prostatectomy. *Clin. Cancer Res*. **2006**, *12*, 2172-2177.

157 Harris, R.E.; Beebe-Donk, J.; Alshafie, G.A. Reduction in the risk of human breast cancer by selective cyclooxygenase-2 (COX-2) inhibitors. *BMC cancer* **2006**, *6*, 27.

158 Steinbach, G.; Lynch, P. M.; Phillips, R. K.; Wallace, M. H.; Hawk, E.; Gordon, G. B.; Wakabayashi, N.; Saunders, B.; Shen, Y.; Fujimura, T.; Su, L. K.; Levin, B. The effect of celecoxib, a cyclo-oxygenase inhibitor, in familial adenomatous polyposis. *N. Engl. J. Med.* **2000**, *342*, 1946-1952.

159 Schiffmann, S.; Maier, T. J.; Wobst, I.; Janssen, A.; Corban-William, H.; Angioni, C.; Geisslinger, G.; Grosch, S. The anti-proliferative potency of celecoxib is not a class effect of coxibs. *Biochem. Pharmacol.* **2008**, *76*, 179-187.

160 Gallego, G. A.; Prado, S. D.; Fonseca, P. J.; Campelo, R. G.; Espinosa, J. C.; Aparicio, L. M. A. Cyclooxygenase-2 (COX-2): a molecular target in prostate cancer. *Clin. Transl. Oncol.* **2007**, *9*, 694-702.

161 Song, X.Q.; Lin, H.P.; Johnson, A.J.; Tseng, P.H.; Yang, Y.T.; Kulp, S.K.; Chen, C.-S. Cyclooxygenase-2, player or spectator in cyclooxygenase-2 inhibitor-induced apoptosis in prostate cancer cells. *J. Nat. Cancer Ins.*. **2002**, *94*, 585-591.

162 Zhu, J.; Song, X.; Lin, H-P.; Young, D.C.; Yan, S.; Marquez, V.E.; Chen, C-S.

Using Cyclooxygenase-2 inhibitors as molecular platforms to develop a new class of apoptosis-inducing agents. *J. Natl. Cancer Ins*. **2002**, *94*, 1745-1757.

[163] Schonthal, A. H. induction of apoptosis by celecoxib in cell culture: uncertain role for cyclooxygenase-2. *Cancer Res.* **2007**, *67*, 5575- 5576.

[164] Arico, S.; Pattingre, S.; Bauvy, C.; Gane, P.; Barbat, A.; Codogno, P.; Ogier-Denis, E. Celecoxib induces apoptosis by inhibiting 3-phosphoinositide dependent protein kinase -1 activity in the human colon cancer HT-219 cell line. *J. Biol. Chem*. **2002**, *277*, 27613-27621.

[165] 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. From the cyclooxygenase-2 inhibitor celecoxib to a novel class of 3-phosphoinositide dependent protein kinase-1 inhibitors. *Cancer Res*. **2004**, *64*, 4309-4318.

[166] Fabbro, D.; Ruetz, S.; Buchdunger, E.; Cowan-Jacob, S.W.; Fendrich, G.; Liebetanz, J.; Mestan, J.; O'Reilly, T.; Traxler, P.; Chaudhuri, B.; Fretz, H.; Zimmermann, J.; Meyer, T.; Caravatti, G.; Furet, P.; Manley, P.W. Protein kinases as targets for anticancer agents: from inhibitors to useful drugs. *Pharmacol. Ther*. **2002**, *93*, 79-98.

[167] Kallblad, P.; Mancera, R.L.; Todorov, N.P. Assessment of multiple binding modes in ligand-protein docking. *J. Med. Chem*. **2004**, *47*, 3334-3337.

[168] Kollman, P.A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D.A.; Cheatham, T.E. 3rd. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000,** *33*, 889-897.

[169] Hamza, A.; Zhan, C.-G. How can (-)-epigallocatechin gallate from green tea prevent HIV-1 virus infection? Mechanistic insights from computational modeling and the implication for rational design of anti-HIV-1 entry inhibitors. *J. Phys. Chem. B* **2006**, *110*, 2910-2917.

[170] Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.;

Shindyalov, I.N.; Bourne, P.E   The protein databank. *Nucleic Acids Res.* **2000**, *28*, 235-242.

[171] Tripos Associates, Inc., 1699 S. Hanley, St. Louis MI 63144, USA.

[172] Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470-489.

[173] Rarey, M.; Wefing, S.; Lengauer, T. Placement of medium sized molecular fragments into active sites of proteins. *J. Comput. Aided Mol. Design* **1996**, *10*, 41-54.

[174] Clark, M.; Cramer III, R. D.; Van Opdenhosch, N. Validation of the General Purpose Tripos 5.2 Force Field. *J. Comp. Chem.* **1989**, *10*, 982-1012.

[175] Shoichet, B.K.; Kuntz, I.D.; Bodian, D.L. Molecular docking using shape descriptors. *J. Comp. Chem.* **1992**, *13*, 380-397.

[176] Meng, E.C.; Shoichet, B.K.; Kuntz, I.D. Automated docking with grid-based energy evaluation. *J. Comp. Chem.* **1992**, *13*, 505-524

[177] Meng, E.C.; Gschwend, D.A.; Blaney, J.M.; Kuntz, I.D. Oriental sampling and rigid-body minimization in molecular docking. *Proteins: Struct. Funct. Genet.* **1993**, *17*, 268-278.

[178] Ewing, T.J.A.; Makino, S.; Skillman, A.G.; Kuntz, I.D. DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *J. Comput. Aided Mol. Des.* **2001,** *15*, 411-428.

[179] Bohm, H. J. The development of simple emprical scoring function to estimate the binding constant for a protein ligand complex of known 3-dimensional structure. *J. Comp. Aided Mol. Des.* **1994**, *8, 243-*256.

[180] Klebe, G. The use of composite crystal-field environments in molecular recognition and the de-novo design of protein ligands. *J. Mol. Biol.* **1994**, *237*, 212-235.

[181] Knegtel, R. M. A.; Bayada, D. M.; Engh, R. A.; Saal, W. V. D.; Geerestein, J. V.; Grootenhius, D. J. Comparison of two implementations of the incremental construction algorithm in flexible docking of thrombin inhibitors. *J. Comp. Aided*

Mol. Des. **1999**, *13,* 167-183.

[182] Moustakas, D. T.; Pegg, S. C. H.; Kuntz, I. D. A practical guide to DOCK 5. In *Virtual screening in drug discovery*. Alvarez, J. Shoichet, B. Eds.; CRC Press, 2005.

[183] Desjarlais, R.L.; Sheridan, R.P.; Seibel, G.L.; Dixon, J.S.; Kuntz, I.D.; Venkataraghavan, R. Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. *J. Med. Chem.* **1988**, *31*, 722-729.

[184] Case, D.A.; Darden, T.A.; Cheatham, III, T.E.; Simmerling, C.L.; Wang, J.; Duke, R.E.; Luo, R.; Merz, K.M.; Wang, B.; Pearlman, D.A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J.W.; Ross, W.S.; Kollman, P.A. AMBER 8, 2004, University of California, San Francisco.

[185] Bayly, C.I.; Cieplak, P.; Cornell, W.D.; Kollman, P.A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges – the RESP model. *J. Phys. Chem.* **1993**. *97*, 10269-10280.

[186] Frisch, M. J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Montgomery, J.A. Jr.; Verven, T.; Kudin, K.N.; Burant, J. C.; Millam, J.M.; Iyengar, S.S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G.A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J.E.; Hratchian, H.P.; Cross, J.B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J.W.; Ayala, P.Y.; Morokumo, K.; Voth, G.A.; Salvador, P.; Dannenberg, J.J.; Zakrzewski, V.G.; Dapprich, S.; Daniels, A.D.; Strain, M.C.; Farkas, O.; Malick, D.K.; Rabuck, A.D.; Raghavachari, K.; Foresman, J.B.; Ortiz, J.V.; Cui, Q.; Baboul, A.G.; Clifford, S.; Cioslowski, J.; Stefanov, B.B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R.L.; Fox, D.J.; Keith, T.; Al-Laham, M.A.; Peng, C.Y.; Nanayakkara, A.; Challacombe, M.; Gill, P.M.; Wong, M.W.; Gonzalez, C.; Pople, J.A. GAUSSIAN 03, Revision A.1, 2003, Gaussian, Inc.,

Pittsburgh.

[187] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D.; Klein, M. L. J. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.***1983,** *79*, 926–935.

[188] Berendsen, H. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684-3690.

[189] Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp. Phys.* **1977**, *23*, 327-341.

[190] Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T. A.; Lee, H., Pedersen; L. G. A smooth particle mesh ewald method. *J. Chem. Phys.* **1995**, *103*, 8577-8593.

[191] Gilson, M. K.; Sharp, K. A.; Honig, B.H. Calculating electrostatic potential of molecules in solution: Method and error assessment. *J. Comput. Chem.* **1988,** *9***,** 327-335.

[192] Jayaram, B.; Sharp, K.A.; Honig, B.H. The electrostatic potential of B-DNA. *Biopolymers* **1989,** *28***,** 975-993.

[193] Sitkoff, D.; Sharp, K.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978-1988.

[194] Cornell, W.; Cieplak, P.; Bayly, C.; Goud, I.; Merz, K.; Ferguson, D.; Spellmayer, D.; Fox, T.; Caldwell, J.; Kollman, P.A. A 2nd generation force-field for the simulation of proteins, nucleic acids and organic-molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

[195] Sanner, M.F.; Olson, A.J.; & Spehner, J.C. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers* **1996,** *38*, 305-320.

[196] McQuarrie, D. A. A. *Statistical Mechanics*; Harper & Row: New York, 1976.

[197] Brooks, B. R.; Janezic, D.; Karplus, M. Harmonic analysis of large systems. 1. Methodology. *J. Comput. Chem.* **1995**, *16*, 1522-1542.

[198] Liu, Y.; Gray, N. S. Rational design of inhibitors that bind inactive kinase

conformations. *Nat. Chem. Biol.* **2006**, *2*, 358-364.

[199] Richardson, C. M.; Nunns, C. L. ; Williamson, D. S.; Parratt, M. J.; Dokurno, P.; Howes, R.; Borgognoni, J.; Drysdale, M. J.; Finch, H.; Hubbard, R. E.; Jackson, P. S.; Kierstan, P.; Lentzen, G.; Moore, J. D.; Murray, J. B.; Simmonite, H.; Surgenor, A. E.; Torrance, C. J. Discovery of a potent CDK2 inhibior with a novel binding mode, using virtual screening and initial, structure-guided lead scoping. *Bioorg. Med. Chem. Lett.* **2007,** 17, 3880-3885.

[200] Cohen, M. S.; Zhang, C.; Shokat, K. M.; Taunton, J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. *Science*, **2005**, *308*, 1318-1321.

[201] AbdulHameed, M. D. M.; Hamza, A.; Liu, J.; Zhan, C-G. Combined 3D-QSAR modeling and molecular docking study on indolinone derivatives as inhibitors of 3-Phosphoinositide dependent kinase-1. *J. Chem. Inf. Mol. Mod.* **2008**, *48*, 1760-1772.

[202] Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959-5967.

[203] Klebe, G.; Abraham, U.; Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* **1994**, *37*, 4130-4146.

[204] Kubinyi, H.; Comparative molecular field analysis (CoMFA), in: *The Encyclopedia of computational chemistry*. Schleyer, P. V. R.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P.A.; Schaefer III, H.F.; Schreiner, P. R. Eds.;Volume 1, John Wiley & Sons: Chichester, 1998, 448-460.

[205] Bohm, M.; Sturzebecher, J.; Klebe, G. Three dimensional quantitative structure activity relationship analyses using comparative molecular field analysis and comparative molecular similarity indices analysis to elucidate selectivity differences of inhibitors binding to trypsin, thrombin, and factor Xa. *J. Med. Chem.* **1999**, *42*, 458-477.

206 AbdulHameed, M. D. M.; Hamza, A.; Liu, J.; Huang, X.; Zhan, C.-G. Human microsomal prostaglandin E synthase -1 (mPGES-1) binding with inhibitors and the quantitative structure-activity correlation. *J. Chem. Inf. Mod.* **2008**, 48, 179-185.

207 Kuo, C. L.; Assefa, H.; Kamath, S.; Brzozowski, Z.; Slawinski, J.; Saczewski, F.; Buolamwini, J. K.; Neamati, N. Application of CoMFA and CoMSIA 3D-QSAR and docking studies in optimization of mercaptobenzenesulphonamides as HIV-1 integrase inhibitors. *J. Med. Chem.* **2004**, *47*, 385-399.

208 Yang, G-F.; Lu, H-T; Xiong, Y.; Zhan, C-G. Understanding the structure activity and structure-selectivity correlation of cyclic guanine derivatives as phosphodiesterase-5 inhibitors by molecular docking, CoMFA and CoMSIA analyses. *Bioorg. Med. Chem.* **2006**, *14*, 1462-1473.

209 Debnath, A. K. Quantitative structure-activity relationship (QSAR) paradigm –Hansch era to new millennium. *Mini Rev. Med. Chem.* **2001**, *1*, 187-195.

210 Faivre, S.; Demetri, G.; Sargent, W.; Raymond, E. Molecular basis for Sunitinib efficacy and future clinical development. Nat. Rev. Drug Discov. 2007, 6, 734-745.

211 Islam, I.; Bryant, J.; Chou, Y. L.; Kochanny, M. J.; Lee, W.; Phillips, G. B.; Yu, H. Y.; Adler, M.; Whitlow, M.; Ho, E.; Lentz, D.; Polokoff, M. A.; Subramanyam, B.; Wu, J. M.; Zhu, D. G.; Feldman, R. I.; Arnaiz, D. O. Indolinone based phosphoinositide-dependent kinase-1 (PDK1) inhibitors. Part 1: Design, synthesis and biological activity. *Bioorg. Med. Chem. Lett.* **2007,** 17, 3814-3818.

212 Islam, I.; Brown, G.; Bryant, J.; Hrvatin, P.; Kochanny, M. J.; Phillips, G. B.; Yuan, S. D.; Adler, M.; Whitlow, M.; Lentz, D.; Polokoff, M. A.; Wu, J.; Shen, J.; Walters, J.; Ho, E.; Subramanyam, B.; Zhu, D. G.; Feldman, R. I.; Arnaiz, D. O. Indolinone based phosphoinositide-dependent kinase-1 (PDK1) inhibitors. Part 2: Optimization of BX-517. *Bioorg. Med. Chem. Lett.* **2007,** 17, 3819-3825.

213 Kubinyi, H. *QSAR: Hansch analysis and related approaches;* VCH Verlag. Weinheim, 1993.

214 Thaimattam, R.; Daga, P. R.; Banerjee, R.; Iqbal, J. 3D-QSAR studies on c-Src kinase inhibitors and docking analices of a potent dual kinase inhibitor of c-Src and c-Abl kinases. *Bioorg. Med. Chem.* **2005**, *13*, 4704-4712.

[215] Besler, B. H.; Merz Jr., K. M.; Kollman, P. A. Atomic charges derived from semiemprical methods. *J. Comp. Chem.* **1990**, *11*, 431-439.

[216] Singh, U. C.; Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comp. Chem.* **1984**, *5*, 129-145.

[217] Wold, S.; Ruhe, A.; Wold, H.; Dunn, W. J. III. The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM J. Sci. Stat. Comput.* **1984**, *5*, 735-743.

[218] Bush, B. L.; Nachbar, R. B. Sample-distance partial least squares: PLS optimized variables, with the application to CoMFA. *J. Comp Aided Mol. Design.* **1993**, 7, 587-619.

[219] Morris, G. M.; Goodsell, D. S.; Halliday, R.S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using Lamarckian genetic algorithm and empirical binding free energy function. *J. Comp. Chem.* **1998**, *19*, 1639-1662.

[220] McGann MR, Almond HR, Nicholls A, Grant JA, Brown FK. Gaussian docking functions. *Biopolymers* **2003**, *68*, 76-90.

[221] OEChem, version 1.4.2, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2005

[222] Case, D.A.; Darden, T.A.; Cheatham, III, T.E.; Simmerling, C.L.; Wang, J.; Duke, R.E.; Luo, R.; Merz, K.M.; Pearlman, D.A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W.S.; Kollman, P.A. Amber9, 2006, University of California, San Francisco.

[223] Bostrom J, Greenwood JR, Gottfries J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *J. Mol. Graph. Model.* **2003**, 21, 449-462.

[224] Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput. Aided. Mol. Des.* **1997**,

11, 425-445.

[225] Gehlhaar DK, Verkhivker GM, Rejto PA, Sherman CJ, Fogel DB, Fogel LJ, Freer ST. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem Biol* **1995,** *2,* 317-324.

[226] Stahl M, Rarey M. Detailed analysis of scoring functions for virtual screening. *J. Med. Chem.* **2001,** 44, 1035-1042

[227] Harvey, S. C. Treatment of electrostatic effects in macromolecular modeling. *Proteins* **1989**, *5*, 78-92.

[228] Sperandiodasilva, G. M.; Sant'Anna, C. M. R.; Barreiro, E. J. A novel 3D-QSAR comparative molecular field analysis (CoMFA) model of imidazole and quinazolinone functionalized p38 MAPkinase inhibitors. *Bioorg. Med. Chem.* **2004**, *12*, 3159-3166.

[229] Xie, A.; Sivaprakasam, P.; Doerksen, R. J. 3D-QSAR analysis of antimalarial farnesyltransferase inhibitors based on a 2,5-diaminobenzophenone scaffold. *Bioorg. Med. Chem.* **2006**, *14*, 7311-7323.

[230] Briens, F.; Bureau, R.; Rault, S.; Robba, M. Applicability of CoMFA in ecotoxicology: A critical study on chlorophenols. *Ecotoxicol. Environ. Saf.* **1995**, *31*, 37-48.

[231] Aronov, A. M.; Mcclain, B.; Moody, C. S.; Murcko, M. A. Kinase-likeness and kinase-privileged fragments: Towards virtual polypharmacology. *J. Med. Chem.* **2008,** *51*, 1214-1222.

[232] Cramer, R. D.; Wendt, B. Pushing the boundaries of 3D-QSAR. *J. Comput. Aided. Mol. Des.* **2007**, *21*, 23-32.

[233] Tonmunphean, S.; Kokpol, S.; Parasuk, V.; Wolschann, P.; Winger, R. H.; Liedl, K. R.; Rode, B. M. Comparative molecular field analysis of artemisinin derivatives: Ab initio versus semiempirical optimized structures. *J. Comp. Aided Mol. Des.* **1998**, *12*, 397-409.

[234] McGauhey G. B., Sheridan R.P.; Bayly C.I., Culberson J.C., Kreatsoulas C.,

Lindsley S., Maiorov V., Truchon J-F., Cornell W.D. Comparison of topological, shape and docking methods in virtual screening. *J. Chem. Inf. Model.* **2007**, *47*,1504-1519.

[235] Bissantz C, Bernard P, Hibert M, Rognam D. Protein-based virtual screening of chemical databases. II. Are homology models of G-protein coupled receptors suitable targets? *Proteins: Struct Funct Genet,* **2003**, *50*, 5-25.

[236] Doman, T.N.; McGovern, S.L.; Witherbee, B.J.; Kasten, T.P.; Kurumbail, R., Stallings, W.C.; Connolly, D.T.; Shoichet, B.K. Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *J. Med. Chem.* **2002**, *45*, 2213-2221.

[237] Gao, Y.; Dickerson, J.B.; Guo, F.; Zheng, J.; Zheng, Y. Rational design and characterization of a Rac GTPase-specific small molecule inhibitor. *Proc Natl Acad Sci USA*, **2004**, *101*, 7618-7623.

[238] Gundla, R.; Kazemi, R.; Sanam, R.; Muttineni, R.; Sarma, JARP.; Dayam, R.; Neamati, N. Discovery of novel small-molecule inhibitors of human epidermal growth factor receptor -2: combined ligand and target-based approach. *J. Med. Chem.* **2008**, *51*, 3367-3377.

[239]Schnecke, V.; Kuhn, L.A. Virtual screening with solvation and ligand-induced complementarity. *Perspect. Drug. Discovery Des.* **2000**, *20*, 171-190.

[240] Vigers, G.P.A; Rizzi, J.P. Multiple active site corrections for docking and virtual screening. *J. Med. Chem.* **2004**, *47,* 80-89.

[241] Miteva, M. A.; Lee, W. H.; Montes, M. O.; Villoutreix, B. O. Fast structure-based virtual ligand screening combining FRED, DOCK, and Surflex. *J. Med. Chem.* **2005**, *48*, 6012-6022.

[242] Wang, J.; Kang, X.; Kuntz, I.D.; Kollman, P.A. Hierarchical database screenings for HIV-1 reverse transcriptase using a pharmacophore model, rigid docking, solvation docking, and MM-PB/SA. *J. Med. Chem.* **2005**, *48,* 2432-2444.

[243] Irwin, J. J.; Shoichet, B. K. ZINC – a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177-182.

[244]ROCS, version 2.3.1, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2007.

[245]Fred, version 2.2.3, OpenEye Scientific Software, Inc., Santa Fe, NM, USA, www.eyesopen.com, 2007.

[246] Fang, X.; Shao, L.; Zhang, H.; Wang, S. CHMIS-C: A comprehensive herbal medicine information system for cancer. *J. Med. Chem.* **2005**, *48,* 1481-1488.

[247] WHO monographs on selected medicinal plants, Vol-I. World Health Organization: Geneva, 1999.

[248] FILTER v.2.0.1 OpenEye scientific software, www.eyesopen.com

[249] Hardcastle, I. R.; Arris, C. E.; Bentley, J.; Boyle, F. T.; Chen, Y.; Whitfield, H. J. N2-Substituted O6-Cyclohexylmethylguanine derivatives: Potent inhibitors of Cyclin-dependent kinases 1 and 2. *J. Med. Chem.* **2004**, *47*, 3710-3722.

[250] Binch, H.; Fraysse, D.; Robinson, D. PCT Intl. Appl. WO2006050249 2006.

[251] Walmsey, D. L.; Drysdale, M. J.; Northfield, C. J.; Fromont, C. PCT Intl. Appl. WO2006134318 2006.

[252] Ramurthy, S.; Lin, X.; Subramanian, S.; Rico, A.; Wang, X. M.; Jain, R.; Murray, J. M.; Bashman, S. E.; Warne, R. L.; Shu, W.; Zhou, Y.; Dove, J.; Aikawa, M.; Amiri, P.; Wang, W.; Jensen, J. M.; Wagman, A. S.; Pfister, K. B.; Ng, S. C. PCT Intl. Appl. WO2007117607 2007.

[253] Arnold, W. D.; Bounaud, P.; Gosberg, A.; Li, Z.; Mcdonald, I.; Steensma, R. W.; Wilson, M. E. PCT Intl. Appl. WO2006015123 2006.

[254] Binch, H.; Brenchley, G.; Golec, J. M.C.; Knegtel, R.; Mortimore, M.; Patel, S.; Rutherford, A. WO2003064397 2003.

[255] Sheridan, R. P.; McGaughey, G. B.; Cornell, W. D. Multiple protein structures and multiple ligands: effects on the apparent goodness of virtual screening results. *J. Comput. Aided Mol. Des.* **2008**, *22*, 257-265.

[256] Omega V2.2 Open Eye Scientific Software www.eyesopen.com

[257] Guenot, J.; Kollman, P. A. Molecular dynamics studies of a DNA-binding protein: 2. An evaluation of implicit and explicit solvent models for the molecular dynamics simulation of the Escherichia coli trp repressor. *Protein Sci*. **1992,** *1*, 1185–1205

[258] http://dtp.nci.nih.gov/screening.html

[259] Monks, A.; Scudiero, D.; Skehan, P.; Shoemaker, R.; Paull, K.; Vistica, D.; Hose, C.; Langley, J.; Cronise, P.; Vaigro-Wolff, A. Feasibility of a high flux anticancer drug screen using a diverse panel of cultured human tumor cell lines. *J. Natl. Cancer. Inst.* **1991,** *86,* 1853-1859.

[260] Hawkins P. C.; Skillman, A. G.; Nicholls, A. Comparison of shape matching and docking as virtual screening tools. *J. Med. Chem.* **2007**, *50*, 74-82.

[261] Venhorst. J.; Nunez, S.; Terpstra, J. W.; Kruse, C. G. Assesment of scaffold hopping efficiency by use of molecular interaction fingerprints. *J. Med. Chem.* **2008**, *51*, 3222-3229.

[262] Franke, L.; Schwarz, O.; Muller-Kuhrt, L.; Hernig, C.; Fischer, L.; George, S.; Tanrikulu, Y.; Schneider, P.; Werz, O.; Steinhilber, D.; Schneider, G. Identification of natural product derived inhibitors of 5-Lipoxygenase activity by ligand-based virtual screening. *J. Med. Chem.* **2007**, *50*, 2640-2646.

APPENDIX I Filter used in virtual screening

FILTER used in PDK1 virtual screening run

#/***

#Copyright (C) 2000-2005 by OpenEye Scientific Software, Inc.

#***

#This file defines the rules for filtering multi-structure files based on properties and substructure

#patterns.

MIN_MOLWT   150   "Minimum molecular weight"

MAX_MOLWT   440   "Maximum molecular weight"

MIN_NUM_HVY   10   "Minimum number of heavy atoms"

MAX_NUM_HVY   35   "Maximum number of heavy atoms"

MIN_RING_SYS   0   "Minumum number of ring systems"

MAX_RING_SYS   3   "Maximum number of ring systems"

MIN_RING_SIZE   0   "Minimum atoms in any ring system"

MAX_RING_SIZE   20   "Maximum atoms in any ring system"

MIN_CON_NON_RING   0   "Minimum number of connected non-ring atoms"

MAX_CON_NON_RING   15   "Maximum number of connected non-ring atoms"

MIN_FCNGRP   0   "Minimum number of functional groups"

MAX_FCNGRP   12   "Maximum number of functional groups"

MIN_UNBRANCHED   0   "Minimum number of connected unbranched non-ring atoms"

MAX_UNBRANCHED   3   "Maximum number of connected unbranched non-ring atoms"

MIN_CARBONS   5   "Minimum number of carbons"

MAX_CARBONS   35   "Maximum number of carbons"

MIN_HETEROATOMS   2   "Minimum number of heteroatoms"

MAX_HETEROATOMS   12   "Maximum number of heteroatoms"

MIN_Het_C_Ratio   0.10   "Minimum heteroatom to carbon ratio"

MAX_Het_C_Ratio   1.1   "Maximum heteroatom to carbon ratio"

MIN_HALIDE_FRACTION   0.0   "Minimum Halide Fraction"

MAX_HALIDE_FRACTION   0.5   "Maximum Halide Fraction"

#count ring degrees of freedom = (#BondsInRing) - 4 - (RigidBondsInRing) - (BondsSharedWithOtherRings)

#must be >= 0, from JCAMD 14:251-265,2000.

ADJUST_ROT_FOR_RING   true   "BOOLEAN for whether to estimate degrees of

freedom in rings"

```
MIN_ROT_BONDS          0          "Minimum number of rotatable bonds"
MAX_ROT_BONDS          10         "Maximum number of rotatable bonds"

MIN_RIGID_BONDS        0          "Minimum number of rigid bonds"
MAX_RIGID_BONDS        25         "Maximum number of rigid bonds"

MIN_HBOND_DONORS   0          "Minimum number of hydrogen-bond donors"
MAX_HBOND_DONORS   4          "Maximum number of hydrogen-bond donors"

MIN_HBOND_ACCEPTORS   0          "Minimum number of hydrogen-bond acceptors"
MAX_HBOND_ACCEPTORS   6          "Maximum number of hydrogen-bond acceptors"

MIN_LIPINSKI_DONORS   0          "Minimum number of hydrogens on O & N atoms"
MAX_LIPINSKI_DONORS   5          "Maximum number of hydrogens on O & N atoms"

MIN_LIPINSKI_ACCEPTORS   0          "Minimum number of oxygen & nitrogen atoms"
MAX_LIPINSKI_ACCEPTORS   10         "Maximum number of oxygen & nitrogen atoms"

MIN_COUNT_FORMAL_CRG      0          "Minimum number formal charges"
MAX_COUNT_FORMAL_CRG      3          "Maximum number of formal charges"

MIN_SUM_FORMAL_CRG    -2          "Minimum sum of formal charges"
MAX_SUM_FORMAL_CRG     2          "Maximum sum of formal charges"

MIN_CHIRAL_CENTERS                     0                     "Minimum
chiral centers"
MAX_CHIRAL_CENTERS                                         4
"Maximum chiral centers"

MIN_XLOGP         -5.0          "Minimum XLogP"
MAX_XLOGP          4.0          "Maximum XLogP"
```

#choices are insoluble<poorly<moderately<soluble<very<highly
```
MIN_SOLUBILITY        moderately        "Minimum solubility"

PSA_USE_SandP      false      "Count S and P as polar atoms"
MIN_2D_PSA         0.0        "Minimum 2-Dimensional (SMILES) Polar Surface Area"
MAX_2D_PSA         150.0      "Maximum 2-Dimensional (SMILES) Polar Surface Area"

AGGREGATORS       true       "Eliminate known aggregators"
#PRED_AGG         true       "Eliminate predicted aggregators"
```

#secondary filters (based on multiple primary filters)

144

#GSK_VEBER   true   "PSA>140 or >10 rot bonds"
MAX_LIPINSKI   1   "Maximum number of Lipinski violations"
MIN_ABS 05   "Minimum probability F>10% in rats"
PHARMACOPIA   true   "LogP > 588 orPSA > 131.6"
ALLOWED_ELEMENTS   H,C,N,O,F,S,Cl,Br
ELIMINATE_METALS Sc,Ti,V,Cr,Mn,Fe,Co,Ni,Cu,Zn,Y,Zr,Nb,Mo,Tc,Ru,Rh,Pd,Ag,Cd
#acceptable molecules must have <= instances of each of the patterns below
#specific, undesirable functional groups
RULE   0   quinone
RULE   0   pentafluorophenyl_esters
RULE   0   paranitrophenyl_esters
RULE   0   HOBT_esters
RULE   0   triflates
RULE   0   lawesson_s_reagent
RULE   0   phosphoramides
RULE   0   beta_carbonyl_quat_nitrogen
RULE   0   acylhydrazide
RULE   0   cation_C_Cl_I_P_or_S
RULE   0   phosphoryl
RULE   0   alkyl_phosphate
RULE   0   phosphinic_acid
RULE   0   phosphanes
RULE   0   phosphoranes
RULE   0   imidoyl_chlorides
RULE   0   nitroso
RULE   0   N_P_S_Halides
RULE   0   carbodiimide
RULE   0   isonitrile
RULE   0   triacyloxime
RULE   0   cyanohydrins
RULE   0   acyl_cyanides
RULE   0   sulfonylnitrile
RULE   0   phosphonylnitrile
RULE   0   azocyanamides
RULE   0   beta_azo_carbonyl
RULE   0   polyenes
RULE   0   saponin_derivatives
RULE   1   cytochalasin_derivatives
RULE   4   cycloheximide_derivatives
RULE   1   monensin_derivatives
RULE   1   squalestatin_derivatives

#functional groups which often eliminate compounds from consideration
RULE   0   acid_halide

RULE   0   aldehyde
RULE   0   alkyl_halide
RULE   0   anhydride
RULE   0   azide
RULE   0   azo
RULE   0   di_peptide
RULE   0   michael_acceptor
RULE   0   beta_halo_carbonyl
RULE   0   nitro
RULE   0   oxygen_cation
RULE   0   peroxide
RULE   0   phosphonic_acid
RULE   0   phosphonic_ester
RULE   0   phosphoric_acid
RULE   0   phosphoric_ester
RULE   0   sulfonic_acid
RULE   0   sulfonic_ester
RULE   0   tricarbo_phosphene
RULE   0   epoxide
RULE   0   sulfonyl_halide
RULE   0   halopyrimidine
RULE   0   perhalo_ketone
RULE   0   aziridine
RULE   1   oxalyl
RULE   0   alphahalo_amine
RULE   0   halo_amine
RULE   0   halo_alkene
RULE   0   acyclic_NCN
RULE   0   acyclic_NS
RULE   0   SCN2
RULE   0   terminal_vinyl
RULE   0   hetero_hetero
RULE   0   hydrazine
RULE   0   N_methoyl
RULE   0   NS_beta_halothyl
RULE   0   propiolactones
RULE   0   nitroso
RULE   0   iodoso
RULE   0   iodoxy
RULE   0   noxide

#groups of molecules
RULE   2   dye

#functional groups which are allowed, but may not be wanted in high quantities
#common functional groups
RULE   6   alcohol
RULE   4   alkene
RULE   4   amide
RULE   4   amino_acid
RULE   2   amine
RULE   4   primary_amine
RULE   4   secondary_amine
RULE   4   tertiary_amine
RULE   2   carboxylic_acid
RULE   6   halide
RULE   0   iodine
RULE   2   ketone
RULE   4   phenol
RULE   1   imine
RULE   1   methyl_ketone
RULE   1   alkylaniline
RULE   4   sulfonamide
RULE   1   sulfonylurea
RULE   0   phosphonamide
RULE   0   alphahalo_ketone
RULE   0   oxaziridine
RULE   1   cyclopropyl
RULE   2   guanidine
RULE   0   sulfonimine
RULE   0   sulfinimine
RULE   1   hydroxamic_acid
RULE   0   phosphoryl
RULE   0   sulfinylthio
RULE   0   disulfide
RULE   0   enol_ether
RULE   0   enamine
RULE   0   organometallic
RULE   0   dithioacetal
RULE   1   oxime
RULE   0   isothiocyanate
RULE   0   isocyanate
RULE   3   lactone
RULE   3   lactam
RULE   1   thioester
RULE   1   carbonate
RULE   0   carbamic_acid
RULE   1   thiocarbamate

RULE   0   triazine
RULE   1   malonic
#other functional groups
RULE   2   alkyne
RULE   4   aniline
RULE   4   aryl_halide
RULE   2   carbamate
RULE   3   ester
RULE   5   ether
RULE   1   hydrazone
RULE   0   nonacylhydrazone
RULE   1   hydroxylamine
RULE   2   nitrile
RULE   2   sulfide
RULE   2   sulfone
RULE   2   sulfoxide
RULE   1   thiourea
RULE   1   thioamide
RULE   1   thiol
RULE   2   urea
RULE   0   hemiketal
RULE   0   hemiacetal
RULE   0   ketal
RULE   1   acetal
RULE   0   aminal
RULE   0   hemiaminal
#protecting groups
RULE   0   benzyloxycarbonyl_CBZ
RULE   0   t_butoxycarbonyl_tBOC
RULE   0   fluorenylmethoxycarbonyl_Fmoc
RULE   1   dioxolane_5MR
RULE   1   dioxane_6MR
RULE   1   tetrahydropyran_THP
RULE   1   methoxyethoxymethyl_MEM
RULE   2   benzyl_ether
RULE   2   t_butyl_ether
RULE   0   trimethylsilyl_TMS
RULE   0   t_butyldimethylsilyl_TBDMS
RULE   0   triisopropylsilyl_TIPS
RULE   0   t_butyldiphenylsilyl_TBDPS
RULE   1   phthalimides_PHT
RULE   2   arenesulfonyl

Mohamed Diwan Mohideen AbdulHameed

Education

Ph.D. in Pharmaceutical sciences, 08/2004 –present
Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky
Lexington, KY
Master of Pharmacy (Pharmaceutical Chemistry), 08/2000 – 03/2002
Birla Institute of Technology, Mesra, India
Bachelor of Pharmacy , 08/1996 – 05/2000
Madurai Medical College, Madurai, India

Awards

- 2nd Best poster award in Rho Chi Research day at College of Pharmacy, University of Kentucky, April 2009
- Bristol Myers Squibb sponsored AAPS Drug Design and Discovery section award for excellence in graduate research work, November 2008
- 'Genentech travel award' for American Association of Pharmaceutical Scientists (AAPS) annual meeting, San Diego, November 2007
- University Grants Commission (India) Research Fellowship, 2000 – 2002

Publications

1. **AbdulHameed, M.D.M.;** Hamza, A.; Zhan, C-G. Microscopic modes and free energies of 3-Phosphoinositide dependent kinase-1 (PDK-1) binding with celecoxib and other inhibitors. *J. Phys. Chem. B.* **2006**, *110*, 26365-26374.
2. **AbdulHameed, M.D.M.;** Hamza, A.; Liu, J.; Huang, X.; Zhan, C-G. Human microsomal prostaglandin E synthase -1 (mPGES-1) binding with inhibitors and the quantitative structure-activity correlation. *J. Chem. Inf. Mod.* **2008***, 48,* 179-185.
3. Hamza, A.; **AbdulHameed, M.D.M.;** Zhan, C-G. Understanding microscopic binding of human microsomal prostaglandin E synthase-1 with substrates and inhibitors by molecular modeling and dynamics simulation. *J. Phys. Chem. B.* **2008**, *112*, 7320-7329. (**Cover Article**)
4. **AbdulHameed, M.D.M.;** Hamza A.; Liu, J.; Zhan, C-G. Combined 3D-QSAR modeling and molecular docking study on indolinone derivatives as inhibitors of 3-Phosphoinositide dependent kinase-1. *J. Chem. Inf. Mod.* **2008,** *9*, 1760-1772.
5. **AbdulHameed, M.D.M.;** Hamza A.; Yang, W.; Zhan, C-G. Hierarchical Virtual Screening and its Application on 3-Phosphoinositide Dependent Protein Kinase-1 (PDK1) **2009**, *submitted to J. Comp. Aided Mol. Des.*

6. **AbdulHameed, M.D.M.**; Liu, J.; Pan, Y.; Silva-Rivera, C.; Zhan, C-G. 3D-QSAR analysis and docking studies on butyrylcholinesterase inhibitors. **2009**, *submitted to J. Phys. Chem. B.*

7. **AbdulHameed, M.D.M.**; Wehenkel, M.; Hamza, A.; Muzyka, J.; Kim. K-B.; Zhan, C-G. The selectivity of the immunoproteasome catalytic subunit LMP2-specific inhibitor revealed by molecular modeling studies. *To be submitted for publication.*

Presentations

1. **AbdulHameed, M.D.M.**; Baskar, M.J.; Harikrishnan, P.; Rameshkumar, K.; Sivasankaran, S.; Kumar, V; Ismail, A.; Verma, S.M. Pharmacognostical, Phytochemical, Larvicidal and Anti- microbial screening of the stem bark of *Crataeva religiosa* Frost. <u>Oral presentation</u> at Indian Pharmaceutical Congress, New Delhi, India, **December 2001**

2. **AbdulHameed, M.D.M.**; Nanga, R.; Allu, J.R.; Sagar, D.; Annabhimoju, R.; Rambabu. G. Computer aided design of Protein tyrosine phosphatase-1B (PTP-1B) inhibitors as anti-diabetic agents. <u>Oral presentation</u> at Indian Pharmaceutical Congress, Chennai, India, **December 2003.**

3. **AbdulHameed, M.D.M.**; Hamza, A.; Zhan, C-G. Computational modeling of microscopic mode of celecoxib derivatives binding with 3-phosphoinositide dependent kinase-1 (PDK-1). <u>Poster</u> presented at American Association of Pharmaceutical Scientists (**AAPS**) annual meeting, San Antonio, **October 2006.**

4. **AbdulHameed, M.D.M.**; Hamza, A.; Zhan, C-G. Microscopic binding of 3-phosphoinositide dependent kinase -1 (PDK-1) with celecoxib and other inhibitors. <u>Poster</u> presented at 233rd American Chemical Society (**ACS**) National Meeting, Chicago, **March 2007.**

5. **AbdulHameed, M.D.M.**; Hamza, A.; Zhan, C-G. Structure-based design and discovery of novel inhibitors of 3-Phosphoinositide dependent protein kinase-1 (PDK-1). <u>Poster</u> presented at American Association of Pharmaceutical Scientists (**AAPS**) annual meeting, San Diego, **November 2007**.

6. Hamza, A.; **AbdulHameed, M.D.M.**; Zhan, C-G. Structure-based modeling of microsomal prostaglandin E synthase -1 (mPGES-1). <u>Poster</u> presented at American Association of Pharmaceutical Scientists (**AAPS**) annual meeting, Atlanta, **November 2008**.

7. **AbdulHameed, M.D.M.**; Hamza, A.; Yang, W.; Wang, W.; Zhan, C-G. Rational design and discovery of novel inhibitors of 3-Phosphoinositide dependent protein kinase-1 (PDK-1) for anticancer therapeutics. <u>Oral Presentation</u> at American Association of Pharmaceutical Scientists (**AAPS**) annual meeting, Atlanta, **November 2008**.

Professional Organizations

- American Chemical Society.
- American Association of Pharmaceutical Scientists.