PKD1 3D Structure Model and Docking Studies for New PKD Inhibitors

by

Qi Xu

Bachelor of Science, Hebei Medical University, 2011

Submitted to the Graduate Faculty of

School of Pharmacy in partial fulfillment

of the requirements for the degree of

Master of Sciences

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

SCHOOL OF PHARMACY

This dissertation was presented

by

Qi Xu

It was defended on

August 7th, 2013

and approved by

Xiangqun Xie, PhD, Professor

Lirong Wang, PhD, Research Assistant Professor

M. Maggie Folan, PhD, Assistant Professor

Dissertation Advisor: M. Maggie Folan, PhD, Assistant Professor

PKD1 Structure and Docking Studies for New PKD Inhibitors

Qi Xu, M.S.

University of Pittsburgh, 2013

Protein kinase Ds (PKDs) are diacylglycerol (DAG)-regulated serine/threonine protein kinases. In intact cells, PKDs are key mediators in cellular processes pertaining to multiple diseases, including cancer, heart diseases, angiogenesis and immune dysfunctions. A number of the novel, potent, and structurally diverse ATP-competitive PKD inhibitors have been reported to selectively modulate the PKD activity and thus, to achieve a potential therapeutic effect on related diseases. Due to a lack of the crystal structure, we have constructed a 3D structure of the human PKD1 protein by using homology modeling. Then, by using our established protein docking protocol, we docked novel PKD inhibitory small molecules and found the hit compounds exhibiting higher binding scores with reasonable binding mode in comparison with the reported active PKD1 inhibitors. Also, we calculated both 2D and 3D molecular similarity between our identified compounds and previously reported PKD1 inhibitors. Moreover, we predicted the possible off-targets of our compounds and our prediction has been validated through a topomer similarity study. In this study, we demonstrated that computational tools, i.e., docking and molecular similarity calculation can be applied to explore the PKD1/inhibitor interactions. In addition, the docking studies and the detailed docking poses provide insight for better understanding of the possible mechanism of a bioactive PKD1 inhibitor in order to guide future optimization for new drug design and discovery.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

1.1 PURPOSE OF STUDY

The purpose of this study is to carry out a computational docking study of new PKD1 inhibitors on the 3D structural homology model of the kinase domain (residues 587 to 835) of human PKD1 in order to explore the interactions of PKD1 inhibitors in assisting future structure-based design.

1.2 PROTEIN KINASE AND SMALL MOLECULE KINASE INHIBITOR

1.2.1 Protein kinases

Protein kinases are kinase enzymes that modify other proteins by chemically adding phosphate groups to them (phosphorylation) which usually results in a functional change of the target protein (substrate) and the subsequent alternation of cellular processes including proliferation or apoptosis[1,2]. Protein kinases can be classified based on substrate specificity or amino acid sequences of the catalytic domains. For the first one, kinases are classified by the amino acids they phosphorylate. The two main classes of kinases are tyrosine kinases (TKs), which phosphorylate tyrosine, and serine-threonine kinases, which phosphorylate serine or

threonine. Tyrosine kinases, the first identified and can, in turn, be classified into receptor or non-receptor tyrosine kinases. The former are transmembrane proteins with a ligand-binding extracellular domain and a catalytic intracellular kinase domain; the later lack a transmembrane domain and are located in the cytosol, the nucleus, and the inner surface of the plasma membrane[3]. The second classification relies on amino acid sequence comparisons of the catalytic domains. The human protein kinase family can thus be divided into seven main groups: the AGC family containing the protein kinases A, G and C; the CAMK family containing Ca2+/CAM-dependent protein kinases; the CK1 family containing the casein kinase 1 group; the CMGC family containing CDK, MAPK, GSK3, and CLKs; the STE family containing homologues of yeast sterile 7, 11, 20 kinases; the TK family containing tyrosine kinases; and the TKL family containing tyrosine kinase-like PKs[4]. (Figure 1)

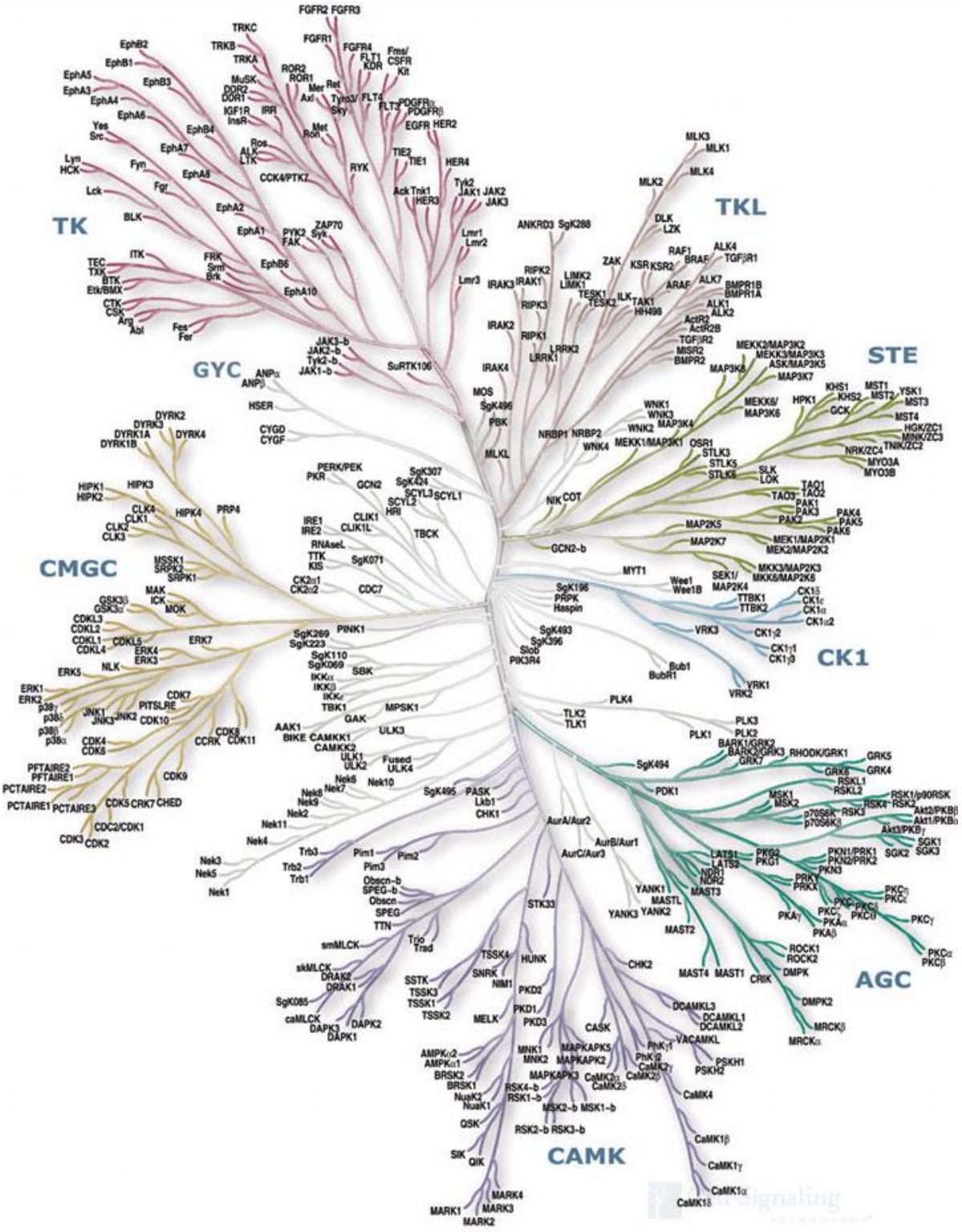

Figure 1. Overview of Human Protein Kinase Protein Family (www.cellsignal.com)

Protein kinases play key roles in intracellular signaling pathways which regulate cell growth, differentiation, development, functions, and death[5]. Mutations dysregulation of protein kinases, and aberrantly-regulated kinase activity are implicated in a wide range of human disorders, particularly several types of cancer, making protein kinases the second most important group of drug targets (after G-protein-coupled receptors[2,6-9]). In the early 1980s, Hiroyoshi Hidaka discovered that Naphthalene sulphonamides, such as N-(6-amino-hexyl)-5-chloro-1-naphthalenesulphonamide (W7), which act as antagonists of the calcium-binding protein calmodulin were also able to inhibit a number of different protein kinases at higher concentration. By replacing the naphthalene ring with isoquinoline, Hidaka identified a group of ATP-competitive protein kinase inhibitors[7]. Scientists have since been making efforts to understand the biology of kinases and their role in disease as well as discovering new kinase inhibitors. In fact, more than twenty drugs target on protein kinases and the receptors that activate them have been launched or are in development. Imatinib mesylate (marked by Novartis as Gleevec)[10], for instance, is a protein-tyrosine kinase inhibitor that inhibits the Bcr-Abl kinase through blocking ATP binding. Imatinib has been approved for the treatment of multiple cancers including Bcr-Abl-positive chronic myeloid leukaemia (CML) and Kit (CD117)-positive gastrointestinal stromal tumor.

1.2.2 The binding sites of small molecule kinase inhibitors

Most of the small molecule kinase inhibitors are ATP-competitive. The binding sites of these kinase inhibitors share certain characteristics: conserved arrangements into 12 subdomains that fold into a bi-lobed catalytic core structure with ATP binding in a deep cleft between the lobes;

conserved activation loop marked by conserved DFG and APE motifs at the start and end of the loop respectively; small molecules form one to three hydrogen bonds to amino acids located in the hinge region, thereby mimicking the hydrogen bonds that are normally formed by the adenine ring of ATP[11]. The majority of ATP-competitive inhibitors belong to the type I inhibitors which recognize the so-called active conformation of the kinase. (Figure 2a) These types of kinase inhibitors bind in and around the region occupied by the adenine ring of ATP (known as the adenine region) and do not require the DFG motif in the activation loop to adopt a 'DFG-out' conformation for binding. Type I inhibitors mimic the exocyclic group of adenine and typically form approximately one to three hydrogen bonds with the kinase hinge residues that link the N- and C-terminal kinase domains[12]. In contrast, type II kinase inhibitors, such as Imatinib recognize the inactive conformation of the kinase. (Figure 2b) In this conformation, the DFG motif is rearranged and the movement of the activation loop exposes an additional hydrophobic biding site near the ATP binding site.

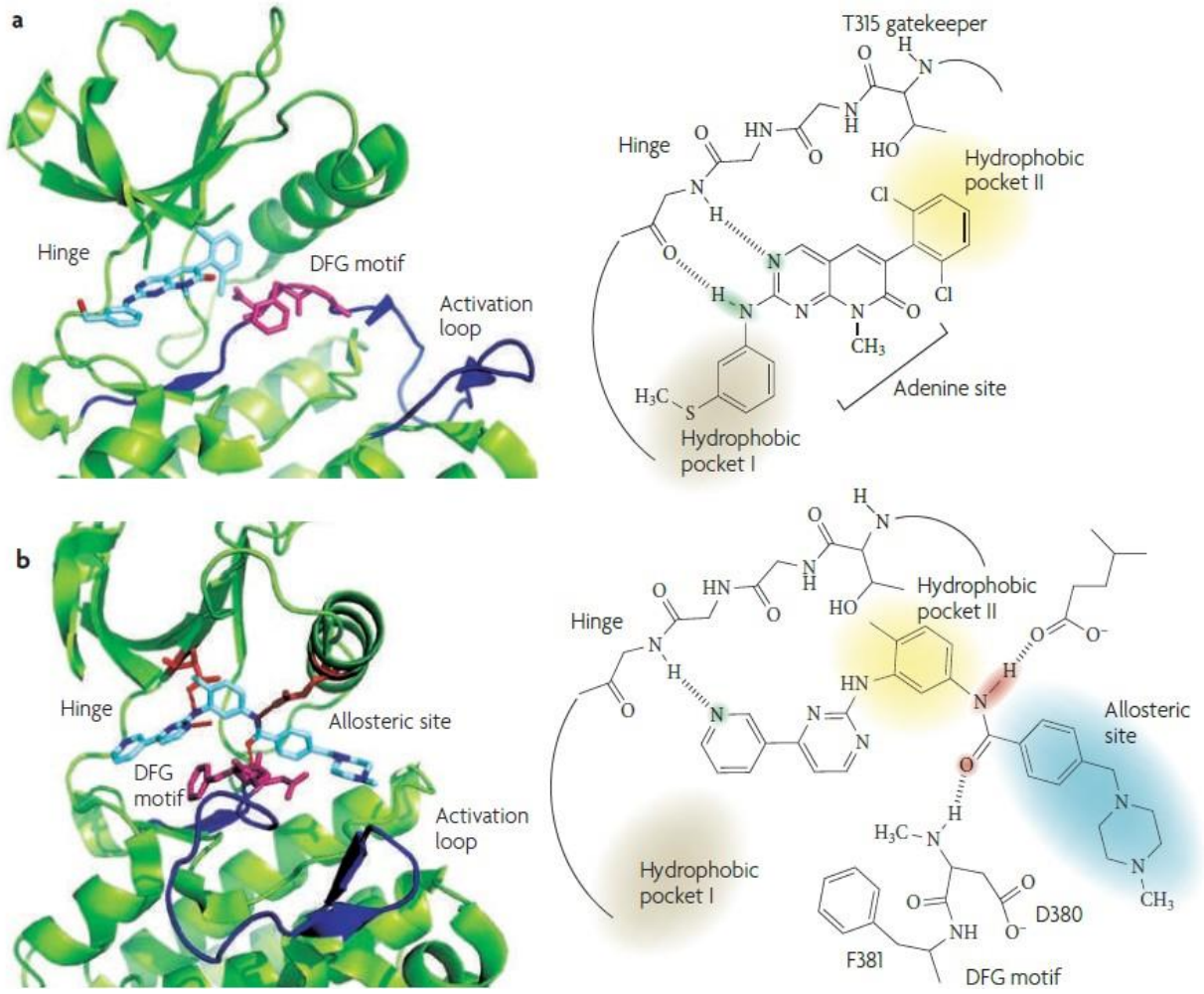

Figure 2: Protein kinase inhibitor binding mode. a. Protein kinase ABL1 in complex with the type 1 ATP-competitive inhibitor PD166326 (PDB ID 1OPK)[11]. b. The DFG-out conformation of the activation loop of ABL1 (dark blue) with the type 2 inhibitor imatinib (PDB ID 1IEP)[13]

1.3 PROTEIN KINASE D1 (PKD1)

Protein kinase Ds (PKDs) constitutes a novel family of diacylglycerol (DAG)-regulated serine/threonine protein kinases that signal downstream of G protein-coupled receptors (GPCRs). Tyrosine kinase receptors belong to a distinct subgroup of the calcium/calmodulin-dependent protein kinase (CAMK) family. Previous studies show that PKDs are activated by DAG responsive PKCs through phosphorylation of S744 and S748 in the activation loop PH domain. This DAG/ PKC/PKD axis has been demonstrated to be a crucial signaling pathway for the regulation of a variety of essential biological events[14].

A previous study showed that all three of the PKD isoforms (PKD1, PKD2 and PKD3) play key roles in cellular processes pertaining to multiple diseases especially in tumor growth, metastasis, and angiogenesis[14,15]. PKDs have been shown to be linked to several major signaling pathways vital to cancer development, most notably the VEGF and MEK/ERK signaling pathway. This gives support to the tumor-associated role of PKD in diverse cancer types including breast cancer, pancreatic cancer, prostate cancer, basal cell, skin, gastric, lung and lymph[15]. Moreover, PKDs target the class IIa histone deacetylases (HDAC 4, 5, 7, 9) which in turn activate the hypertrophic response of the heart. Gain- and loss-of-function approaches were used in assessing the role of PKD1 in heart muscles which indicate an enhanced PKD1 catalytic activity in pathological cardiac hypertrophy animal models[16]. Mice lacking cardiac PKD1 exhibit a decreased response to stress signals that normally lead to cardiac hypertrophy, fibrosis, and fetal gene activation, indicating a critical role of PKD in this pathological process[17]. As such, PKD has emerged as a potential therapeutic target for cancer, cardiac hypertrophy, and other diseases[18].

2.0 MATERIAL AND METHOD

2.1 HOMOLOGY MODEL GENERATION

Due to the lack of a PKD1 crystal structure, we generated the 3D homology models for PKD1 by using the I-TASSER server. The I-TASSER server is an internet service for protein structure and function predictions; I-TASSER was ranked as the number 1 server in the recent Critical Assessment of Techniques for Protein Structure Prediction (CASP9, 2010) competition for homology modeling and threading[19,20]. Academic users input amino acid sequences which the program translates into a 3D structure with high-quality predictions of biological function of protein molecules. Using this process, several template proteins of similar folds were retrieved by the server from the PDB library by LOMETS, a locally installed meta-threading approach. These templates go further into fragmentation and reassemble into full-length models, with the threading unaligned regions built by *ab initio* modeling and identified for lower energy states. A second simulation was done to cluster the decoys and select the lowest energy models. (Figure 3)

The PKD1 kinase domain sequence, which started from Glu587 to Ser835 was submitted to the I-TASSER 3D structure prediction server, producing five similar models for PKD1 kinase domain. Protein structures 1ql6_A (rabbit, phosphorylase kinase), 2bdw_A (caenorhabditis elegans, calcium/calmodulin activated kinase II), 3mfr_A (human, calcium/calmodulin (CaM)-activated serine-threonine kinase), 2jam_B (human, calcium/calmodulin-dependent protein

kinase type 1G), and 2y7j_A (human, phosphorylase kinase, Gamma 2) were chosen by I-TASSER as the templates in the modeling[19].

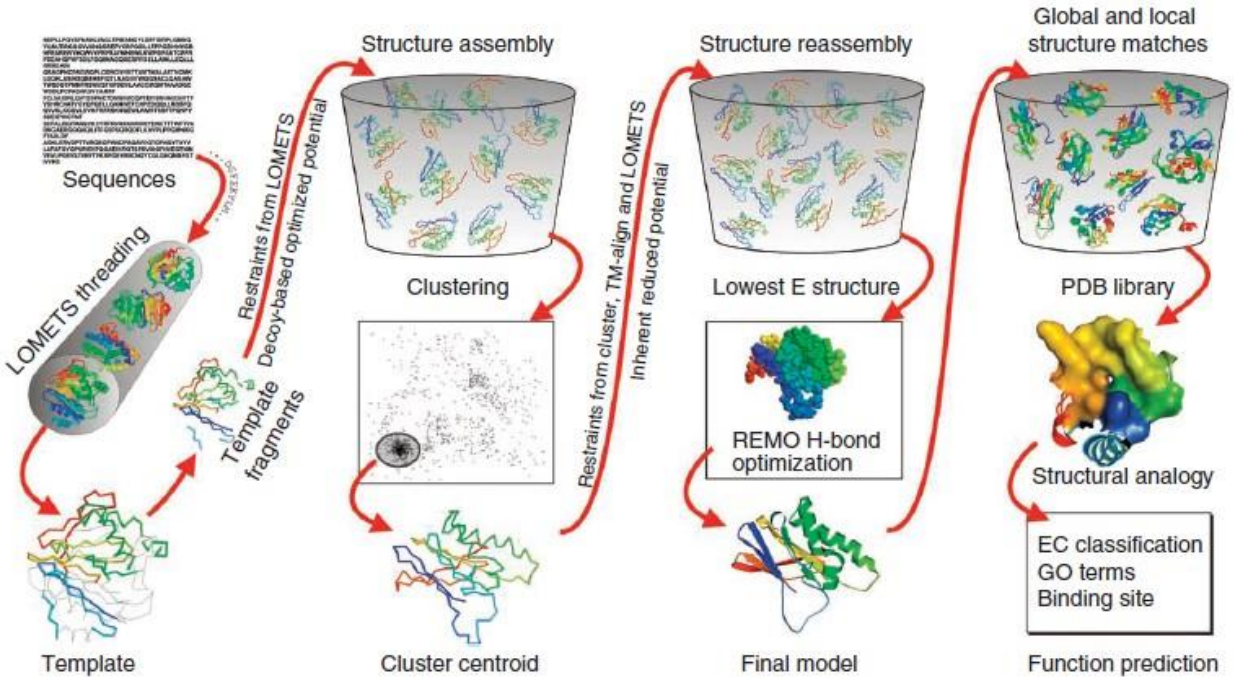

Figure 3. Protein 3D structure models generation by I-TASSER server. The protein chains are colored from blue at the N-terminus to red at the C-terminus[19].

2.2 PARAMETER SETTINGS OF DOCKING PROGRAM

All the docking calculations were performed on Sybyl x1.3 by using the Surflex-Dock Method followed by our established protocols[21,22]. Surflex-Dock uses an empirical scoring function from the Hammerhead docking system and a patented search engine to dock ligands into a protein's binding site. It is particularly successful at eliminating false positive results and can, therefore, be used to narrow down the screening pool significantly, while still retaining a large number of active compounds.

2.2.1 The Surflex-Dock Protomol

The Surflex-Dock Protomol is referred to as the "binding pocket", a computational representation of the proposed binding site in which target small molecules are aligned. Generally, the protomol was constructed by the following processes. First, the selected protein surface is coated with certain types of probes representing potential hydrogen bonds and favorable hydrophobic interactions with protein atoms. The probes are positioned and oriented by a score function representing the binding contribution of a similar atom on a ligand. Furthermore, the probes are filtered by score resulting in a cluster of high-scoring probes which identify the "sticky parts" of the protein's surface. Disconnected sticky spots are discarded with the rest form spheres on a 1 Å cubical grid. Each spheres grows until it reaches the van der Waals surface of a protein atom; sphere with a radius less than 0.5 Å are discarded. Finally, the sticky spots are merged into a pocket by accretion on the set of remaining protein-free spheres[23].

In our study, the protein kinase domain models were first modified by adding all hydrogens. The docking area was defined by a pocket set to cover the ATP binding domain of the protein. Protomol construction was based both on protein residues Ala610, Lys612, Met659, Glu660, Lys661, Leu662, His663, Glu710, Leu713, Cys726, and on parameters set to produce a small and buried docking target[24]. (Figure 4) Similar to the type 1 inhibitor, the alkylaminopyridine shown in the figure forms two hydrogen bonds with the Leu662 hinge residue; the naphthyridine 6-nitrogen contacts with the catalytic Lys612, and the piperazine nitrogen participates a salt bridge with a pendent Glu710 of the sugar pocket. Two parameters (protomol_bloat and protomol_threshold) determine the extent of the protomol and the docking performance depends on the binding site (protomol). We chose the parameter: protomol_threshold of 0.6 and protomol_bloat of 3.0. This was sufficient for generating a small

and adequate pocket for docking. (Figure 5) We treated all five models with the same protomol and docking was run with default settings for all other parameters.

2.2.2 The Surflex-Dock Docking Procedure

Surflex-Dock GeomX (SFXC) was chosen to be the docking mode in which twenty additional starting conformation per molecule were added[25]. In our study, we allowed for flexibility of protein atoms whose van der Waals distance from ligand atoms were less than 4 Å and adapted the active site conformation to the docked ligand. (Figure 6) PKD1 small molecule inhibitors were translated into SDF files and minimized to keep the lowest energy state. We entered both the protomol and the prepared 3D small molecule ligand into the Surflex-Dock program and started the calculation. The Surflex-Dock scoring function is a weighted sum of non-linear functions involving van der Waals surface distances between the appropriate pairs of exposed protein and ligand atoms. The list of atom pairs of interest is established by pruning out all protein-ligand atom pairs for which the distance between their van der Waals surfaces is greater than 2 Å. Each atom in the remaining protein-ligand pairs is labeled as being non-polar (e.g. H in CH3) or polar (e.g. H in N-H or O in C=O)[26].

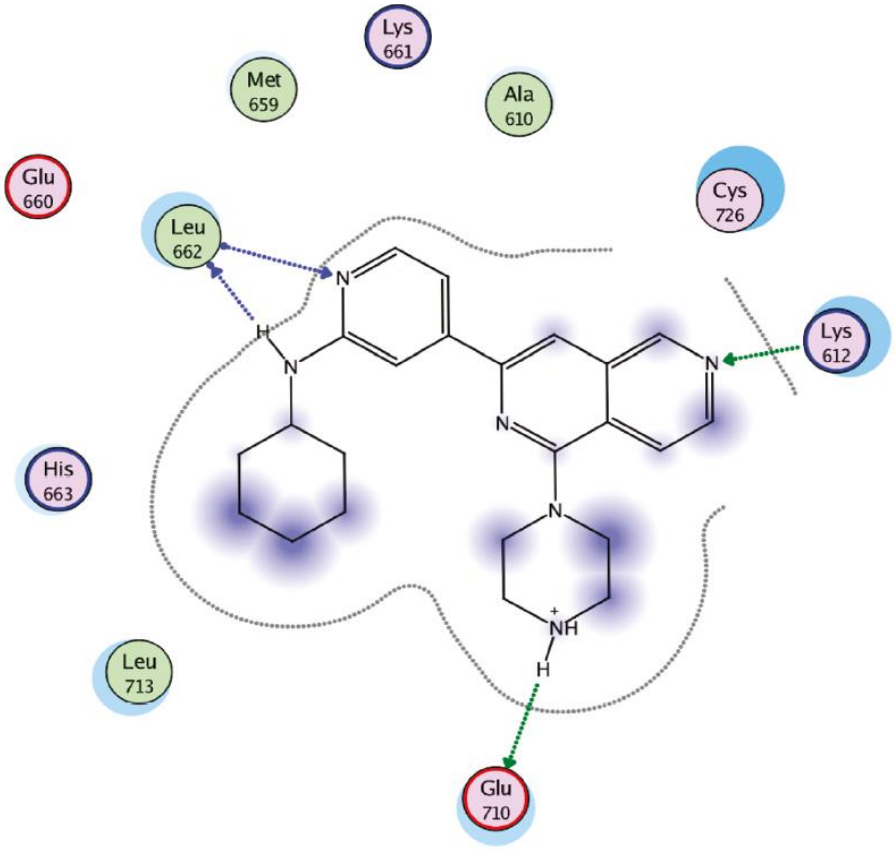

Figure 4: Proposed key contacts for PKD inhibitor 13c in PKD1[24].

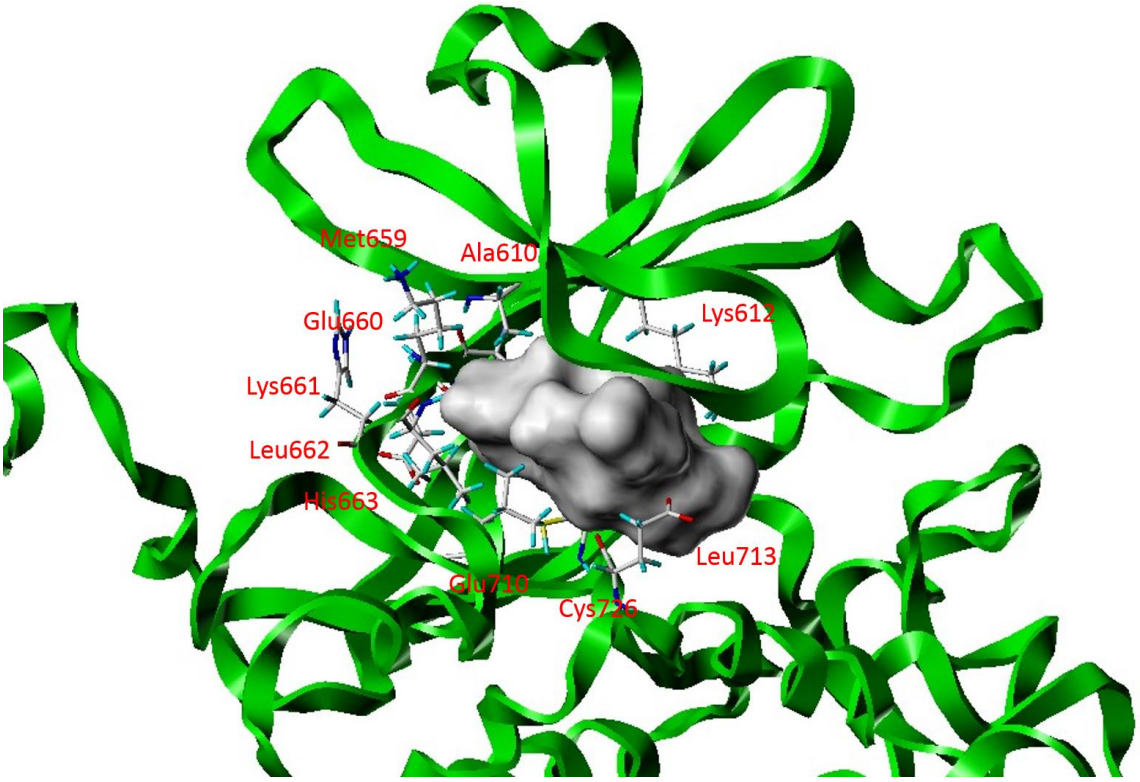

Figure 5. Surflex-Dock Protomol for PKD1 3D homology model 3.

2.3 STRUCTURAL COMPARISON BETWEEN OUR COMPOUNDS AND KNOWN PDK1 INHIBITORS

In order to evaluate the molecular similarity between our active compounds and known PKD1 inhibitors, both 2D and 3D similarity methodologies were employed to compare these 8 compounds against 18 previously reported PKD1 inhibitors as well as ATP[27].

2.3.1 2D molecular similarity study

UNITY 2D search was used to calculate the Tanimoto scores based on the computation and comparison between the fragment bitmaps of two structures. These fragment bitmaps are often referred to as "fingerprints," and have bits set according to the fragments found in the structure. The Tanimoto coefficient scores (Sim_T) can be calculated as:

$$Sim_T = C/(A+B-C)$$

where C is the count of bits set in both fingerprints, A is the count of bits set in fingerprint #1, and B is the count of bits set in fingerprint #2. The TS (Tanimoto score) range from 0.0 to 1.0, where a large TS score implies two chemicals similar in their 2D structures[28-31].

2.3.2 3D molecular similarity study

The Surflex-Sim 3D similarity program was used to calculate the morphological similarity score (MSS). Scores ranged from 0.0 to 10.0, with large values indicating two compounds with similar 3D shapes[32]. Surflex-Sim is a 3D molecular similarity optimization and searching program which applies a morphological similarity function and fast pose generation techniques to generate alignments of molecules. Surflex-Sim bases similarity on a molecule's shape, hydrogen bonding and electrostatic properties. The program considers molecular surfaces, not volumes, thus molecules of different sizes are handled easily by Surflex-Sim.

3.0 RESULTS

3.1 STRUCTURE MODELING OF PKD1 KINASE DOMAIN

Five models defined as models 1 through 5 were chosen based on the best C scores which were calculated to show that their 3D structures present similar topologies and overall shapes. (Figure 5)[19] Although the sequence identities between PKD1 and their templates are moderate (approximately 30% to 37%), their 3D structures present similar topologies and overall shapes. Specifically, conserved structure elements fold into a bi-lobed catalytic core structure with ATP binding in a deep cleft located between the lobes. These observations support our strategy to take advantage of the structural conservation in the PKD1 kinase domain to identify the key residues for inhibitor-protein interactions and then search for equivalent residues.

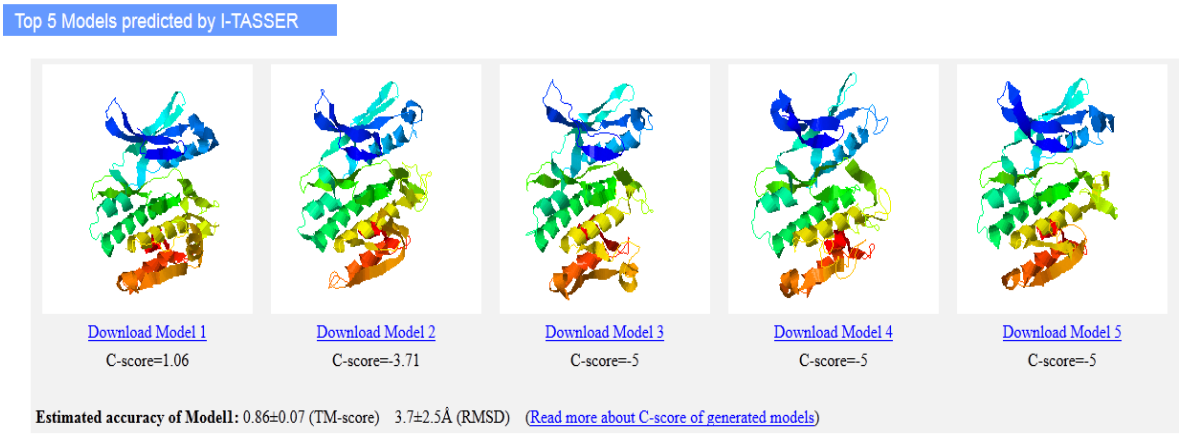

Figure 6. Top five models generated by I-TASSER. C-score was calculated to show that their 3D structures present similar topologies and overall shapes[18].

A total of 28 bioactive PKD1 inhibitors selected from 235 compounds were docked into the PKD kinase domain binding sites for all five models. Surflex-Dock scores are expressed in -log10(Kd) units to represent binding affinities. Here we show the resulting model 3 and 4, which have the highest docking scores. (Table 2) The docking scores of the bioactive molecules together with their inhibition are shown in Table 1. All docking scores ranged from 4 to 9, which can be correlated to Kd values of 100 mM to 1 nM, respectively[33]. For docking programs and scoring functions, a number of false positives would usually appear in the top ranking list. It is necessary to manually check and analyze the binding mode of each compound to determine if it has reasonable interaction and geometry fitting. Those molecules with binding modes exhibiting large differences with the corresponding existing compounds were excluded from the list after this inspection. Taking the binding score, inhibition, and binding mode into consideration, a representative hit, RO3202312-001 (renamed as compound 139) was found to dock well with the PKD kinase domain binding site in models 3 and 4. The structure of this validated hit was identified and its interaction with PKD kinase domain binding site is shown in Figure 6. Note that, the NH_2 and OH group form hydrogen-bonding interactions with the back bone of Leu662 and Gly664, located in the hinge region. The fluorine atom on the benzene ring and the nitrogen atom of the pyridine were observed to interact with the charged NH3+ of Lys612. These interactions are in conformity with the reported binding mode of experimentally tested previous PKD1 inhibitors[24]. The hydrogen-bonding interactions between the hydroxyl group of the hit and Leu662 as well as Glu660 are acknowledged as common to all of the previously known PKD1 inhibitors. However, we did not observe another common hydrogen-bonding interaction between our hit molecule and Clu710. Instead, an unoccupied pocket next to the indole nitrogen atom suggests the substitution at this position may be well tolerated.

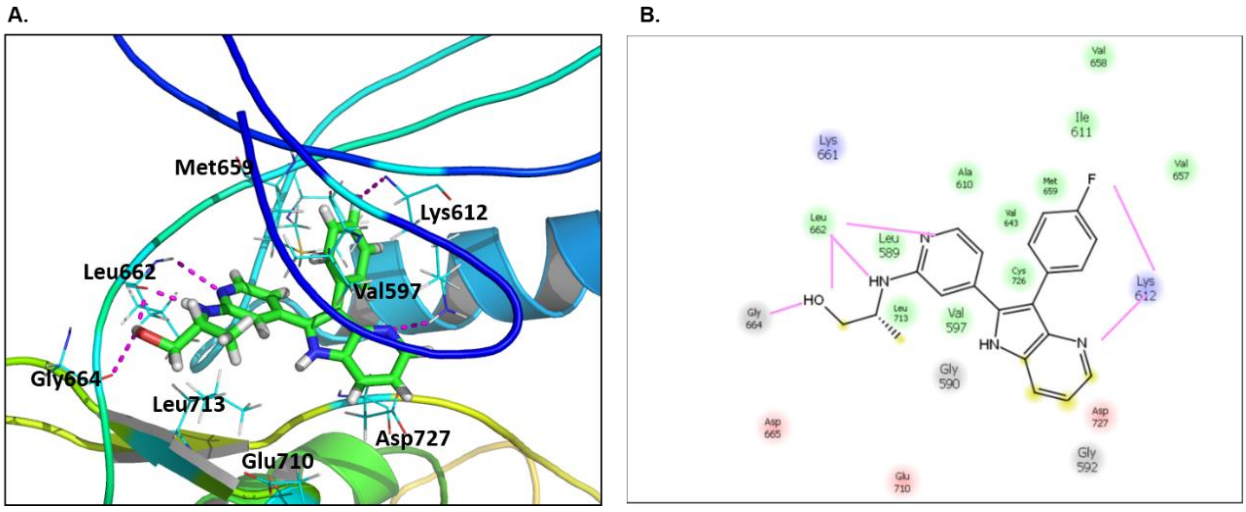

Figure 7. Molecular modeling of compound 139 in the active site of a PKD1 homology model. A. The docking result of the bioactive compound 139 in the ATP binding site of the PKD1 kinase domain. Carton ribbon and thick line, PKD1; ball and stick, Compound 139; thin line, residues in the binding pocket; magenta line, hydrogen bond. B. The proposed key contacts in the active site. Purple line, hydrogen bond; residues in different colors: purple, basic; pink, acidic; green, hydrophobic; gray, hydrophilic[18].

Compound ID	Inhibition	Model3			Model4		
		Total score	Crash	Polar	Total score	Crash	Polar
ro0272159-000	55%	4.8858	-0.6717	0.5727	5.9596	-0.9128	3.0941
ro0281601-001	94%	6.9009	-1.3660	1.9981	7.2726	-1.2405	3.8065
ro0282155-000	80%	8.1487	-1.2844	2.1006	7.2622	-1.9783	2.7245
ro0282986-001	97%	7.8100	-1.3037	1.7776	6.7615	-0.8366	1.4902
ro0283049-001	95%	7.8903	-1.9141	5.5857	7.6094	-0.8638	3.0703
ro0283120-000	97%	7.5244	-2.7031	3.3006	8.0268	-3.4905	4.9235
ro0317253-000	54%	5.5374	-0.2949	1.2629	5.5452	-1.9602	2.2220
ro0317340-000	65%	6.5484	-1.4297	0.5104	5.7383	-3.7040	2.1839

ro0317377-000	65%	6.8866	-1.5538	2.7313	7.6009	-0.9385	1.7922
ro0480500-002	77%	8.3315	-3.9738	2.0981	8.1414	-3.5176	4.7950
ro0504833-000	86%	6.3595	-0.6336	3.4251	6.7998	-0.6228	3.2615
ro0504985-000	93%	7.6159	-1.7532	5.4655	6.7209	-0.6299	5.9762
ro0506220-000	65%	6.5284	-2.7029	2.4351	7.8633	-1.0306	2.4698
ro1153853-000	86%	5.2845	-0.9685	1.9311	4.9812	-0.9037	1.0894
ro1155240-000	55%	5.3839	-0.4834	0.2696	4.7941	-0.2906	0.0153
ro1155697-000	62%	5.0076	-0.8801	0.8249	5.3993	-1.2469	2.3886
ro1155798-000	58%	7.3376	-0.8577	2.2108	6.5969	-2.7463	3.1254
ro3202312-001	94%	9.0223	-0.8286	2.2653	8.3194	-1.4282	3.7485
ro3206145-001	80%	7.7077	-1.7656	4.1853	8.4016	-1.6166	2.0628
ro4241967-000	62%	3.5125	-0.3827	0.9004	6.0081	-0.9309	2.7081
ro4367842-001	78%	7.5907	-1.1975	2.3250	6.7506	-1.3341	2.2950
ro4442080-000	55%	5.9590	-1.3500	1.0243	6.2918	1.0300	1.5459
ro4503319-000	49%	6.8369	-1.3919	0.0914	7.1934	-1.5664	3.7593
ro4509200-000	62%	7.2031	-1.6132	0.2601	7.5887	-1.1101	2.4558
ro4554339-000	85%	6.1197	-1.3313	0.1721	5.7010	-1.9621	1.7506
ro4569139-000	82%	5.4102	-1.0296	2.5235	4.8094	-0.9669	1.9553
ro4595949-000	64%	4.9175	-0.7034	1.7907	4.7002	-1.1703	2.2313

Table 1. Docking result of the novel PKD1 inhibitors in model 3 and model 4. Total score indicates the binding score between the small molecule ligand and the PKD model. Compound ro3202312-001 is the same as hit compound 139 (renamed afterwards).

3.2 STRUCTURAL DIVERSITY ANALYSIS

A structural diversity analysis was done between eight PKD1 selective compounds and 18 known PKD1 inhibitors as well as ATP. Here, we show the results for the three most promising compounds: 122, 139, and 140. The 3D MSS ranged from 3.29 to 7.10 and the 2D Tanimoto scores ranged from 0.174 to 0.544. (Table 2) Indicating that the three novel compounds increased the structural diversity of PKD1 inhibitors.

Compound ID	122		139		140	
	UPCMLDRO1155697-000		UPCMLDRO3202312-001		UPCMLDRO3206145-001	
	MSS	TS	MSS	TS	MSS	TS
CID_2876479	5.41	0.196	6.53	0.245	6.79	0.238
CID_755673	5.57	0.201	6.96	0.223	7.08	0.228
CID_646236	5.58	0.287	6.14	0.297	6.35	0.302
CID_5086667	5.92	0.231	6.78	0.276	6.58	0.277
CID_2958734	5.90	0.267	4.94	0.340	5.17	0.349
CID_16230	4.97	0.199	6.45	0.239	6.34	0.236
CID_4438738	4.98	0.232	7.04	0.219	7.05	0.217
CID_663844	4.98	0.319	4.98	0.359	5.32	0.371
CID_5389142	4.60	0.354	6.80	0.434	7.10	0.456
CID_9549170	4.70	0.304	5.39	0.341	5.63	0.349
CID_2011756	3.46	0.219	4.23	0.289	4.38	0.291
CID_1893668	3.29	0.174	4.39	0.234	4.61	0.235
ATP	5.85	0.265	5.47	0.290	5.76	0.305
kb-NB142-70	4.83	0.177	5.46	0.207	5.46	0.212
BPKDi	5.77	0.362	5.67	0.513	5.49	0.523

13c	5.61	0.379	5.97	0.544	5.51	0.530
24c	4.91	0.273	5.94	0.334	5.91	0.340
CRT5	5.63	0.316	5.32	0.404	5.73	0.410
12a	4.90	0.362	5.19	0.513	4.91	0.523

Table 2: Structural diversity analysis results. 2D and 3D similarity search between the three most promising new PKD1 inhibitors and 18 known PKD1 inhibitors as well as ATP.

4.0 DISCUSSION

In this study, a 3D PKD1 protein structure was generated through homology modeling and verified by newly synthesized small molecule PKD1 inhibitors in the subsequent docking studies. All ligands were docked into the assumed ATP binding pocket of the kinase domain, resulting in relatively high docking scores. Among them, our lead compound 139 formed a reasonable binding mode not different with the previous reports. Furthermore, structural diversity analysis demonstrated an increase in diversity of three representative compounds.

We also performed the Topomer similarity study between new PKD1 inhibitors and the PDB ligand library. Possible off-target effects of new PKD1 inhibitors were identified as a result. The identified compounds were similar to an inhibitor of CDK2 (FMD), which was co-crystallized with CDK2 in PDB 1R78[34]. Compound 16 was docked into the kinase domain of CDK2 with a docking score of 10.92 and a nearly overlapped binding mode with FMD. (Figure 7) Additionally, our lead compound 139 was found to be similar to the p38 kinase inhibitor (FPH) which was co-crystallized with P38 kinase in PDB 1OZ1[35]. Docking studies give a docking score as high as 13.3 for compound 139 with p38 kinase. These possible off-target effects give further direction to the study of new PKD1 inhibitors to discover their role in other protein kinases and the PKD signaling transduction pathway. Taken together, the computational tools proved to be useful in many aspects of small molecule discovery processes including exploring the possible

mechanism of a bioactive compound, guiding future optimization, or finding possible off-targets of small molecules.

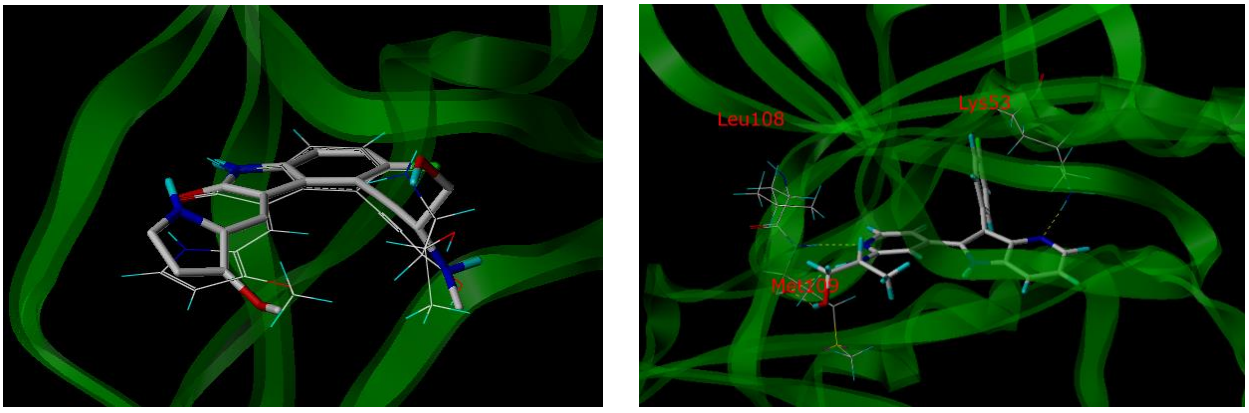

Figure 8. Possible off-target effect of PKD1 inhibitors. In the left figure compound 16 was docked into CDK2 PDBID: 1R78 (CDK2 co-crystallized with FMD). Ball and sticks represent compound16; lines represent FMD. In the right figure, compound 139 was docked into P38 PDB: 1OZ1 (P38 co-crystallized with FPH, FPH is not shown here)

5.0 FUTURE SPECULATION

One of the biggest obstacles in developing protein kinase inhibitors is their PKD1 selectivity in that the ATP binding site of protein kinases share similar arrangements and shapes. The possible off-target effects of new PKD1 inhibitors indicates that these small molecules may require structural optimization in order to increase their selectivity. In comparison with the proposed key contact or binding residues of PKD1 in the reported study, we found that our screened hit compounds lack the important interaction with Glu 710, while retaining an unoccupied pocket between the pyrrole ring and Glu 710. Hence, we hypothesize that introducing a piperazine ring on the pyrrole ring may increase the binding affinity and enhance the selectivity of the PKD1 inhibitors. This could be an interesting direction for future search.

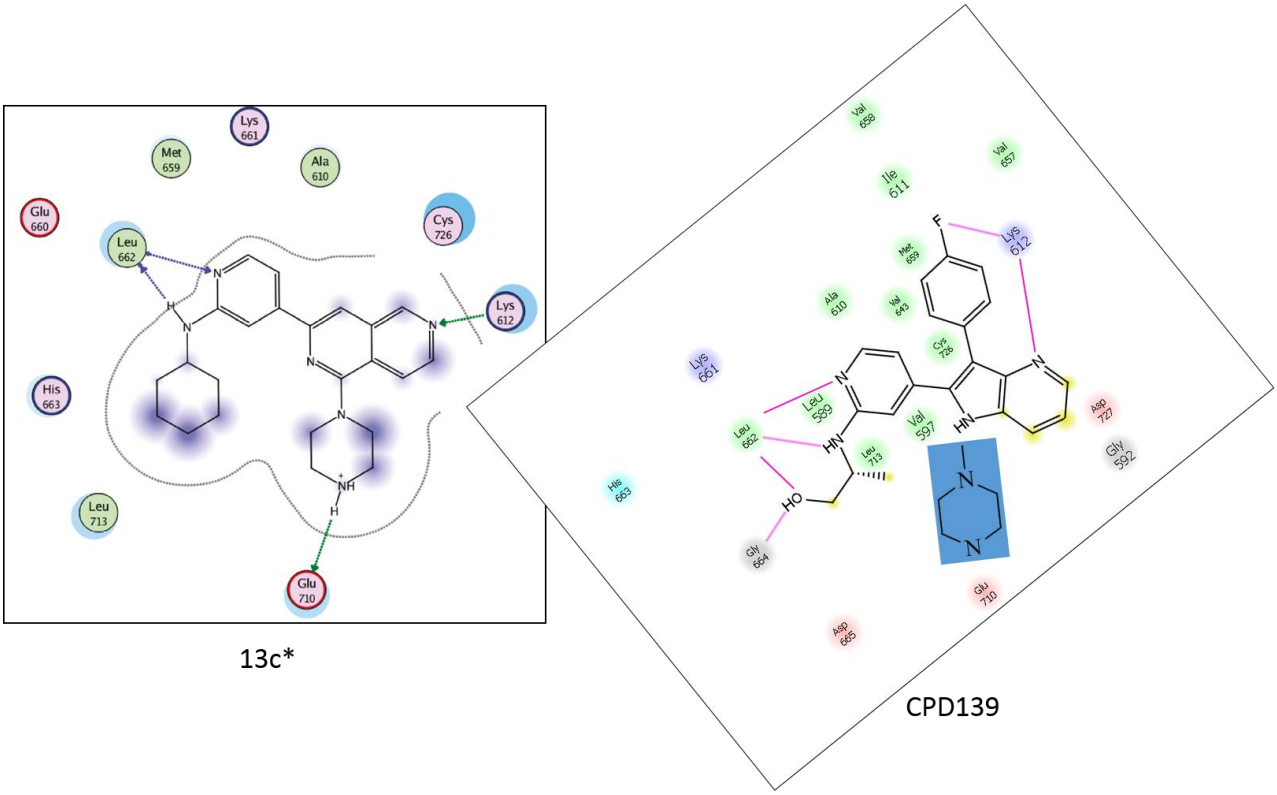

Figure 9. Comparison between compound 139 and the reported PKD1 inhibitor 13c. Proposed structural modification was shown in the blue rectangle.

APPENDIX A

PROTOCOL

A.1 HOMOLOGY MODELING

1. Open I-TASSER online server.
2. Write the PKD1 sequence in FASTA format. Load the sequence onto the server.
3. The models will be sent by e-mail the following day.

A.2 DOCKING STUDY BY SURFLEX-DOCK.

A.2.1 Initiating Surflex-Dock

1. Open Sybyl X 1.3.
2. File -> Import File: select the correct file type and choose the PKD1 model3.
3. Edit -> Hydrogens -> Add All Hydrogens
4. View -> Surface and Ribbons -> Quick Ribbons -> Ribbon. This step is not necessary for Docking study, but will give a better vision on both the protein and the kinase domain.

5. Application -> Docking Suites -> Dock ligands
6. Docking parameters setting: Choose the Docking Mode as Surflex-Dock GeomX (SFXC). Different docking mode gives different criteria. In “Options”, click Surflex-Dock. Allow both Hydrogen and Heavy Atoms movement. Change the “Additional Starting Conformations per Molecule” into 20. Click “OK” to save the modification.
7. On the right side of “Filename”, click “Define”.
8. Binding pocket generation: Select the “Protein Structure” as “Mol Area” and then choose “M1: model3”. Define the “Mode” as “Residues”, and then choose the residues as shown in section 2.2.1. Set the “Threshold” as 0.6 and “Bloat” as 3. Click “Generate” to generate a new model. Change the file name and click “OK” to return to the Surflex-Dock main menu.
9. In “Ligand Source”, select the corresponding file type and choose the file which contains 28 bioactive PKD1 inhibitors.
10. Change the Jobname and click OK to start the Docking process which takes approximately one day’s time.

A.2.2 Analyzing Docking Result

1. Application -> Docking Suite -> Analyze Results
2. Click Jobname to start the analysis, all the 28 molecules with their total binding score were shown.
3. Left click on molecule “RO3202312-001”, the interaction between the molecule and the protein was shown.

4. Close the result analyzing panel. Left click to select the residues which form hydrogen bond with the molecule. Then, right click and select atom display -> label -> substructure. The residues will be labeled.

A.3 STRUCTURAL DIVERSITY ANALYSIS

A.3.1 UNITY 2D search

1. File -> Import File. Open the file contains 18 know PKD1 inhibitors and ATP.
2. Select and click File -> Put Rows into UNITY database. Close the dialog.
3. UNITY -> UNITY tools -> translate molecular files. Translate the 8 new PKD1 inhibitors into SLN file.
4. UNITY -> Start UNITY Search. Select “Other Database” in the “Open Database” section. Click “Open” to load the new database. Open one of the translated SLN files, copy the content and paste under “Search Query in SLN”. Select “2D Similarity” in the “Query Type”. Make a Job name and click OK to run the search.
5. UNITY -> Search Management: the searching status is shown in the panel. Click “Load into table” to see the results.
6. Perform same process to complete UNITY 2D similarity search for all 8 new inhibitors.

A.3.2 Surflex-Sim 3D similarity search

1. Applications -> Similarity Suite -> Surflex-Sim.

2. Select “Surflex-Sim Flexible Superposition” as Similarity Mode.
3. Translate 8 new inhibitors into Mol2 Files. Use one as Template.
4. Put all 18 known inhibitors and ATP in a SDF file. Load this file on “Ligand Source” and select “SD file”.
5. Make a Jobname and click OK to run the search.
6. Applications -> Similarity Suite -> Analyze Results. Find the search in Jobname and see the results in the table. Select the new inhibitor in “View” and the known inhibitors or ATP in the table to compare their 3D similarity.
7. Perform the same research for all 8 new inhibitors.

BIBLIOGRAPHY

1 Jeffrey R. Simard & Mattha¨us Getlik, e. a. Fluorophore Labeling of the Glycine-Rich Loop as a Method of Identifying Inhibitors That Bind to Active and Inactive Kinase Conformations. *J. AM. CHEM. SOC.* **132**, 4152-4160 (2010).

2 Kirkland, L. O. & McInnes, C. Non-ATP competitive protein kinase inhibitors as anti-tumor therapeutics. *Biochemical pharmacology* **77**, 1561-1571, doi:10.1016/j.bcp.2008.12.022 (2009).

3 Daniela S. Krause & Etten., R. A. V. Tyrosine Kinases as Targets for Cancer Therapy. *N Engl J Med* **353**, 172-187 (2005).

4 Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912-1934, doi:10.1126/science.1075762 (2002).

5 I. SHCHEMELININ, L. ŠEFC & NEČAS, E. Protein Kinases, Their Function and Implication in Cancer and Other Diseases. *Folia Biologica (Praha)* **52**, 81-101 (2006).

6 Vlahovic, G. e. a. Activation of tyrosine kinases in cancer. *The Oncologist* **8**, 531–538 (2003).

7 Cohen, P. Protein kinases, the major drug targets of the 21st century? . *Nature Rev. Drug Discov* **1**, 309-316 (2002).

8 Jeffrey R. Simard & Christian Gru¨tter, e. a. High-Throughput Screening To Identify Inhibitors Which Stabilize Inactive Kinase Conformations in p38a. *J. AM. CHEM. SOC.* **131**, 18478-18488 (2009).

9 L. Garuti, M. Roberti & Bottegoni., G. Non-ATP Competitive Protein Kinase Inhibitors. *Current Medicinal Chemistry* **17**, 2804-2821 (2010).

10 AC Backes, B Zech, B Felber, B Klebl & Mu¨ller., G. Small-molecule inhibitors binding to protein kinases. Part I: exceptions from the traditional pharmacophore approach of type I inhibition. *Expert Opinion on Drug Discovery* **3**, 1409-1425 (2008).

11 Zhang, J., Yang, P. L. & Gray, N. S. Targeting cancer with small molecule kinase inhibitors. *Nature reviews. Cancer* **9**, 28-39, doi:10.1038/nrc2559 (2009).

12 Liu, Y. & Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nature chemical biology* **2**, 358-364, doi:10.1038/nchembio799 (2006).

13 Nagar, B. e. a. Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). *Cancer Res* **62**, 4236-4243 (2002).

14 Rozengurt, E. Protein kinase D signaling: multiple biological functions in health and disease. *Physiology* **26**, 23-33, doi:10.1152/physiol.00037.2010 (2011).

15 LaValle, C. R. *et al.* Protein kinase D as a potential new target for cancer therapy. *Biochimica et biophysica acta* **1806**, 183-192, doi:10.1016/j.bbcan.2010.05.003 (2010).

16 Harrison, B. C. *et al.* Regulation of cardiac stress signaling by protein kinase d1. *Molecular and cellular biology* **26**, 3875-3888, doi:10.1128/MCB.26.10.3875-3888.2006 (2006).
17 Fielitz, J. *et al.* Requirement of protein kinase D1 for pathological cardiac remodeling. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 3059-3063, doi:10.1073/pnas.0712265105 (2008).
18 Manuj Tandon, Lirong Wang & Qi Xu, e. a. A Targeted Library Screen Reveals a New Inhibitor Scaffold for Protein Kinase D. *PLoS One* **7**, 1-14, doi:10.1371/journal.pone.0044653.g001 (2012).
19 Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC bioinformatics* **9**, 40 (2008).
20 Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature protocols* **5**, 725-738, doi:10.1038/nprot.2010.5 (2010).
21 Jian-Zhong Chen, Junmei Wang & Xie, X.-Q. GPCR Structure-Based Virtual Screening Approach for CB2 Antagonist Search. *J. Chem. Inf. Model.* **47**, 1626-1637 (2007).
22 Xiang-Qun Xie, Jian-Zhong Chen & Billings, E. M. 3D Structural Model of the G-Protein-Coupled Cannabinoid CB2 Receptor. *PROTEINS: Structure, Function, and Genetics* **53**, 307-319 (2003).
23 Ruppert, J., Welch, W. & Jain, A. N. Automatic identification and representation of protein binding sites for molecular docking. *Protein Sci* **6**, 524-533 (1997).
24 Meredith, E. L. *et al.* Identification of orally available naphthyridine protein kinase D inhibitors. *Journal of medicinal chemistry* **53**, 5400-5421 (2010).
25 Jain, A. N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *Journal of medicinal chemistry* **46**, 499-511 (2003).
26 Jain, A. N. Scoring noncovalent protein-ligand interactions: A continuous differentiable function tuned to compute binding affinities. *J. Comput. Aided-Mol. Des.* **10**, 427-440 (1996).
27 Sharlow, E. R. *et al.* Discovery of diverse small molecule chemotypes with cell-based PKD1 inhibitory activity. *PLoS One* **6**, e25134, doi:10.1371/journal.pone.0025134 (2011).
28 P. Willett & V. Winterman. A Comparison of Some Measures for the Determination of Inter-Molecular Structural Similarity Measures of Inter-Molecular Structural Similarity. *QSAR* **5**, 18-25 (1986).
29 J.D. Holliday, S.S. Ranade & Willett, a. P. *QSAR* **14**, 501-506 (1996).
30 C. Cheng, G. Maggiora, M. Lajiness & Johnson, a. M. *J. Chem. Inf. Comput. Sci.* **36**, 909-915 (1996).
31 Ma, C., Wang, L. & Xie, X. Q. GPU accelerated chemical similarity calculation for compound library comparison. *Journal of chemical information and modeling* **51**, 1521-1527, doi:10.1021/ci1004948 (2011).
32 Yera, E. R., Cleves, A. E. & Jain, A. N. Chemical structural novelty: on-targets and off-targets. *Journal of medicinal chemistry* **54**, 6771-6785, doi:10.1021/jm200666a (2011).
33 Meurice, N. *et al.* Structural conservation in band 4.1, ezrin, radixin, moesin (FERM) domains as a guide to identify inhibitors of the proline-rich tyrosine kinase 2. *Journal of medicinal chemistry* **53**, 669-677, doi:10.1021/jm901247a (2010).
34 Luk, K. C. *et al.* A new series of potent oxindole inhibitors of CDK2. *Bioorganic & medicinal chemistry letters* **14**, 913-917, doi:10.1016/j.bmcl.2003.12.009 (2004).

35 Alejandra Trejo & Humberto Arzeno, e. a. Design and synthesis of 4-azaindoles as inhibitors of p38 MAP kinase. *J. Med. Chem* **46**, 4702-4713 (2003).