Combined 3D-QSAR Modeling and Molecular Docking Study on Indolinone Derivatives as Inhibitors of 3-Phosphoinositide-Dependent Protein Kinase-1

Mohamed Diwan M. AbdulHameed, Adel Hamza, Junjun Liu, and Chang-Guo Zhan*

Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, 725 Rose Street, Lexington, Kentucky 40536

3-Phosphoinositide-dependent protein kinase-1 (PDK1) is a promising target for developing novel anticancer drugs. In order to understand the structure−activity correlation 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 study has resulted in two types of satisfactory 3D-QSAR models, including the CoMFA model ($r^2 = 0.907$; $q^2 = 0.737$) and CoMSIA model ($r^2 = 0.991$; $q^2 = 0.824$), for predicting the biological activity of new compounds. The detailed microscopic structures of PDK1 binding with inhibitors have been studied by molecular docking. We have also developed docking-based 3D-QSAR models (CoMFA with $q^2 = 0.729$; CoMSIA with $q^2 = 0.79$). The contour maps obtained from the 3D-QSAR models in combination with the docked binding structures help to better interpret the structure−activity relationship. All of the structural insights obtained from both the 3D-QSAR contour maps and molecular docking are consistent with the available experimental activity data. This is the first report on 3D-QSAR modeling of PDK1 inhibitors. The satisfactory results strongly suggest that the developed 3D-QSAR models and the obtained PDK1-inhibitor binding structures are reasonable for the prediction of the activity of new inhibitors and in future drug design.

1. INTRODUCTION

As an important member of the survival signaling pathway,[1] 3-phosphoinositide-dependent protein kinase-1 (PDK1) is a protein kinase belonging to the cAMP-dependent, cGMP-dependent, protein kinase C (AGC) kinase family.[1] PDK1 has an N-terminal kinase domain and a C-terminal pleckstrin homology (PH) domain.[2] The PH domain is essential for interaction with the cell membrane, and the kinase domain is involved in the phosphorylation and activation of down-stream kinases.[2] PDK1 has been found to be essential for the activation of several important kinases. Hence, PDK1 is known as the master regulator of AGC kinases.[2] In particular, PDK1 is known to phosphorylate and activate a number of downstream kinases like protein kinase B (PKB, also known as Akt), p70 ribosomal S6 kinase, serum- and glucocorticoid-induced protein kinase, and protein kinase C isoforms.[2] Accumulating pharmacologic and genetic evidence supports the potential role of PDK1 as a promising anticancer target.[3−6] It has been reported that overexpression of PDK1 in mammary cells resulted in their transformation in vitro and tumor formation in vivo.[4] Elevated levels of PDK1 phosphorylation were also reported in metastasized breast tumors.[7] Another study in which PDK1 is targeted with antisense oligonucleotides has reported a marked reduction of cell proliferation and survival and also an increased rate of apoptosis than that observed in phosphoinositide-3-kinase or PKB inhibition.[5] PDK1-expressing cells were shown to have enhanced invasion and cell growth.[8] PDK1 was also reported as a potential target for sensitizing breast cancer cells to chemotherapeutic agents.[3] Knock down of PDK1 was recently reported to enhance the antitumor effect of an epidermal growth factor receptor inhibitor.[9] This also underlines the role of PDK1 in the development of resistance to current cancer therapies. PDK1 hypomorphic mice expressing only 10% of the normal level of PDK1 were reported to be viable and fertile.[10] This finding has shown 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$^\pm$ mice markedly protected the animal from a wide range of tumors.[6] Thus, PDK1 has become a well-validated anticancer target.

Kinase inhibitors, such as Imatinib (Gleevec) and Erlotinib (Tarceva), have recently been approved for clinical use in the market as anticancer agents.[11] The success of these new drugs has given a new impetus for developing better anticancer agents that target particular pathways with a specific kinase overexpressed and critical to tumor progression.[12] Because PDK1 was found to be overexpressed in cancers[7] and has a crucial role in the survival pathway,[1] it is an ideal target for drug development. The development of PDK1 inhibitors could pave the way for the development of better anticancer therapeutics. Compared to other kinases, only a few inhibitors of PDK1 have been reported in the literature as potent nanomolar inhibitors,[13,14] including staurosporine, UCN-01,[13] and aminopyrimidine derivatives.[14] More recently, a new series of indolinone derivatives were reported as PDK1 inhibitors.[15,16] This new series of compounds was developed using an initial hit identified from a high-throughput-screening study.[15,16]

Computational approaches like structure- and ligand-based ones have been found to be valuable in further optimization

* To whom correspondence should be addressed. Tel: 859-323-3943. Fax: 859-323-3575. E-mail: zhan@uky.edu.

3D-QSAR and Docking for PDK1 Inhibitors

J. Chem. Inf. Model., Vol. 48, No. 9, 2008 **1761**

and the development of novel inhibitors. Ligand-based three-dimensional quantitative structure−activity relationship (3D-QSAR) approaches, including comparative molecular field analysis (CoMFA)[17] and comparative molecular similarity indices analysis (CoMSIA),[18] were reported to be effective for understanding the structure−activity relationships.[19] The 3D-QSAR modeling is useful to predict the activity of new molecules to be synthesized.[19] 3D-QSAR methods serve as an important complement to the structure-based methods.[19] CoMFA and CoMSIA are two 3D-QSAR methods that have been successfully employed in drug design.[20,21] These methods were useful in the lead optimization and also in understanding the drug−target interaction.[22−24] In CoMFA, the biological activity of molecules is correlated with their steric and electrostatic interaction energies.[25] The steric and electrostatic interaction energies are calculated using Lennard-Jones and Coulombic potentials, respectively.[17] In CoMSIA, similarity indices are calculated at regularly placed grid points for the aligned molecules.[21] CoMSIA includes additional molecular descriptors like hydrophobic fields and hydrogen-bond donor and acceptor fields.[21] 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.[24]

We previously carried out a structure-based computational study on celecoxib derivatives binding with PDK1.[26] In this present work, we report the first 3D-QSAR study on PDK1 inhibitors using CoMFA and CoMSIA methods. 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. The 3D-QSAR models obtained from both the ligand- 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 further help the lead optimization and future rational drug design and discovery efforts.

2. COMPUTATIONAL DETAILS

2.1. Data Sets and Alignment. All compounds used in the present study were reported recently by Islam and co-workers[15,16] as inhibitors of PDK1. The IC_{50} values of these inhibitors were reported to be measured using a PDK1-mediated AKT2 activation assay (cAKT2).[15] 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., $-\log IC_{50}$) values. 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 1.

Identification of the bioactive conformation and molecular alignment of compounds are two important steps in a 3D-QSAR study.[27] The X-ray crystal structures of this class of inhibitors bound with PDK1 are available from the protein

data bank (PDB).[28] The bound conformation of compound **53** (PDB-ID: 2PE2)[16] had a relatively better resolution (2.13 Å) and was used as a template for alignment molecular structures. The 3D structures of all compounds were built using *SYBYL* software (Tripos, Inc.). The geometries of all compounds were optimized by using the 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-fitted atomic charges, i.e., the ESP charges, that fit to the ESP at points selected according to the Merz−Singh−Kollman scheme.[29] Two alignment methods were used. One method, denoted by Alignment-I, is a substructure-based alignment. In this method, all compounds in the data set were aligned to a common substructure (Figure 1) using the "align database" command in *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-square (rms) fit method. The determined substructure and the reference atoms/centroids used for the alignment are shown in Figure 1.

2.2. 3D-QSAR Analysis. The CoMFA and CoMSIA models were generated by using *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, and 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 $+1.0$ was used to calculate the respective fields. The attenuation factor α was set to 0.3. CoMSIA similarity indices (A_F) for molecule j with atom i at grid point q are calculated by eq 1

$$A_{F,k}{}^q(j) = -\sum \omega_{\text{probe},k}\omega_{ik}e^{-\alpha r_{iq}{}^2} \qquad (1)$$

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

The partial-least-squares (PLS)[30] analysis was used to derive the 3D-QSAR models. The sample−distance PLS (SAMPLS) algorithm[31] was used for the leave-one-out (LOO) cross-validation. The optimum number of 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 five and two groups in the training set. Because only the number of groups and the number of validation times can be controlled in the *SYBYL* "cross-validation" process, the process was repeated 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 $r_{\text{pred}}{}^2$) for molecules in the test set was calculated by using eq 2

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

where SD is the sum of the squared deviations of the individual biological activity values for the test-set molecules

Table 1. Molecular Structures of Compounds Used in the Training and Test Sets and Their PDK1 Inhibitory Activity[a]
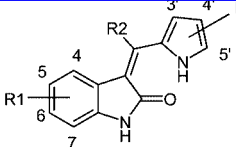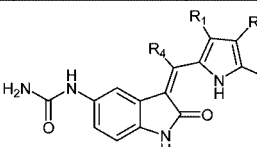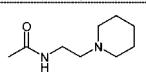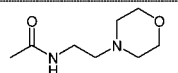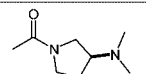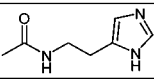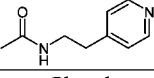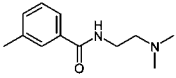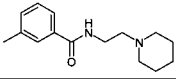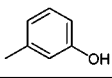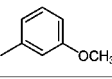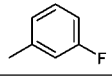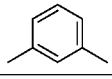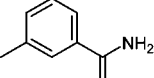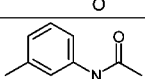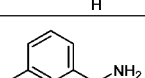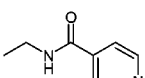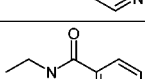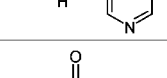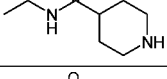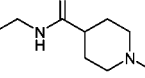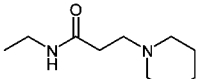

Compd.	R1	R2	R3	pIC50
1	H	H	H	5.74
2	4-Me	H	H	6.29
3[b]	5-Me	H	H	5.57
4	7-Me	H	H	4.58
5	4-OH	H	H	6.00
6[b]	5-OH	H	H	6.47
7	6-OH	H	H	6.05
8	7-OH	H	H	4.58
9	5-OH	Me	H	7.10
10[b]	5-OH	H	3'-Me	6.55
11	5-OH	H	4'-Me	6.55
12	5-OH	H	5'-Me	6.17
13	5-OH	H	5'-Et	5.64
14	5-OH	H	3',5'-Me	5.96
15	5-OMe	H	H	6.24
16	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
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

Table 1. Continued

Compd.	R1	R2	R3	R4	pIC50
47	H		H	H	7.85
48	H		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		H	H	8.40
53	H		H	H	8.40
54[b]	H		H	H	8.22
55	H		H	H	7.72
56[b]	H		H	H	7.29
57	H		H	H	7.43
58	H		H	H	8.30
59[b]	H		H	H	8.10
60	H		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		H	CH$_3$	8.30
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 15 and 16. [b] Compounds used in the test set were based on random selection.
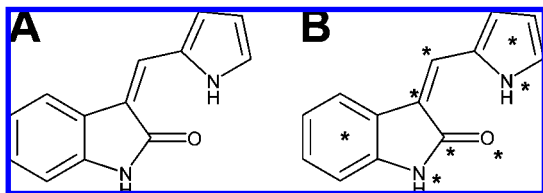
Figure 1. (A) Substructure used for the "common substructure-based alignment" (Alignment-I). (B) Atoms/centroids used in "atom/centroid-based alignment" as shown by asterisks (Alignment-II).
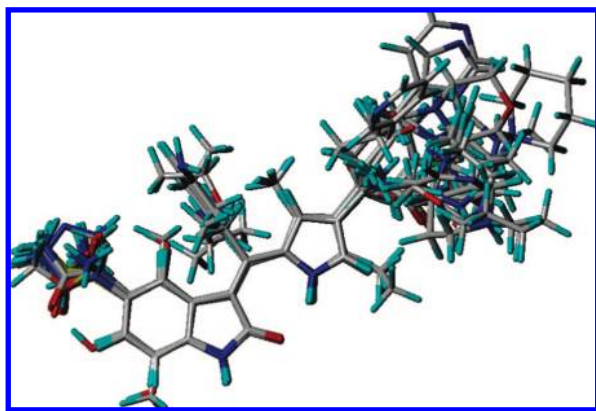

Figure 2. Superposition of all compounds in the training and test sets based on Alignment-I.

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.

2.3. Molecular Docking. Molecular docking was carried out to understand the detailed binding modes of PDK1 binding with this series of inhibitors and to 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, and 2PE2, respectively.[15,16] 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 was 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,[32] AutoDock,[33] and FlexX.[34] The fourth is an approach combining FRED docking[35,36] with energy minimization calculations on the docked binding structures, denoted by FRED-EM for convenience. On the basis of 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 an incremental construction algorithm for flexible docking.[32] This program uses the anchor and grow strategy for flexible ligand docking. 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 the SPHGEN program.[37] The spheres were selected such that they cover the entire ligand-binding region. These spheres served to orient ligands in the active site.[37] A box was created to enclose the spheres, and the energetic grid was created by the GRID module of the DOCK program. Grid calculations were carried out by using a 0.3 Å grid spacing. Both the anchor and ligand were minimized. A total of 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.[33] The AutoDock Tools were used in the preparation of protein and ligands. The molecular docking was performed using the Lamarkian genetic algorithm. The size of the grid, in which both the enzyme and ligand were embedded, was set to be 60 Å × 60 Å × 60 Å along the *x*, *y*, and *z* directions of the Cartesian coordinate system. The size of the grid is large enough to cover all of 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 the FlexX module of *SYBYL*.[34] FlexX uses an incremental construction algorithm called pose clustering.[34] In the 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 the core fragment is first placed according to the given scoring function.[38] The molecule is built up by performing a conformational search of the ligand in the active site.[34] Only the best conformation was built upon, and the others were discarded. 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, Ser160, and Ala162 was defined as the core subpocket. The ligands were docked using the multiple-ligand docking option of 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,[39] 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 500−1000 output conformations.[40] 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-dihydroindol-2-one) similar to the one found in the X-ray complex. FRED (Open Eye Scientific Software) docking calculations were carried out using protein structures with all hydrogen atoms and with the binding site definitions provided by the 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

3D-QSAR and Docking for PDK1 Inhibitors

J. Chem. Inf. Model., Vol. 48, No. 9, 2008 **1765**

Table 2. Summary of the Results Obtained from the CoMFA and CoMSIA Analyses[a]

	Alignment-I				Alignment-II			
	CoMFA-1a	CoMSIA-1b	CoMSIA-1c	CoMSIA-1d	CoMFA-1e	CoMSIA-1f	CoMSIA-1g	CoMSIA-1h
PLS statistic	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_{pred}^2	0.812	0.834	0.911	0.883	0.743	0.819	0.909	0.877
$r_{bs}^{2\,b}$	0.942			0.994				
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			0.402	0.230			0.412	0.233
donor				0.235				0.223
acceptor				0.217				0.224

[a] CoMFA and CoMSIA analyses were performed 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] A total of 100 runs of bootstrap analysis.

was placed into a 0.5-Å-resolution grid box encompassing all active-site atoms (including hydrogen atoms) using a smooth Gaussian potential.[35] 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,[35] ChemScore,[41] PLP,[42] and ScreenScore.[43] 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 the AMBER9 program. During the energy minimization in vacuum, the nonbonded cutoff and the dielectric constant were set to group-based (20 Å cutoff distance) and distance-dependent ($\epsilon = 4r$), respectively.[44] The pose with the lowest interaction energy was selected as the best binding mode.

The quantum mechanical calculations using *Gaussian03* were performed on an IBM X-series cluster with 340 nodes and 1360 processors at the University of Kentucky Center for Computational Sciences. The 3D-QSAR analyses and FlexX docking were performed by using SYBYL 7.0 software[45] on a Silicon Graphics Fuel workstation, and the other computations were carried out on a 34-processor IBM x335 Linux cluster in our own laboratory.

3. RESULTS AND DISCUSSION

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 2. The structural alignment of the compounds is one of the important steps in the development of a 3D-QSAR model.[17] 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 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 greater than 0.6. A q^2 value of greater than 0.5 is usually considered significant.[46] 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, and hydrogen-bond 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_{pred}^2 should also be used to choose the predictive QSAR models.[47] Hence. the CoMSIA model (denoted by CoMSIA-1d) from Alignment-I was selected for final analysis based on the larger r_{pred}^2 value and used to predict the activity of the compounds (see Table 2).

The CoMFA-1a model (see Table 2) has a q^2 value of 0.737 and an r^2 value of 0.907. The model was developed with five 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 nine 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, the hydrophobic field has ~23% contribution, the donor feature has 24% contribution, and the 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-
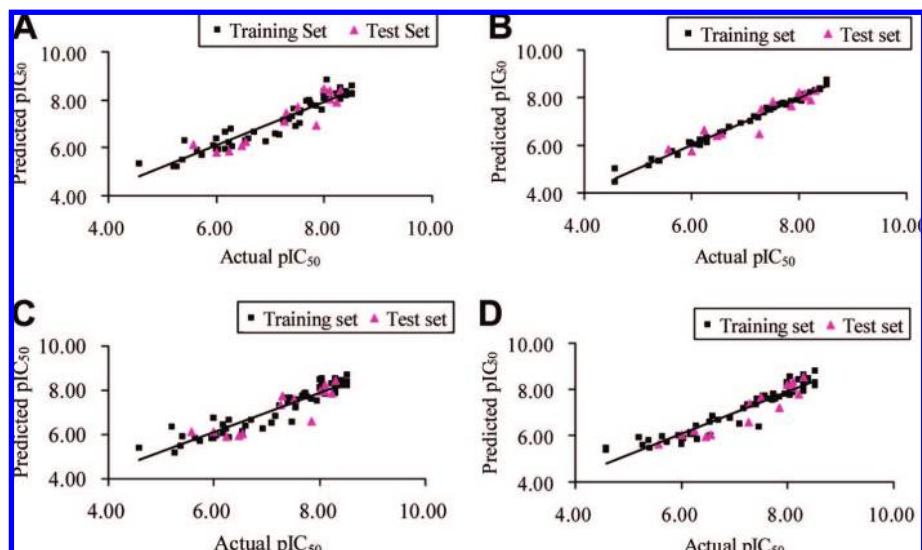
Figure 3. Plots of the predicted pIC$_{50}$ values versus the actual pIC$_{50}$ values using the training set of 56 compounds and the test set of 14 compounds: (A) CoMFA-1a model using substructure-based alignment; (B) CoMSIA-1d model using substructure-based alignment; (C) CoMFA-2a model using docking-based alignment; (D) CoMSIA-2b model using docking-based alignment.

Table 3. Result of the Cross-Validation Analyses Using Five and Two Groups

| | $q^{2\ a}$ | | | |
| | 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.

and test-set compounds (see the Supporting Information). Both the CoMFA-1a and CoMSIA-1d models have the larger r_{pred}^2 values, i.e., 0.812 and 0.883, respectively. The predicted activity of the compounds and their residuals (deviations) is given in the Supporting Information, and the plots obtained are depicted in Figure 3. 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. In the case of cross-validation using two groups, the training set is divided into two groups, 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 five 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, the average q^2 values obtained in this way were 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 of bootstrap analyses were also
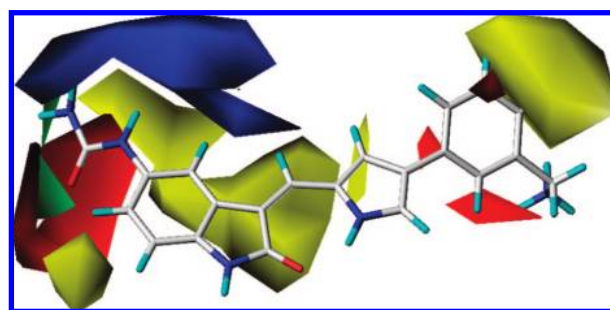

Figure 4. CoMFA steric and electrostatic contour maps around compound **60**. Green isopleths enclose areas where steric interaction is favored. Yellow contours are areas where the steric interaction is disfavored. The blue region represents the area where an electropositive group is favorable for binding. The red region refers to the area where an electronegative group is favorable for binding.

carried out. The bootstrapping results are considered as good indicators concerning whether there exist possible chance correlations.[48] Our analyses gave bootstrap r^2 values 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.

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 the steric group is not favored. The red contours show favorable electronegative regions, and the blue contour shows the region where the electropositive region is favored. These contour maps (as depicted in Figures 4 and 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 4. The yellow contour
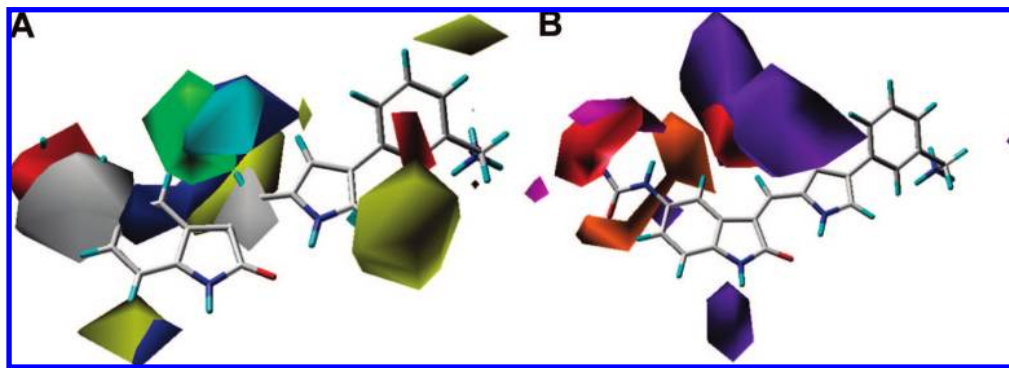
3D-QSAR and Docking for PDK1 Inhibitors

J. Chem. Inf. Model., Vol. 48, No. 9, 2008 **1767**

Figure 5. (A) CoMSIA steric, electrostatic, and hydrophobic contour maps around compound **60**. Green and yellow contours represent regions with favorable and unfavorable steric interactions, respectively. Blue and red contours represent regions that favor electropositive and electronegative groups, respectively. Cyan isopleths enclose areas where hydrophobic groups could enhance the activity. White contours represent areas where hydrophobic groups are disfavored. (B) CoMSIA acceptor and donor contour maps around compound **60**. Orange isopleths indicate regions where an acceptor group is favored, and red contours indicate regions where an acceptor group is not favored. Magenta and purple contours represent favorable and unfavorable hydrogen-bond donor regions, respectively.

region near the 7 position of the 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 methyl and hydroxyl groups, respectively, at the 7 position, and both compounds have lower activity. The yellow contours are present below the plane of the indolinone group. This shows that there would 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).[16] The indolinone group is sandwiched between hydrophobic residues. If bulky groups are introduced below the plane of the indolinone group, they will have an unfavorable steric interaction with Leu212.

The green region near the urea group shows a favorable steric interaction in this position. The red contour near the carbonyl of the urea group shows that electronegative groups are favored in this region. This explains the lower activity of compound **23**. Compound **23** has a $-CH_2NH_2$ substituent at the 5 position of the indolinone group and has an IC_{50} value of 1000 nM. Similarly, a red contour region (favorable electronegative group) is found in 4′ substituents at the pyrrole group. Compounds **39−48** have a carbonyl group at this position and have an IC_{50} value in the low nanomolar range. The blue contour region shows that electropositive groups are favored in this region. The $-NH_2$ of the urea group is present near this region and explains the increased activity of compound **26** ($IC_{50} = 18$ nM) compared to that of 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 5a shows compound **60** superimposed on the CoMSIA steric, electrostatic, and hydrophobic contour plots.

The yellow contour near the 7 position of the 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 the 5′ position of the pyrrole ring has an unfavorable steric interaction. This agrees with the reported experimental results. Compounds **12−14** and **42** have a methyl or ethyl substituent at this position and have lower activity. A green contour is present between the

4 position of the 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 **27−29** have bulky substituents in this region and have higher activity. Similarly, when we compare compound **2** ($IC_{50} = 510$ nM) with **5** ($IC_{50} = 1000$ nM), we find that the 4-methyl group (compound **2**) is slightly more favored than the 4-OH group (compound **5**) in this region, which is consistent with the relative IC_{50} values. The red contour near the carbonyl of the urea group at the 5 position of the indolinone ring shows that electronegative groups are favored in this region. The blue contours occur immediately adjacent to it. This shows the favorable electropositive region. Another favorable electronegative region is present near the pyrrole ring. A favorable hydrophobic region (cyan contour) is present near the methylidenyl ($>C=CH-$) group at the 3 position of the indolinone ring (Figure 5a). This is consistent with the increased activity observed for compounds **27−29**. In Figure 5a, a white contour showing an unfavorable hydrophobic interaction region is present near the 5 position of the indolinone group. This explains the lower activity of compound **3** ($IC_{50} = 2700$ nM) compared to that of compound **6** ($IC_{50} = 340$ nM). The hydrogen-bond donor and acceptor contours superimposed on compound **60** are shown in Figure 5b. The orange contour shows a favorable hydrogen-bond acceptor region, and the red contour shows regions in which the 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 5a. The favorable hydrogen-bond donor contour is present near the $-NH_2$ of the urea group. This is in agreement with the reported X-ray crystal structure of PDK1 bound with compound **53** (PDB ID: 2PE2).[16] In the crystal structure, the carboxylate side chain of Glu130 is present near the $-NH_2$ of the urea group and could have a favorable interaction. There are unfavorable hydrogen-bond donor and acceptor contours (Figure 5b, red and magenta contours) above the plane of the methylidenyl ($>C=CH-$) group (at the 3 position of the indolinone ring). Taking into
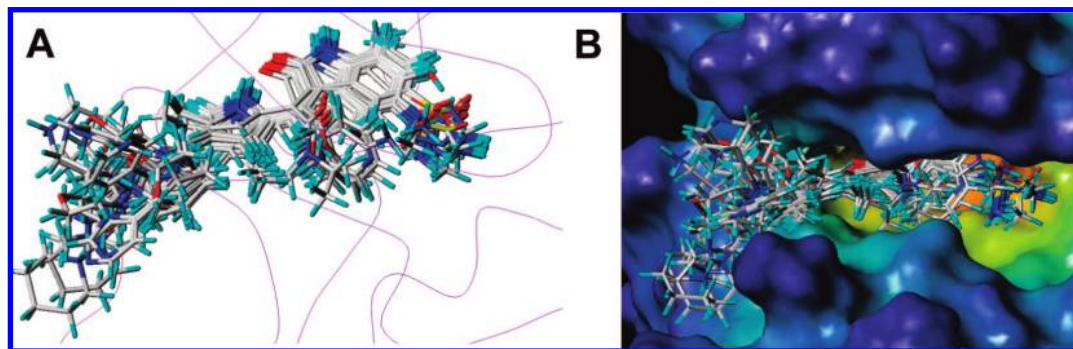
Figure 6. Molecules docked using the FRED-EM approach, i.e., combining FRED docking with energy minimization using the Amber9 program. (A) Docked poses of all 70 inhibitors in the active site of PDK1; the trace view of the PDK1 backbone is shown in magenta. (B) Different view of the 70 docked molecules; PDK1 is shown in the Connolly surface, and the surface is colored according to cavity depth.

account the hydrophobic contour from Figure 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 **31−34** because they have either a donor or acceptor group in this region. There is an additional unfavorable hydrogen-bond donor contour (purple) near the >NH of the indolinone group. The X-ray crystal structure (2PE2) shows that the −NH− of the 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 the carbonyl group of the serine backbone is present. So, this purple contour represents that the −NH− of the indolinone ring is the optimal or ideal group at this position. If a hydrogen-bond donor substitution is made at the 1 position of the indolinone ring (instead of −NH− on the ring), there will be a clash with the backbone atom, and such a hydrogen-bond donor will not be favorable for the activity.

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.[15,16] 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 those

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, 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 the pyrrole ring is flipped by ∼180° compared to that of 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 are shown in Figure 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 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_{pred}^2 values. Hence, CoMFA-2a and CoMSIA-2b were used in our further analysis below. The results are summarized in Table 4.

CoMFA. The CoMFA-2a model developed with the docked poses has a q^2 value of 0.729 and a r^2 value of 0.884. The validity of the model was tested with the external test set of 14 compounds. The 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 7b). The predicted activity of the compounds and their residuals is provided as Supporting Information, and the plots obtained are depicted in Figure 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 the actual and predicted values for compounds **28** and **33** are −1.25 and −1.12, respectively. A yellow contour present near the 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

3D-QSAR and Docking for PDK1 Inhibitors

J. Chem. Inf. Model., Vol. 48, No. 9, 2008 **1769**

Table 4. Summary of the Results Obtained from the Docking-Based CoMFA and CoMSIA Analyses

PLS statistic	FRED-EM[a]		FlexX[b]	
	COMFA-2a SE	COMSIA-2b SEHDA	CoMFA-3a SE	CoMSIA-3b 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_{pred}^2	0.736	0.840	0.726	0.806
r_{bs}^2[c]	0.911	0.926		
Fraction				
steric	0.428	0.079	0.459	0.086
electrostatic	0.572	0.283	0.541	0.248
hydrophobic		0.245		0.229
donor		0.237		0.271
acceptor		0.156		0.167

[a] CoMFA and CoMSIA analyses were performed using the training set of 56 compounds, and the alignment is based on the docked poses from the FRED-EM method. [b] CoMFA and CoMSIA analyses were performed 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] A totall of 100 runs of bootstrap analysis.

in the regions of Leu159 and Val143. Hence, bulky substitutions at the 7 position of the 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 rings. Our docked model shows that a bulky substituent at this position will have an unfavorable steric interaction with the backbone of Lys163 and the side chain of Tyr161. The red contour represents a favorable electronegative region and is present near the carbonyl of the urea group. According to the docked structure, this favorable electronegative region is surrounded by Lys111, Thr222, and the 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 the Gly165 residue. Thus, it is expected to have a favorable interaction with electronegative groups like carbonyl. The amino group on the 3″ position of the phenyl ring is present near the backbone carbonyl group of Asn164. This explains the increased activity of compounds with electropositive groups in this region. The blue contour is present in the region around the 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.

CoMSIA. The CoMSIA model developed with the docked poses has a q^2 value of 0.79 and a r^2 value of 0.909. The validity of the model was tested with the external test set of 14 compounds. The 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 is provided as Supporting Information, and the plots obtained are depicted in Figure 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 (deviation) between actual and predicted values for compound **30** is 1.09. In Figure 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 because this group is surrounded by Thr222, Lys111, and Asp 223. Compound **26** with a 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 the 3 position of the indolinone ring). This is in agreement with the experimental observation that compounds **36−38** with a hydrophobic susbtituent in this region have 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 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) contours in this region. The unfavorable hydrogen-bond donor contour near the 2-carbonyl group of the indolinone ring corresponds to the −NH− backbone of Ala162. This shows that a hydrogen-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 1 of the 10 most frequently encountered substructures in kinase inhibitors.[49] Reasonable 3D-QSAR models will be of great help in the further rational design of new ligands of the indolinone series as PDK1 inhibitors. Through a combined use of 3D-QSAR modeling and molecular docking, in this present study, we have developed ligand- and docking-based 3D-QSAR models for 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 a 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. Because the docked pose gives the bioactive conformation of the ligands, this method helps to overcome the error that 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
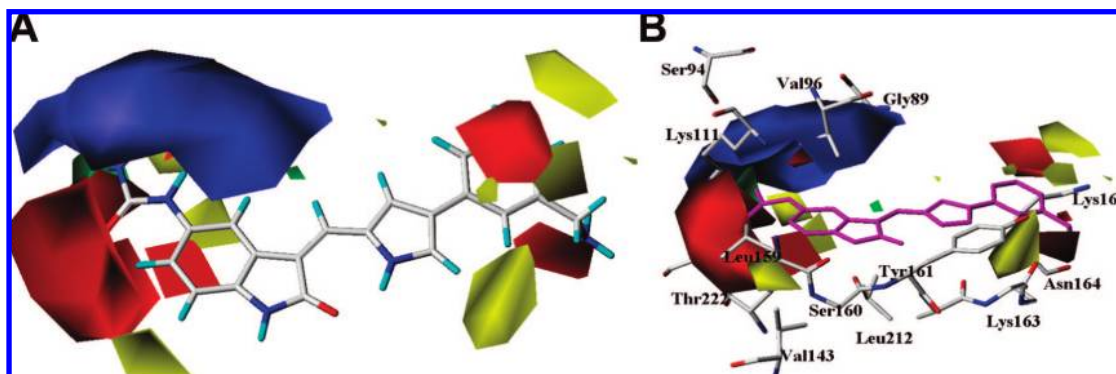
Figure 7. (A) CoMFA steric and electrostatic contour maps around compound **60**. Green isopleths enclose areas where a steric bulk could enhance the activity. Yellow contours are areas where the steric interaction is disfavored. Blue regions represent areas where a positive charge is favorable for binding. Red regions refer to areas where a negative charge is favorable for binding. (B) CoMFA contour plots shown superimposed with PDK1 active-site residues. Compound **60** is shown in magenta. Hydrogen atoms and some side chains are not shown for better clarity of this figure.
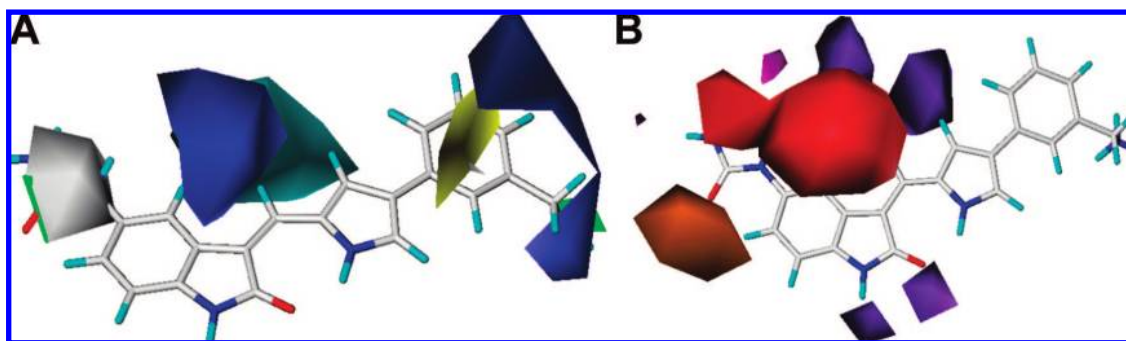

Figure 8. (A) CoMSIA steric, electrostatic, and hydrophobic contour maps around compound **60**. Green isopleths indicate regions where there is favorable steric interaction. The yellow contour is an area where the steric interaction is disfavored. Blue regions represent areas where a positive charge is favorable for binding. Red regions refer to the area where a negative charge is favorable for binding. Cyan isopleths enclose areas where hydrophobic groups could enhance the activity. The white region represents an area where hydrophobic groups are disfavored. (B) CoMSIA acceptor and donor contour maps around compound **60**. Orange isopleths indicate regions where acceptor groups are favored, and red regions indicate where acceptor groups are not favored. The magenta region indicates an area where hydrogen-bond donor groups are favored. The purple region indicated an area where hydrogen-bond donors are disfavored.

reasonable and that the developed 3D-QSAR models are reliable. One possible disadvantage of the docking-based 3D-QSAR method is that the small variation in the 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 PLS to detect the real "signal".[50] 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- and docking-based 3D-QSAR models could help us to overcome some possible error that may occur if we use one particular method.

4. CONCLUSION

The combined 3D-QSAR modeling and molecular docking results in valuable insights into PDK1 binding with 70 inhibitors and their structure−activity correlation. The 3D-QSAR modeling leads to the development of two types of satisfactory 3D-QSAR models, including the CoMFA ($r^2 = 0.907$; $q^2 = 0.737$) and CoMSIA ($r^2 = 0.991$; $q^2 = 0.824$) models, for predicting the biological activity of new compounds. Molecular docking reveals 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 3D contour maps obtained from the CoMFA and CoMSIA models in combination with the detailed PDK1 inhibitor binding structures obtained from molecular docking help to better interpret the structure−activity relationship of these PDK1 inhibitors and provide valuable insights into rational drug design for further improvement of the biological activity of the PDK1 inhibitors.

Supporting Information Available: Two tables for the detailed numerical results (pIC$_{50}$ values) predicted for all of the 70 compounds by the CoMFA-1a, CoMFA-1d, CoMFA-2a, and CoMFA-2b models in comparison with the corresponding experimental data. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Toker, A.; Newton, A. C. Cellular signaling: Pivoting around PDK-1. *Cell* **2000**, *103*, 185–188.

(2) 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.

(3) 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.

(4) Zeng, X.; Xu, H.; Glazer, R. I. Transformation of mammary epithelial cells by 3-phosphinositide dependent protein kinase-1 (PDK-1) is associated with the induction of protein kinase C. *Cancer Res.* **2002**, *62*, 3538–3543.

(5) 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.

(6) Bayascas, J. R.; Leslie, N. R.; Parsons, R.; Fleming, S.; Alessi, D. R. Hypomorphic mutation of PDK1 suppresses tumorigenesis in PTEN$^\pm$ mice. *Curr. Biol.* **2005**, *15*, 1839–1846.

(7) Lin, H.-J.; Hsieh, F.-C.; Song, H.; Lin, J. Elevated phosphorylation and activation of PDK-1/AKT pathway in human breast cancer. *Br. J. Cancer* **2005**, *93*, 1372–1381.

(8) Xie, Z.; Yuan, H.; Yin, Y.; Zeng, X.; Bai, R.; Glazer, R. I. 3-Phosphoinositide-dependent protein kinase-1 (PDK1) promotes invasion and activation of matrix metalloproteinases. *BMC Cancer* **2006**, *6*, 77.

(9) 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. U.S.A.* **2006**, *103*, 6901–6906.

(10) 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.

(11) Garber, K. The second wave in kinase cancer drugs. *Nat. Biotechnol.* **2006**, *24*, 127–130.

(12) Kim, J. A. Targeted therapies for the treatment of cancer. *Am. J. Surg.* **2003**, *186*, 264–268.

(13) 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.

(14) 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.

(15) 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.

(16) 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.

(17) 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.

(18) 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.

(19) Martin, Y. C. 3D QSAR: Current state, scope, and limitations. *Perspect. Drug Discovery Des.* **1998**, *12*, 3–23.

(20) 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, H. F., III, Schreiner, P. R. Eds.; John Wiley & Sons: Chichester, U.K., 1998; Vol. 1, pp 448−460.

(21) 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.

(22) 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. Modeling* **2008**, *48*, 179–185.

(23) Kuo, C. L.; Assefa, H.; Kamath, S.; Brzozowski, J.; 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.

(24) 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.

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

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

(27) 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.

(28) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(29) (a) Besler, B. H., Jr.; Kollman, P. A. Atomic charges derived from semiemprical methods. *J. Comput. Chem.* **1990**, *11*, 431–439. (b) Singh, U. C.; Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, *5*, 129–145.

(30) Wold, S.; Ruhe, A.; Wold, H., III *Sci. Stat. Comput.* **1984**, *5*, 735–743.

(31) Bush, B. L.; Nachbar, R. B. Sample-distance partial least squares: PLS optimized variables, with the application to CoMFA. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 587–619.

(32) 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.

(33) 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. Comput. Chem.* **1998**, *19*, 1639–1662.

(34) Rarey, M.; Kramer, B.; Lengauer, T.; Kleb, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.

(35) McGann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K. Gaussian docking functions. *Biopolymers* **2003**, *68*, 76–90.

(36) *OEChem*, version 1.4.2;Open Eye Scientific Software, Inc.: Santa Fe, NM, 2005; www.eyesopen.com.

(37) 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 3-dimensional structure. *J. Med. Chem.* **1988**, *31*, 722–729.

(38) Rarey, M.; Wefing, S.; Lengauer, T. Placement of medium sized molecular fragments into active sites of proteins. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 41–54.

(39) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; 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*; University of California: San Francisco, 2006.

(40) Bostrom, J.; Greenwood, J. R.; Gottfries, J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *J. Mol. Graphics Modell.* **2003**, *21*, 449–462.

(41) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. 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.

(42) Gehlhaar, D. K.; Verkhivker, G. M.; Rejto, P. A.; Sherman, C. J.; Fogel, D. B.; Fogel, L. J.; Freer, S. T. Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chem. Biol.* **1995**, *2*, 317–324.

(43) Stahl, M.; Rarey, M. Detailed analysis of scoring functions for virtual screening. *J. Med. Chem.* **2001**, *44*, 1035–1042.

(44) Harvey, S. C. Treatment of electrostatic effects in macromolecular modeling. *Proteins* **1989**, *5*, 78–92.

(45) Tripos Associates, Inc., 1699 S. Hanley Street, St. Louis, MI 63144.

(46) 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.

(47) 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.

(48) 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.

(49) 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.

(50) Cramer, R. D.; Wendt, B. Pushing the boundaries of 3D-QSAR. *J. Comput.-Aided. Mol. Des.* **2007**, *21*, 23–32.