Computational Analysis of PKA−Balanol Interactions

Chung F. Wong,*[,†,‡] Philippe H. Hünenberger,[§] Pearl Akamine,[‡,||] Narendra Narayana,[⊥] Tom Diller,[‡,||] J. Andrew McCammon,[†,‡,||] Susan Taylor,[‡,||] and Nguyen-Huu Xuong[||,#]

Department of Pharmacology, Howard Hughes Medical Institute, School of Medicine, Department of Chemistry and Biochemistry, and Departments of Biology and Physics, University of California at San Diego, La Jolla, California 92093, Laboratorium für Physikalische Chemie, ETH-Zentrum, Universitãtstrasse 6, CH-8092 Zürich, Switzerland, and Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106

Received October 11, 2000

Protein kinases are important targets for designing therapeutic drugs. This paper illustrates a computational approach to extend the usefulness of a single protein−inhibitor structure in aiding the design of protein kinase inhibitors. Using the complex structure of the catalytic subunit of PKA (cPKA) and balanol as a guide, we have analyzed and compared the distribution of amino acid types near the protein−ligand interface for nearly 400 kinases. This analysis has identified a number of sites that are more variable in amino acid types among the kinases analyzed, and these are useful sites to consider in designing specific protein kinase inhibitors. On the other hand, we have found kinases whose protein−ligand interfaces are similar to that of the cPKA−balanol complex and balanol can be a useful lead compound for developing effective inhibitors for these kinases. Generally, this approach can help us discover new drug targets for an existing class of compounds that have already been well characterized pharmacologically. The relative significance of the charge/polarity of residues at the protein−ligand interface has been quantified by carrying out computational sensitivity analysis in which the charge/polarity of an atom or functional group was turned off/on, and the resulting effects on binding affinity have been examined. The binding affinity was estimated by using an implicit-solvent model in which the electrostatic contributions were obtained by solving the Poisson equation and the hydrophobic effects were accounted for by using surface-area dependent terms. The same sensitivity analysis approach was applied to the ligand balanol to develop a pharmacophoric model for searching new drug leads from small-molecule libraries. To help evaluate the binding affinity of designed inhibitors before they are made, we have developed a semiempirical approach to improve the predictive reliability of the implicit-solvent binding model.

Introduction

Protein kinases play an important role in cell signaling, and they represent interesting targets for developing therapeutic agents. For example, a new approach for treating cancer by stopping blood supply to cancer cells may be realized by developing drugs inhibiting proangiogenic protein kinases. These protein kinases include the receptors for the vascular endothelial growth factor,[1] the basic fibroblast growth factor,[2] the platelet-derived growth factor,[3] angiopoietin-1,[4] and epidermal growth factor.[5] Several drug candidates designed upon this principle are currently undergoing clinical trials.

The protein kinase inhibitors that are most advanced in drug development bind to the catalytic domain of protein kinases and compete with ATP for binding.[6,7] With the determination of the crystal structures of a number of catalytic domains of protein kinases and of their complexes with inhibitors,[8−11] structure-based drug design is currently underway in many laboratories. However, given the similarity of the ligand or ATP-binding sites observed in these structures, it is dubious whether specific kinase inhibitors targeting the ATP-binding site can be designed. On the other hand, screening data have already demonstrated that some inhibitors can achieve specific binding. These include lavendustin A,[12] 3-substituted quinoline derivatives,[13,14] dianilinophthalimides,[15−17] phenylaminopyrimidines,[18,19] and phenylaminoquinazolines.[20,21] Also, several balanol analogues present binding specificity even among isoforms of the same kinase.[22−27] Since the ATP-binding pocket is very similar among protein kinases, what then are the key determinants for specific binding?

One possible explanation is that there may be sites, in the ATP binding pocket, that are more variable in amino acid types for different kinases. If this were the case, the identification of these sites would provide a valuable basis for the design of specific protein kinase inhibitors. In the present study, we use the crystal structure of the cPKA−balanol complex as a guide to define a protein−ligand interface and analyze the distribution of amino acid types near the interface using almost 400 protein kinases. We also quantify the role of these amino acids in balanol recognition by carrying

* To whom correspondence should be addressed. Phone: 858-534-2971. Fax: 858-534-7042. E-mail: c4wong@ucsd.edu.
† Department of Pharmacology, University of California at San Diego, La Jolla.
‡ Howard Hughes Medical Institute, School of Medicine, University of California at San Diego, La Jolla.
§ Laboratorium für Physikalische Chemie, ETH-Zentrum.
|| Department of Chemistry and Biochemistry, University of California at San Diego, La Jolla.
⊥ Case Western Reserve University.
Departments of Biology and Physics, University of California at San Diego, La Jolla.

10.1021/jm000443d CCC: $20.00
Published on Web 04/06/2001

out extensive computational "mutagenesis" experiments employing an implicit-solvent model.

Cocrystal structures of inhibitor−receptor complexes have generated many useful insights into drug−receptor interactions. One can often identify key functional groups of a ligand interacting with the protein by graphically examining these structures. Yet, studying drug−receptor interaction can be more complicated. For example, when a hydrogen bond is formed between a ligand and its receptor, this does not automatically imply that the hydrogen bond aids ligand−receptor recognition. To form such an intermolecular hydrogen bond, the favorable hydrogen bonds that the donor and acceptor form with water molecules are disrupted upon complexation. Therefore, without a quantitative tool for estimating binding affinity, it is difficult to conclude whether a hydrogen bond is significant for binding. Continuum electrostatic calculations, for example, have already demonstrated that hydrogen bonds may or may not stabilize proteins, depending on the specific context.[28]

To address this type of problem, we proposed to dissect the determinants of ligand−receptor recognition by perturbing the charge, polarity, size, etc. of different functional groups to examine their possible significance in binding.[29] In the present work, this idea is applied to identify some of the key features of balanol responsible for binding to cPKA. Once these features are identified, they can help us search new leads of protein kinase inhibitors by scanning small-molecule databases for compounds possessing these features. Such an approach can increase the amount of information a single structure of a receptor−ligand complex can provide in structure-based drug design.

Once new leads are found, a more challenging computational problem is to rank their derivatives, which may not have yet been synthesized, according to their binding affinity to their target. For rapid computational screening, fast but reliable computational models need to be developed. Although explicit-solvent models have a firmer theoretical basis, they are too expensive to use in day-to-day drug discovery efforts. On the other hand, simpler quantitative structure−activity relationship (QSAR) models are cheaper to use but do not have as sound a theoretical basis. Many QSAR models still assume a linear relationship between free energy and QSAR descriptors, although more rigorous explicit-solvent models have suggested nonlinear behavior.[29] Nevertheless, when good descriptors are chosen, the parameters of a linear model may effectively include some nonlinear effects when these parameters are fitted to experimental data. A popular application of QSAR models has been to relate the activity of ligands to their molecular properties. When the structure of a ligand−receptor complex is known or can be modeled, it is beneficial to use quantities pertaining to receptor−ligand interaction rather than ligand characteristics alone because ligand−receptor interaction energy tells more about ligand−receptor recognition than ligand properties. Some encouraging semiempirical models using interaction energies have already been developed. For example, several semiempirical linear response theories have been used to correlate averaged ligand−receptor interaction energy obtained from explicit-solvent molecular dynamics simulations with binding affinity.[30−33] Semiempirical implicit-solvent models have also been used in docking programs such as Autodock.[34] Among different kinds of implicit-solvent models, those that describe the physics more faithfully should yield more reliable results. For computational efficiency, many implicit-solvent models draw on simple electrostatic descriptions, such as using distance-dependent dielectric functions. With the increasing availability of powerful computers and the development of new methods for solving the Poisson−Boltzmann equation for molecules with complex shapes, it should soon become feasible to use these more elaborate continuum electrostatics models in day-to-day drug discovery efforts. Here, using the interaction between the catalytic subunit of PKA and a number of its inhibitors as an example, we examine whether a semiempirical approach utilizing the finite-difference Poisson−Boltzmann solver implemented in the program UHBD[35,36] can improve the predictive reliability of the binding affinity of protein kinase inhibitors to cPKA.

Methods

Database Analysis. There are probably about 2000 kinases in the human genome.[37] It will be some time before the crystal/NMR structures of all the kinases are determined, not to mention that some of them have not yet been identified or disclosed. Here, we use an existing database of 390 proteins to examine the conserved and variable features of the ligand-binding site in order to gain insights into designing specific protein kinase inhibitors. This database contains proteins whose sequences have been aligned by Hanks and Quinn[38] and is available in the Protein Kinase Resource maintained at the University of California, San Diego (http://www.sdsc.edu/kinases/pkr/pk_catalytic/pk_hanks_seq_align_long.html). The database also includes protein kinases found in species other than human. However, since there are many protein kinases in the human genome that have not yet been identified or released to the public, we kept in our analysis protein kinases from other species as an approximation to yet unknown or undisclosed human protein kinases.

The definition of a ligand-binding site depends on the ligand considered. The region of the catalytic domain for binding ATP and peptide substrate is quite extensive. Different small-molecule inhibitors also vary somewhat in the extent and region of their interaction with their targets. In addition, some small-molecule inhibitors interact with regions not targeted by natural ligands. Here, we focus on the part of the catalytic domain where balanol recognizes cPKA, as an illustration for studying specific binding by balanol and its derivatives.

Starting from the cPKA−balanol crystal structure,[11] polar hydrogens were added using CHARMm22(39). For neutral histidines, the hydrogen was added at either the δ or ϵ position depending on which site was more likely to act as a hydrogen bond donor. The CHARMm22 force field[39] was used in the calculations. For balanol, we used the same potential parameters as in Hünenberger et al.'s work.[40] The protein−ligand complex was allowed to relax by 100 steps of conjugate gradient energy minimization using a distance dependence dielectric of $\epsilon = 5R$ where R is an interatomic distance. The ligand was then subjected to extensive conjugate gradient energy minimization using the same distance dependent dielectric until a gradient norm less than 0.001 kcal/mol/Å was reached.

Protein atoms at the protein−ligand interface were then identified by finding those atoms lying within a given distance R_{cut} from any ligand heavy atom. Residues associated with these atoms were considered to be lying at the protein−ligand interface. It is then possible to analyze, from the sequence alignment by Hanks and Quinn,[38] the variability of amino acid residues at these sites. Sites characterized by a large vari-

Table 1: Protein and Ligand Functional Groups Whose Charges Are Modified in Sensitivity Analysis Calculations

functional group	charge modifications[a]
C−O−H of balanol	H, 0.40 → 0; O, −0.65 → 0; C, 0.25 → 0
C_2NH_2 of balanol	N, −0.3 → 0; H, 0.35 → 0; C, 0.3 → 0
COO^- of balanol	C, 0.14 → 0; O, −0.57 → 0
CO of benzophenone moiety of balanol	C, 0.4 → 0; O, −0.4→ 0
CO group of amide in balanol	C, 0.55 → 0; O, −0.55 → 0
NH group of amide in balanol	N, −0.35 → −0.1; H, 0.25 → 0
CO of ester linkage in balanol	C, 0.7 → 0.3; O, 0.4 → 0
COC (C4−2O1A−C1″) near ester linkage in balanol	C4, 0.7 → 0.3; 2O1A, −0.55 → 0; C1″, 0.25 → 0
extended carbon atoms on aromatic rings of balanol	0 → 0.3
CO group of peptide linkage	C, 0.6 → 0.05; O,−0.55 → 0
NH_3 group of LYS	NH_3, 1 → 0
OH of SER or THR	O, −0.65 → −0.25; H, 0.4 → 0
COO^- of GLU or ASP	COO^-, −1 → 0
NH_2 of GLN	N,−0.6 → 0; H, 0.3 → 0
charge of nonpolar carbons	0 → 0.3

[a] Charges are expressed in electrostatic unit.

ability are of special interest because they represent sites to consider for designing specific kinase inhibitors. The identification of protein residues at the protein−ligand interface depends on the choice of R_{cut}. In the present study, we have considered a range of R_{cut} varying between 3 and 4 Å.

Binding Free Energy Calculations. Three calculations were carried out to obtain a single binding free energy: one calculation on the kinase−inhibitor complex, one on the kinase alone, and one on the inhibitor alone. We carried out the calculations in the same way as in Hünenberger et al.'s work[40] to facilitate comparison with earlier results. The energy-refined structure of cPKA−balanol was placed in a 240 × 240 × 240 grid with a 0.3 Å grid spacing. The electrostatic energy of the complex was calculated by solving the Poisson equation on this grid. The electrostatic energy consisted of two terms: one resulting from the assembling of atomic charges in a dielectric medium characterized by the dielectric constant of the complex (ΔG_{coul}); the other resulting from the creation of a dielectric boundary when the complex was immersed in a high-dielectric solvent (ΔG_{PB}). These two terms were calculated using the UHBD program.[35,36] Similar electrostatic calculations were carried out separately for the protein and the ligand using the same conformation as in the complex. The protein or the ligand was positioned in the grid exactly the same way as it was in the complex to minimize computational differences arising from different placement of the molecule onto the grid. In the electrostatic calculations, an internal dielectric constant of 2 and an external dielectric constant of 78 were used. The difference between the electrostatic free energy of the complex and the sum of the electrostatic free energies of the protein and the ligand was used to estimate the electrostatic contribution to the binding energy in the single-conformation approximation.

The contribution of hydrophobic effects to the binding energy was estimated by multiplying the change in solvent accessible surface area upon binding by 25 cal mol^{-1} $Å^{-2}$,[41,42] leading to an additional term (ΔG_{SA}). The solvent accessible surface was determined by using a probe sphere of 1.4 Å radius, and the van der Waals radii of protein atoms were taken from CHARMm22.[39]

Sensitivity Analysis. To study the significance of the polarity or charge of a functional group in affecting binding, we turned off the polarity or charge of specific groups and repeated the binding energy calculation as discussed above. We then calculated ΔG_{bind}(after) − ΔG_{bind}(before) where ΔG_{bind}(after) and ΔG_{bind}(before) were respectively the binding energies after and before the charge or polarity was turned off. Table 1 describes how the switching off of charge or polarity of different functional groups was done. Some atoms/extended atoms possessing zero charges are also included. Since one can no longer turn off their charges, we turned on the charge of each of these atoms/extended atoms by adding 0.3 e to it and repeated the binding energy calculation. These calculations probe the electrostatic environment around these atoms and suggest whether substituting these nonpolar atoms by polar groups may improve binding.

Semiempirical Model. In a semiempirical model, we scaled different contributing terms to a binding energy before summing them together. In a two-parameter model, we estimated the binding energy as $\Delta G_{bind} = a + b(\Delta G_{coul} + \Delta G_{PB} + \Delta G_{SA})$. Although $\Delta G_{coul} + \Delta G_{PB} + \Delta G_{SA}$ can in principle be directly compared to corresponding experimental binding free energy,[40] in practice this suffers from a number of deficiencies: (i) the single-conformation approximation does not account for conformational fluctuation; (ii) the choice of model parameters such as atomic partial charges, van der Waals radii, microscopic "surface tension", and dielectric constant is not optimal; (iii) the loss of translational−rotational entropy is not accounted for. Scaling $\Delta G_{coul} + \Delta G_{PB} + \Delta G_{SA}$ by a constant and adding a constant to the result could partially correct these deficiencies in an empirical sense. The constants a and b were adjusted to obtain the least root-mean-square difference between ΔG_{bind} and the corresponding experimental value for a series of inhibitors.

If adequate experimental data are available for determining additional empirical parameters, one may use a model with three or more parameters. In this paper, we also examine the models $\Delta G_{bind} = a + b\Delta G_{SA} + c(\Delta G_{coul} + \Delta G_{PB})$ and $\Delta G_{bind} = a + b\Delta G_{SA} + c\Delta G_{PB} + d\Delta G_{Coul}$. In this work, 17 compounds from Hünenberger et al.'s study[40] were used to determine the model parameters in a least-squares manner.

Results and Discussions

Database Analysis. To gain insight into what other protein kinases may also be inhibited by balanol, a good starting point is to identify protein kinases having the same amino acid residues at the protein−balanol interface as the energy-refined cPKA−balanol complex. In this complex, seven residues possess at least one atom lying within 3 Å of a balanol heavy atom. Two of them (GLY50 and GLU91) contact balanol only through backbone atoms. Since these residues are not likely to account for binding specificity, they are ignored in the following analysis. The remaining five residues use their side chains to recognize balanol, and examining the variability of amino acid types in these five positions among many kinases can provide insights into specific binding by balanol. These five residues are SER53 that interacts with the O4″ of balanol (see Figure 1) through its OH group, LYS72 that recognizes 2O10 with its positive terminal ammonium group, GLU91 that interacts with the hydrogen attached to 2O10 with its carboxylate group, THR183 that forms a hydrogen bond with O1′ through its OH group, and ASP184 that interacts with the hydrogen attached to O6″ with one of its carboxylate oxygens. Among the kinases analyzed by Hanks and Quinn,[38] 22 possess exactly these five

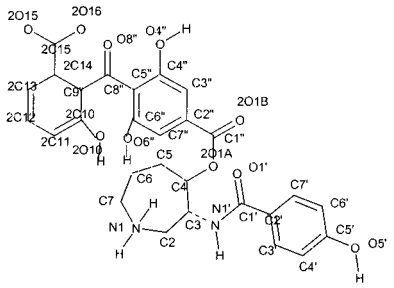

Figure 1. Labeling of atoms, as in the PDB entry 1BX6 (http://www.rcsb.org/pdb/), in balanol.

Table 2: Kinases Having the Same Residues as PKA at the Protein−Balanol Interface[a]

	class	$R_c = 3.0$, $n = 5$	$R_c = 3.4$, $n = 8$	$R_c = 3.6$, $n = 13$	$R_c = 3.8$, $n = 8$
cAPKa	AGC group I	x	x	x	x
cAPKb		x	x	x	x
cAPKg		x	x	x	x
EcAPKa		x			
DC0		x	x	x	
DC1		x			
AplC		x			
SAK		x			
DdPK1		x			
TPK1		x	x		
TPK2		x	x		
TPK3		x	x		
B.emC		x	x	x	x
PKCg	AGC group II	x	x		
PKCz		x			
PKCi		x			
RSK1N	AGC group VI	x			
RSK2N		x			
Sgk	AGC group other	x	x		
DdK2		x			
PSK−H1	CaMK group I	x			
SER	PTK group X	x			

[a] Marked by crosses. Residues having an atom within R_c of an atom of balanol are considered residues at the protein−ligand interface. R_c is in units of angstrom. Kinases studied by Hanks and Quinn[38] are included in this analysis. n = number of matching residues.

residues at the protein−ligand interface (Table 2, 3 Å column). If these five residues provide the dominant contribution to binding affinity, these 22 kinases may bind balanol. This illustrates how this type of analysis, based on a single kinase−inhibitor complex, can suggest what other kinases may bind the inhibitor. This information is very valuable in drug design because it may lead to the discovery of new drug targets for compounds that are already well characterized in terms of their chemical, physical, biological, pharmacological, and toxicological properties.

For the 22 kinases identified, not all of them are in the same group as PKA. Three of them (PKCg, PKCz, PKCi) are in the protein kinase C (PKC) family. Two of them (RSK1N, RSK2N) are protein kinases that phosphorylate ribosomal protein S6. In addition, a serum/glucocorticoid-regulated kinase (Sgk), a protein kinase regulated by calcium/calmodulin (PSK−H1), and a protein kinase in the epidermal growth factor receptor family (SER) have also been identified. It is interesting to see that this kind of analysis can identify targets outside the family toward which a class of compounds is originally aiming for.

However, this type of qualitative analysis also has limitations. For example, balanol is known to inhibit the PKC isozymes α, β, γ, δ, ϵ, and η,[22−25,27] but they were not picked up by this analysis. This is because these isozymes only have three or four residues matching the corresponding ones in cPKA. The residues that do not match may enhance, diminish, or not change the binding affinity. But it is difficult to assess the role of these residues without doing quantitative measurements or calculations. Nevertheless, for identifying new drug targets, missing some is acceptable because it is already profitable to come up with just a few new potential targets for which new drugs can be developed based on a company's existing compound library. In the example just given, a protein kinase regulated by calcium/calmodulin has been identified. There are already data showing that balanol binds proteins in this family, although with an inhibition constant (∼70 nM) an order of magnitude less favorable than that for PKA. But a 70 nM compound can already be a useful lead for developing effective inhibitors for a target.

The above analysis is based on residues closest to the ligand. Since long-range interactions, arising from electrostatic interactions for example, may also influence binding, it is useful to examine the effect of residues more remote from the inhibitor. To this end, the analysis was repeated for several other cutoff values (Table 2). By use of a 3.4 Å cutoff, eight residues of cPKA are found to interact with balanol through their side chains. Only 10 protein kinases in the database possess exactly these same residues at the protein−ligand interface. When the cutoff is increased to 3.6 Å, 13 residues are found at the interface and 5 proteins have these same 13 residues at the interface. For a 3.8 Å cutoff, there are 18 residues at the interface and only 4 protein kinases hold these residues at the interface. Therefore, longer-ranged interactions, if significant, could help different protein kinases with very similar folds exhibit distinct binding affinities toward the same ligand such as balanol.

Table 3 gives the distribution of amino acid types at each of the 18 positions identified using the 3.8 Å cutoff. Conserved residues occupy several of these positions. K72 and E91 are a conserved salt-bridge present in many kinases. D184 is the conserved catalytic aspartate. G55 is also quite conserved, although 37 kinases in the database possess a serine at this position and 19 kinases an alanine. However, the β-carbon of this residue points away from the ligand-binding pocket so that its side chain does not interact directly with the ligand. Therefore, the substitution of this glycine by an alanine or a serine may only affect binding indirectly, such as by altering the local conformation, if at all. V57, A70, and G186 are also largely conserved and therefore may not be useful residues to consider in designing specific kinase inhibitors.

Other positions display more variability in amino acid types. Many other polar, nonpolar, or charged amino acid types in other protein kinases replace S53 in cPKA. This residue is close to the carboxylate of balanol, and replacing the carboxylate group by other functional groups may yield balanol derivatives selective toward certain desired kinases. T183 is replaced by nonpolar or polar, but not charged, residues in other kinases. As indicated by binding energy calculations (see below), the polarity of the hydroxyl group of the threonine side chain does play a role in recognizing balanol. Thus, this

Table 3: Distribution of Amino Acid Types at the PKA−Balanol Interface[a]

	S53	K72	E91	T183	D184	F54	G55	V57	L74
G	38	0	1	47	0	7	326	0	0
A	75	0	1	123	1	0	19	6	4
V	5	0	1	11	0	0	0	358	46
L	1	0	0	3	0	0	1	2	174
I	1	0	0	33	0	8	0	6	97
S	86	0	0	59	0	12	37	1	2
T	45	0	0	53	0	4	1	4	0
D	8	1	2	0	379	0	0	0	0
N	25	0	0	0	0	3	0	0	0
K	14	384	0	0	0	0	0	0	0
E	27	0	375	0	0	0	0	0	7
Q	27	0	2	0	0	2	0	0	2
R	14	1	1	0	0	0	0	0	1
H	8	0	1	1	0	1	1	0	1
F	3	0	0	0	0	260	0	0	20
C	7	0	0	54	0	1	0	8	4
W	0	0	0	0	0	0	0	0	0
Y	0	0	0	0	1	82	0	0	0
M	1	0	0	0	0	5	0	0	28
P	0	0	0	0	0	0	0	0	0

	Y122	A70	V123	E127	L49	G186	L173	M120	Q84
G	0	0	3	0	0	372	0	1	18
A	4	358	29	7	1	3	0	0	17
V	1	2	73	1	31	0	7	13	18
L	85	2	51	0	214	0	314	65	18
I	9	0	20	0	131	0	5	9	12
S	1	0	2	98	1	2	1	5	25
T	0	1	1	14	0	0	0	80	11
D	0	0	0	148	0	7	0	0	64
N	0	0	0	30	1	0	0	1	7
K	2	0	1	1	3	0	0	0	28
E	2	0	0	66	0	0	0	1	52
Q	0	0	2	3	0	0	0	22	58
R	2	0	0	0	0	0	0	2	9
H	12	0	11	0	0	0	0	0	7
F	81	1	0	1	2	0	20	66	12
C	7	3	59	8	0	0	0	0	2
W	7	0	0	0	0	0	0	0	1
Y	167	0	7	0	0	0	0	18	14
M	4	1	124	1	1	0	37	113	9
P	0	0	1	5	0	0	0	0	3

[a] 390 protein kinases in the sequence alignment by Hanks and Quinn[38] were analyzed. The first row lists the amino acids lying at the protein−ligand interface. The first five residues, S53, K72, E91, T183, and D184, were identified by using $R_c = 3.0$ Å (see Table 2 and text). More residues were identified when R_c was increased to 3.4 Å, and these residues were added to the previous list after the double line. This is repeated for $R_c = 3.6$ and 3.8 Å. The first column lists the 20 amino acid types that can replace a particular amino acid of cPKA at the protein−ligand interface.

position may be worth considering for designing specific kinase inhibitors. F54 is quite conserved but is replaced by Y in some kinases. Since F is nonpolar but Y contains a polar hydroxyl group, suitable modification of the benzophenone moiety of balanol might distinguish F-containing kinases from Y-containing kinases. There is some variability in position 74, although nonpolar residues, either small ones such as L and I or large ones such as F, are more common. Thus, it may be possible to use these size differences to achieve binding specificity. V123 in cPKA points into the binding pocket to interact with the *p*-hydroxylbenzamide moiety of balanol. This position is somewhat variable in amino acid type, although more often occupied by a nonpolar residue, and is thus also worth considering for designing balanol derivatives targeted toward certain kinases. Y122 is also near the *p*-hydroxylbenzamide moiety of balanol, and this position is somewhat variable in amino acid type. However, the side chain of this residue is pointing away from the balanol-binding pocket, making it more difficult to utilize this position to design specific kinase inhibitors. Nevertheless, some indirect effects may be achieved with the C-terminal tail of some protein kinases when this tail can interact with the residue at position 122. The amino acid sequences of the C-terminal tail are somewhat variable, and this variability can be transmitted into the binding site through residue 122. Leucine occurs frequently in position 49 but is replaced by I and occasionally V in some protein kinases. Because of the similarity of these residues, it may be hard to utilize this position to design specific inhibitors. E127 of cPKA interacts with the quaternary ammonium group of balanol. Many kinases have either a D or an E in this position. Balanol may not be able to distinguish these kinases because both D and E are negatively charged under physiological conditions. However, since this is the helix-capping position of helix D, helix-capping residues besides D and E are found in other protein kinases. Therefore, one may still be able to design balanol anologues that can distinguish kinases possessing D/E in this position from those that do not. M120 in cPKA is replaced by a number of different residues of various sizes in other kinases and can be a useful site to consider in designing balanol derivatives targeted toward specific protein kinases. Q84 of PKA is near the carboxylate group of balanol. This position is quite variable, including both charged and uncharged amino acids and small and large residues and may be another useful site to consider in specific inhibitor design.

Role of Residue Charge or Polarity on Balanol Recognition. Since there may not be a simple correlation between the distance of a residue from the ligand and the significance of the residue in recognizing balanol, we performed a quantitative sensitivity analysis by successively turning off/on the charge or polarity of different functional groups (see Table 1) near the ligand to examine how significant this change can affect binding affinity. Table 4 lists the results obtained for residues lying within 3 Å of the ligand and in the ranges 3−3.4, 3.4−3.6, and 3.6−3.8 Å. First consider the five residues closest to balanol. Turning off the charge of the ammonium group of K72 or of the carboxylate group of D184 has the largest unfavorable influence on binding. Although these two residues are important for binding balanol, they are conserved among protein kinases. Thus, inhibitors targeted toward these residues may distinguish protein kinases from nonprotein kinases but may not differentiate specific members of the protein kinase family. The polarity of T183 also appears significant in influencing binding. Turning off the polarity of the hydroxyl group decreases the binding affinity by 1.1 kcal/mol. As discussed above, either polar or nonpolar residues may occupy this position in other kinases. The functional group of balanol closest to the hydroxyl group of T183 is the carbonyl group of the amide linkage of balanol. By suitably modifying the linkage, one may be able to design balanol derivatives targeted toward specific protein kinases. Although E91 also possesses a carboxylate as does D184, turning off its charge has less effect on binding. This glutamate is a conserved residue that often forms a salt bridge with K72. Turning off the

Table 4: Sensitivity[a] of Binding Energy to Perturbing Polarity or Charge

	sensitivity (kcal/mol)	3.0 A	3.4 A	3.6 A	3.8 A
53SER OG	0.72	x	x	x	x
72LYS NZ	3.5	x	x	x	x
91GLU OE1	0.81	x	x	x	x
183THR OG1	1.1	x	x	x	x
184ASP OD1	2.2	x	x	x	x
54F CD1	−1.2		x	x	x
54F CE2	−0.86		x	x	x
55G CA	0.21		x	x	x
57V CG2	0.68		x	x	x
74L CD2	−0.83		x	x	x
122Y CD1	−0.067		x	x	x
54F CE1	−0.92			x	x
54F CD2	−0.067			x	x
70A CB	0.59			x	x
54F CD2	−1.0			x	x
123V CG1	0.27			x	x
127E OE	1.5			x	x
183T CG2	1.0			x	x
327F CZ	0.38			x	x
54F CG	−1.0				x
54F CZ	−0.68				x
49L CD1	0.39				x
186G CA	0.045				x
173L CD1	0.39				x
120M CE	−0.064				x
84Q NE2	0.025				x

[a] Sensitivity measures the change in calculated binding affinity, ΔG_{bind}(after) − ΔG_{bind}(before), when the polarity/charges of a charge group is modified according to Table 1.

polarity of the hydroxyl group of S53 also has some effect on binding, although not as significant as turning off the polarity of the hydroxyl group of T183. This hydroxyl group interacts with the carboxylate group of balanol. Since this serine is replaced by a number of other amino acids in other kinases, replacing the carboxylate group of balanol by other functional groups may produce balanol derivatives that bind specifically to certain protein kinases.

Between 3 and 3.4 Å, several additional residues appear as proximal residues to balanol: F54, G55, V57, V74, and Y122. The atoms of these residues near balanol are all nonpolar. However, changing the polarity of some could affect binding, as indicated by the sensitivity of the binding affinity to placing charges of 0.3 e on some of these atoms. The calculations suggest that making the phenylalanine ring of F54 more positive could improve binding. However, this position is largely occupied by aromatic residues in other kinases, although a few kinases have polar residues such as serine in this position. Therefore, this may not be a very good site to consider for designing specific kinase inhibitors. On the other hand, placing positive charges on atoms of residue 57 diminishes binding. However, this position is conserved and thus less useful to consider for inhibitor design. Making the L74 side chain more positive could improve binding. Although most kinases have a leucine in this position, a small number of kinases have charged residues and these protein kinases may bind balanol differently than PKA. Placing charges on atoms of G55 and Y122 has smaller effects on binding.

Extending the cutoff distance to 3.6 Å, residue E127 appears to play a key role in binding. Turning off the charge of its carboxylate group decreases the binding affinity by ∼1.5 kcal/mol. This is probably a consequence

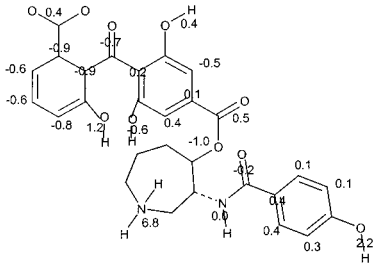

Figure 2. Binding energy change as a result of perturbing the charge or polarity of different functional groups of balanol, as described in Table 1.

of the favorable electrostatic interaction between the ammonium group of balanol and the carboxylate oxygens situated at a distance between 4 and 4.6 Å away. Although many protein kinases have a glutamate or an aspartate residue in this position, some contain uncharged residues. These uncharged residues are mostly polar ones and are residues commonly found in a helix-capping motif because this is a helix-cap position of the D helix. Therefore, the ammonium group of balanol probably helps to distinguish between kinases containing negatively charged residues and those containing uncharged residues at position 127. Replacing the ammonium group by a negatively charged group can provide a mechanism to switch the selective binding between these two groups of protein kinases.

Sensitivity Analysis of Ligand Functional Groups. One can apply the same idea to analyze significant functional groups of balanol for recognizing PKA. Figure 2 illustrates the change in binding affinity when the charge or polarity of different atoms or atom groups of balanol is changed (see Table 1). The positively charged ammonium group of balanol plays a very important role in recognition. Removing its charge reduces the binding affinity by 6.8 kcal/mol. As discussed earlier, the ammonium group is close to a negatively charged residue E127 and the favorable electrostatic interactions between this residue and the ammonium group of balanol may account for the significance of the ammonium group in binding. This calculation result is consistent with the experimental activity measurements[43] that replacing the positively charged ammonium group with neutral and electron-rich groups ($NH_2^+ \rightarrow CH_2$, O, S, or SO_2) diminishes binding to PKC (0.074 → 0.27, 6.7, 3.3, or 9.6, respectively, in IC_{50} for PKC_α). The experimental data for PKC should be relevant for checking the calculation result on PKA here because both PKC and PKA have a negatively charged residue at position 127 (PKA has an E and PKC has a D).

The polarity of the *p*-hydroxylbenzamide also appears to be important. The binding affinity is reduced by 2.2 kcal/mol when the polarity of the hydroxyl group is turned off. This is consistent with the decrease in PKC inhibition observed when a methyl group replaces this hydroxyl group. The significance of the polarity of this hydroxyl group in recognizing kinases parallels the previous hypothesis that many kinase inhibitors, as ATP, utilize two hydrogen bonds to recognize the linker region between the N-terminal and C-terminal lobes of the catalytic domain.[7] Figure 3 illustrates the formation of these two hydrogen bonds in the cPKA−balanol complex. In the energy-refined structure, the distance

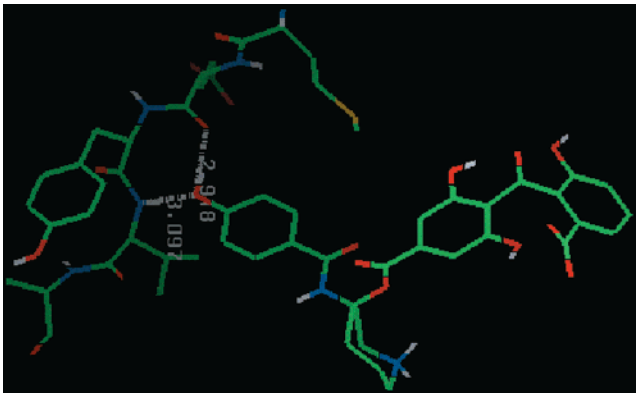

Figure 3. Formation of two hydrogen bonds between the OH group of the *p*-hydroxylbenzamide moiety of balanol and the linker region between the N-terminal and C-terminal lobes of PKA.

from the hydroxyl oxygen to the carbonyl oxygen (hydrogen bond acceptor) of residue 121 is 2.9 Å and the distance to the amide nitrogen (hydrogen bond donor) of residue 123 is 3.1 Å. However, the manner in which balanol forms two hydrogen bonds with the linker region is somewhat different from that of ATP[44−46] and protein kinase inhibitors such as SU5402 (9). In ATP and SU5402, the hydrogen bond donor and acceptor come from two different functional groups rather than from a single one. There are three other hydroxyl groups in balanol. These three groups behave differently. The polarity of one (associated with 2O10) is quite significant for binding, one (associated with O4″) is not very important for binding, and the other (associated with O6″) appears to reduce the binding affinity. This analysis provides some hints on modifying balanol to improve its binding to PKA. For example, replacing the OH group associated with O6″ by a nonpolar group with a similar size may improve binding.

Although the carboxylate group of balanol is charged, its charge does not enhance binding as much as that of the ammonium group. Turning off its charge only reduces the binding affinity by 0.4 kcal/mol. As discussed earlier, this carboxylate group is close to S53, which is replaced by a number of amino acids—nonpolar, positively charged, or negatively charged—in other protein kinases. Suitable replacement of this group may yield balanol derivatives having better selectivity toward desired kinases. The calculation results here is consistent with the experimental finding[22] that replacing the carboxylate group by its methyl ester reduces activity in PKC, which also has a serine in position 53.

The polarity of the amide linkage of balanol does not seem to affect binding significantly. If necessary, this linkage could be replaced by a nonpolar one to improve membrane permeability without sacrificing binding affinity. The nitrogen atom of this amide linkage is relatively far away from the protein. The closest protein heavy atom is 4.3 Å away. The closest four residues are V57, E127, L173, and T183. The atoms of V57 that are closest to the nitrogen are CB (4.8 Å), CG1 (5.0 Å), and CG2 (4.5 Å). For E127, the closest atom is OE2 that is 4.5 Å away. The CD1 and CD2 of L173 are also 4.8 or 4.9 Å from the nitrogen. For T183, OG1 and CG2 are respectively 4.7 and 4.3 Å from the nitrogen. One way to make this part of the inhibitor interact directly with the protein may be to replace the hydrogen of the amide NH group by a larger one such as a methyl group. However, since nonpolar, polar, and negatively charged residues would interact with this group, the effects of the replacement on binding are hard to determine without doing a quantitative calculation or carrying out an experiment.

Changing the polarity of the aromatic rings on the two ends of the inhibitor has different effects on binding. Introducing positive charges of 0.3 e to atoms in the aromatic ring of the *p*-hydroxylbenzamide moiety slightly diminishes binding, whereas doing the same on the aromatic ring on the other end of the molecule enhances binding. This suggests that the protein produces a negative potential in the aromatic ring of the benzophenone moiety. Therefore, replacing the carboxylate of balanol by an electron-withdrawing group may enhance the interaction between this aromatic ring and cPKA.

Generally, this type of sensitivity analysis for determining the key functional groups of an inhibitor can help develop a pharmacophoric model for identifying new drug leads for the target. For example, one may use the spatial arrangement of the key functional groups to construct a three-dimensional pharmacophoric model for searching small molecule libraries for compounds containing these features. Commercial packages, such as the Catalyst,[47] are already available for carrying out such a database search.

Semiempirical Model. Although solving the Poisson equation using atomic partial charges and a detailed description of molecular shape provides a sophisticated description of electrostatic contributions to binding affinity, this approach still relies on a few approximations that may impair its predictive reliability. For example, one typically divides space into an inner solute region of low dielectric surrounded by an outer solvent region of high dielectric. The choice of this boundary can have a significant influence on the calculation result. Atomic partial charges commonly used in molecular mechanics calculations are also imperfect, and electronic polarization effects are largely ignored in these calculations. In addition, entropy changes associated with overall rotation and translation and with internal rotation about single bonds are often ignored or estimated by simplified models. Therefore, empirically scaling different components obtained from an implicit-solvent model may effectively correct some of these deficiencies and improve the predictive power of the model.

Table 5 shows the results obtained by using semiempirical two-parameter, three-parameter, and four-parameter implicit-solvent models. Regression analysis between experimental and predicted results has also been carried out for each model, and a *t*-test of the regression coefficient (Table 5) demonstrates that the correlation between the predicted and experimental results is significant for each model. In the two-parameter model, the total free energy of binding ($\Delta G_{coul} + \Delta G_{PB} + \Delta G_{SA}$) is scaled by a constant b and a second constant a is added to the result. These two constants are estimated by requiring the calculated results from Hünenberger et al.'s work[40] to agree with corresponding experimental data in a least-squares-fit manner, using 17 ligands in Table 1 of their paper (compounds 2 and 3 were excluded from this analysis because only experi-

Table 5: Performance of Several Semiempirical Models in PKA−Balanol Calculations[a]

model	$\Delta G_{predicted}$ (kcal/mol)	rms difference (kcal/mol)	correlation coeff	regression coeff[c]
two-parameter model	$7.8 + 0.72(\Delta G_{coul} + \Delta G_{PB} + \Delta G_{SA})$	1.6 (1.8[b])	0.78	0.62 (0.25[d])
three-parameter model	$6.6 + 0.68\Delta G_{SA} + 0.85(\Delta G_{coul} + \Delta G_{PB})$	1.5 (1.8[b])	0.81	0.66 (0.25[d])
four-parameter model	$2.0 + 0.58\Delta G_{SA} + 0.84\Delta G_{PB} + 0.86\Delta G_{Coul}$	1.1 (1.4[b])	0.90	0.81 (0.21[d])

[a] ΔG_{SA}, ΔG_{Coul}, and ΔG_{PB} are respectively the contribution to the binding energy from the hydrophobic effects (assuming to be proportional to the change in accessible surface area of the molecules upon binding), from assembling charges in a medium having the solute's dielectric constant (estimated from Coulomb's law), and from formation of a dielectric boundary due to hydration (estimated by numerically solving the Poisson equation). The prefactors and the additive constant were determined by minimizing the difference between $\Delta G_{predicted}$ (kcal/mol) and the corresponding experimental data in a least-squares manner. The 17 inhibitors studied in Hünenberger et al.'s work[40] were used in the fit (see text). [b] From leave-one-out cross-validation studies. [c] Regression coefficient (R) in $Y = RX + C$, where Y is an experimental measurement, X is a prediction from the two-parameter, three-parameter, or four-parameter model, and C is a regression constant. [d] 95% confidence limit.

mental bounds were provided for their binding constants). The scaling factor is found to be 0.72, and the additive constant is 7.8 kcal/mol. The root-mean-square difference between calculated and experimental results is 1.6 kcal/mol. When the surface-area-dependent hydrophobic term is separated from the electrostatic contributions in the three-parameter model, the root-mean-square difference is only slightly improved to 1.5 kcal/mol. When the electrostatic contributions are further split into terms arising from Coulombic interactions and from solvation effects, one obtains a more substantial improvement with a root-mean-square difference of 1.1 kcal/mol. To check that this improvement is not simply due to the addition of more parameters to the model to fit the same data, we also carried out leave-one-out cross-validation studies. In a leave-one-out cross-validation study, all except one point are used to determine the parameters of the model and the resulting model is tested in its ability to predict the binding constant of the complex excluded from the fitting. This calculation is repeated excluding each point in turn and the root-mean-square difference between predicted and experimental results computed. As one can see from Table 5, the root-mean-square difference also decreases in going from the three-parameter to the four-parameter model. The correlation coefficient between experimental and predicted results for each model was also calculated. One can see from Table 5 that the results from the four-parameter model correlates substantially better with experimental data than the other two models. Therefore, all these additional tests also illustrate that the four-parameter model performs better than the two- or three-parameter models. A major change in going from the three- to four-parameter model is in splitting the electrostatic contributions into two parts. One part involves bringing the charges together from infinity in a dielectric medium having a dielectric constant of 2, i.e., that of the solute. This part can be calculated easily with Coulomb's law. The other one involves creating a dielectric boundary between the solute and the solvent as a result of solvating the solute. There are extra uncertainties in choosing this boundary and in assuming space to be divided into only two parts having different dielectric constants. Separating this term from the Coulombic one allows it to be scaled differently to best-fit the experimental data. One needs to scale down the two electrostatic terms by ~25% to get more reliable predictions from the models.

To obtain better agreement with experimental results, the contribution from the surface-area-dependent hydrophobic term needs to be reduced by ~30% in the two- and three-parameter models and by ~40% in the four-parameter model. In these calculations, we used the van der Waals radii from CHARMm22[39] to calculate the accessible surface area and a microscopic "surface tension" of 25 cal mol^{-1} Å^{-2}. The present results indicate that a microscopic "surface tension" of 15−18 cal mol^{-1} Å^{-2} would be more appropriate for this application using the CHARMm22[39] van der Waals radii. This value is somewhat smaller than but within the error bars of the "optimal" value of 19 ± 6 cal mol^{-1} Å^{-2} obtained from calculating the cyclohexane-to-water transfer energy of amino acid analogues using a similar implicit-solvent model.[42] By applying this semiempirical approach to more protein−substrate systems, one may be able to conclude whether a smaller microscopic "surface tension" of 15−18 cal mol^{-1} Å^{-2} is generally appropriate for studying the binding between proteins and small-molecule ligands.

Conclusions

As protein kinases become important targets for drug design, the determination of the structure of more kinase−inhibitor complexes will be useful for aiding the design process. However, it takes time to determine many structures. It is therefore useful to draw as much information from a kinase−inhibitor structure as possible. In this paper, we have combined information available from a single kinase−inhibitor structure with that from sequence alignment of a large number of protein kinases to generate new hypotheses to aid the design of specific protein kinase inhibitors. These hypotheses can be further evaluated by experimental means or by more elaborate computations. For example, we found that the ligand-binding pocket is variable among almost 400 kinases if the definition of the protein−ligand interface includes amino acid residues not directly in contact with the ligand. Therefore, if medium or/and long-range interactions are significant for ligand recognition, binding specificity can be achieved to a certain degree by designing inhibitors targeted toward sites, near the protein−ligand interface, that are more variable in amino acid types. This sequence/structure analysis approach was reinforced by computational studies quantifying intermolecular interactions. By perturbing the charge or polarity of different functional groups in the inhibitor, one could examine the significance of these features in affecting protein−inhibitor recognition. This information provides additional hints for drug designers to decide what compounds to make or/and screen for efficacy. The identification of the important features of the inhibitor contribut-

ing to binding can also generate pharmacophoric models for fishing out new drug leads from small-molecule libraries. Complementary to other approaches employing a large number of inhibitors, the present approach can be based on a single inhibitor, provided a crystal/NMR structure of the receptor−inhibitor complex is available. The sensitivity analysis carried out in this work focuses on electrostatic effects and assumes fixed protein and ligand conformation, but this approach can be easily generalized. For example, sensitivity analysis based on molecular dynamics simulations can treat the effects of changing molecular size on binding affinity and can take into account structural relaxation.[29] However, since the latter approach requires lengthy molecular dynamics simulations to be finished before hypotheses can be generated, the approach presented in this paper can serve as a preliminary screening tool before expensive molecular dynamics simulations are completed. To further evaluate computationally designed compounds, we have introduced a semiempirical approach to improve the predictive reliability of implicit-solvent models for calculating binding affinity, using suitable experimental data to scale different terms of these models.

Acknowledgment. This research has been supported in part by the NIH, NSF, and Molecular Simulations Inc. We thank Professor Lawrence Brunton for sending useful reprints on the study of the interaction between balanol and its derivatives on several protein kinases.

References

(1) Ferrara, N.; Henzel, W. J. Pituitary follicular cells secrete a novel heparin-binding growth factor specific for vascular endothelial cells. *Biochem. Biophys. Res. Commun.* **1989**, *161*, 851−858.
(2) Rifkin, D. B.; Moscatelli, D. Recent developments in the cell biology of basic fibroblast growth factor. *J. Cell Biol.* **1989**, *109*, 1−6.
(3) Nicosia, R. F.; Nicosia, S. V.; Smith, M. Vascular endothelial growth factor, platelet-derived growth factor, and insulin-like growth factor-1 promote rat aortic angiogenesis in vitro. *Am. J. Pathol.* **1994**, *145*, 1023−1029.
(4) Suri, C.; McClain, J.; Thurston, G.; McDonald, D. M.; Zhou, H.; et al. Increased vascularization in mice overexpressing angiopoietin-1. *Science* **1998**, *282*, 468−471.
(5) Gleave, M. E.; Hsieh, J. T.; Wu, H. C.; Hong, S. J.; Zhau, H. E.; et al. Epidermal Growth Factor Receptor-Mediated Autocrine and Paracrine Stimulation of Human Transitional Cell Carcinoma. *Cancer Res.* **1993**, *53*, 5300−5307.
(6) Lawrence, D. S.; Niu, J. K. Protein kinase inhibitors: The tyrosine-specific protein kinases. *Pharmacol. Ther.* **1998**, *77*, 81−114.
(7) Traxler, P. M. Protein tyrosine kinase inhibitors in cancer treatment. *Expert Opin. Ther. Pat.* **1997**, *7*, 571−588.
(8) Engh, R. A.; Girod, A.; Kinzel, V.; Huber, R.; Bossemeyer, D. Crystal Structures of Catalytic Subunit of Camp-Dependent Protein Kinase in Complex with Isoquinolinesulfonyl Protein Kinase Inhibitors H7, H8, and H89−Structural Implications for Selectivity. *J. Biol. Chem.* **1996**, *271*, 26157−26164.
(9) Mohammadi, M.; McMahon, G.; Sun, L.; Tang, C.; Hirth, P.; et al. Structures of the tyrosine kinase domain of fibroblast growth factor receptor in complex with inhibitors. *Science* **1997**, *276*, 955−960.
(10) Mohammadi, M.; Froum, S.; Hamby, J. M.; Schroeder, M. C.; Panek, R. L.; et al. Crystal structure of an angiogenesis inhibitor bound to the FGF receptor tyrosine kinase domain. *EMBO J.* **1998**, *17*, 5896−5904.
(11) Narayana, N.; Diller, T. C.; Koide, K.; Bunnage, M. E.; Nicolaou, K. C.; et al. Crystal structure of the potent natural product inhibitor balanol in complex with the catalytic subunit of cAMP-dependent protein kinase. *Biochemistry* **1999**, *38*, 2367−2376.
(12) Onoda, T.; Iinuma, H.; Sasaki, Y.; Hamada, M.; Isshiki, K.; et al. Isolation of a Novel Tyrosine Kinase Inhibitor, Lavendustin-a, from Streptomyces-Griseolavendus. *J. Nat. Prod.-Lloydia* **1989**, *52*, 1252−1257.
(13) Dolle, R. E.; Dunn, J. A.; Bobko, M.; Singh, B.; Kuster, J. E.; et al. 5,7-Dimethoxy-3-(4-Pyridinyl)Quinoline Is a Potent and Selective Inhibitor of Human Vascular Beta-Type Platelet-Derived Growth Factor Receptor Tyrosine Kinase. *J. Med. Chem.* **1994**, *37*, 2627−2629.
(14) Maguire, M. P.; Sheets, K. R.; McVety, K.; Spada, A. P.; Zilberstein, A. A New Series of Pdgf Receptor Tyrosine Kinase Inhibitors−S-Substituted Quinoline Derivatives. *J. Med. Chem.* **1994**, *37*, 2129−2137.
(15) Trinks, U.; Buchdunger, E.; Furet, P.; Kump, W.; Mett, H.; et al. Dianilinophthalimides−Potent and Selective, Atp-Competitive Inhibitors of the Egf-Receptor Protein Tyrosine Kinase. *J. Med. Chem.* **1994**, *37*, 1015−1027.
(16) Buchdunger, E.; Trinks, U.; Mett, H.; Regenass, U.; Muller, M.; et al. 4,5-Dianilinophthalimide−a Protein−Tyrosine Kinase Inhibitor with Selectivity for the Epidermal Growth Factor Receptor Signal Transduction Pathway and Potent in Vivo Antitumor Activity. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2334−2338.
(17) Buchdunger, E.; Mett, H.; Trinks, U.; Regenass, U.; Muller, M.; et al. 4,5-Bis(4-Fluoroanilino)Phthalimide−a Selective Inhibitor of the Epidermal Growth Factor Receptor Signal Transduction Pathway with Potent in Vivo Antitumor Activity. *Clin. Cancer Res.* **1995**, *1*, 813−821.
(18) Buchdunger, E.; Zimmermann, J.; Mett, H.; Meyer, T.; Muller, M.; et al. Selective Inhibition of the Platelet-Derived Growth Factor Signal Transduction Pathway by a Protein−Tyrosine Kinase Inhibitor of the 2-Phenylaminopyrimidine Class. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2558−2562.
(19) Buchdunger, E.; Zimmermann, J.; Mett, H.; Meyer, T.; Muller, M.; et al. Inhibition of the Abl Protein−Tyrosine Kinase in Vitro and in Vivo by a 2-Phenylaminopyrimidine Derivative. *Cancer Res.* **1996**, *56*, 100−104.
(20) Fry, D. W.; Kraker, A. J.; McMichael, A.; Ambroso, L. A.; Nelson, J. M.; et al. A Specific Inhibitor of the Epidermal Growth Factor Receptor Tyrosine Kinase. *Science* **1994**, *265*, 1093−1095.
(21) Rewcastle, G. W.; Denny, W. A.; Bridges, A. J.; Zhou, H. R.; Cody, D. R.; et al. Tyrosine Kinase Inhibitors. 5. Synthesis and Structure−Activity Relationships for 4-[(Phenylmethyl)amino]- and 4-(Phenylamino)quinazolines as Potent Adenosine 5′-Triphosphate Binding Site Inhibitors of the Tyrosine Kinase Domain of the Epidermal Growth Factor Receptor. *J. Med. Chem.* **1995**, *38*, 3482−3487.
(22) Crane, H. M.; Menaldino, D. S.; Jagdmann, G. E.; Darges, J. W.; Buben, J. A. Increasing the Cellular Pkc Inhibitory Activity of Balanol−a Study of Ester Analogues. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2133−2138.
(23) Jagdmann, G. E.; Defauw, J. M.; Lai, Y. S.; Crane, H. M.; Hall, S. E.; et al. Novel Pkc Inhibitory Analogs of Balanol with Replacement of the Ester Functionality. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2015−2020.
(24) Jagdmann, G. E.; Defauw, J. M.; Lampe, J. W.; Darges, J. W.; Kalter, K. Potent and Selective Pkc Inhibitory 5-membered Ring Analogs of Balanol with Replacement of the Carboxamide Moiety. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1759−1764.
(25) Hu, H.; Mendoza, J. S.; Lowden, C. T.; Ballas, L. M.; Janzen, W. P. Synthesis and protein kinase C inhibitory activities of balanol analogues with modification of 4-hydroxybenzamido moiety. *Bioorg. Med. Chem. Lett.* **1997**, *5*, 1873−1882.
(26) Gustafsson, A. B.; Brunton, L. L. Differential and selective inhibition of protein kinase A and protein kinase C in intact cells by balanol congeners. *Mol. Pharmacol.* **1999**, *56*, 377−382.
(27) Setyawan, J.; Koide, K.; Diller, T. C.; Bunnage, M. E.; Taylor, S. S.; et al. Inhibition of protein kinases by balanol: Specificity within the serine/threonine protein kinase subfamily. *Mol. Pharmacol.* **1999**, *56*, 370−376.
(28) Honig, B.; Yang, A. S. Free Energy Balance in Protein Folding. *Adv. Protein Chem.* **1995**, *46*, 27−58.
(29) Wong, C. F.; Thacher, T.; Rabitz, H. Sensitivity analysis in biomolecular simulation. *Rev. Comput. Chem.*; Wiley-VCH: New York, 1998; pp 281−326.
(30) Carlson, H. A.; Jorgensen, W. L. An extended linear response method for determing free energies of hydration. *J. Phys. Chem.* **1995**, *99*, 10667−10673.
(31) McDonald, N. A.; Carlson, H. A.; Jorgensen, W. L. Free energies of solvation in chloroform and water from a linear response approach. *J. Phys. Org. Chem.* **1997**, *10*, 563−576.
(32) Åqvist, J.; Medina, C.; Samuelsson, J.-E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **1994**, *7*, 385−391.
(33) Paulsen, M. D.; Ornstein, R. L. Binding free energy calculations for P450cam-substrate complexes. *Protein Eng.* **1996**, *9*, 567−571.
(34) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.

(35) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; et al. Electrostatics and Diffusion of Molecules in Solution−Simulations with the University of Houston Brownian Dynamics Program. *Comput. Phys. Commun.* **1995**, *91*, 57−95.
(36) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. Electrostatics and Diffusion of Molecules in Solution−Simulations with the University-of-Houston-Brownian Dynamics Program. *Comput. Phys. Commun.* **1991**, *62*, 187−197.
(37) Hunter, T. 1001 protein kinases redux−towards 2000. *Semin. Cell Biol.* **1994**, *5*, 367−376.
(38) Hanks, S.; Quinn, A. M. Protein kinase catalytic domain sequence database: Identification of conserved features of primary structure and classification of family members. *Methods Enzymol.* **1991**, *200*, 38−62.
(39) *CHARMm*, version 22; Molecular Simulations, Inc.: San Diego, CA, 1992.
(40) Hünenberger, P. H.; Helms, V.; Narayana, N.; Taylor, S. S.; McCammon, J. A. Determinants of ligand binding to cAMP-dependent protein kinase. *Biochemistry* **1999**, *38*, 2358−2366.
(41) Honig, B.; Nicholls, A. Classical Electrostatics in Biology and Chemistry. *Science* **1995**, *268*, 1144−1149.
(42) Simonson, T.; Brünger, A. T. Solvation Free Energies Estimated from Macroscopic Continuum theory: An Accuracy Assessment. *J. Phys. Chem.* **1994**, *98*, 4683−4694.
(43) Lai, Y. S.; Stamper, M. Heteroatom Effect in the Pkc Inhibitory Activities of Perhydroazepine Analogs of Balanol. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2147−2150.
(44) Zheng, J. H.; Knighton, D. R.; Xuong, N. H.; Taylor, S. S.; Sowadski, J. M.; et al. Crystal Structures of the Myristylated Catalytic Subunit of Camp-Dependent Protein Kinase Reveal Open and Closed Conformations. *Protein Sci.* **1993**, *2*, 1559−1573.
(45) Bossemeyer, D.; Engh, R. A.; Kinzel, V.; Ponstingl, H.; Huber, R. Phosphotransferase and Substrate Binding Mechanism of the Camp-Dependent Protein Kinase Catalytic Subunit from Porcine Heart as Deduced from the 2.0 Angstrom Structure of the Complex with Mn2+ Adenylyl Imidodiphosphate and Inhibitor Peptide Pki(5−24). *EMBO J.* **1993**, *12*, 849−859.
(46) Hubbard, S. R. Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J.* **1997**, *16*, 5572−5581.
(47) *Catalyst*, version 4.5; Molecular Simulations, Inc.: San Diego, CA, 1999.

JM000443D