

Quantum Mechanical Pairwise Decomposition Analysis of Protein Kinase B Inhibitors: Validating a New Tool for Guiding Drug Design

Xiaohua Zhang,[†] Alan C. Gibbs,[‡] Charles H. Reynolds,^{*,§} Martin B. Peters,^{†,#} and Lance M. Westerhoff^{*,†}

QuantumBio Inc., 200 Innovation Boulevard, State College, Pennsylvania 16803 and Johnson & Johnson Pharmaceutical Research and Development, L.L.C., Welsh and McKean Roads, P.O. Box 776, Spring House, Pennsylvania 19477

Received September 2, 2009

Quantum mechanical semiempirical comparative binding energy analysis calculations have been carried out for a series of protein kinase B (PKB) inhibitors derived from fragment- and structure-based drug design. These protein–ligand complexes were selected because they represent a consistent set of experimental data that includes both crystal structures and affinities. Seven scoring functions were evaluated based on both the PM3 and the AM1 Hamiltonians. The optimal models obtained by partial least-squares analysis of the aligned poses are predictive as measured by a number of standard statistical criteria and by validation with an external data set. An algorithm has been developed that provides residue-based contributions to the overall binding affinity. These residue-based binding contributions can be plotted in heat maps so as to highlight the most important residues for ligand binding. In the case of these PKB inhibitors, the maps show that Met166, Thr97, Gly43, Glu114, Ala116, and Val50, among other residues, play an important role in determining binding affinity. The interaction energy map makes it easy to identify the residues that have the largest absolute effect on ligand binding. The structure–activity relationship (SAR) map highlights residues that are most critical to discriminating between more and less potent ligands. Taken together the interaction energy and the SAR maps provide useful insights into drug design that would be difficult to garner in any other way.

INTRODUCTION

Structure-based drug design (SBDD) and fragment-based drug design (FBDD) play increasingly important roles in drug discovery,¹ as more protein structures become available and as the computational tools for exploiting those structures become more capable. Ultimately, the success or failure of these efforts rests on the ability to accurately compute protein–ligand interaction energies. This is a difficult problem because of the complexity of the molecular structures involved and the very significant challenge of computing energy differences to sufficient accuracy to provide useful predicted binding affinities. There are many approaches to this problem that vary greatly in terms of their accuracy, generality, and efficiency. At one extreme are simulation-based approaches, such as free energy perturbation (FEP).² FEP provides a theoretically rigorous estimate of the free energy change for permuting one ligand into another. In particular, FEP addresses the problems of sufficient sampling and the computation of true free energies.^{3,4} Even so, this approach is limited by the quality of the force field and by other limitations inherent in classical molecular models. At the other extreme are highly empirical scoring functions, such

as are commonly employed in docking and scoring programs.^{5–9} These models are designed to be fast and, therefore, inevitably sacrifice theoretical rigor and accuracy.

In recent years, there has been significant progress in the development of fast quantum mechanical methods for computing protein-size molecular systems.^{10,11} These linear-scaling approaches have made quantum calculations for protein–ligand complexes tractable, and they have provided an important new tool for computing protein–ligand interaction energies. In particular, quantum methods offer the prospect of a much more accurate representation of electronic effects in proteins and ligands.^{12–14} Indeed, previous work has shown that there are significant charge transfer and polarization effects in protein–ligand complexes that are not captured in classical models.¹⁵ In addition, methods have long been available for partitioning quantum energies into pairwise contributions.^{16,17} The pairwise decomposition (PWD) method divides the electrostatic interaction energy into self- and cross-components between atoms. PWD has successfully been applied to the investigation of the effect of binding in a series of fluorine-substituted ligands to human carbonic anhydrase II.¹⁷

A receptor-based QSAR method, comparative binding energy analysis (COMBINE) formalism, was proposed by Ortiz and co-workers.^{18,19} COMBINE obtains descriptors from the intermolecular interactions between the receptor and the ligand, which are calculated by using a pairwise molecular mechanics (MM) potential energy function. Based on the MM descriptors, QSAR models were built by

* Corresponding author. E-mail: lance@quantumbioinc.com.

[†] QuantumBio Inc.

[‡] Johnson & Johnson Pharmaceutical Research and Development.

[§] Ansaris Biopharma, Four Valley Square, 512 East Township Line Rd., Blue Bell, PA 19422.

[#] Molecular Design Group, School of Biochemistry & Immunology, Trinity College Dublin, Dublin 2, Ireland.

multivariate statistical tools, such as partial least-squares (PLS).^{20,21} Semiempirical pairwise decomposition, along with COMBINE, have been integrated into a new approach for computing protein–ligand interaction energies (SE-COMBINE) on a residue-by-residue basis.²² This SE-COMBINE approach offers the potential to provide new mechanistic insight into the factors governing these interactions as well as to improve overall accuracy.

A series of 45 inhibitors (Table 1) for protein kinase B (PKB) were selected to test the SE-COMBINE method.^{23–27} These compounds were chosen for two reasons: First, both structures and affinities are available for many of these ligands. This provides a unique opportunity to compare our computational results to high-quality experimental data for both structure and activity. Second, the ligands can be grouped into structurally related classes, in many cases being the product of a fragment-based design. This simplifies interpretation and validation of individual ligand–residue interactions computed by SE-COMBINE. QM-PWD was used to compute all of the pairwise ligand–residue interactions between the 45 ligands and the protein kinase A (PKA)–PKB chimera. These computed interaction energies were converted to heat map representations using SE-COMBINE. Analysis of the interaction energies, heat maps, and structures show that SE-COMBINE provides a predictive model for affinity, a clear indication of the most critical residues for binding, and a meaningful indication of the many structural trade-offs necessary to design potent ligands for PKB.

METHODS

Ligand Selection. In a series of papers, Astex Pharmaceuticals has reported 16 X-ray structures of PKA–PKB chimera complexes as well as the binding affinity data for more than 60 additional ligands (see PDB ID in Figure 1). We aligned these sequences using ClustalX (Figure 1).²⁸ This alignment shows that the 16 X-ray sequences are fully conserved except at the N-terminus regions that vary in length and at position 47 of one chimera (2UW4). For this paper, residue numbers were assigned according to the sequential number label beneath the aligned sequences in Figure 1. This facilitates comparison of residues across structures. All of the active site residues fall within the conserved range so that the active sites for all 16 X-ray structures are identical, at least with regard to sequence. These 16 well-defined protein–ligand complexes were selected as the training set for model building.

The other ligands with reported affinities were reserved for the prediction set.^{23–27} Of the 76 ligands considered, 30 contain hypervalent sulfurs (e.g., sulfonamides). Since the semiempirical Hamiltonians employed in this study are not suitable for computing hypervalent species, these compounds were removed from consideration. This left 30 ligands (Table 1) for the prediction set and 15 protein–ligand complexes for the training set. A histogram of ligand binding affinities is provided for the training and test data sets in Figure 2. This histogram shows that the $\log(\text{IC}_{50})$ s exhibit a wide range of values (from -9 to -3) and importantly that both the training and test data sets contain members spanning this range of potencies. It should be noted that for this model, the training set is smaller than the prediction set. This size

distribution closely mimics the situation often experienced in drug discovery where one may have access to a small representative sample of experimental structures.

It is important to demonstrate that the affinities are not highly correlated with molecular size. The number of heavy atoms range from 7 to 30 atoms, and the molecular weights are spread over 300 amu. The potency shows a poor correlation (Figure 3) with the number of heavy atoms ($R^2 = 0.1843$) and with the molecular weight ($R^2 = 0.1458$). The ligand efficiencies (i.e., $\text{LE} = p\text{IC}_{50}/\text{HA}$) and fit quality (FQ) scores were also evaluated. The FQ concept was first introduced by Reynolds et al.^{29,30} The FQ is a scaled ligand efficiency that takes ligand size into account. These values are also reported in Figure 3. FQ scores near 1.0 (scores can exceed 1) indicate near optimal ligand binding, while low scores are indicative of suboptimal binding. The values of FQ for the combined training and prediction sets range from 0.4 to 1.0, regardless of size.

Structure Preparation. X-ray Structure Preparation. The 15 X-ray structures, corresponding to the 15 ligands in the training set, were downloaded from the PDB. The PDB identifications are given in Figure 1. The Reduce program³¹ was employed to protonate the X-ray structures using the definitions in the het dictionary provided by the PDB. This het dictionary was edited to account for correct, bidentate protonation of the ligand pyrazoles. The generalized AMBER force field (GAFF)³² was used to represent the ligands, while the standard amino acids were treated using the parm99 force field.³³ Upon protonation, all protons were minimized, while restraining the heavy atoms, using 500 steps of steepest descent followed by 1000 steps of conjugate gradient minimization using the Sander module of AMBER 8.³⁴ Once this initial minimization was complete, all atoms in the complex were subjected to limited minimization with a constraint weight of 10.0 on heavy atoms (500 steps of steepest descent followed by 1000 conjugate gradient steps).

Protein Complex Structure Preparation. Since the prediction set ligands have no X-ray structures available, it was necessary to build models for them. These models were built using an in-house flexible three-dimensional structure alignment program that utilizes the MTK++ library,³⁵ along with a number of additional routines in order to align a ligand to a structurally similar template ligand. The 15 ligand poses extracted from the X-ray structures served as templates for this process. Each test-set ligand was assigned to one of the 15 ligand templates for alignment, according to maximum similarity. After alignment, the ligands were fitted into the active site of the PKA–PKB chimera in the same way as their templates. The templates for each of the model-built complexes were shown in Table 1 by square brackets with PDB identifications of the templates inside. Due to structural variations among ligands, it was expected that some bad contacts would be introduced during model building. These “hot spots” were allowed to relax in the active site using a restrained AMBER minimization of 1500 steps (500 steepest descent followed by 1000 conjugate gradient) followed by a full minimization of all atoms in the system (500 steepest descent followed by 1000 conjugate gradient steps).

SE-COMBINE Calculations. Linear scaling, semiempirical divide and conquer (D&C) single-point calculations¹⁰ were carried out on each PKA–PKB complex using the PM3³⁶ and AM1¹³ Hamiltonians, as implemented in the

Table 1. PKB Inhibitors and Affinities^a

No.	PDB ID	Structure	Class	log(IC ₅₀)	No.	PDB ID	Structure	Class	log(IC ₅₀)	No.	PDB ID	Structure	Class	log(IC ₅₀)			
1	[2VO6]		1	-6.080	12	[2UVZ]		2	-6.397	24	[2VNY]		3	-6.460	35	[2UW3]	
2	[2VO6]		1	-6.035	13	[2UVY]		2	-6.365	25	[2VO0]		3	-8.153	36	[2UW3]	
3	[2VO3]		2	-4.228	14	[2VO0]		2	-8.153	26	[2VNY]		3	-7.601	37	[2UW4]	
4	[2VNY]		2	-6.567	15	[2UVY]		2	-8.044	27	[2VO6]		3	-8.220	38	[2UW4]	
5	[2VNW]		2	-6.536	16	[2VO0]		2	-7.601	28	[2VO0]		3	-8.095	39	[2UW4]	
6	[2VO6]		2	-3.958	17	[2VNY]		2	-7.794	29	[2VO3]		3	-8.300	40	[2UW8]	
7	[2UVY]		2	-5.160	18	[2VNY]		2	-7.147	30	[2JDV]		4	-9.299	41	[2UW5]	
8	[2UVZ]		2	-5.795	19	[2VO0]		2	-8.521	31	[2UVX]		5	-3.999	42	[2UW5]	
9	[2UVZ]		2	-5.494	20	[2UVZ]		2	-7.999	32	[2VNW]		5	-6.054	43	[2UW6]	
10	[2UVY]		2	-5.021	21	[2UVZ]		2	-7.698	33	[2VNY]		5	-7.919	44	[2UW7]	
11	[2VO3]		2	-4.386	22	[2VNW]		3	-6.744	34	[2VNY]		5	-7.601	45	[2UW7]	
					23	[2VNW]		3	-6.112								

^a Class refers to the particular scaffold, or chemotype family, to which each compound belongs. The class numbers define groups of ligands that share common structural features. Although compound **40** does not contain a pyrazole, it is included with class 6 compounds as it was originally synthesized with them. Compounds with Protein Data Bank (PDB) identifications (protein databank accession number) are members of the training set. In column with title of "PDB ID", all 15 X-ray structures give their PDB identifications, while the rest of the 30 model-built complexes are labeled by their templates PDB identifications with square brackets.

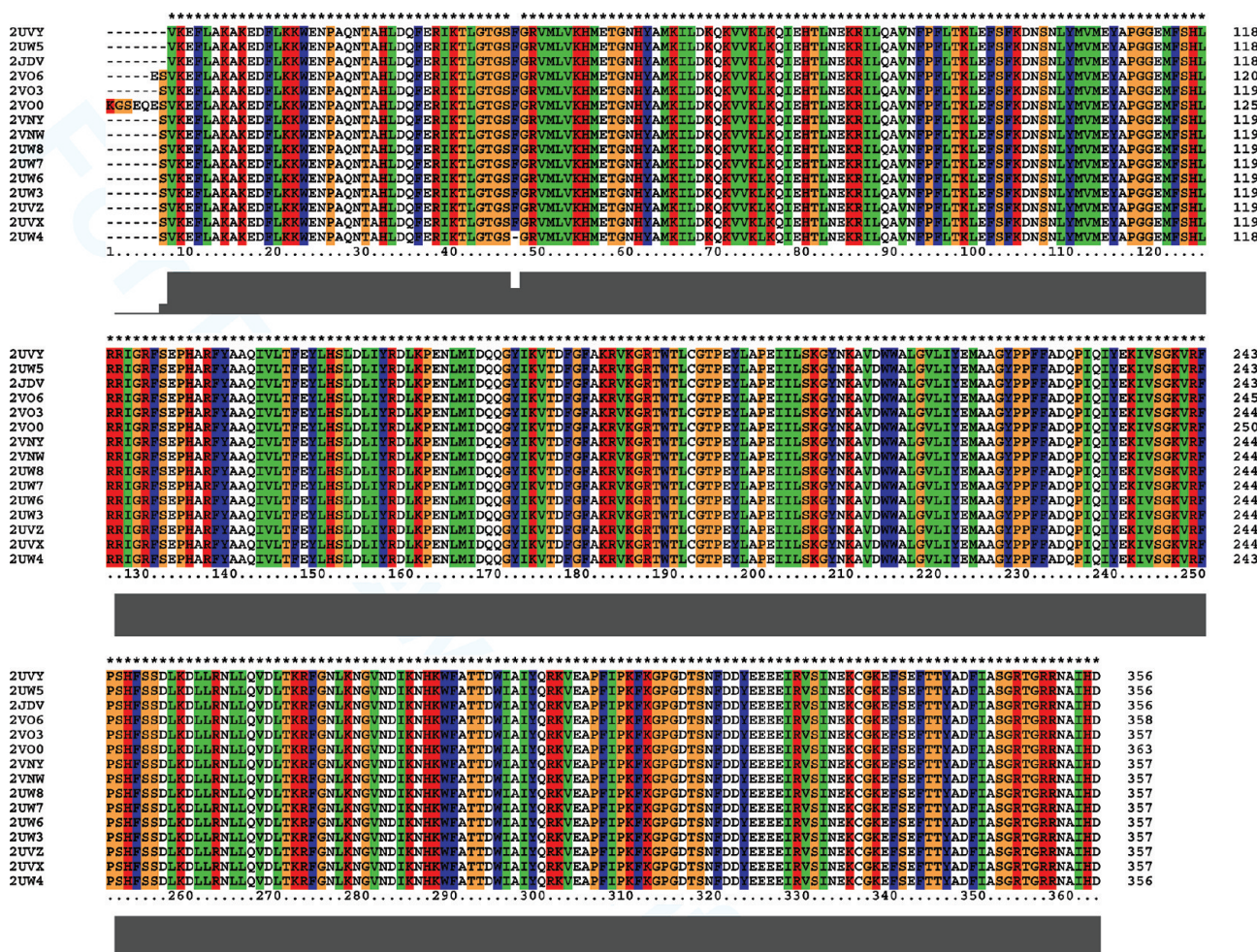


Figure 1. Sequence alignment of the PKA–PKB X-ray structures. Sequences are labeled by the PDB identification. The total number of residues in the sequence are labeled in the last column. Residues are colored according to amino acid type. The histogram (colored in gray) below the ruler indicates the degree of similarity. The * character indicates positions that have been fully conserved.

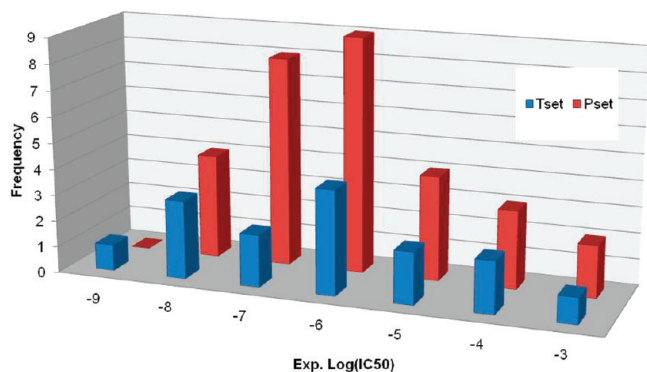


Figure 2. Histogram of inhibitor binding affinity for the training set (blue; 15 ligands from the X-ray structures) and test set (red; 30 ligands).

DivCon 5.x program.³⁷ DivCon calculates the quantum mechanical atom-to-atom, pairwise energy decomposition (QM-PWD)¹⁶ and the overall QMScore¹⁵ for each complex. A CUTOFF for the Fock matrix of 20 Å and a D&C buffering scheme of 4.2/2 Å was employed throughout the simulation. The D&C atom-to-atom pairwise interaction energies were calculated and then converted to residue-based pairwise interaction energies (i.e., pairwise interactions between each protein residue and the complete ligand structure). In addition to the QM calculations, DivCon was used to assign a generalized AMBER force field (GAFF) or

a parm99 atom-type to each atom and to calculate the molecular mechanical atom-by-atom, pairwise energy decomposition (MM-PWD) terms.

Seven scoring functions (from SF-1 to SF-7) using various combinations of the QM-PWD, the QMScore, and the MM-PWD energy terms were evaluated using PLS (Table 2).³⁸ The QM-PWD energy terms consist of E_{AB} , E'_{AB} , and E_{AB}^{core} energy terms.²² E_{AB} contains the exchange between atoms and makes a small negative contribution to the total energy. E'_{AB} represents electron–electron repulsion and contributes positively to the total energy. Finally, E_{AB}^{core} is the core–core repulsion between atoms. The computed molecular mechanics (MM) energy terms include dispersion, van der Waals (summation of dispersion and repulsion terms), and electrostatics. The composition of each scoring function is listed in the Table 2. SF-1 includes only the pairwise energies of the complex in vacuum and is a close approximation to the original version of SE-COMBINE,²² while SF-2, -3, -4, -6, and -7 are extensions of the original SE-COMBINE. SF-5 only contains MM energy terms and is analogous to the conventional implementation of COMBINE.³⁹ SF-2 contains pairwise energies of the complex in solvent. SF-3 and -4 include both QM pairwise energies of complex and dispersion energy, while SF-6 and -7 include both QM pairwise energies of complex and van der Waals energy terms. The

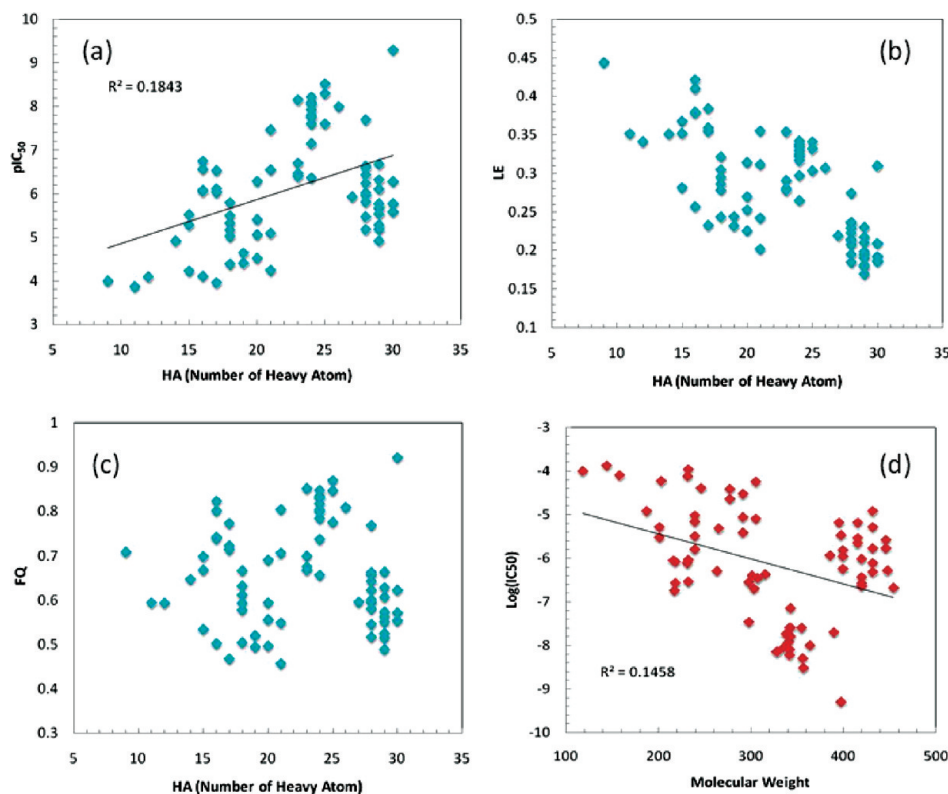


Figure 3. Ligand properties. (a) Plot of pIC_{50} versus the number of heavy atom (HA). (b) The original ligand efficiency as a function of HA. (c) Fit quality scores as a function of HA. (d) $\text{Log}(IC_{50})$ versus molecular weight.

Table 2. Scoring Functions Used in SE-COMBINE^a

scoring function	QM energy terms		MM energy terms		
	QM-PWD in vacuum	QM-PWD in solvent	dispersion	VDW	ele
SF-1	+	-	-	-	-
SF-2	-	+	-	-	-
SF-3	+	-	+	-	-
SF-4	-	+	+	-	-
SF-5	-	-	-	+	+
SF-6	+	-	-	+	-
SF-7	-	+	-	+	-

^a The “+” sign designates that the energy term is included in the scoring function, while the “-” sign indicates the energy is not included.

$\Delta\Delta G_{\text{solv}}$ energy term was originally evaluated via QMScore but was dropped since it was found to contribute little.

Statistical/Chemometric Analysis. The R program was used for all statistical analysis of the energy terms. A descriptor matrix was built where each descriptor contains one column of data from the matrix. Descriptors were pruned initially by removing those with near zero values or with standard deviations less than 0.05. The descriptors were autoscaled (i.e., each value was subtracted by the average and divided by the standard deviation for that descriptor). After scaling, each set of descriptors has a mean of zero and a standard deviation of one. This scaling ensures that certain descriptors do not dominate due to their magnitude.

PLS³⁸ models were built to describe the structure–activity relationship of the inhibitors. Internal validation was carried out using leave-one-out (LOO) cross-validation for the training set, and the optimal dimensionality of each model was assigned from its cross-validated predictive ability. External validations were also carried out with the prediction

set. The models were evaluated for their predictive quality using a range of standard statistical measures, such as the correlation (R^2) and cross-validated correlation (Q^2) coefficients, the standard deviation of error of calculations (SDEC), and the standard deviation of error prediction (SDEP). SDEP was defined as the root-mean-squared error of the dependent variables in a LOO scheme. Similarly, SDEC is calculated for those variables used to build the model or training set. The optimal PLS model was determined by the Q^2 values. The PLS model associated with biggest Q^2 is the optimal model.

Residue-Based Interactions. Two distinctly different heat maps have been constructed. The first map is based on the absolute E_{AB} terms from PWD and is referred to as an interaction map. In this heat map, all of the interaction terms are negative, since the repulsive energy terms are excluded. The second map consists of the most important residue–ligand interaction terms derived from the PLS analysis of all the interaction terms (attractive and repulsive). This is called a structure–activity (SAR) map. These quantities can be either positive or negative. In both maps, the key protein residues, as defined below, are given on the x -axis, and the ligands are given on the y -axis.

The interaction map highlights residues that have favorable interactions with the ligand. This is useful when trying to identify common critical residues across a series of ligands. An example of this type of interaction is the hydrogen-bonding network between residues Glu114 and Ala116 and the heterocycles found in almost all of the ligands examined in this study. The SAR map is scaled based on the variance across the series of ligands and tends to highlight protein residues that are responsible for the changes in potency across a series of ligands. For example, Met166 varies greatly

between the most potent ligand (**30**) and one of the weakest ligands (**40**). Compound **40** also has a very poor (red) interaction relative to any of the other ligands with the hinge residues mentioned above (Glu114 and Ala116). The hinge is located between the N- and C-terminal lobes of the kinase domain.⁴⁰ It is the canonical catalytic site of phosphorylation by ATP. A quick look at **40** explains this result, since it is essentially the only ligand in the series that lacks any hydrogen-bonding partners for these two residues. Ligand **10** on the other hand has a SAR map value of 0.0. This does not mean that there is no favorable interaction; it only means that **10** has an average interaction energy (relative to the series) for residues Glu114 and Ala116.

Although the interaction map is good at characterizing attractive interaction energy terms, it does not take into account repulsive energy terms. An algorithm derived from the PLS method has been developed by QuantumBio to address this. A PLS-based score function was defined as:

$$\text{Res}_l^{(i)} = \sum b_{lk} \left(\frac{x_{lk} - x_k^{\text{mean}}}{x_k^{\text{std}}} \right)$$

where x_{lk} represents a element from the descriptor matrix, x_j^{mean} is the column average, x_j^{std} is the column standard deviation, and b_{ij} is the associated PLS coefficient. Thus the activity, y_i , can be rewritten as:

$$y_i = \sum \text{Res}_l^{(i)}$$

Therefore, the scores $\text{Res}_l^{(i)}$ are exactly the contribution of each residue to the binding affinity. The more negative the value of $\log(\text{IC}_{50})$, the stronger the binding. A negative $\text{Res}_l^{(i)}$ value means the residue contributes favorably to the binding affinity, while a positive value of $\text{Res}_l^{(i)}$ suggests it reduces affinity.

RESULTS AND DISCUSSION

QMScore Results. QMScore calculations were performed with the PM3 Hamiltonian on the modeled (aligned) and X-ray structures. A plot of experimental $\log(\text{IC}_{50})$, as a function of QMScore, is shown in Figure 4. The QMScore values for all 45 ligands range from -833 to -307 kcal/mol. The regression line for the QMScores from the training set is plotted as a blue solid line in the Figure 3. The QMScores achieve good agreement with the experimental $\log(\text{IC}_{50})$ values of the training set, and the R^2 is high at 0.84. The R^2 for the prediction set is also statistically significant at 0.68. Data for the prediction set distribute themselves evenly along the regression line of the training set. These results demonstrate that the PM3 calculations can reproduce the overall trends in affinity even without any complex statistical analysis or descriptor selection. The QMScore is composed of gas-phase heat of formation, electrostatic solvation, attractive Lennard-Jones, solvation entropy, and vibrational entropy energy terms. A multiple linear regression analysis has been performed on these five energy terms in order to “tune up” the QMScore. The R^2 values of the model-fitting QMScore and the experimental $\log(\text{IC}_{50})$ for the training set and all 45 ligands are 0.90 and 0.82, which are improved from that of the directly calculated

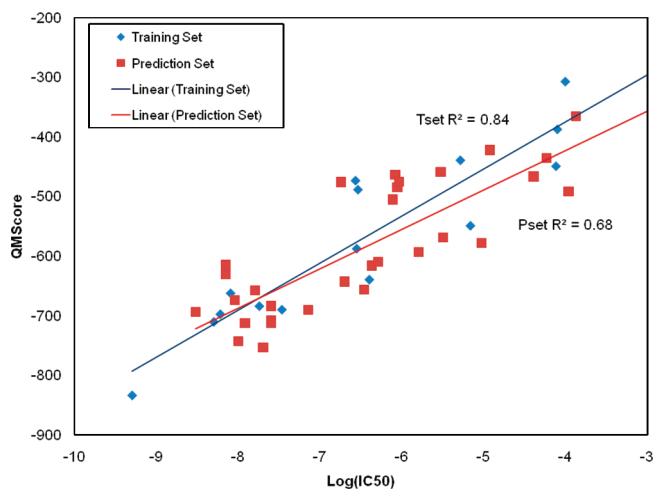


Figure 4. QMScore versus $\log(\text{IC}_{50})$ plot. The QMScores of the training set (X-ray structures) are labeled as blue diamonds, and the prediction set are red squares. Unit for QMScore is in kcal/mol. The regression lines were plotted for the training/prediction sets, and their associated R^2 values are included.

QMScore ($R^2 = 0.84$ for training set and $R^2 = 0.76$ for all 45 ligands).

Data Preparation. The descriptor matrices associated with different scoring functions were obtained from the SE-COMBINE calculations. The descriptor matrices were pruned and scaled. For example, the scoring function SF-1 consisted only of QM-PWD energy terms, E_{AB} , E'_{AB} , and E_{AB}^{core} . There are 355 residues in the descriptor matrix, and each residue contributes three descriptors, E_{AB} , E'_{AB} , and E_{AB}^{core} terms. Thus SF-1, has a total of 1065 descriptors derived from the ligand–receptor interaction energies. As mentioned above, near zero descriptors were eliminated, as were descriptors with a standard deviation of less than 0.05. The number of descriptors was reduced to 732 after pruning, and the remaining descriptors were autoscaled because the E_{AB} , E_{AB}' and E_{AB}^{core} terms span different ranges.

PLS Analysis Results. PLS models were constructed for each of the scoring functions SF-1 through SF-7. In each case, the R^2 values gradually increase with every additional latent variable. The Q_2 behaves differently, reaching a peak at the optimal number of latent variables. Thus, the optimal PLS model is determined by the standard deviation of prediction (SDEP) and by Q_2 . The optimal PLS models are outlined in Table 3. The number of latent variables range from 2 to 6. Values higher than 0.5 for R^2 were considered statistically significant. Values greater than 0.4 for Q^2 were viewed as significant. All of the QM-based PLS models explain 97–100% of the variance, while the pure MM (SF-5) model explains only 79% variance. The QM models have universally higher Q^2 values (as high as 0.81) relative to the MM Q^2 of only 0.33. Likewise, SDEP values from the QM models are significantly lower than that of the MM model. Overall, the pure-QM and mixed-QM/MM optimal models characterize the training set better than the pure-MM model. Evaluating two different semiempirical QM Hamiltonians reveals the training set of PLS results from PM3 are comparable to those of AM1.

As an example, the $\log(\text{IC}_{50})$ values were modeled using SF-1. The plot of computed and experimental affinities for this model are given in Figure 5. The correlation for the training set ($R^2 = 0.92$) and the external test data set ($R^{2\text{ext}}$

Table 3. PLS Results for the Astex Set^a

method	SF	descriptors	LV	X-var. explained (%)	R ²	Q ²	SDEC	SDEP	SDEP ^{ext}	R ^{2ext}
PM3	SF-1	732	2	96	0.92	0.79	0.44	0.67	0.77	0.71
PM3	SF-2	732	2	97	0.92	0.79	0.45	0.75	0.78	0.71
PM3	SF-3	765	2	96	0.93	0.77	0.44	0.78	0.78	0.70
PM3	SF-4	765	2	96	0.93	0.78	0.44	0.77	0.78	0.70
PM3	SF-6	790	5	96	1.00	0.80	0.11	0.73	0.98	0.50
PM3	SF-7	790	5	96	1.00	0.80	0.11	0.73	0.98	0.50
AM1	SF-1	723	2	98	0.92	0.81	0.47	0.72	0.80	0.70
AM1	SF-2	723	2	98	0.92	0.81	0.47	0.72	0.80	0.69
AM1	SF-3	756	2	96	0.92	0.78	0.45	0.77	0.79	0.69
AM1	SF-4	756	2	96	0.92	0.78	0.45	0.77	0.79	0.69
AM1	SF-6	781	6	97	1.00	0.81	0.06	0.72	1.00	0.49
AM1	SF-7	781	6	97	1.00	0.81	0.06	0.72	1.00	0.49
MM	SF-5	337	3	79	0.99	0.42	0.14	1.25	1.15	0.64

^a SF represents different scoring functions. The total number of descriptors in each model is presented. LV designates the number of latent variables in the model. The optimal number of LVs is given. R² and Q² represent the correlation coefficient of the training set and the LOO validation. SDEC and SDEP are the standard deviation of the calculated and predicted values for the internal validation. SDEP^{ext} and R^{2ext} are similar definitions, while applied to the external set.

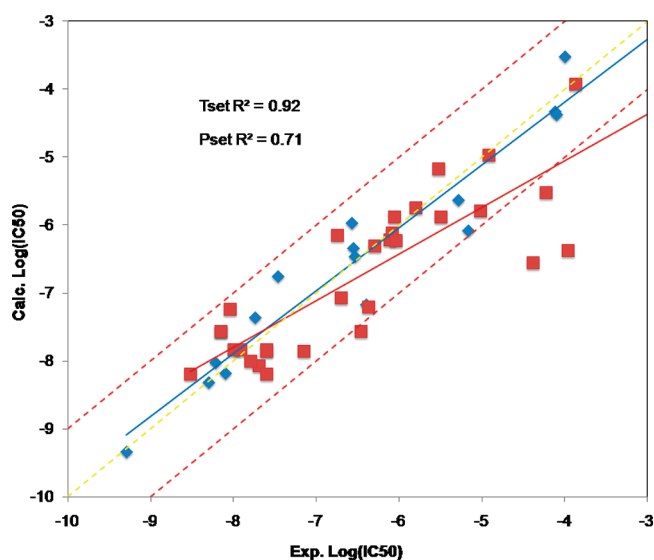


Figure 5. Plot of calculated versus experimental Log(IC₅₀) values from the training (blue) and prediction (red) sets. The blue and red lines are trend lines for the training and prediction sets, respectively and, their corresponding R² values are labeled. Three dashed lines define the regions of the plot: the yellow line is the diagonal line and the two red lines are ± 1 offsets to the diagonal line.

= 0.71) are good. Most of the data points in Figure 5 fall between the two yellow-dashed lines, indicating errors of less than 1 order of magnitude. There are two ligands, **6** and **11**, that lie far off the line. A careful examination of these structures suggests that the large deviations may be partly due to errors in building the protein–ligand complexes. If these two compounds were omitted, then the R² increases from 0.71 to 0.83.

Key Residues. The relative contribution of descriptors to the optimal PLS model can be determined from the PLS coefficients. A PLS coefficient plot for SF-1, using the PM3 Hamiltonian, is shown in Figure 6. These residues are defined as key residues, all of which are in the active site of the receptor. This analysis reveals that the model is dominated by the E_{AB} terms. In particular, the E_{AB} energy terms for Glu114, Ala116, Val50, and other residues highlighted in Figure 6 contribute heavily to the best PLS model. These key active site residues were used to construct the interaction map in Figure 7.

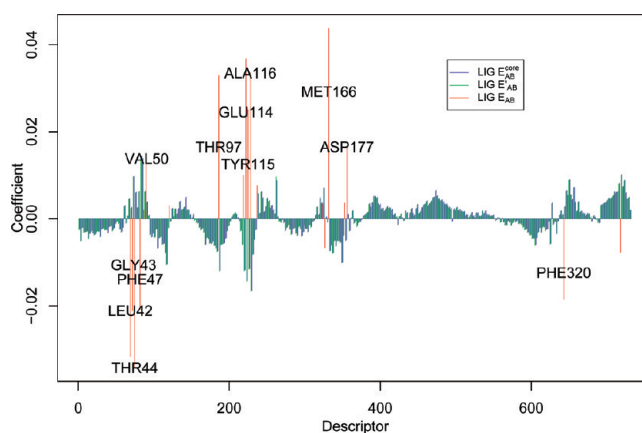


Figure 6. PLS coefficient plot for optimal model (LV = 2) of scoring function SF-1. Key residues are labeled.

Interaction Map. After the determination of key active site residues, an intermolecular interaction map is generated based on the interaction energies (Figure 7). The ligands on the y-axis of Figure 7 are ordered from most to least potent (based on IC₅₀'s) with the most potent at the top of the map. The interaction map in Figure 7 highlights the dominant E_{AB} terms between the ligands and the receptors. The E_{AB} interactions are always attractive, and therefore, must be negative or zero. Of course, the more negative values indicate stronger attractive interactions between the residue and the ligand. The interaction maps allow one to easily identify the most attractive and important residue–ligand interactions. For example, the strong interaction between **30** and Lys65 stands out by virtue of the negative interaction energy and the corresponding dark blue color.

The interaction map in Figure 7 shows two columns of consistently dark-blue data points corresponding to Glu114 and Ala116. These residues have very favorable interactions with all but one ligand. This is presumably an important and highly conserved interaction across the series. Examination of the complex structures shows that there are multiple cooperative hydrogen bonds between these residues and the heterocyclic moieties found in most of the ligands (Figure 8a). Of course, this is a well-known, and important, interaction for kinase inhibitors that target the adenosine triphosphate (ATP) domain. Thus, it is reassuring, and validates the methodology, to see that this critical interaction is starkly

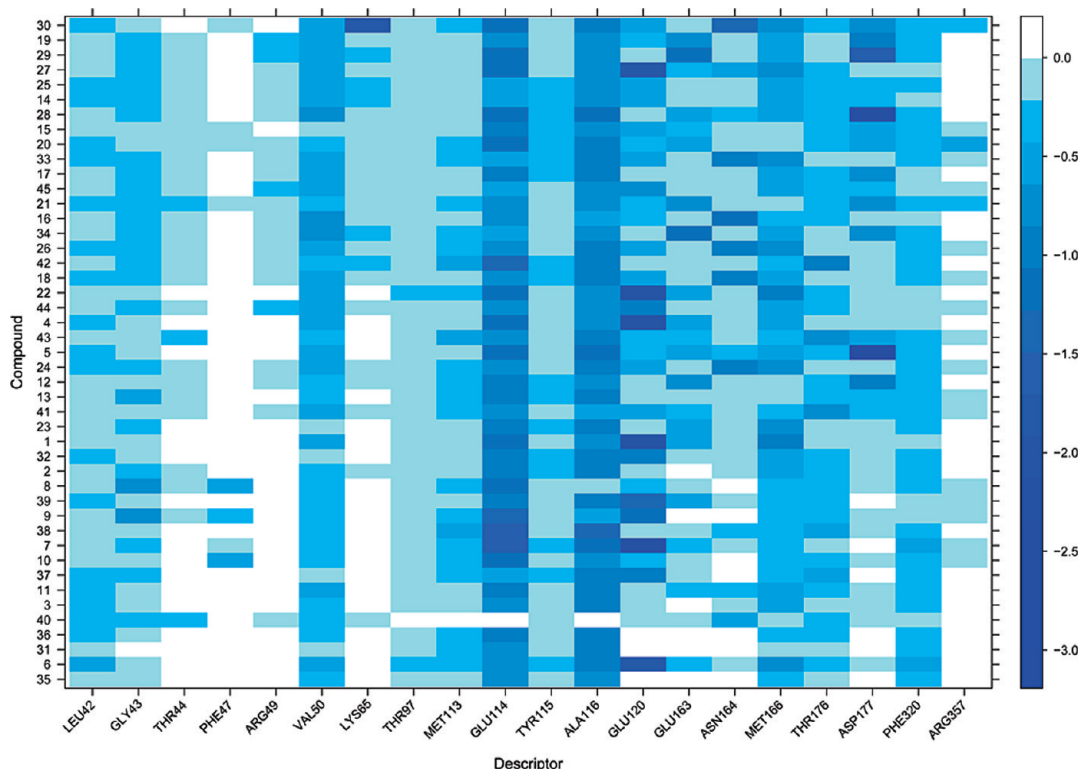


Figure 7. Interaction map for the key residues (x-axis). The compounds on the y-axis are arranged with respect to the activity. The binding affinity decreases from top to bottom. The legend designates the magnitude of the unscaled descriptor in eV.

visible in the interaction map. The exception to this trend in Figure 7 is compound **40** (white cells). This is reasonable since **40**, unlike the other ligands, has no such hydrogen-bonding functionality (Figure 8b). A look at the structure of **40**, and its complex, shows that **40** is not capable of acting as a hydrogen-bond partner to either Glu114 or Ala116. This stands out in the interaction map (white cells) and clearly identifies **40** as being structurally unique relative to the other ligands.

The interaction map shows a favorable interaction between **30** and residues Asp177 and Asn164. This interaction is due to the basic nitrogen in **30** that forms a series of hydrogen bonds with the carboxylate in Asp177 and the backbone carbonyl of Asn164 (Figure 8c). This type of interaction is present in most of the more active ligands but does not appear to be as dominant as might appear, if we were using classical point charges.

Since the semiempirical Hamiltonians utilized in the present study are known to have limitations concerning the treatment of dispersion and other nonbonding interactions⁴¹ relative to higher level QM methods, one might be concerned that only purely electrostatic interactions would be seen in the interaction map. But there are examples of favorable interactions that are consistent across the series and would not be expected to be strictly electrostatic. One example is the negative interaction, likely hydrophobic, between Val50 and the majority of ligands in this study. The E_{AB} values for Val50 range from -0.1 to -0.8 eV across all 45 ligands (Figure 7). As can be seen in (Figure 7), there are 9 nonpolar residues (Leu42, Gly43, Gly45, Phe47, Val50, Met113, Ala116, Met166, and Phe320) in a total of 21 key residues. Most of these nonpolar key residues show consistent negative E_{AB} values for each ligand. While in some cases the interactions with these nonpolar residues may be due to polar

backbone atoms, such as is the case with Ala116, many of these interactions are primarily nonpolar. Taking Val50 as a specific example, three ligand complexes (**31**, **43**, and **28**) are shown in Figure 9 (the binding affinities increase from a to c). As can be seen in Figure 9, the potency increases as the number of nonpolar interactions between the ligand and Val50 increases. Thus, E_{AB} becomes more negative going from **31** to **28** due to these nonpolar interactions.

SAR Map. The PLS-based $Res^{(j)}$ values for SF-1 (PM3 Hamiltonian) are plotted as an SAR map in Figure 10. This SAR map is distinct from the interaction map in Figure 7, in that, the PLS coefficients are not driven by the absolute interaction energies but instead are ranked by their contribution to the variance in potency across the series of ligands. For example, residues Ala116 and Glu114 are consistently dark blue in the interaction map (Figure 7) because these interactions are strongly negative and are present in all but one of the ligands. In the case of the SAR map (Figure 10), the energies, and corresponding colors, range from unfavorable positive interactions (red) to favorable negative interactions (green), reflecting the range of interactions with different ligands.

A residue that illustrates use of the SAR map is Met166. The interaction between Met166 and **30** is highly negative (dark blue), consistent with the high potency of **30**. In contrast, the positive interaction (orange) between Met166 and **40** is likely a contributing factor to the poor affinity of **40**. The molecular structures are consistent with this trend. Met166 has many favorable contacts with **30** (Figure 11a). It is not surprising that these interactions would improve potency. Once again, this demonstrates that the underlying QM calculations are able to reproduce trends in essentially hydrophobic interactions. By comparison, examination of the complex with **40** shows very little contact between Met166

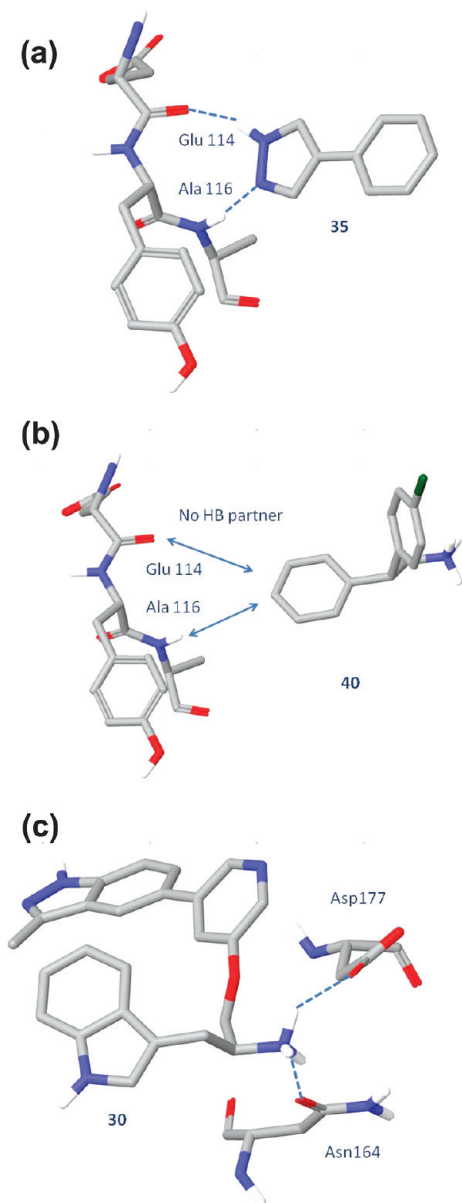


Figure 8. Hydrogen-bond interactions between receptor and ligands: (a) ligand **35** hydrogen bond to Glu114 and Ala116; (b) ligand **40** has no hydrogen bonds with Glu114 and Ala116; and (c) ligand **30** hydrogen bond to Asn 164 and Asp 177.

and the ligand trends in affinity across chemotypes (Figure 11b). Previous research has suggested that π -stacking interaction could be significant between the side chain of methionine and the phenyl group.⁴² However, the SAR map shows that the interaction of phenyl group of **40** and Met166 contributes little to the binding affinity. This is also observed from the interaction map, the E_{AB} value between Me166 and **40** is only -0.017 kcal/mol. Checking the X-ray structure of 2UW8, the shortest distance of heavy atoms between Met166 (sulfur atom) and the phenyl group of **40** (carbon atom from the benzene ring) is 4.41 Å, which is considered to be too large to achieve good electronic overlapping between two atoms. Thus, the E_{AB} value is relatively smaller than other ligands.

The SAR Map in Figure 10 can also be used to identify broad trends in affinity not only within but across chemotypes. For example, while there is a great deal of variation in the heat maps for the most potent compounds on the left

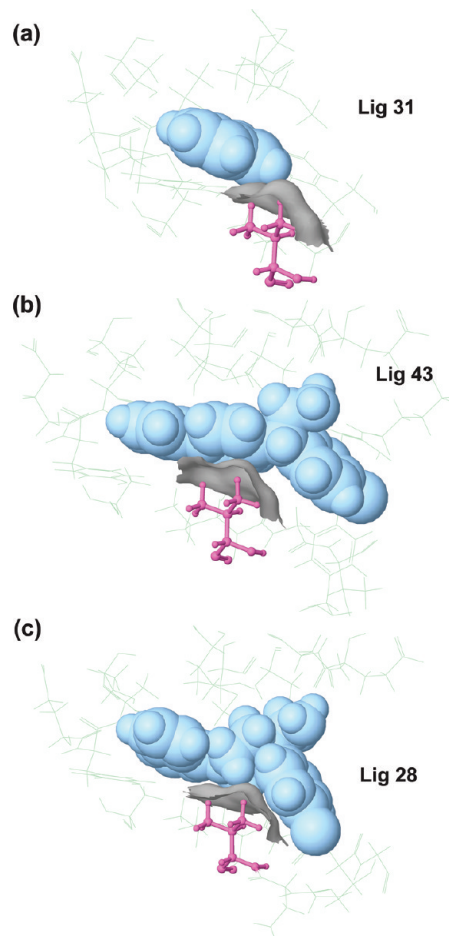


Figure 9. Hydrophobic interactions between key residue Val50 (magenta) and ligands (light blue): (a) ligand **31**; (b) ligand **43**; and (c) ligand **28**.

side of the map (i.e., Met166 through Gly48), most of these potent ligands appear to benefit from favorable interactions with residues Met113–Lys174 and Tyr62–Gln170. Upon closer inspection, it becomes apparent that the increased affinity is due to interactions with ligands that contain two aromatic rings (one being the hinge binding heterocycle) separated by a third ring. These compounds clearly light up green on the right side of the SAR Map, and the trend holds over all chemotypes in the study. One more interesting trend apparent on the left side of the SAR map is the yellow grids (between Met113–Lys174 and Tyr62–Gln170), for compounds with the lowest affinity. These compounds, from multiple chemotypes, are devoid of a basic amine. Thus, compounds containing a primary amine in the correct location are clearly visible in the interaction map as blue grids for Asp177 and/or Glu120. Conversely, the SAR map shows that these same residues are colored yellow, in the case of ligands, that lack this basic amine.

CONCLUSION

We have used 15 crystal structures for a series of PKB inhibitors to train an SE-COMBINE model. This model was then applied to a larger, external test set of 30 ligands. The QM energies and the COMBINE-optimized models were both reasonably predictive for the training and, most importantly, the test set. The goal of this work was not just to build a model of potency but to use the recently developed



Figure 10. PLS-based SAR map. The key residues (*x*-axis) are ordered according to their average scores. The scores decrease from left to right. The compounds on the *y*-axis are arranged with respect to the activity. The binding affinity decreases from top to bottom. Scores higher than 0.5 or lower than -0.5 were plotted as 0.5 and -0.5 , respectively.

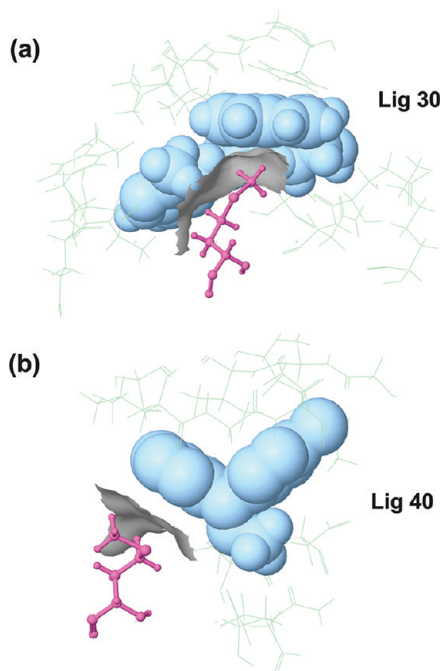


Figure 11. Residue Met166 (magenta) interacting with ligands (light blue): (a) interacting with ligand 30; and (b) having limited contacts with ligand 40.

pairwise decomposition scheme in DivCon to analyze the contributions of individual residues to ligand binding. This was done using two formalisms: First, the absolute E_{AB} values from PWD were used to construct a heat map (interaction map), where residues were coded by their absolute pairwise interaction energies. Second, a SAR map was constructed from the PLS pairwise terms that are scaled

by variance. The first map highlights favorable interaction energies in an absolute sense. An example is the hydrogen-bond network between the ligand heterocyclic cores and the hinge region of the kinase. This region “lights up” in the interaction map with all but one ligand (a ligand that lacks this functionality). The second map (SAR map) highlights residues that contribute the most to discriminating between potent and weak ligands. In this case, the hydrogen-bonding interactions between the ligands and the hinge are less striking because there is less variance across the series of ligands. On the other hand, other interactions that play a major role in the variation of potency across the series are more apparent in the SAR map. Both heat maps provide valuable insights into the role of specific residues in conveying activity. They provide detailed and quantitative information not otherwise available.

ACKNOWLEDGMENT

National Institutes of Health Small Business Innovation Research (SBIR) Grant no. R44GM080115 for financial support for X.Z. and L.W.

Supporting Information Available: 3D structures of all proteins, ligands, and complexes discussed within this publication are provided. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES AND NOTES

- (1) Whittle, P. J.; Blundell, T. L. *Protein Structure-Based Drug Design. Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 349–375.
- (2) Zwanzig, R. W. High-temperature equation of state by a perturbation method 0.1. nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420–1426.

- (3) Debolt, S. E.; Kollman, P. A. Investigation of structure, dynamics, and solvation in 1-octanol and its water-saturated solution - molecular dynamics and free-energy perturbation studies. *J. Am. Chem. Soc.* **1995**, *117*, 5316–5340.
- (4) Best, S. A.; Merz, K. M.; Reynolds, C. H. Free energy perturbation study of octanol/water partition coefficients: Comparison with continuum GB/SA calculations. *J. Phys. Chem. B* **1999**, *103*, 714–726.
- (5) Halperin, I.; Ma, B. Y.; Wolfson, H.; Nussinov, R. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Struct., Funct., Genet.* **2002**, *47*, 409–443.
- (6) Ewing, T. J. A.; Kuntz, I. D. Critical evaluation of search algorithms for automated molecular docking and database screening. *J. Comput. Chem.* **1997**, *18*, 1175–1189.
- (7) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
- (8) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (9) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004**, *47*, 1739–1749.
- (10) Dixon, S. L.; Merz, J. K. M. Semiempirical Molecular Orbital Calculations with Linear System Size Scaling. *J. Chem. Phys.* **1996**, *104*, 6643–6649.
- (11) Dixon, S. L.; Merz, J. K. M. Fast, Accurate Semiempirical Molecular Orbital Calculations for Macromolecules. *J. Chem. Phys.* **1997**, *107*, 879–893.
- (12) Stewart, J. J. P. Optimization of parameters for semiempirical methods I. Method. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (13) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (14) Stewart, J. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (15) Raha, K.; Merz, J. K. M. A Quantum Mechanics Based Scoring Function: Study of Zinc-ion Mediated Ligand Binding. *J. Am. Chem. Soc.* **2004**, *126*, 1020–1021.
- (16) Raha, K.; Merz, K. M. Large Scale Validation of a Quantum Mechanics Based Scoring Function: Predicting the Binding Affinity and the Binding Mode of a Diverse set of Protein-Ligand Complexes. *J. Med. Chem.* **2005**, *48*, 4558–4575.
- (17) Raha, K.; van der Vaart Arjan, J.; Riley, K. E.; Peters, M. B.; Westerhoff, L. M.; Kim, H.; Merz, K. M. Pairwise decomposition of residue interaction energies using semiempirical quantum mechanical methods in studies of protein-ligand interaction. *J. Am. Chem. Soc.* **2005**, *127*, 6583–6594.
- (18) Ortiz, A. R.; Pisabarro, M. T.; Gago, F.; Wade, R. C. Prediction of Drug-Binding Affinities by Comparative Binding-Energy Analysis. *J. Med. Chem.* **1995**, *38*, 2681–2691.
- (19) Wade, R. C.; Ortiz, A. R.; Gago, F. Comparative binding energy analysis. *Perspect. Drug Discovery Des.* **1998**, *9–11*, 19–34.
- (20) Lindgren, F.; Geladi, P.; Wold, S. The Kernel Algorithm for PLS. *J. Chemom.* **1993**, *7*, 45–59.
- (21) Wold, S.; Sjostrom, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130.
- (22) Peters, M. B.; Merz, K. M. Semiempirical Comparative Binding Energy Analysis (SE-COMBINE) of a Series of Trypsin Inhibitors. *J. Chem. Theory Comput.* **2006**, *2*, 383–399.
- (23) Caldwell, J. J.; Davies, T. G.; Doiiald, A. R. R.; McHardy, T.; Rowlands, M. G.; Ahern, G. W.; Hlinter, L. K.; Taylor, K.; Ruddie, R.; Raynaud, F. I.; Verdonk, M.; Workman, P.; Garrett, M. D.; Collins, I. Identification of 4-(4-aminopiperidin-1-yl)-7H-pyrrolo[2,3-d]pyrimidines as selective inhibitors of protein kinase B through fragment elaboration. *J. Med. Chem.* **2008**, *51*, 2147–2157.
- (24) Collins, I.; et al. Structure-based design of isoquinoline-5-sulfonamide inhibitors of protein kinase B. *Bioorg. Med. Chem.* **2006**, *14*, 1255–1273.
- (25) Davies, T. G.; Verdonk, M. L.; Graham, B.; Saalau-Bethell, S.; Hamlett, C. C. F.; McHardy, T.; Collins, I.; Garrett, M. D.; Workman, P.; Woodhead, S. J.; Jhoti, H.; Barford, D. A structural comparison of inhibitor binding to PKB, PKA and PKA-PKB chimera. *J. Mol. Biol.* **2007**, *367*, 882–894.
- (26) Donald, A.; McHardy, T.; Rowlands, M. G.; Hunter, L. J. K.; Davies, T. G.; Berdini, V.; Boyle, R. G.; Ahern, G. W.; Garrett, M. D.; Collins, I. Rapid evolution of 6-phenylpurine inhibitors of protein kinase B through structure-based design. *J. Med. Chem.* **2007**, *50*, 2289–2292.
- (27) Saxty, G.; Woodhead, S. J.; Berdini, V.; Davies, T. G.; Verdonk, M. L.; Wyatt, P. G.; Boyle, R. G.; Barford, D.; Downham, R.; Garrett, M. D.; Carr, R. A. Identification of inhibitors of protein kinase B using fragment-based lead discovery. *J. Med. Chem.* **2007**, *50*, 2293–2296.
- (28) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentini, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. Clustal W and clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948.
- (29) Charles, H.; Reynolds, S. D. B.; Tounge, B. A. The role of molecular size in ligand efficiency. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4258–4261.
- (30) Reynolds, C. H.; Tounge, B. A.; Bembenek, S. D. Ligand binding efficiency: Trends, physical basis, and implications. *J. Med. Chem.* **2008**, *51*, 2432–2438.
- (31) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **1999**, *285*, 1735–1747.
- (32) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (33) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
- (34) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
- (35) Peters, M. B.; Williams, D. E.; Wollacott, A. M.; Merz, J. K. M. MTK++, version 0.1.4.; 2007.
- (36) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. PDDG/PM3 and PDDG/MNDO: Improved semiempirical methods. *J. Comput. Chem.* **2002**, *23*, 1601–1622.
- (37) *DivCon Discovery Suite*, version 5.x; QuantumBio, Inc.: State College, PA, 2009.
- (38) Geladi, P. K. B. Partial least square regression: A tutorial. *Anal. Chim. Acta* **1986**, *35*, 1–17.
- (39) 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.
- (40) Liao, J. J.-L. Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors. *J. Med. Chem.* **2007**, *50*, 409–424.
- (41) Resac, J.; Fanfrlik, J.; Salahub, D.; Hobza, P. Semiempirical Quantum Chemical PM6 Method Augmented by Dispersion and H-Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes. *J. Chem. Theory Comput.* **2009**, *5*, 1749–1760.
- (42) Tatko, C. D.; Waters, M. L. Investigation of the nature of the methionine-pi-interaction in beta-hairpin peptide model systems. *Protein Sci.* **2004**, *13*, 2515–2522.

CI9003333