Charge Optimization of the Interface between Protein Kinases and Their Ligands

PETER A. SIMS,[1] CHUNG F. WONG,[2,3,*] J. ANDREW McCAMMON[1,2,3]
[1]*Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0365*
[2]*Howard Hughes Medical Institute, University of California, San Diego, La Jolla, California 92093-0365*
[3]*Department of Pharmacology, University of California, San Diego, La Jolla, California 92093-0365*

Received 2 February 2004; Accepted 19 April 2004
DOI 10.1002/jcc.20067
Published online in Wiley InterScience (www.interscience.wiley.com).

Abstract: Examining the potential for electrostatic complementarity between a ligand and a receptor is a useful technique for rational drug design, and can demonstrate how a system prioritizes interactions when allowed to optimize its charge distribution. In this computational study, we implemented the previously developed, continuum solvent-based charge optimization theory with a simple, quadratic programming algorithm and the UHBD Poisson–Boltzmann solver. This method allows one to compute the best set of point charges for a ligand or ligand region based on the ligand and receptor shape, and the receptor partial charges, by optimizing the binding free energy obtained from a continuum-solvent model. We applied charge optimization to a fragment of the heat-stable protein kinase inhibitor (PKI) of protein kinase A (PKA), to three flavopiridol inhibitors of CDK2, and to cyclin A which interacts with CDK2 to regulate the cell cycle. We found that a combination of global (involving every charge) and local (involving only charges in a local region) optimization can give useful hints for designing better inhibitors. Although some parts of an inhibitor may already contribute significantly to binding, we found that they could still be the most important targets for modifications to obtain stronger binders. In studying the binding of flavopiridol inhibitors to CDK2, comparable binding affinity could be obtained regardless of whether the net charges of the inhibitors were constrained to −2, −1, 0, 1, or 2 during the optimization. This provides flexibility in inhibitor design when a certain net charge of the inhibitor is desired in addition to strong binding affinity. For the study of the PKA–PKI and CDK2–cyclin A interfaces, we identified residues whose charge distributions are already close to optimal and those whose charge distributions could be refined to further improve binding.

 J Comput Chem 25: 1416–1429, 2004

Key words: charge optimization; sensitivity analysis; inhibitor design; Poisson–Boltzmann solver; binding free energy

Introduction

Optimization of the electrostatic complementarity between a ligand and a receptor by balancing their interaction energy with the desolvation penalty is a problem that has been studied extensively by Bruce Tidor's group at MIT.[1–8] The theory is well established, and has been applied to a few biological systems including the barstar–barnase interface[6,7] and chorismate mutase.[8] The theory allows one to compute the electrostatically optimal charge distribution for a given ligand with respect to charges on a given receptor, the ligand's own spatial parameters (position and shape), and the spatial parameters of the receptor within a continuum-solvent electrostatic model.[1,3,4]

Our interests involve utilizing results from charge optimization in molecular design. Previously, we used sensitivity analysis as a way to guide drug design. Therefore, we first examined which charges were already useful for binding and kept those portions in a lead optimization process. We then focused on modifying other portions of the ligand to search for new inhibitors with improved

Correspondence to: C. F. Wong; e-mail: C4wong@ucsd.edu

Contract/grant sponsors: NIH, NSF, Howard Hughes Medical Institute, Accelrys, Inc., National Biomedical Computational Resource, Center for Theoretical Biological Physics at UCSD, and the W. M. Keck Foundation.

binding affinity. Another way to utilize sensitivity analysis is to find out which regions can provide significant improvement in binding affinity if their charge distributions are modified according to the analysis. We have also utilized mathematical tricks to speed up such an analysis.[9,10] However, such analysis does not usually lead to the charge distribution that gives the maximum improvement in binding affinity. This can be achieved if one is willing to invest additional computational effort such as using the charge optimization method introduced by Tidor's group. In this study, we try to learn how to better utilize results from charge optimization to design compounds that can actually be made. A charge optimization only rearranges the charge distribution of an existing inhibitor. Finding a compound that fits this optimal charge distribution and can actually be chemically synthesized is not always straightforward. Here, we try to gain some insights by carrying out charge optimization on two systems that we studied before: the interactions between protein kinase A and the residue 5–24 fragment of its inhibitor PKI (PKI-(5–24)), and the interactions between cyclin-dependent kinase 2 (CDK2) and the flavopiridol class of inhibitors. In addition, we include a larger system in this study: the interface between CDK2 and cyclin A.

The PKA–PKI System

The PKA–PKI complex has been studied with extensive mutagenesis experiments,[11–19] crystallography,[20–22] and computational sensitivity analysis.[23] The system is of interest for a variety of reasons, including its function as a useful model for studying protein kinase–substrate recognition, and for the design of small molecule inhibitors of protein kinases.[20,23] Numerous derivatives and fragments of PKI-(5–24) have been synthesized and evaluated for binding affinity, and the determinants of ligand binding are well understood.[17] Likewise, crystallographic data has revealed the structural properties of PKA that promote recognition.[20] A recent study by Gould and Wong applied computational sensitivity analysis to PKI-(5–24) to propose a small-molecule scaffold fragment of the peptide for inhibitor design.[23]

Because of the volume of biological data and previous computational experience with this system, we found the complex to be an ideal subject for charge optimization experiments. The peptide inhibitor is quite large (306 atoms), and so we focused on fast methods for charge optimization and efficient, useful ways to express the resulting data. Additionally, PKI-(5–24) is a very potent inhibitor of PKA (K_i = 2.3 nM),[17] and substantially improving its binding affinity may be quite challenging. Identifying electrostatically vulnerable regions of the peptide could prove particularly valuable to an efficient design effort. One of our primary goals was to use charge optimization to quickly target residues, functional groups, and even atoms on the large inhibitor where modifications could be most profitable. We examine the use of two optimization procedures: (1) global optimization in which the charge of every atom is allowed to vary during the charge optimization, and (2) local optimization in which only the charges in a local region are allowed to vary. We try to learn how these optimization results can help to suggest chemical modifications that can improve binding.

The CDK2–Cyclin A System

Cyclin-dependent kinase 2 (CDK2) and the CDKs in general are popular targets for antiproliferative drug design because of their vital role in cell cycle regulation.[9,24] As their name suggests, another family of proteins known as the cyclins are required for CDK activation. In this study, we examine the protein–protein interface that forms between CDK2 and cyclin A during the S-phase (DNA replication) of the cell cycle.[9,24]

The crystal structure used for this part of the study includes CDK2 with ATP in its active site along with a fragment of cyclin A containing residues 173–432.[24] This CDK2–cyclin A fragment complex has nearly the same histone phosphorylation activity as the full CDK2–cyclin A complex.[24] We will present preliminary results from a fixed conformation model of cyclin A (which is well justified due to its relatively inflexible interfacial region) from an ongoing and thorough study of both the unphosphorylated and phosphorylated systems. The interfacial region (547 atoms) that we consider on cyclin A includes the 32 amino acids whose atoms are closest to CDK2 in the unphosphorylated complex crystal structure. In one optimization, we considered all interfacial atoms simultaneously. Additionally, we optimized each residue in the interface individually with respect to the remaining system (which was held constant at its original CHARMM27 parameters).[25] We hoped to glean information on how these results can guide inhibitor design.

Methods

Binding Free-Energy Model

In all instances, binding free-energy changes and electrostatic potentials arising from the Poisson Equation were computed with UHBD.[26,27] Coulombic contributions were computed with a separate script that accounted for all nonbonded interactions. In most cases, partial charges were taken from the CHARMM27 all-atom force field[28] (if they were not available, they were obtained by performing quantum chemical calculations as discussed in the Molecular Modeling section), and in all cases, van der Waals radii were taken from this same force field. The Poisson equation was solved using a 240-Å^3 grid with 0.4-Å grid spacing for the PKA–PKI system and the CDK2–Cyclin A system. However, for systems involving CDK2 bound to deschloroflavopiridol or other model structures, a 175 × 250 × 215 Å^3 grid was used with 0.3-Å grid spacing.[9] The internal or solute dielectric was set at 2, while the external or solvent dielectric was set at 78. In computing the binding free-energy changes, we first considered three systems separately with respect to their absolute free energies—the complex (bin), the ligand (lig), and the receptor (rec). Our model dictates that:

$$G_i = G_{\mathrm{COUL},i} + G_{P,i} + G_{\mathrm{SA},i} \quad (1)$$

where G_i is the free energy of system i with contributions from Coulomb's Law (G_{COUL}), the reaction field from solution of the Poisson equation (G_P), and a hydrophobic term (G_{SA}), which we assume to be proportional to the solvent accessible surface area.

We computed the hydrophobic term according to $G_{SA} = \gamma(\text{SASA})$, where SASA is the solvent-accessible surface area of the system and γ is a surface tension. We let $\gamma = 25$ cal/mol · A^2 as in our previous work.[9] After computing these terms, one can approximate the binding free energy change by:

$$\Delta_{bind}G = G_{bin} - (G_{rec} + G_{lig}) \quad (2)$$

Our model of binding affinity uses a fixed conformation approximation in which it is assumed that the ligand and receptor have conformations identical to those observed in the complex crystal structure.[9,23,29]

Charge Optimization Theory

We now briefly review portions of the charge optimization theory developed by Tidor et al. that are relevant to our study.[1,3,30] In the previous section, we presented a specific treatment of the following, more general model for binding free-energy changes:

$$\Delta_{bind}G = \Delta_{elect}G + \Delta_{np}G \quad (3)$$

The binding free-energy change is simply divided into an electrostatic contribution that is a function of charge, and a nonpolar contribution that is not a function of charge but only of spatial parameters. One can write the electrostatic contribution to the binding free-energy change as the sum of separable Coulombic and reaction field terms as we demonstrated in the previous section. Alternatively, one can express the electrostatic term as the sum of desolvation penalty and interaction energy terms using matrix–vector notation:[1,3,30]

$$\Delta_{elect}G = \frac{1}{2}\mathbf{Q}_\mathbf{L}^T L \mathbf{Q}_\mathbf{L} + \frac{1}{2}\mathbf{Q}_\mathbf{R}^T R \mathbf{Q}_\mathbf{R} + \mathbf{Q}_\mathbf{R}^T C \mathbf{Q}_\mathbf{L} \quad (4)$$

The vectors $\mathbf{Q_L}$ and $\mathbf{Q_R}$ contain the ligand and receptor charges, respectively, while the symmetric matrices $\mathbf{L}$ and $\mathbf{R}$ are independent of the charges.[1,3,30] The matrix C, which again is independent of the charges, help to calculate the electrostatic interactions between the ligand and the receptor.[1,3,30] The elements of these matrices are defined as follows:

$$L_{ij} = \phi_i^{bound}(r_j^L) - \phi_i^{unbound}(r_j^L)$$

$$R_{ij} = \phi_i^{bound}(r_j^R) - \phi_i^{unbound}(r_j^R)$$

$$(C^T Q_R)_i = \sum_{j=1}^{m} q_j^R \phi_i^{bound}(r_j^L) \quad (5)$$

where m is the number of receptor atoms. bound and unbound label the bound and unbound states, respectively. L and R label the ligand and receptor, respectively. Thus, $\phi_i^{bound}(r_j^L)$ is the electrostatic potential on atom j of the ligand located at r_j^L when a charge of $+1$ is put at atom i and when the ligand is bound to the receptor.[1,3,30] The binding affinity can be minimized with respect to the ligand charge distribution (in this case with point charges as a basis) by taking the gradient of eq. (4):[1,3,30]

$$\nabla(\Delta_{elect}G) = L\mathbf{Q_L} + C^T\mathbf{Q_R} = \mathbf{0}$$

$$L\mathbf{Q}_{\mathbf{L},opt} = -C^T\mathbf{Q_R} \quad (6)$$

Equation (6) can be solved directly by LU decomposition, but the addition of constraints to the optimal charge distribution requires quadratic programming. Constraints are usually added because there is no guarantee that the optimized charges will sum to an integral molecular net charge or that they will have physically reasonable magnitudes if results are taken directly from eq. (6).

A region of a molecule can also be charge optimized in a similar manner. For example, one could optimize a functional group on a ligand while holding the remaining atoms at their initial charges. Here, we make a distinction between the total ligand (L) charges, the fragment of the ligand to be optimized (Lfrag), and the remaining fragment of the ligand (Lrem) which we hold constant. We define the matrices somewhat differently:[7]

$$L_{ij} = \phi_i^{bound}(\mathbf{r}_j^{L\text{frag}}) - \phi_i^{unbound}(\mathbf{r}_j^{L\text{frag}})$$

$$(C^T\mathbf{Q}_{\mathbf{R}+L\text{rem}})_i = \sum_{j=1}^{m} q_j^R \phi_j^{bound}(\mathbf{r}_i^{L\text{frag}}) + \sum_{j=1}^{m} (\phi_j^{bound}(\mathbf{r}_i^{L\text{frag}}) - \phi_j^{unbound}(\mathbf{r}_i^{L\text{frag}}))q_j^{L\text{rem}} \quad (7)$$

where m is still the number of receptor atoms and n is the number of ligand atoms held constant (the number of Lrem atoms). If the potential matrices are defined as in eq. (7), they can be substituted into eq. (6) just as in the global optimization case to obtain local optimization results that we will demonstrate to be indispensable for addressing large systems.

Frank–Wolfe and Interior Point Algorithms

To perform constrained charge optimization, we treated the minimization problem with quadratic programming. The Frank–Wolfe Method is a sequential linear approximation algorithm in which a quadratic function is linearized and the linearization is minimized along with some further adjustment at each iteration until the current solution is negligibly different from the previous solution.[31] The minimization of the linearized function can be accomplished by the revised simplex method or the faster interior point method. The Frank–Wolfe algorithm can be summarized (in terms of the constrained charge optimization problem) as follows where (i) is the iteration number:[31]

1. Initialize the algorithm by generating a ligand charge distribution $\mathbf{Q}_\mathbf{L}^{(i)} = \mathbf{Q}_\mathbf{L}^{(0)}$ that satisfies the desired constraints.
2. Compute the gradient of $-\Delta_{elect}G(\mathbf{Q}_\mathbf{L}^{(i)})$.
3. Let $F(\mathbf{Q}_\mathbf{L}^{(i)}) = \mathbf{grad}(-\Delta_{elect}G(\mathbf{Q}_\mathbf{L}^{(i)}))^T\mathbf{Q_L}$. Maximize F using the interior point method (to be described) such that the ligand charge magnitudes are less than or equal to $0.85e$, and such that the charges sum to a specified integral molecular net charge. The resulting vector shall be called $\mathbf{Q}_\mathbf{L}^{\text{lin}(i+1)}$. Note that we limited the magnitude of an atomic partial charge to be

$0.85e$. Otherwise, very large unphysical charges can result from an unconstrained charge optimization.

4. Parameterize $-\Delta_{\text{elect}}G$ such that $P(t) = -\Delta_{\text{elect}}G(\mathbf{Q_L})$ where $\mathbf{Q_L} = \mathbf{Q_L}^{(i)} + t(\mathbf{Q_L}^{\text{lin}(i+1)} - \mathbf{Q_L}^{(i)})$ and $t\varepsilon[0{:}1]$. Maximize $P(t)$ and set the resulting $\mathbf{Q_L}$ to $\mathbf{Q_L}^{(i+1)}$.
5. If the length of $\mathbf{Q_L}^{(i+1)} - \mathbf{Q_L}^{(i)}$ is small enough (in this work, 10^{-5} for big systems or 10^{-6} for small systems) stop. If the length is not small enough, return to step 2 where $i = i + 1$.

A version of the interior point algorithm for linear programming was chosen, and can be initialized with any feasible starting vector (one that satisfies the constraints but is not necessarily optimal or close to optimal). The method produces solutions by iteratively moving through the feasible region at the maximum rate (in the direction that is orthogonal to the function being minimized).[31] By computing the gradient of the function being minimized projected onto the feasible region (the projected gradient), the algorithm identifies the best direction for minimization in the current iteration.[31] The interior point algorithm operates efficiently on linear programming problems that are presented in "augmented form."[31] In this form, one expresses the problem as a maximization subject only to equality constraints with the exception of the default constraint, which requires that all elements of the solution vector be greater than or equal to zero.[31] In this case, we want to minimize the linearization of the electrostatic contribution to binding affinity where all of the optimized charges (as multiples of e) have magnitudes less than or equal to 0.85 and sum to a specified integer:

$$\text{Maximize } (-L\mathbf{Q_L}^{(i)} - C^T\mathbf{Q_R})^T \cdot \mathbf{Q_L}$$

$$\text{where } |q_{L,i}| \leq 0.85$$

$$\sum_{i=1}^{n} q_i = c \text{ for } c\varepsilon\{\ldots, -1, 0, 1, \ldots\} \quad (8)$$

This can be written in augmented form as follows:

$$\text{Define } q'_{L,i} = q_{L,i} + 0.85$$

$$\text{Maximize } (-L\mathbf{Q'_L}^{(i)} + L(0.85, \ldots, 0.85)^T - C^T\mathbf{Q_R})^T \cdot \mathbf{Q'_L}$$

$$\text{where } q'_{L,i} + \lambda_i = 1.7; \quad q'_{L,i} \geq 0;$$

$$\sum_{i=1}^{n} q'_{L,i} = c + n(0.85) \text{ for } c = \text{formal charge},$$

$$n = \text{number of ligand atoms.} \quad (9)$$

The variable λ is a slack variable that changes along with the q' variables during optimization. Slack variables are used in optimization problems to convert inequality constraints into equality constraints.[32] For example, constraining q' to be less than or equal to 1.7 can also be expressed as $q' + \lambda = 1.7$, where we allow λ to change with q'. These methods are well suited to large quadratic programming problems with many constraints, and give fast initial rates of convergence.[31] The above algorithm was coded in a variety of versions for different purposes (e.g., optimization of a ligand charge distribution, optimization of a protein–protein interfacial region, and optimzation of a single side chain or group of atoms).

Coupling Sensitivity Analysis with Charge Optimization

To be effective in using charge optimization data for molecular design, one needs to go beyond simply examining deviations between original and optimized charge distributions. For example, the original charge distribution may be far away from the optimized charge distribution but replacing the original charge distribution with the optimized one may not necessarily improve binding significantly. In the other extreme, an optimized charge distribution may be only slightly different from the original one but may drastically improve binding. It is useful to estimate the improvement in binding affinity that comes with the optimized charge distribution. Here, we discuss this in light of our earlier sensitivity analysis approach.[9,23,29]

In our earlier work, we calculated the change in binding affinity $\Delta\Delta_{\text{bind}}G$ when a charge (or a set of atomic charges in the general case) is changed from q_i^{init} to q_i^{fin} by using a Taylor series expansion:[9]

$$\Delta\Delta_{\text{bind}}G \approx \left(\frac{\partial\Delta_{\text{bind}}G}{\partial q_i}\right)_{q_i=q_i^{\text{init}}}(q_i^{\text{fin}} - q_i^{\text{init}}) + \frac{1}{2}\left(\frac{\partial^2\Delta_{\text{bind}}G}{\partial q_i^2}\right)_{q_i=q_i^{\text{init}}}(q_i^{\text{fin}} - q_i^{\text{init}})^2 \quad (10)$$

In the fixed-conformation continuum-solvent model used here, this second-order approximation in eq. (10) becomes exact and allows us to calculate $\Delta\Delta_{\text{bind}}G$ defined by:

$$\Delta\Delta_{\text{bind}}G = \Delta_{\text{bind}}G(q_{L,j} = q_{\text{lig},j}) - \Delta_{\text{bind}}G(q_{L,j,j\neq i} = q_{\text{lig},j}, q_{L,i} = q_{L,i}^{\text{con}}) \quad (11)$$

Equation (11) expresses the original charge of the ligand as q_{lig} and the constrained, optimized charge on a given ligand atom i as q_i^{con}. This $\Delta\Delta_{\text{bind}}G$ measures the sensitivity of $\Delta_{\text{bind}}G$ when the charge of interest goes from the optimized to the original charge. Here, we recast the sensitivity in a different form in terms of the quantities obtained from solving the Poisson–Boltzmann equation. First, we rewrite eq. (4) as

$$\Delta_{\text{elect}}G = \frac{1}{2}\sum_{i=1}^{n} q_{L,i}\sum_{j=1}^{n} L_{ij}q_{L,j} + \frac{1}{2}\sum_{i=1}^{m} q_{R,i}\sum_{j=1}^{m} R_{ij}q_{R,j} + \sum_{i=1}^{m} q_{R,i}\sum_{j=1}^{n} C_{ij}q_{L,j} \quad (12)$$

In eq. (12), n is the number of ligand (L) atoms and m is the number of receptor (R) atoms. The derivatives can now be evaluated as follows:

$$\left(\frac{\partial \Delta_{\text{bind}}G}{\partial q_{L,k}}\right)_{q_{L,k}=q_{L,k}^{\text{con}}} = \sum_{j=1,j\neq k}^{n}(L_{kj}q_{L,j}) + \sum_{j=1}^{m}(C_{kj}^{T}q_{R,j}) + L_{kk}q_{L,k}^{\text{con}}$$

$$\left(\frac{\partial^2 \Delta_{\text{bind}}G}{\partial q_{L,k}^2}\right)_{q_{L,k}=q_{L,k}^{\text{con}}} = L_{kk} \quad (13)$$

We can then substitute eq. (13) into eq. (10) such that the quantity described in eq. (11) can be evaluated:

$$\Delta\Delta_{\text{bind}}G_k = \left(\sum_{j=1,j\neq k}^{n}(L_{kj}q_{L,j}) + \sum_{j=1}^{m}(C_{kj}^{T}q_{R,j}) + L_{kk}q_{L,k}^{\text{con}}\right)(q_{L,k} - q_{L,k}^{\text{con}}) + \frac{1}{2}L_{kk}(q_{L,k} - q_{L,k}^{\text{con}})^2 \quad (14)$$

When all of the ligand charges are held constant at their unconstrained, optimized charges and one charge k is changed back to its original value as in:

$$\Delta\Delta_{\text{bind}}G_k = \Delta_{\text{bind}}G(q_{L,j,j\neq k} = q_{\text{lig}}^{\text{opt}}, q_{L,k} = q_{\text{lig},k}) - \Delta_{\text{bind}}G(q_{L,j} = q_{\text{lig}}^{\text{opt}}) \quad (15)$$

the first-order term becomes zero such that:[7]

$$\Delta\Delta_{\text{bind}}G_k = \frac{1}{2}L_{kk}(q_{L,k} - q_{L,k}^{\text{opt}})^2 \quad (16)$$

This is because the first-order derivative is zero when evaluated under the assumption that all of the charges are at their unconstrained, optimized values as evidenced by the second part of eq. (6). Equation (16) reveals that for a ligand atom k, the sensitivity of the binding free energy to deviations from the optimized charge, $\Delta\Delta_{\text{bind}}G_k$, is directly proportional to L_{kk}. The larger L_{kk} is, the more binding affinity is lost as an atomic charge departs from its optimized value.

Another useful method for probing charge sensitivity (especially when dealing with large systems) is to compute $\Delta\Delta_{\text{bind}}G$ for the perturbation of a group of charges (such as an amino acid or side chain) while holding the remainder of the ligand at its original charge distribution:

$$\Delta\Delta_{\text{bind}}G = \Delta_{\text{bind}}G(q_{L,i} = q_{\text{lig},i}) - \Delta_{\text{bind}}G(q_{L,i\in L\text{rem}} = q_{\text{lig},i}, q_{L,i\in L\text{res}} = q_{L,i}^{\text{con}})$$

$$\Delta\Delta_{\text{bind}}G = \frac{1}{2}Q_{L\text{res}}^{T}LQ_{L\text{res}} + Q_{R+L\text{rem}}^{T}C(Q_{L\text{res}} - Q_{L\text{res}}^{\text{con}}) - \frac{1}{2}Q_{L\text{res}}^{T\text{con}}LQ_{L\text{res}}^{\text{con}} \quad (17)$$

where the potential matrices and Lrem are defined as in eq. (7) and Lres refers to the charges on the residue or group of atoms being perturbed.

Using the diagonal elements of the matrix L together with the Taylor series implementation of sensitivity analysis, one can determine how important the optimization of a ligand atom or region is to the overall binding affinity.[7,8] With this information, one can select target areas of a ligand where optimization yields the greatest improvement in binding affinity and then proceed to suggest structural changes to those areas based on the constrained, optimized charge distribution. We have found that it is useful to first look at the constrained and unconstrained optimizations of the total ligand to determine a region's "preference" for positive or negative charge. We could then constrain that region to have a similar charge, with the constraint that the whole molecule has an integral net charge, while holding the remaining charges of the ligand at their original values.[6] Optimization of a specific region of a ligand with constraints that are guided by the total ligand optimization is a useful method to improve electrostatic complementarity within this model.[8] Using the locally optimized charge distribution guarantees that the binding affinity is improved when the original charges are replaced by the optimized ones. On the other hand, the binding affinity may not necessarily improve when one uses the charges obtained from a global optimization, because the binding affinity can improve only when all the charges, including those outside the localized regions, are all replaced by the optimized charges.

Molecular Modeling

All molecular modeling of ligand derivatives and hypothetical structures was done with the InsightII Builder, Biopolymer, and Discover Modules.[33] In all cases, hydrogens were added to crystal structures using the Biopolymer Module.[33] Modifications to ligands were constructed from the coordinates of the crystal structure using the Builder Module.[33] Energy minimization was then used to relax the new or modified functional groups in their environment while the remainder of the system was constrained to its original crystal structure coordinates.[9] Using the Discover Module, the functional groups were energy minimized with the InsightII CVFF force field and the steepest descent algorithm with a distance-dependent dielectric such that the RMS derivative was less than or equal to 0.001 kcal/mol ∗ Å.[33]

For the free-energy calculations on the hypothetical modifications to PKI, the CHARMM27 force field parameters were used for the unmodified ligand atoms.[28] However, we employed *ab initio* quantum chemical methods to obtain partial charges for the entire side chain of any modified amino acid because they are not available in the CHARMM27 force field. The charges were generated by Gaussian 98 using the Merz–Kollman method and the 6-31G* basis set.[34–36]

Results and Discussion

Charge Optimization Trends in the CDK2–Deschloroflavopiridol System

We first applied charge optimization to three small structures from our previous study of the CDK2–deschloroflavopiridol system and found several important trends that would also arise in the more complicated PKA–PKI system. The structures involved in this analysis are shown in Figure 1. Although we have a crystal

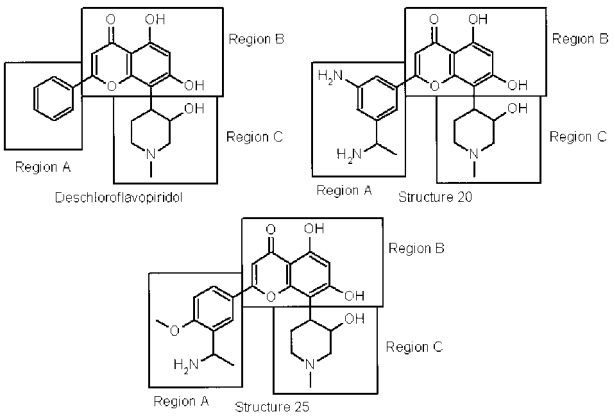

Figure 1. Structures of the deschloroflavopiridol, model Structure 20, and model Structure 25 each divided into three regions for local optimization.

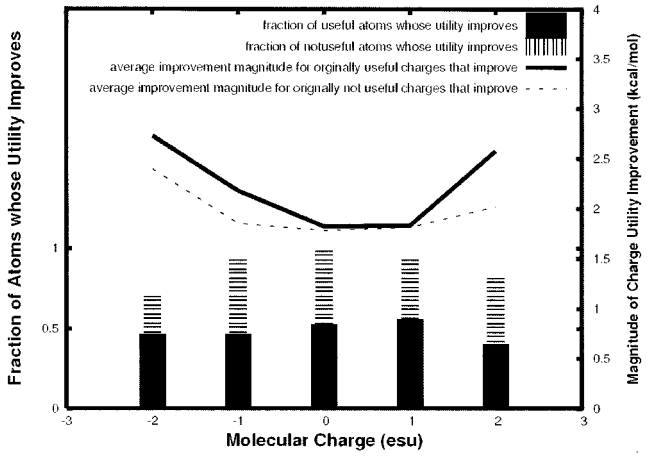

Figure 2. Charge utility analysis for the CDK2–deschloroflavopiridol system.

structure for the CDK2–deschloroflavopiridol system, the other two structures were modeled as outlined in Sims et al.,[9] using a procedure that is nearly identical to that presented in the Molecular Modeling section of this article. In our earlier work, we estimated the binding affinity of deschloroflavopiridol for CDK2 to be about -14 kcal/mol, and we gave estimates of -35 and -63 kcal/mol for Structure 20 and Structure 25, respectively.[9]

To aid the optimization of the flavopiridol class of compounds, we previously relied on gauging the importance of each charge to binding affinity by turning its charge off and examining the extent that the binding affinity was modified. We referred to atoms that lost more binding affinity as having greater charge utility. In other words, putting charges there enhance binding. We then focused on modifying the parts of the lead compound that have low charge utility to see whether stronger binders could be found. This strategy seemed to work as we had designed structures with greater binding affinity with just a small number of trials. Here, we carried out similar charge utility analysis using the charges obtained from global optimization rather than the original quantum mechanically derived charges. Five global charge optimizations were done by imposing five different net charge constraints: $-2e$, $-e$, 0, e, and $2e$ on the molecule. Each charge was constrained to a magnitude of at most $0.85e$ as before. The charge redistribution after optimizations improved the overall binding affinity within our model. However, the relative importance of different atoms in contributing to binding was altered, with some became better and some worse contributors. A consistent observation across all three systems with all five formal charge constraints was that the fraction of atoms in a system whose original charges were not electrostatically useful but whose charge utility improved upon optimization was substantially higher than the fraction of atoms whose original charges were deemed electrostatically useful and still improved in utility (Figs. 2, 3, and 4). So, our earlier strategy that focused on modifying the parts of the lead compound that have low charge utility appears to be a reasonable one for inhibitor design. On the other hand, we found that the charge utility could also be further improved for some atoms that were originally deemed useful because further significant improvement can be achieved by optimizing their charges. Previously, we did not target these atoms for improvement.

Figures 5, 6, and 7 compare the change in binding affinity that occurs upon changing from the constrained, optimized charge distribution to the original charge distribution for five different molecular net charge constraints in all three systems. We see here that $\Delta\Delta_{\mathrm{bind}}G$ is strikingly insensitive to the net charge constraint, suggesting that similar maximal binding affinity can be achieved by placing a positive, negative, or neutral net charge on the ligand. This suggests that one can choose the most convenient charge state for the situation at hand.

These same three figures also contain data from local optimizations of the three ligand regions defined in Figure 1. Each of these three regions was optimized subject to the same five molecular net charge constraints along with the $0.85e$ individual charge magnitude constraint. For deschloroflavopiridol, optimizing region A alone could only improve binding affinity when the total net charge of the ligand was 1 or 2 and the improvement was very

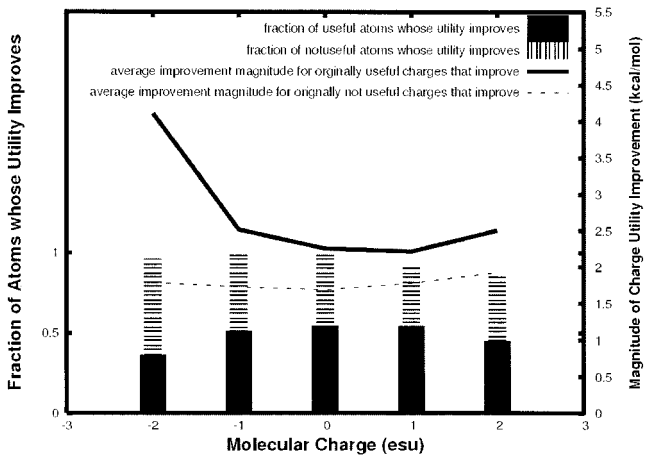

Figure 3. Charge utility analysis for the CDK2–Structure 20 system.

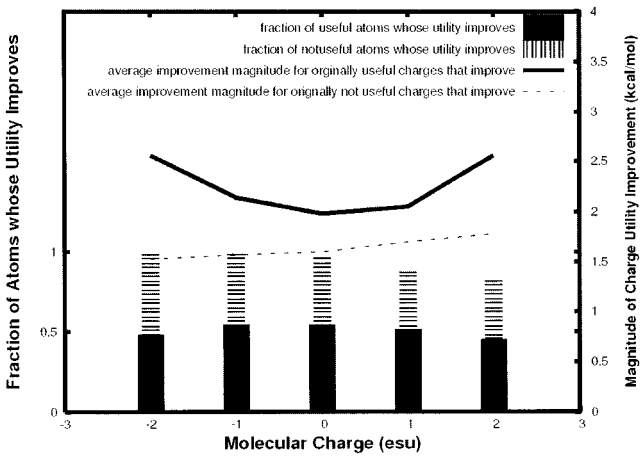

Figure 4. Charge utility analysis for the CDK2–Structure 25 system.

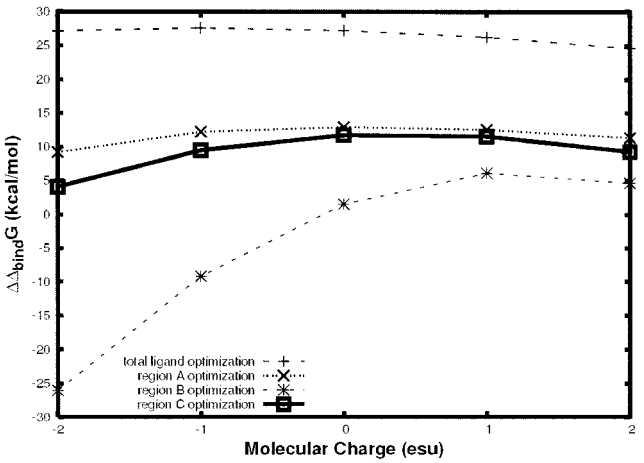

Figure 6. Binding affinity gain when the charges on Structure 20 were optimized in the CDK2–Structure 20 system. The ligand was constrained to five different net charges. Results when only a part of the three ligand regions diagrammed in Figure 1 were optimized are also shown.

small. Optimizing region B was a bit better, especially when the net charge of the ligand was 0 or 1. Optimizing region C provided the best improvement for all the net charge constraints (−2, −1, 0, 1, and 2). For structure 20, optimizing region B did not provide much improvement except when the net charge of the ligand was 1 and 2. On the other hand, optimizing regions A and C could lead to significant improvement no matter at what total net charge the ligand was constrained. Structure 25 produced similar results to structure 20.

Therefore, in all three structures, the optimized charges did not change much in region B. As discussed earlier, this region interacts with the linker region between the N- and C-terminal lobes and is relatively buried.[9] To keep favorable interactions in this region, charge optimization only changed the charges in the other two regions to achieve better binding while satisfying a given net molecular charge constraint. The feasibility of obtaining comparable optimal binding affinity with different net charge constraints gives a drug designer some flexibility in controlling other important properties of the ligand, besides binding affinity, that determine its usefulness as a drug.

Finally, one should also note from these three figures that the structures whose binding affinities are estimated to be higher have $\Delta\Delta_{bind}G$ values that are also much higher. For example, Structure 25 has an estimated $\Delta_{bind}G$ of −63 kcal/mol, whereas deschloroflavopiridol has an estimated $\Delta_{bind}G$ of −14 kcal/mol.[9] Nonetheless, the $\Delta\Delta_{bind}G$ values for the global, constrained optimiza-

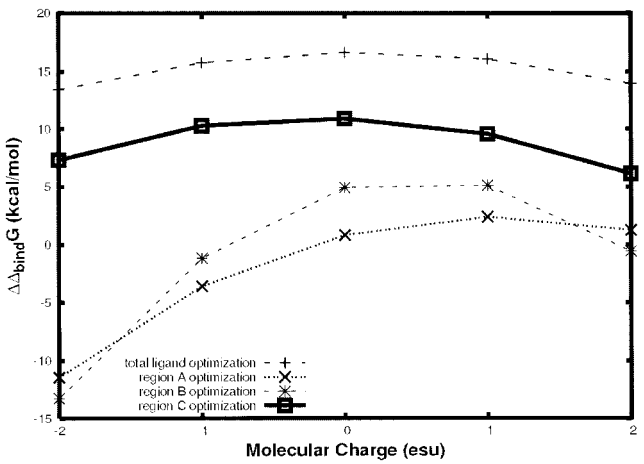

Figure 5. Binding affinity gain when the charges on deschloroflavopiridol were optimized in the CDK2–deschloroflavopiridol system. The ligand was constrained to five different net charges. Results when only a part of the three ligand regions diagrammed in Figure 1 were optimized are also shown.

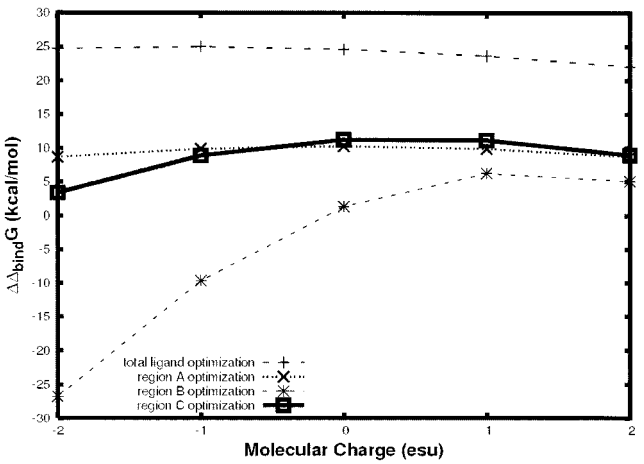

Figure 7. Binding affinity gain when the charges on Structure 25 were optimized in the CDK2–Structure 25 system. The ligand was constrained to five different net charges. Results when only a part of the three ligand regions diagrammed in Figure 1 were optimized are also shown.

ANALOG SEQUENCES

5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
	T	Y	A	D	F	I	A	S	G	R	T	G	R	R	N	A	I			PEPTIDE 1
	T	Y	A	D	F	I	A	S	G	R	T	G	R	R	N	A	I	H	D	MODEL 1
T	T	Y	A	D	F	I	A	S	G	R	T	G	R	R	N	A	I	H	D	PEPTIDE 2*
T	T	Y	A	D	F	I	A	S	G	R	T	G	R	R	N	A	I	H	D	MODEL 2
T	**A**	Y	A	D	F	I	A	S	G	R	T	G	R	R	N	A	I			PEPTIDE 3
T	**A**	Y	A	D	F	I	A	S	G	R	T	G	R	R	N	A	I	H	D	MODEL 3
T	T	Y	A	D	F	**A**	A	S	G	R	T	G	R	R	N	A	I			PEPTIDE 4
T	T	Y	A	D	F	**A**	A	S	G	R	T	G	R	R	N	A	I	H	D	MODEL 4
T	T	**A**	A	D	F	I	A	S	G	R	T	G	R	R	N	A	I			PEPTIDE 5
T	T	**A**	A	D	F	I	A	S	G	R	T	G	R	R	N	A	I	H	D	MODEL 5
T	T	Y	A	D	F	I	A	**L**	**I**	R	T	G	R	R	N	A	I			PEPTIDE 6
T	T	Y	A	D	F	I	A	**L**	**I**	R	T	G	R	R	N	A	I	H	D	MODEL 6

* wildtype

MODEL	PEPTIDE K_i (nM)	MODEL $\Delta_{bind}G$ (kcal/mol)
1	1.7	-79.7
2	2.3	-71.2
3	7.1	-70.4
4	8.1	-70.2
5	14.0	-69.7
6	130.0	-69.1

Figure 8. Sequences of the PKI-(5–24) analogs (referred to by Peptide #) and the corresponding sequences of the computationally modeled analogs (referred to by Model #) along with the experimental binding constant of each peptide and computed binding affinity of each model.

tions of Structure 25 are around 10 kcal/mol higher on average than those of deschloroflavopiridol. This suggests that Structure 25, which has already a greater estimated binding affinity, has a less optimal charge distribution than deschloroflavopiridol relative to its constrained, optimized distribution. Therefore, the potential for improving Structure 25 even further is higher.

Binding Energy Calculations for the PKA–PKI System

In our previous study of the CDK2–deschloroflavopiridol system, we modeled analogs of deschloroflavopiridol that had been synthesized and whose inhibitory activity had been measured experimentally.[9] By comparing the calculated binding affinity of these compounds using the fixed conformational continuum-solvent model as described above, one can gauge the reliability of the model in reproducing experimental trend before one uses it to make predictions. We pursued a similar kind of analysis here.

Many of the biological assays of the PKA–PKI system were actually performed with truncated versions of the PKI-(5–24) ligand that only contained residues 5–22 or 6–22.[17] To make it easier to computationally estimate the relative binding affinity of polypeptides with different mutations, we modeled all of the polypeptides with the same length without truncation. Because we were using a rigid conformational model here, we only attempted to study those mutated peptides for which there was sufficient space in the complex to accommodate the mutation without significant structural change to PKA and PKI. As a result, six mutated PKIs from Glass et al.'s experiments[17] were selected for comparison, and we found that our model ordered their potencies correctly (Fig. 8). It should be noted that we added a penalty relative to the wildtype PKI-(5–24) of 0.3 kcal/mol for each sp^3 rotatable bond in the modeled peptides. For example, if a given peptide has two more sp^3 rotatable bonds than PKI-(5–24), we would add 0.6 kcal/mol to the binding free energy obtained as outlined in the Binding Free Energy Model section. This penalty is meant to approximate the effects of rotational entropy on our binding affinity ranking, and has been employed in past computational studies including Sims et al., Gidofalvi et al., Bohm, and Morris et al.[9,37–39]

Charge Optimization of PKI-(5–24)

After computing the potentials as described in the Charge Optimization Theory section with UHBD[26,27] for the 306 atoms of PKI-(5–24) with respect to PKA, we obtained unconstrained charge optimization data by solving eq. (6) directly and fully constrained optimization data via our implementation of the Frank–Wolfe algorithm. During the fully constrained optimization, magnitudes of all 306 point charges could not exceed $0.85e$, and the sum of the 306 point charges had to retain the original ligand charges of $+1e$.

For each atom in each amino acid of PKI-(5–24), the absolute value of the deviation between the fully constrained, optimized charge and the original CHARMM27 charge was computed, and the average value of this quantity for each amino acid appears in Figure 9. The fully constrained optimization suggests that Phe10 has the smallest overall deviation form its optimized charge distribution. Interestingly, a 1989 article by Glass et al.[18] focuses exclusively on the effects that Phe10 has on the binding affinity of PKI-(6–22). It is clear from the crystal structure that Phe10 fits in a hydrophobic pocket of PKA consisting of residues 235–239 (Tyr-Pro-Pro-Phe-Phe).[20] This is a highly aromatic pocket that requires a nonpolar charge distribution for optimal interaction. The charge optimization recognizes this property, and PKI itself exploits the aromatic pocket for tight binding. Other noteworthy residues appear in the top 10 lowest average deviation list includ-

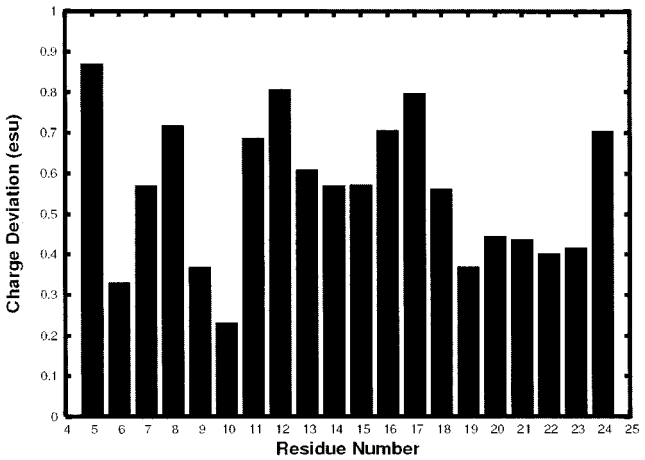

Figure 9. Deviations of optimized charges from their original CHARMM27 charges for PKI-(5–24). The absolute average deviation over atoms in each residue is shown. Results from global, fully constrained charge optimization.

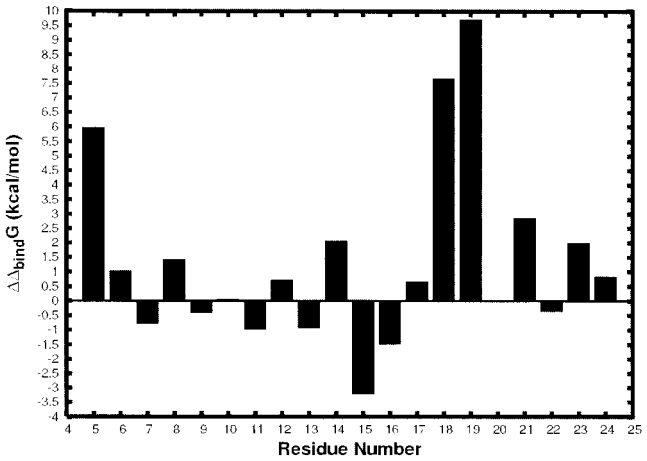

Figure 10. Change in binding free energy, $\Delta\Delta_{bind}G$ from eq. (17), produced by optimized charges obtained from the global, fully constrained optimization of PKI-(5–24).

ing Arg18 and Arg19, which are the two most important basic residues for peptide recognition as was discovered by Kemp et al.[11] and was also evident in computational studies.[23] The P-site residue (Ala21) also appears in this list with Ile22, which is particularly noted for its contribution to peptide recognition.[20]

Again, one cannot evaluate how optimized the charge distribution of a residue is by simply examining the deviation of optimized from original charges because the sensitivity of the binding affinity to charge variations is not uniform along the length of the polypeptide. A slight deviation from the optimized charge distribution can produce a large change in binding affinity in some parts of the polypeptide, whereas a large deviation may not produce significant change in the other parts. A more direct quantity to look at is the deviation between the binding affinity presented by the original charges and that presented by the optimized charges, which can be calculated in a way similar to sensitivity analysis.

Sensitivity Analysis and the Global Charge Optimization of PKI-(5–24)

We found that our ability to efficiently propose enhancements to PKI-(5–24) was sharpened by a sensitivity analysis approach to interpreting charge optimization data. In Figure 10, we show the $\Delta\Delta_{bind}G$ values for the PKI global, fully constrained optimization. In this case, eq. (17) was used to compute $\Delta\Delta_{bind}G$. This measures the extent to which the binding affinity is diminished when the optimized charge of an atom is changed to that of the original charge, keeping all other charges at their original values. Although $\Delta\Delta_{bind}G$ for the total ligand must necessarily be positive for the optimized charges to result in a more favorable $\Delta_{bind}G$, the corresponding values when only the charges on a specific residue were changed to the optimized ones need not be. For example, we see that the very important Arg18 and Arg19 can be improved substantially by optimization, but the third important basic residue, Arg15, would actually harm binding affinity if it had the charges proposed by global optimization while the remaining ligand had its original charges. This is because one can only obtain a more favorable binding affinity when all the charges are modified according to the charge optimization. Changing only a subset of the charges may not give more favorable binding. Because it is usually difficult to find a real molecule that can give a charge distribution suggested by a global optimization, a local optimization that focuses on only a part of the molecule keeping everything else fixed may be a better strategy. Nonetheless, global optimization can give preliminary insights into which regions may be profitably modified to achieve better binding, as we will show later.

The global, fully constrained optimization also reveals that Phe10, which has the lowest average deviation from its optimized charge distribution, has a relatively low $\Delta\Delta_{bind}G$, indicating that the charge distribution on this residue is close to giving optimal binding, confirming again the importance of having a hydrophobic residue here. Figure 11 shows the average diagonal element of L for the amino acids in PKI-(5–24). From eq. (16), we know that this number gives some hint as to how much binding affinity is lost if the charge distribution were not close to optimal. The residues of PKI-(5–24) that present the highest L_{kk} are Phe10, Arg15, Arg18, Arg19, Ala21, and Ile22, which are shown to be quite important in binding.[20] As discussed earlier, Phe10 already has charge distribution that is close to optimal. On the other hand, it seems that the charge distribution on Arg18 and Arg19 can be further modified to gain significant improvement in binding, as their large $\Delta\Delta_{bind}G$ values in Figure 10 indicate. However, as discussed earlier, local optimization is necessary to provide more reliable predictions if one only modifies a local charge distribution without changing the other parts of the polypeptide.

Sensitivity Analysis and the Local Charge Optimization of PKI-(5–24)

Figure 12 shows the average absolute deviation between the original CHARMM27 charge distribution and the fully constrained, optimized charge distribution of PKI-(5–24) residues where only the side chains of PKI were optimized with respect to the remaining ligand atoms which were held constant. Although all of the

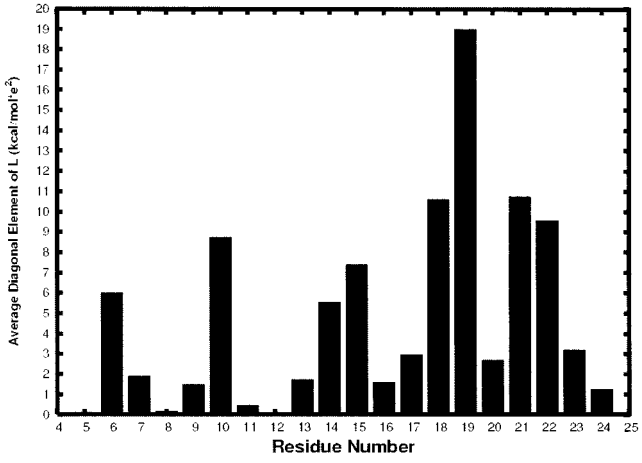

Figure 11. Average diagonal element of the matrix L for each amino acid in PKI-(5–24).

remaining ligand charges were held at their original values, the charges on a single PKI side chain were allowed to change as long as no charge magnitude exceeded $0.85e$ and the net charge of the residue remained unchanged. Figure 12 presents data that are analogous to Figure 9 but for local optimization of PKI side chains. The Gly14 and Gly17 side chains were omitted from this analysis because they consist of only a single atom. Phe10, Thr6, Arg15, Arg19, Ala21, and Ile22 are the residues with the smallest deviations.

Just as in global optimization, we can consider $\Delta\Delta_{bind}G$ for local optimization data. Figure 13 shows results that are analogous to Figure 10 for the local charge optimization where Gly14 and Gly17 are once again omitted. It should be noted that these $\Delta\Delta_{bind}G$ values are necessarily positive because the charges that were modified are also those that were optimized. These data reveal the amount of binding affinity (within this model) by which local modification of a given, individual PKI residue can improve the overall ligand. For example, if the charges on Arg18 match those of its local, fully constrained optimized charge distribution while the remaining ligand maintains its original CHARMM27 charges and the shape of Arg18 is unaltered, the free energy change on binding PKI-(5–24) to PKA in our model would improve by about 7.5 kcal/mol. It is clear from this analysis that within this model, Arg15, Arg18, Arg19, and Tyr7 offer the best opportunities for local modifications to yield significant, overall improvements to binding affinity. It is interesting to see that although previous experimental and computational studies have shown that Arg15, Arg18, and Arg19 already contribute significantly to binding, these residues might be modified further to gain substantial improvement in binding affinity. Also, the small deviations of their charges from the optimal ones indicate that only small modifications of the charges on these residues are needed to improve binding significantly.

The global charge optimization data suggested that Arg18 prefers to have higher net charge (data not shown), so we attempted a second, local optimization of this side chain in which the

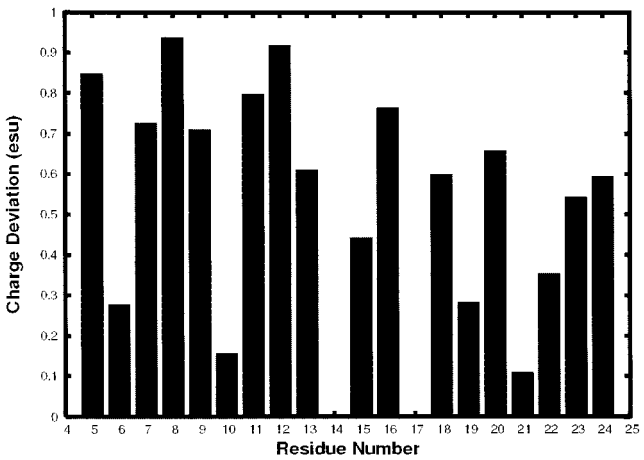

Figure 12. Deviations of optimized charges from their original CHARMM27 charges for PKI-(5–24). The absolute average deviation over atoms in each residue is shown. Results from local, fully constrained charge optimization of individual side chains.

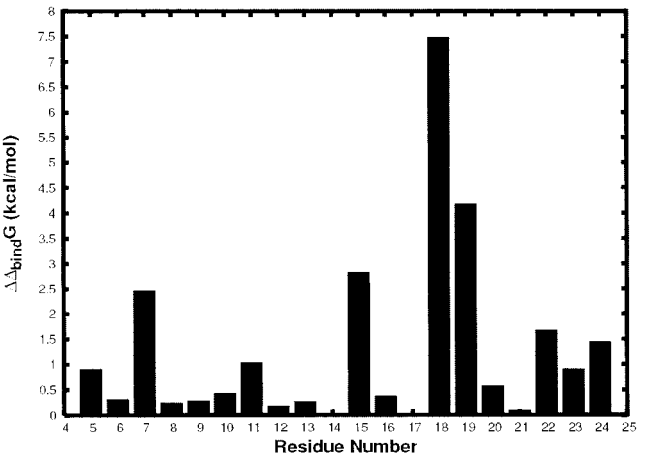

Figure 13. Change in binding free energy, $\Delta\Delta_{bind}G$ from eq. (17), produced by optimized charges obtained from the local, fully constrained optimization of individual side chains of PKI-(5–24).

net charge of the residue was constrained to be +2 rather than +1. Recalculation of $\Delta\Delta_{bind}G$ showed that a potential improvement of about 9.7 kcal/mol was possible for Arg18 with a +2 charge as opposed to the 7.5 kcal/mol $\Delta\Delta_{bind}G$ available to Arg18 with a +1 charge. Similarly, a $\Delta\Delta_{bind}G$ of over 6 kcal/mol was found possible for Arg19 with a +2 charge as opposed to the 4.2 kcal/mol $\Delta\Delta_{bind}G$ value when a +1 charge constraint was in place. Global optimization data also suggested that Tyr7 could benefit from having a lower net charge, so local optimization was recalculated using a net charge constraint of −1 on Tyr7. A potential improvement of 4.4 kcal/mol becomes available as opposed to the 2.5 kcal/mol $\Delta\Delta_{bind}G$ from the zero net charge constraint. These data combined with the actual local and global charge distributions can be used to propose chemical modifications to PKI-(5–24) as long as the suggestions do not drastically alter the spatial parameters of the ligand. A binding energy calculation using the same continuum-solvent model used in the charge optimization studies can then further evaluate the effectiveness of these proposed modifications.

Designed Peptide Mimics—Suggested Modifications to PKI-(5–24)

We propose 10 designed modifications that perturb the charge distribution of PKI-(5–24) in accordance with our charge optimization data. We targeted Arg18, Arg19, and Tyr7 for modifications. Some of the proposed structures reflect the aforementioned changes in net charge of the residues. The modifications and the computed effects on their possible binding affinities are recorded in Figure 14.

Structure 1 involves a modification to the Arg19 side chain that requires very little change in the overall shape of the residue (and much less when the ligand is considered in total), but presents charges closer to the optimized ones. The hybridizations of the atoms remain unchanged, although some radii, and of course, the charge distribution are perturbed. However, the increase in binding affinity is less than 2 kcal/mol within our model. Further modifi-

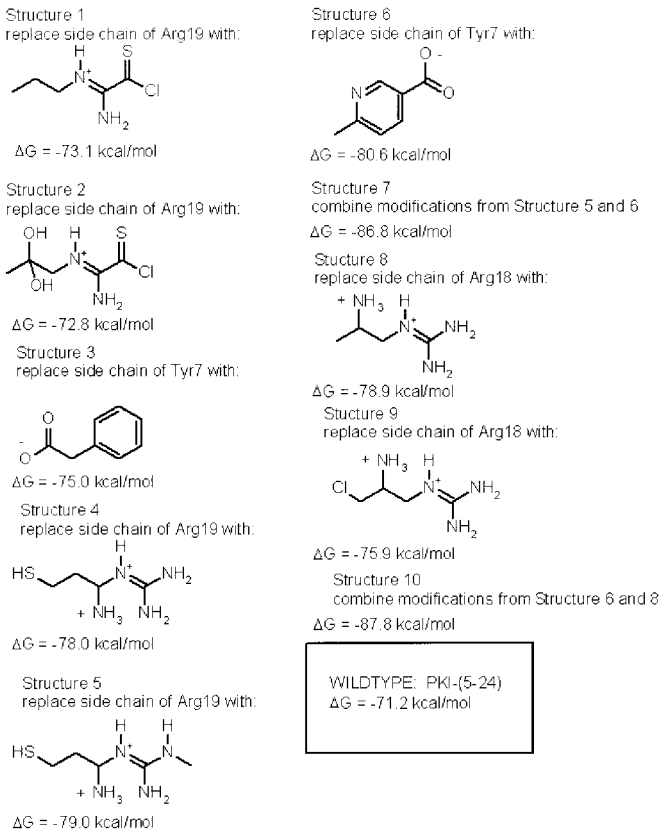

Figure 14. Structures of the modified side chains in each of our proposed, hypothetical structures for PKI derivatives and the corresponding computed binding free energy changes (for association with PKA).

cation in Structure 2 involved placing two hydroxyl groups on the beta-carbon, but even less is gained here (~1.6 kcal/mol).

Placing a −1 charge on the originally neutral Tyr7 by acetylating the beta-carbon while removing the hydroxyl group improves the computed $\Delta_{bind}G$ by about 3.8 kcal/mol (Structure 3). Similarly, when we changed the formal charge of Arg19 from +1 to +2 in accordance with the previous section by the addition of an ammonium group and also a mercapto group, we computed an improvement in $\Delta_{bind}G$ of 6.8 kcal/mol (Structure 4). Further perturbation by replacing one of the guanidine amines with a methylamine while including the modifications from Structure 4 gave us an improvement in $\Delta_{bind}G$ of 7.8 kcal/mol for Structure 5.

Careful analysis of local charge optimization data along with the original quantum mechanically derived charges suggested that the carboxyl group used in Structure 3 would be better placed on the phenyl ring *para* to the amino acid backbone. In addition, the local optimization suggested that a more electronegative atom should be placed *ortho* to the amino acid backbone to polarize the bonds at that position somewhat. A heterocyclic nitrogen was modeled at that position to complete Structure 6, which had a computed improvement in $\Delta_{bind}G$ of about 9.4 kcal/mol. Combining the modification in Structure 6 with that in Structure 5 yielded Structure 7, which has the same overall net charge as the wild-type PKI-(5–24) but with a computed 15.6 kcal/mol improvement in $\Delta_{bind}G$. We applied alterations to Arg18 that were similar to those used in modifying Arg19 based on global and local optimization data. The modifications in Structure 8 (adding an ammonium group to the gamma-carbon) and Structure 9 (adding both an ammonium group to the gamma-carbon and a chloro group to the beta-carbon) gave computed improvements in binding affinity of 7.7 and 4.7 kcal/mol, respectively. Finally, the combination of the modifications from Structure 6 and Structure 8 formed Structure 10. This structure, like Structure 7, has the same overall net charge as the original polypeptide but improved the binding affinity by 16.6 kcal/mol (−87.8 kcal/mol vs. −71.2 kcal/mol).

Although the shape of the ligand needs to be modified somewhat in designing molecules that can actually be made, the gain in binding affinity of our designed compounds are close to those suggested by charge optimization without changing the original ligand structure. In the designed compounds, the hydrophobic term described in our Binding Free Energy Model may also change somewhat. However, these changes in ΔG_{SA} are relatively small because we focused on designing peptide mimics without significantly changing the shape of the original peptide so that the results from charge optimization with constant shape can be used more readily. Here, we show that although the PKI fragment already has binding affinity in the nM range, it may be feasible to improve it further by making chemical modifications such as those shown in Figure 14.

Charge Optimization at the Cyclin A/CDK2 Interface

A 32-amino acid (547 atom) region of the aforementioned cyclin A fragment was designated as the interfacial region with respect to CDK2 based on a simple distance criterion applied to the unphosphorylated complex crystal structure.[24] We defined any amino acid in cyclin A containing an atom within 3 Å of any CDK2 atom to be an interfacial amino acid. We then employed global and local charge optimization using a fixed conformation model as in studying the other two systems.

During global optimization, the individual charge magnitudes were restricted just as in all previous constrained optimizations, and the net charge was constrained at the original value of $+3e$. However, during global charge optimization, the charges on individual residues can rearrange significantly that can still yield the same net charge. On the other hand, in local optimization, only the side chain atoms were considered and the net charge of each residue was restricted to its original value of −1, 0, or 1. As described earlier in PKA–PKI interactions, it is easier to combine results from local optimization and chemical intuition to come up with chemical modifications that can improve binding.

In this analysis, we will focus on a few amino acids in the interface that exhibited interesting transitions upon optimization. The total results of both optimization types are summarized in three figures. Figure 15 graphs the average diagonal element of the matrix L for each amino acid. Figures 16 and 17 show the average absolute charge deviation and sensitivity analysis data for the global and local optimizations, respectively.

One particularly noteworthy amino acid is Phe267, which has the fifth highest average diagonal element of L (8.10 kcal/mol $*$ e^2). This signifies large deviations from optimal binding affinity if

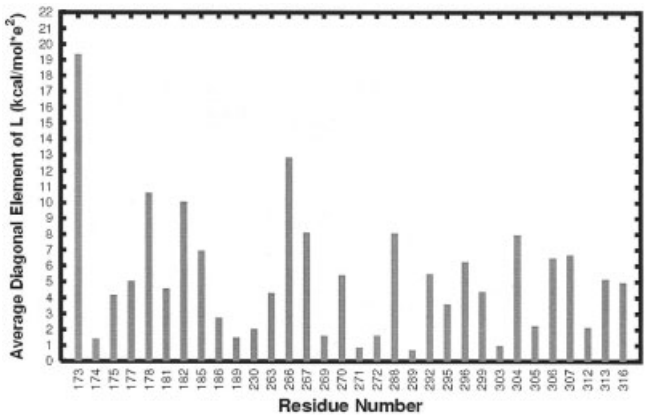

Figure 15. Average diagonal element of the matrix L for each amino acid in the Cyclin A-CDK2 interfacial region.

the atomic charges on this residue are not optimized, as implied by eq. (16). We found that this amino acid has the smallest average absolute deviation from the local, optimized distribution, and the seventh smallest such deviation from the global, optimized distribution. More importantly, Phe267 has the lowest local $\Delta\Delta_{\text{bind}}G$ (only 0.28 kcal/mol where the average $\Delta\Delta_{\text{bind}}G$ is 1.82 kcal/mol) and the eighth lowest $\Delta\Delta_{\text{bind}}G$ from the global analysis (1.00 kcal/mol where the average $\Delta\Delta_{\text{bind}}G$ is 1.81 kcal/mol). These results indicate that the charge distribution on this residue is already rather optimized. In fact, its average charge deviation in global optimization is far below average (by about $0.12e$ where the range of average values is $0.17e$–$0.55e$). This is consistent with an analysis of the crystal structure, which shows that this residue forms part of a hydrophobic pocket around the CDK2 amino acid Ile49 along with cyclin A residues Leu263, Leu299, Leu306, and part of Lys266 (carbon chain).[24] Any significant deviation from nonpolar charges could significantly diminish binding. Thus, it will be better to leave this residue alone in search of cyclin A mimics that can bind better to CDK2.

The original charge distribution of Phe304 appears to be comparable in optimality to that of Phe267. This amino acid has the seventh highest average diagonal element of L at 7.95 kcal/mol $*$ e^2. It also has the lowest and second lowest deviations from its optimized distribution in the global and local optimizations, respectively. The global $\Delta\Delta_{\text{bind}}G$ value for Phe304 is -1.02 kcal/mol (seventh lowest) and 0.59 kcal/mol for the local optimization (eighth lowest). This residue is another example of cyclin A utilizing a hydrophobic pocket to achieve binding affinity. Both Phe304 and Phe267 are major components of the hydrophobic cyclin box portion of the interface that forms major buried hydrophobic interactions with one of CDK2's flexible helices.[24] Phe304 makes noteworthy van der Waals contacts with Ile52 on CDK2.[24] Ile52 is solvent inaccessible in the free crystal structure of CDK2[40] but undergoes a positional change upon binding to cyclin A, which facilitates this and other interactions.[24] The optimality of these two hydrophobically interacting phenylalanines is reminiscent of Phe10 on PKI, which fits into an aromatic pocket on PKA and exhibits a relatively optimized charge distribution.

The three residues that give the largest diagonal elements of L are Asn173, Lys266, and Tyr178. These residues could lose a lot of binding affinity if their charge distributions are far from their optimal values. Examining the average $\Delta\Delta_{\text{bind}}G$s from global and local optimization shows that modifying Asn173 and Tyr178 in such a way to yield charge distributions closer to their optimal value could produce large gain in binding affinity. On the other hand, it seems that Lys266 already has a charge distribution quite close to their optimal values that no significant improvement in binding affinity is gained when its charge distribution is set to its optimal value. Another residue that stands out in the average $\Delta\Delta_{\text{bind}}G$ plots is Lys288. Modifying its charge distribution to its optimal one could significantly improve binding. Thus, these calculations suggest that Asn173, Tyr178, and Lys288 are three of the residues that are worthwhile to consider modifying for a better binder to CDK2.

The cyclin A interface appears to be comparable in optimality to PKI-(5–24) with respect to its target receptor PKA. The average L_{ii} values for the two systems are very similar and range from less

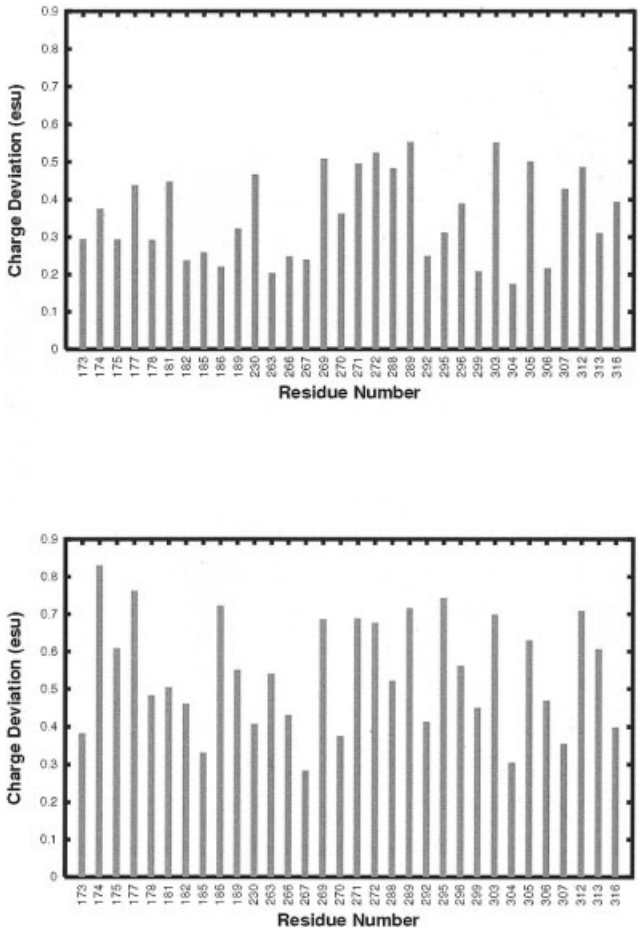

Figure 16. Charge deviation analysis for the global, fully constrained optimization of the cyclin A–CDK2 interfacial region (top) and for the local, fully constrained optimization of the side chains in the cyclin A–CDK2 interfacial region (bottom). Absolute average deviation over atoms in each residue is shown.

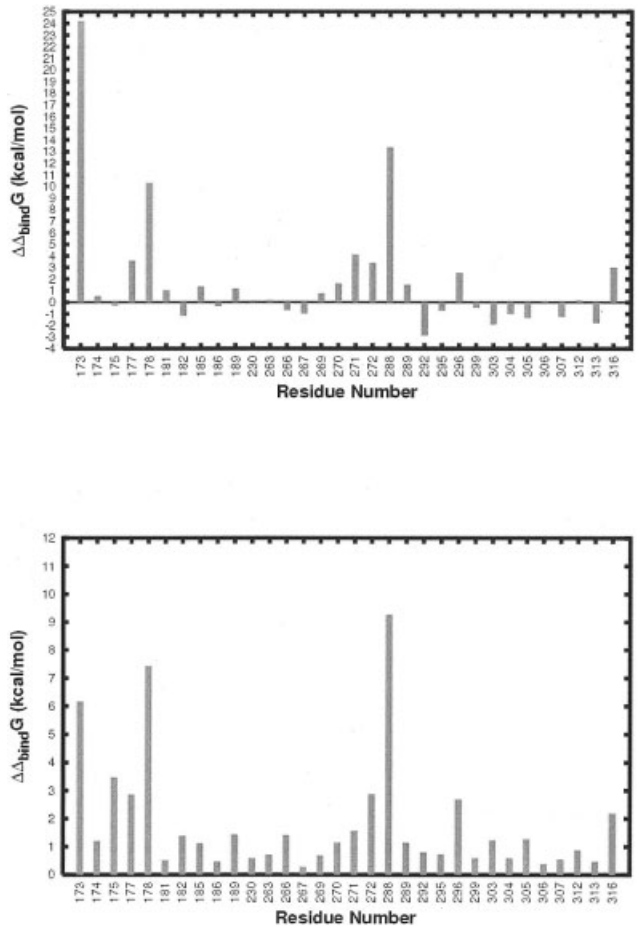

Figure 17. Change in binding free energy, $\Delta\Delta_{bind}G$ from eq. (17), produced by optimized charges obtained from the global, fully constrained optimization of the cyclin A–CDK2 interfacial region (top) and for the local, fully constrained optimization of the side chains in the cyclin A–CDK2 interfacial region (bottom).

than 1 kcal/mol $*$ e^2 to about 19 kcal/mol $*$ e^2. The average $\Delta\Delta_{bind}G$s for the two global optimizations are also fairly close: 1.81 kcal/mol for the cyclin A interface and 1.35 kcal/mol for PKI-(5–24). For the two local optimizations, the average $\Delta\Delta_{bind}G$ values are 1.82 kcal/mol and 1.44 kcal/mol for the cyclin A interface and PKI-(5–24), respectively.

Conclusions

The charge optimization approach to rational, computational drug design appears to have great potential due to the wealth of information that can be obtained from one calculation and its relatively low computational cost. Despite its methodological similarity to our earlier sensitivity analysis studies, the data presented by charge optimization and sensitivity analysis yield different insights, and can be profitably used together to aid molecular design. In our previous computational design attempt with the CDK2–deschloroflavopiridol system, we used sensitivity analysis to determine which regions of the ligand lacked a significant contribution to the electrostatic binding free energy change.[9] That is, we tried to gauge the utility of a partial charge. In the design phase that followed, we kept regions of the ligand that are already useful in binding and focused on improving the region that lacked utility.[9] As long as the modeling space in the crystal structure allowed it, we made significant changes to the overall ligand shape.[9] In the current charge optimization studies of the PKA–PKI-(5–24) and CDK2–deschloroflavopiridol systems, we found that ligand regions whose contributions to binding are already substantial can still have great potential for improvement. This suggests another strategy for inhibitor design.

As noted in this and other studies, there are numerous limitations to such a continuum solvent model of binding free energy.[9,23,29] In this particular case, we applied a fixed conformation model that assumes that the ligand and receptor undergo no conformational change upon binding. This is a reasonable assumption in guiding inhibitor design as long as the designed compounds are reasonably similar to their parent compounds or peptides. As we illustrated here and in our earlier work, chemical modifications that do not significantly alter the conformation of the targets, and the lead inhibitors can be designed that could significantly improve binding.

Our analysis of the protein–protein interface between CDK2 and cyclin A, although preliminary, has already pointed out that charge optimization can suggest which parts of the inhibitor may be profitably modified to improve binding affinity. Despite their already impressive optimal binding, we still found great potential for further improving the binding affinity between these two proteins.

Acknowledgments

We thank Professor Sung Hou Kim for providing the coordinates of the CDK2-L868276 complex. P.A.S. is the recipient of a Goldwater Scholarship.

References

1. Lee, L. P.; Tidor, B. J Chem Phys 1997, 106, 8681.
2. Chong, L. T.; Dempster, S. E.; Hendsch, Z. S.; Lee, L. P.: Tidor, B. Protein Sci 1998, 7, 206.
3. Kangas, E.; Tidor, B. J Chem Phys 1998, 109, 7522.
4. Kangas, E.; Tidor, B. Phys Rev E 1999, 59, 5958.
5. Kangas, E.; Tidor, B. J Chem Phys 2000, 112, 9120.
6. Lee, L. P.; Tidor, B. Nat Struct Biol 2001, 8, 73.
7. Lee, L. P.; Tidor, B. Protein Sci 2001, 10, 362.
8. Kangas, E.; Tidor, B. J Phys Chem B 2001, 105, 880.
9. Sims, P.; Wong, C. F.; McCammon, J. A. J Med Chem 2003, 46, 3314.
10. Wong, C. F.; Thacher, T.; Rabitz, H. In Reviews in Computational Chemistry; Lipkowitz, K. B.; Boyd, D. B., Eds.; Wiley–VCH: New York, 1998, p. 281.
11. Kemp, B. E.; Graves, D. J.; Benjamini, E.; Krebs, E. G. J Biol Chem 1977, 252, 4888.
12. Feramisco, J. R.; Glass, D. B.; Krebs, E. G. J Biol Chem 1980, 255, 4240.

13. Cheng, H.-C.; Kemp, B. E.; Pearson, R. B.; Smith, A. J.; Misconi, L.; Van Patten, S. M.; Walsh, D. A. J Biol Chem 1986, 261, 989.
14. Scott, J. D.; Glaccum, M. B.; Fischer, E. H.; Krebs, E. G. Proc Natl Acad Sci USA 1986, 83, 1613.
15. Reed, J.; De Ropp, J. S.; Trewhella, J.; Glass, D. B.; Liddle, W. K.; Bradbury, E. M.; Kinzel, V.; Walsh, D. A. Biochem J 1989, 264, 371.
16. Reed, J.; Volker, K.; Cheng, H.-C.; Walsh, D. A. Biochemistry 1987, 26, 7641.
17. Glass, D. B.; Cheng, H.-C.; Mende-Mueller, L.; Reed, J.; Walsh, D. A. J Biol Chem 1989, 264, 8802.
18. Glass, D. B.; Lindquist, L. J.; Katz, B. M.; Walsh, D. A. J Biol Chem 1989, 264, 14579.
19. Glass, D. B.; Trewhella, J.; Mitchell, R. D.; Walsh, D. A. Protein Sci 1995, 4, 405.
20. Knighton, D. R.; Zheng, J. H.; Teneyck, L. F.; Xuong, N. H.; Taylor, S. S.; Sowadski, J. M. Science 1991, 253, 414.
21. Knighton, D. R.; Bell, S. M.; Zheng, J. H.; Teneyck, L. F.; Xuong, N. H.; Taylor, S. S.; Sowadski, J. M. Acta Crystallogr D Biol Crystallogr 1993, 49, 357.
22. Narayana, N.; Cox, S.; Shaltiel, S.; Taylor, S. S.; Xuong, N. H. Biochemistry 1997, 36, 4438.
23. Gould, C.; Wong, C. F. Pharmacol Ther 2002, 93, 169.
24. Jeffrey, P. D.; Russo, A. A.; Polyak, K.; Gibbs, E.; Hurwitz, J.; Massague, J.; Pavletich, N. P. Nature 1995, 376, 313.
25. MacKerell, A. D.; Banavali, N.; Foloppe, N. Biopolymers 2000, 56, 257.
26. Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. Comput Phys Commun 1991, 62, 187.
27. Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosiewicz, J.; Gilson, M. K.; Bagheri, B.; Scott, L. R.; McCammon, J. A. Comput Phys Commun 1995, 91, 57.
28. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph–McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz–Kuczera, J.; Yin, D.; Karplus, M. J Phys Chem B 1998, 102, 3586.
29. Wong, C. F.; Hünenberger, P. H.; Akamine, P.; Narayana, N.; Diller, T.; McCammon, J. A.; Taylor, S.; Xuong, N. H. J Med Chem 2001, 44, 1530.
30. Narayana, N.; Diller, T. C.; Koide, K.; Bunnage, M. E.; Nicolaou, K. C.; Brunton, L. L.; Xuong, N. H.; Ten Eyck, L. F.; Taylor, S. S. Biochemistry 1999, 38, 2367.
31. Hillier, F. S.; Lieberman, G. J. Introduction to Mathematical Programming; McGraw-Hill: New York, 1995.
32. Dixon, L. C. W. Nonlinear Optimisation; Crane, Russak: London, UK, 1972.
33. InsightII.; Accelrys Inc.: San Diego, 2000.
34. Singh, U. C.; Kollman, P. A. J Comp Chem 1984, 5, 129.
35. Besler, B. H.; Merz, K. M.; Kollman, P. A. J Comput Chem 1990, 11, 431.
36. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pornelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, J. B.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, A.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head–Gordon, M.; Replogle, E. S.; Pople, J. A. Gaussian Inc.: Pittsburgh, PA, 1998.
37. Gidofalvi, G.; Wong, C. F.; McCammon, J. A. J Chem Ed 2002, 79, 1122.
38. Böhm, H. J. J Comput Aided Mol Des 1994, 8, 243.
39. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. J Comput Chem 1998, 19, 1639.
40. Debondt, H. L.; Rosenblatt, J.; Jancarik, J.; Jones, H. D.; Morgan, D. O.; Kim, S. H. Nature 1993, 363, 595.