JCTC Journal of Chemical Theory and Computation

Sensitivity Analysis and Charge-Optimization for Flexible Ligands: Applicability to Lead Optimization

Michael K. Gilson*

Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, Maryland 20850

Received September 9, 2005

Abstract: Sensitivity analysis and charge-optimization have been suggested as methods to guide the optimization of lead compounds in early-stage drug discovery. However, applications to date have been restricted by the simplifying assumption of a rigid ligand. The present study applies both formalisms to the case of a flexible ligand in a model application to an HIV-protease inhibitor. The results suggest that sensitivity analysis is a fast and robust method for guiding charge changes in both a rigid and a flexible ligand, although its accuracy is limited by the fact that it represents a linear approximation. The more complete quadratic analysis provided by charge-optimization produces unexpected results when the ligand is considered to be flexible. For example, it can yield atomic charges which powerfully stabilize the bound conformation of the ligand relative to the conformation assumed for the free state, thus markedly destabilizing the assumed free conformation. Such results are traceable to the fact that the energy matrix possesses negative eigenvalues. However, optimizing charges under the assumption that the ligand does not change conformation upon binding leads to a set of charges that robustly improve affinity, even when the free conformation is later allowed to vary. Thus, both sensitivity analysis and charge-optimization appear to be useful techniques.

1. Introduction

Structure-based drug discovery frequently begins with identification of a lead compound, a small molecule with moderate affinity for a targeted protein of known structure. This is followed by modification of the lead compound in order to arrive at a high-affinity drug candidate. This second step, improving on the lead compound, remains a significant challenge because it is rarely clear what chemical changes will lead to greater affinity for the target. One rather obvious approach is to make changes that will increase the favorable Coulombic interactions between the ligand and the protein. However, any charges that are added to the ligand and that come to lie in the ligand−protein interface are stripped of water during the process of binding and thus incur a desolvation free energy penalty which can be substantial. In fact, calculations frequently indicate that the desolvation penalty exceeds the attractive interaction, making for an unfavorable net electrostatic interaction even when the binding interface appears to possess good electrostatic complementarity; see, e.g. refs 1 and 2.

In recent years, a series of studies (see, e.g., refs 3−5) has addressed this issue, pointing out that the variation of the electrostatic contribution to the binding energy as a function of the atomic charges of the ligand has the form of a parabola with upward curvature, when the ligand is considered to be rigid. As a consequence, there is a set of ligand charges that minimizes the change in electrostatic energy and hence maximizes affinity, all other things being equal. Moreover, when optimal or near-optimal charges are assumed, the net change in electrostatic energy upon binding can be strongly favorable. These observations have led to the exploration[5−7] and successful use[6,8] of charge-optimization methods as a basis for lead optimization.

One potential limitation of the charge optimization approach is that the values of the optimal charges in one part of a ligand are influenced by charges assumed to exist at

* Corresponding author phone: (240)314-6217; fax: (240)314-6255; e-mail: gilson@umbi.umd.edu.

10.1021/ct050226y CCC: $33.50
Published on Web 01/27/2006

other ligand atoms. As a consequence, if one is interested in identifying only parts of the ligand whose charges can be changed to improve affinity, the charges of a fully optimized ligand may not faithfully indicate the changes needed for just a part, as recently noted[9] and further discussed in this paper. Another potential drawback of the method is that it tends to be time-consuming, at least as originally formulated, since the method has required solving the linearized Poisson−Boltzmann (LPB) equation at least once for every atom whose charge is to be optimized. This means on the order of 100 LPB calculations for a druglike ligand. On the other hand, recent algorithmic advances promise to markedly reduce the computational cost of the method.[10,11]

Sensitivity analysis represents another promising approach to guiding the electrostatic optimization of ligands; see, e.g. refs 9 and 12−14. In the present context, this method involves computing the first derivative of the binding free energy with respect to the partial atomic charges of the ligands; affinity can then be improved by raising the charge of atoms with negative derivatives or lowering the charge of atoms with positive derivatives. Sensitivity analysis requires fewer numerical solutions of the LPB equation than does standard charge-optimization and therefore should be less time-consuming. On the other hand, it is a linear approximation and thus does not account for the parabolic curvature of the electrostatic energy with respect to atomic charges. As a consequence, it may be less accurate.

Thus, both charge-optimization and sensitivity analysis are theoretically interesting methods that could be quite useful during the lead optimization stage of drug discovery. To date, however, these methods have been limited to applications in which the ligand is assumed to be rigid. In fact, it has not been clear whether charge-optimization could be generalized to the more realistic case of a flexible ligand. Moreover, we are not aware of a direct comparison of the two approaches, for either a rigid or a flexible ligand. The present paper thus discusses how both methods generalize to the case of a flexible ligand and compares their properties for a model system in which the ligand is treated first as rigid and then as flexible.

2. Theory

This section derives the equations of sensitivity analysis and charge-optimization from statistical thermodynamic expressions for the binding affinity of a ligand and a protein, allowing for molecular flexibility. The resulting expressions prove to be essentially the same as for a rigid ligand, except that Boltzmann-averaged electrostatic potentials replace the potentials computed for a single conformation of the ligand. The rigid ligand thus represents a special case in which the Boltzmann average includes only a single conformation having a probability of 1. However, accounting for ligand flexibility can strongly affect the results obtained with these methods, especially charge-optimization.

2.1. Derivatives of the Binding Free Energy. The standard free energy of association of a ligand L and receptor R to form a noncovalent complex can be written as the difference between the standard chemical potentials μ° of the respective molecular species

$$\Delta G^\circ = \mu^\circ_{RL} - \mu^\circ_R - \mu^\circ_L \quad (1)$$

where R, L, and RL indicate the receptor, the ligand, and their complex, respectively. The standard chemical potential of each molecular species in turn can be written as[15,16]

$$\mu^\circ_X = -RT \ln\left(\frac{8\pi^2}{C^\circ}\int e^{-\beta E(\mathbf{s},\mathbf{r})} d\mathbf{r}\right) \quad (2)$$

Here C° is the standard concentration which, combined with the factor of $8\pi^2$, accounts for the positional and orientational mobility of the free molecule at standard concentration;[15] $\beta \equiv 1/kT$, k being Boltzmann's constant and T the absolute temperature; $\mathbf{r}$ represents the internal coordinates of the molecular species and thus defines its three-dimensional conformation; and $E(\mathbf{s}, \mathbf{r})$ is the energy of the molecule as a function of its conformation and a set of computational parameters $\mathbf{s}$. In a typical force field-based calculation, $\mathbf{s}$ will include such solute parameters as atomic partial charges and Lennard-Jones parameters, along with solvent parameters such as the atomic partial charges of an explicit water model like TIP3P or the dielectric constant of an implicit solvent model. The derivative of the chemical potential with respect to atomic parameter s_i is

$$\frac{\partial \mu^\circ_X}{\partial s_i} = \frac{\int \frac{\partial E(\mathbf{s},\mathbf{r})}{\partial s_i} e^{-\beta E(\mathbf{s},\mathbf{r})} d\mathbf{r}}{\int e^{-\beta E(\mathbf{s},\mathbf{r})} d\mathbf{r}} = \left\langle \frac{\partial E(\mathbf{s},\mathbf{r})}{\partial s_i} \right\rangle \quad (3)$$

The quantity in angle brackets is the Boltzmann average of the derivative of the energy function with respect to the parameter of interest. This formula is consistent with prior expressions for free energy derivatives from Cieplak and coworkers.[17]

In the present application, we are interested in the case where the energy model includes a solvent-screened charge−charge interaction term and a solvent reaction field term, both linear with respect to the N atomic charges $\mathbf{q} \equiv (q_1, q_2, \ldots, q_N)$, along with other energy contributions that do not depend directly upon atomic charges. The other contributions can be lumped together as $E_{other}(\mathbf{s},\mathbf{r})$, where $\mathbf{s}$ now refers exclusively to noncharge parameters. The energy thus can be written as

$$E(\mathbf{q},\mathbf{s},\mathbf{r}) = \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{q_i q_j}{D^{eff}_{ij} r_{ij}} + \frac{1}{2}\sum_{i=1}^{N} q_i \phi^{RF}_i + E_{other}(\mathbf{s},\mathbf{r}) \quad (4)$$

where i and j index the molecule's atoms, D^{eff}_{ij} is the effective dielectric constant[18] of the interaction between atoms i and j, and ϕ^{RF}_i is the part of the solvent reaction field at atom i that is induced by the charge at atom i. The assumption of linearity implies furthermore that

$$\phi^{RF}_i \propto q_i \quad (5)$$

Note that the proportionality constant and the values of D^{eff}_{ij}

and r_{ij} depend on the atomic coordinates $\boldsymbol{r}$; i.e., upon the conformation of the molecule.

Substituting eq 5 into eq 4 and then using eq 3 to take the derivative of the chemical potential with respect to q_i yields that

$$\frac{\partial \mu^\circ}{\partial q_i} = \left\langle \sum_{j \neq i}^{N} \frac{q_j}{D_{ij}^{\mathrm{eff}} r_{ij}} + \phi_i^{\mathrm{RF}} \right\rangle = \langle \phi_i \rangle \tag{6}$$

Thus, the derivative of the molecule's chemical potential with respect to the charge of atom i is the Boltzmann-averaged electrostatic potential at atom i. This observation is not surprising, but it is useful because it generalizes what is commonly known for a rigid molecule to the case of one that is flexible. The rigid molecule becomes a special case of eq 6, in which the Boltzmann average of the potential equals the potential computed for a single conformation.

Finally, eq 6 can be combined with eq 1 to show that the derivative of the binding free energy with respect to the charge of atom i, which can belong to either the ligand or the receptor, is simply

$$\frac{\partial G^\circ}{\partial q_i} = \langle \phi_i^{\mathrm{b}} \rangle - \langle \phi_i^{\mathrm{f}} \rangle \tag{7}$$

where the superscripts "b" and "f" indicate respectively the bound (RL) and free (R or L) states of the system. Sensitivity analysis can then be used to generate a first-order prediction of the change in binding energy for a small change in an atomic charge, Δq_i:

$$\Delta G \approx \frac{\partial \Delta G}{\partial q_i} \Delta q_i \tag{8}$$

The present analysis is consistent with an earlier and more general discussion of the application of sensitivity analysis to free energies.[12]

2.2. Electrostatic Optimization. The theory of electrostatic optimization for rigid molecules has been elegantly laid out and explored by its originators; see, e.g., refs 3 and 4. Here, the binding energy derivatives discussed in section 2.1 are employed to generalize the formalism of electrostatic optimization to the case of flexible molecules.

Consider a ligand L with N atoms of charge q_i, $i \in [1...N]$ and a receptor R with M atoms of charge q_i, $i \in [N+1...N+M]$. We wish to find the values of the ligand charges that minimize the binding energy and thus maximize affinity, subject to the physically reasonable constraint that the total charge of the ligand, Q, equals a user-specified integer; i.e., that

$$g(\boldsymbol{q}) = \sum_{i}^{N} q_i - Q = 0 \tag{9}$$

Equation 9 introduces the function $g(\mathbf{q})$ which is used to define the constraint on total charge. A set of charges that minimizes the binding energy subject to this constraint can be found by the method of Lagrangian multipliers (e.g., ref 19 p 946), which leads to a system of N equations, one for each atomic charge q_i

$$\frac{\partial \Delta G^\circ}{\partial q_i} + \lambda \frac{\partial g}{\partial q_i} = \frac{\partial \Delta G^\circ}{\partial q_i} + \lambda = 0 \tag{10}$$

where λ, the undetermined multiplier, represents an additional unknown. Supplementing eqs 10 with the constraint on the total charge, eq 9, yields a system of $N+1$ equations in $N+1$ unknowns. It is important to note that, although eq 10 is a necessary condition for a set of charges to minimize the binding free energy while meeting the constraint on total charge, it is not sufficient to guarantee an energy-minimum, because a set of charges that satisfies eq 10 could also be a maximum or a saddle, at least in principle. The nature of the stationary point will be determined by the specific molecular problem and the parameters of the calculation.

Interestingly, eq 10 can be rewritten with the aid of eq 7 as

$$\langle \phi_i^{\mathrm{b}} \rangle - \langle \phi_i^{\mathrm{f}} \rangle = -\lambda \tag{11}$$

This implies that the charges from the Lagrangian procedure also cause all atoms of the ligand to experience the same change in mean potential upon binding and that this change in potential is $-\lambda$. This equation generalizes the concept of a residual potential at each atom[4] which goes to zero for a ligand which minimizes the energy in the absence of any constraint on the total charge. Equation 11 shows that the residual potential goes to $-\lambda$ rather than zero when the constraint is imposed.

Assuming the classical properties of linearity and reciprocity allows eqs 10 to be rewritten as a set of N linear equations. Working from eqs 5 and 7, we define

$$a_{ii} \equiv \frac{\langle \phi_i^{\mathrm{RF}} \rangle}{q_i} \tag{12}$$

$$a_{ij} \equiv \langle (D_{ij}^{\mathrm{eff}} r_{ij})^{-1} \rangle \tag{13}$$

where a_{ii}^{b}, a_{ij}^{b} and a_{ii}^{f}, a_{ij}^{f} refer to values for the bound and free states of the ligand, respectively. Substituting these terms into eq 6 and then using eq 11 yields

$$q_i(a_{ii}^{\mathrm{b}} - a_{ii}^{\mathrm{f}}) + \sum_{j \neq i}^{N} q_j(a_{ij}^{\mathrm{b}} - a_{ij}^{\mathrm{f}}) + \sum_{j=N+1}^{N+M} q_j(a_{ij}^{\mathrm{b}} - a_{ij}^{\mathrm{f}}) + \lambda = 0 \tag{14}$$

for $i = (1, 2, \ldots, N)$. The first term is the change upon binding of the reaction field at atom i due to charge q_i; the second term is the change upon binding of the screened Coulomb potential at atom i due to other ligand atoms j; and the third term is the change in the potential at atom i of the ligand due to the charges j of the receptor, ϕ_i^{b}. Note that the ligand does not feel the receptor when they are not bound, so $a_{ij}^{\mathrm{f}} = 0$.

Equations 9 and 14 can be expressed together in matrix form

$$\begin{pmatrix} a_{11}^{b} - a_{11}^{f} & a_{12}^{b} - a_{12}^{f} & \cdots & a_{1N}^{b} - a_{1N}^{f} & 1 \\ a_{21}^{b} - a_{21}^{f} & a_{22}^{b} - a_{22}^{f} & \cdots & a_{2N}^{b} - a_{2N}^{f} & 1 \\ & & \cdots & & \\ a_{N1}^{b} - a_{N1}^{f} & a_{N2}^{b} - a_{N2}^{f} & \cdots & a_{NN}^{b} - a_{NN}^{f} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \cdots \\ q_N \\ \lambda \end{pmatrix} = \begin{pmatrix} -\phi_1^{b} \\ -\phi_2^{b} \\ \cdots \\ -\phi_N^{b} \\ Q \end{pmatrix} \tag{15}$$

or

$$\mathbf{Aq} = \mathbf{B} \tag{16}$$

The equation represented in the lowest row of the matrices is the constraint on the total charge (eq 9). A typical druglike ligand possesses on the order of 100 atoms, so the dimensions of $\boldsymbol{A}$ are roughly 100 × 100. This matrix equation can be readily solved by matrix inversion or by iterative methods to yield a stationary point $\mathbf{q}^\circ$ on the hyperplane of net charge Q, along with the corresponding value of λ which, as noted earlier in this section, equals the change in potential on binding when the charges equal $\mathbf{q}^\circ$.

The change in the solvation and the intramolecular charge−charge interactions of the ligand upon binding is $\mathbf{q}^{\circ T}\mathbf{A}\mathbf{q}^\circ$, while the interaction of the ligand with the protein is $\mathbf{q}^{\circ T}\mathbf{B}$, where $\mathbf{q}^\circ \equiv (q_1, q_2, \ldots, q_N, 0)$. The change in the protein solvation energy upon binding also contributes to the overall electrostatic binding energy, but this quantity is independent of the ligand charges. Therefore, charge optimization focuses on the change in ligand−ligand and ligand−protein energies:

$$\Delta G_{ll,lp} = \mathbf{q}^{\circ T}(\mathbf{A}\mathbf{q}^\circ + \mathbf{B}) \tag{17}$$

The full change in electrostatic energy can be obtained by separately computing the change in protein solvation energy, ΔG_{pp}, and adding it to $\Delta G_{ll,lp}$.

3. Methods

3.1. Molecular Systems and Parameters. Charge optimization and sensitivity analysis were studied for a model system, the association of HIV-1 protease with the cyclic urea inhibitor XK263.[20] The protein was fixed in its crystal conformation,[21] and a single conformation of the bound ligand was considered. However, three conformations were considered for the free ligand: a conformation identical to the bound conformation and two alternate conformations. Thus, the binding processes considered are as follows:

Rigid Ligand: $\text{LigConf}^0 + \text{Protein} \rightarrow \text{LigConf}^0\cdot\text{Protein}$

Flexible Ligand 1: $\text{LigConf}^1 + \text{Protein} \rightarrow \text{LigConf}^0\cdot\text{Protein}$

Flexible Ligand 2: $\text{LigConf}^2 + \text{Protein} \rightarrow \text{LigConf}^0\cdot\text{Protein}$

Here LigConf^0 is the bound conformation of the ligand, and LigConf^1 and LigConf^2 are the two alternate conformations of the free ligand, which will be referred to as Flex 1 and Flex 2.

The molecular models were prepared as follows. Polar hydrogen atoms were added to the crystal structure, and CHARMM[22] force-field parameters were assigned with the program Quanta.[23] The cyclic urea inhibitor was relaxed by redocking it to the original crystal structure of the protease with the program Vdock,[24,25] which allows continuous variation of nonring single-bonds and of the overall position and orientation of the compound within the binding site. The resulting conformation (LigConf^0) was used in calculating electrostatic terms for the receptor−ligand system and for the free ligand when it was considered as rigid. The two alternative conformations of the free ligand, Flex 1 and Flex 2, were generated by running on the order of 10−100 ps of stochastic dynamics[26] at 300 K for the ligand alone, with a time-step of 1 fs. During the MD calculations, screening of electrostatic interactions by solvent was accounted for in an approximate fashion via the distance-diependent dielectric model with a coefficient of 4.

Electrostatic terms were obtained from finite-difference solutions of the linearized Poisson−Boltzmann equation carried out with the program UHBD,[27] as detailed in section 3.2. The dielectric constants of the solutes and solvent were set to 1 and 78.5, respectively, with solvent ionic strength of 150 mM and an ion-exclusion radius (Stern layer) of 2 Å thickness. The boundary between the low dielectric interior and the high dielectric solvent was defined by the molecular surface[28] with a probe radius of 1.4 Å and atomic radii set to their CHARMM Lennard-Jones R_{min} values.

3.2. Calculation of the Electrostatic Terms. To set up the electrostatic optimization problem for a given system, it is necessary to obtain the values of a_{ij}^{b}, a_{ij}^{f}, and ϕ_i^{b}, where i, $j \le N$. (See eq 14.) These terms were obtained here with a series of finite-difference Poisson−Boltzmann[29−31] (FDPB) calculations. In each case, potentials were computed via an initial FDPB calculation with a coarse grid spacing of 0.5 Å and then a second "focusing"[31] FDPB calculation with a grid spacing of 0.2 Å, where the grid encompassed the entire ligand, and boundary conditions were drawn from the initial coarse grid run. These quantities were computed either under the assumption of a rigid ligand or a flexible ligand whose conformation changed from one bound conformation to a single different free conformation. However, as discussed in the Theory section, it would also be possible to compute Boltzmann averages over multiple bound and/or free conformations.

The **B** vector was filled with values of ϕ_i^{b} by a single FDPB calculation for the ligand−receptor complex in which all ligand charges were artificially set to zero, but all protein charges were kept at their normal values. The resulting electrostatic potential at atom i is ϕ_i^{b}.

The **A** matrix was filled by computing the values of a_{ij}^{b} and a_{ij}^{f}; see eqs 12 and 13. The calculations for a_{ij}^{b} are the same as those for a_{ij}^{f} except that the former use the bound conformation of the ligand in the presence of the protein, which is treated as electrically neutral, while the latter use the free conformation of the ligand in the absence of the protein. (As noted above, the free conformation is the same as the bound conformation when the ligand is assumed to be rigid.) Hence, the following description refers to a_{ij}^{b} and

a_{ij}^{f} generically as a_{ij}. The values of a_{ij} were computed as the sum of Coulombic potentials ϕ_{ij}^{C} based on the dielectric constant of the protein interior and reaction field potentials produced by the solvent. The Coulombic potentials were computed by placing a unit charge on atom i and zeroing all other charges, setting the dielectric constant of the solvent region to the dielectric constant of the molecular interior, and evaluating the resulting potentials at every other atom. Self-interactions ($i = j$) were omitted, along with interactions between atoms directly bonded to each other and interactions between atoms in a 1−3 bonding relationship. These exclusions are standard practice in force field calculations, and here they avoid allowing nonphysical short-ranged interactions to influence the charge optimization procedures. Interactions across a dihedral angle (1−4 interactions) were included without any special scaling factor. The solvation parts of a_{ij} were computed by carrying out an additional FDPB calculation with a unit charge on atom i and all other charges zeroed but now with the solvent dielectric constant set to the solvent value. The solvation parts of a_{ij} were then set to the difference between the solvated and the unsolvated potential at atom j. Note that this same procedure gives the correct value for a_{ii}.

3.3. Numerical Methods. Matrices were diagonalized with dsyev and associated subroutines, and matrix equations were solved with dgesv and associated subroutines, all drawn from LAPACK.[32] In prior applications of Lagrangian charge optimization, the values of q_i have been constrained to lie within a range that is typical for current empirical force fields; e.g., $q_i \leq 0.85$. No such constraint was applied here, however.

When the top left $N \times N$ submatrix of the **A** matrix possesses negative eigenvalues, the charges provided by the method of Lagrangian multipliers may not correspond to a minimum of the electrostatic energy; they could also represent a saddle point or a local maximum. In such cases, charges that minimize the electrostatic energy were sought with the minimization program PRAXIS,[33] obtained from the Netlib repository of mathematical software.[34] More particularly, PRAXIS was used to minimize a quantity consisting of $\Delta G_{ll,lp}$ (eq 17) supplemented with a pseudoenergy term which restrains the absolute value of the total charge to zero. In some calculations, additional pseudoenergy terms were included to keep individual atomic charges in the range −0.85 to 0.85, as previously proposed.[6] Note that $\Delta G_{ll,lp}$ can be computed very quickly for a given set of ligand charges by using the precalculated **A** and **B** matrices in eq 17, so these minimizations are not overly time-consuming. The results of minimization with the PRAXIS algorithm can depend on the initial guess for the values of the charges.

4. Results

This section describes the properties of sensitivity analysis and charge-optimization when the free conformation of the ligand is assumed to be the same as that of the bound conformation and when the free ligand is considered to adopt a different conformation.

4.1. Rigid Ligand. 4.1.1. Sensitivity Analysis. Table 1, columns 2 and 4, shows the derivatives $\partial\Delta G/\partial q_i$ of the binding energy with respect to each atomic charge i when the free ligand conformation is considered to be the same as the bound conformation; i.e., for the assumption of a rigid ligand. Atom labels are listed in Figure 1. Nearly all the derivatives are negative, indicating that making the atomic charge more positive will favor binding. This is a conse-

Table 1. Derivatives of Binding Free Energy with Respect to Partial Atomic Charges of Ligand Atoms for Rigid Ligand (kcal/mol/au), i.e., with Free Conformation of Ligand Same as Bound Conformation

atom i	$\partial\Delta G°.\partial q_i$	atom i	$\partial\Delta G°.\partial q_i$	atom i	$\partial\Delta G°.\partial q_i$
C1	−26.58	C37	−42.51	H11	−37.28
O1	−7.63	C61	−59.54	H12	−28.18
N2	−38.44	C62	−39.58	H13	−16.18
C2	−23.81	C63	−36.87	H14	−8.10
C3	−67.09	C64	−11.40	H15	−7.12
C4	−102.04	C65	−10.09	H16	−12.06
O4	−129.84	C66	−12.74	H17	−27.51
C5	−97.85	C67	−21.12	H18	−47.88
O5	−120.70	C70	−31.19	H19	−88.61
C6	−63.66	C71	−29.86	H20	−10.71
N7	−37.24	C72	−26.47	H21	2.03
C7	−23.48	C73	−18.29	H22	−3.00
C20	−32.88	C74	−14.76	H23	−7.53
C21	−36.87	C75	−15.77	H24	−60.23
C22	−32.66	C76	−18.91	H25	−47.69
C23	−27.48	C77	−26.61	H26	−85.14
C24	−17.72	C78	−29.91	H27	−55.48
C25	−11.92	C79	−32.81	H28	10.81
C26	−13.10	H1	−11.06	H29	−2.74
C27	−18.08	H2	−15.40	H30	−3.06
C28	−27.12	H3	−67.54	H31	−12.00
C29	−30.08	H4	−103.80	H32	−29.57
C31	−61.58	H5	−150.63	H33	−12.10
C32	−41.25	H6	−94.64	H34	−8.23
C33	−19.72	H7	−142.94	H35	−19.07
C34	−10.02	H8	−62.18	H36	−16.80
C35	−11.24	H9	−11.18	H37	−25.66
C36	−19.85	H10	−16.86	H38	−31.35

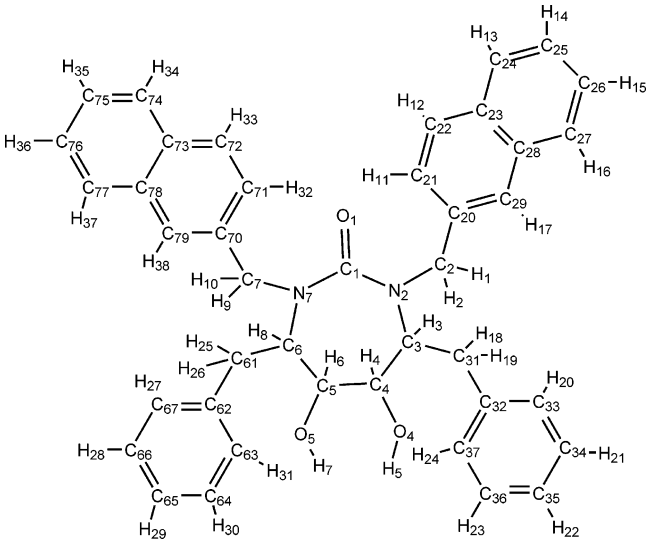

Figure 1. Diagram of ligand XK263 with atom codes used in tables.

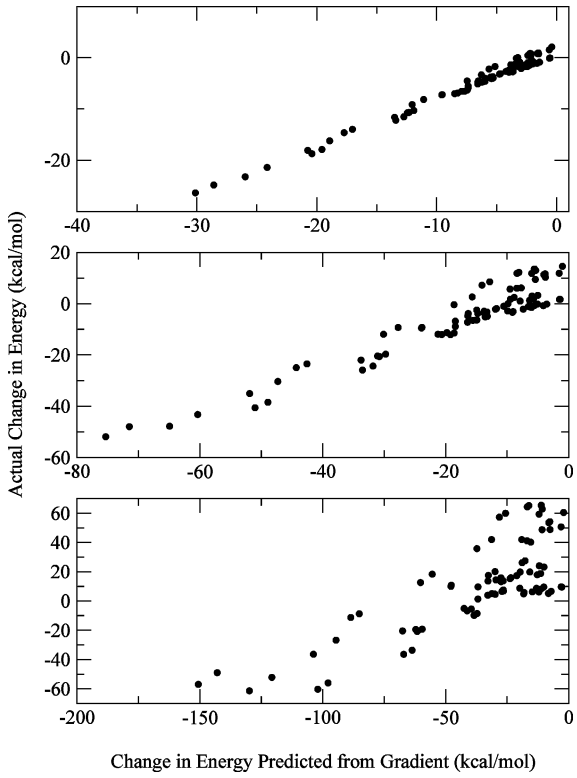

Figure 2. Accuracy of energy predictions from sensitivity analysis when the ligand is assumed rigid, shown as scatter plots of the change in electrostatic energy (kcal/mol) computed with the full parabolic energy surface versus the change predicted by sensitivity analysis, for charge changes of 0.2 au (top), 0.5 au (middle), and 1.0 au (bottom).

quence of the dominant influence of the two negatively charged aspartyl groups in the active site of the protease, which produce a positive potential at virtually every atom of the ligand.

The accuracy of sensitivity analysis is assessed here by changing each ligand charge by a small amount Δq_i in a direction opposite to the local derivative and using eq 8 to estimate the resulting change in binding energy. Figure 2 compares these first-order predictions with the actual value of $\Delta\Delta G_{ll,lp}$ computed with eq 17. Comparisons are shown for charge changes of 0.2 au (top), 0.5 au (middle), and 1.0 au (bottom). It is evident that the smaller charge changes almost always improve the binding energy, and the linear predictions from the gradients correlate well with the actual results. However, for charges changes of 1.0 au, many of the charge changes make the binding energy more positive, and the energy predictions are quite poor. This is a consequence of the quadratic dependence of energy upon charge: large charge changes often cross the energy minimum and climb the far side of the parabola.

4.1.2. Charge Optimization. The eigenvalues associated with the top left $N \times N$ submatrix of **A** obtained when the ligand is assumed to be rigid are plotted in ascending order in Figure 3 (black line). All eigenvalues are positive, indicating that the charges from the method of Lagrangian multipliers will represent not only a stationary point of the binding energy but also a minimum. (A maximum is mathematically possible but unlikely physically because the

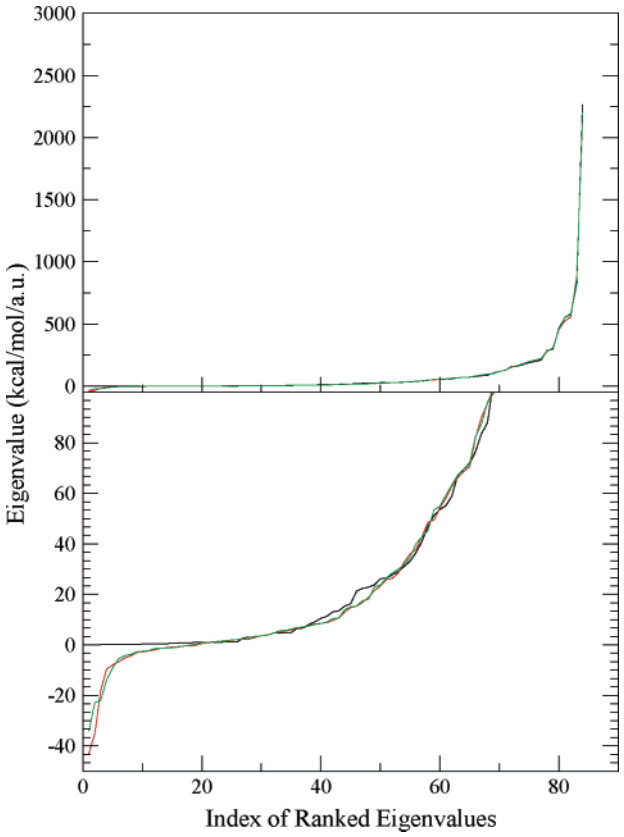

Figure 3. Rank-ordered eigenvalues of top left $N \times N$ submatrix of **A** matrix when ligand is assumed rigid (black), and when two variant conformations of the free ligand are assumed (red, green). (See Flex 1 and Flex 2 data in Table 2.) The bottom graph is the same as the top graph except for the scale of the ordinate.

desolvation energy of the ligand upon binding, which depends quadratically upon charge, is expected to be unfavorable, leading to a parabolic energy function with upward curvature.) Table 2 (columns 3 and 10) lists these optimal charges. Although the individual charges are not constrained, only a few are larger than 1 au in magnitude and none are greater than 1.6 au. Table 2 (columns 2 and 9) lists the ligand charges assigned by Quanta/CHARMM, for comparison. Not surprisingly, there is no evident correlation with the optimal charges in columns 2 and 9.

Table 2 also compares $\Delta G_{ll,lp}$, the change in electrostatic energy upon binding less the protein desolvation energy, for the optimized charges and the CHARMM charges. Optimizing the charges dramatically lowers the energy, from −15.1 kcal/mol to −113 kcal/mol. The full change in electrostatic energy can be obtained by adding the cost of desolvating the protein, ΔG_{pp}; a separate calculation yields a value of 98.4 kcal/mol for this final part of the energy, yielding a net change in electrostatic energy of −14.6 kcal/mol. These results are consistent with previous work noting that, although continuum electrostatics models tend to yield an unfavorable binding energies, the electrostatic energy can contribute favorably to binding if the charges are right.[3,4]

Like sensitivity analysis, the charge optimization methodology can be used to guide chemical modifications of a ligand aimed at improving its affinity. Figure 4 evaluates

Table 2. Partial Atomic Charges (au, in Top 42 lines) and Associated Electrostatic Contributions to the Binding Free Energy (kcal/mol) for Three Assumptions about the Conformation of the Ligand in the Free State (Bottom 3 Lines on Right-Hand Side)[a]

atom	CHARMM	optimized (rigid)	stationary point (Flex 1)	stationary point (Flex 2)	min. 1 (Flex 1)	min. 2 (Flex 1)	atom	CHARMM	optimized (rigid)	stationary point (Flex 1)	stationary point (Flex 2)	min. 1 (Flex 1)	min. 2 (Flex 1)
C1	0.600	1.559	−0.008	−1.647	−0.850	0.850	C78	−0.129	−0.121	8.005	−2.198	0.336	0.360
O1	−0.550	−0.723	−1.086	0.154	−0.846	−0.061	C79	−0.129	0.173	−1.478	2.410	−0.846	0.148
N2	−0.250	−1.699	3.928	2.835	−0.845	0.239	H1	0.051	−0.004	−0.181	−0.051	0.791	−0.850
C2	0.091	0.566	−3.503	−0.640	−0.281	0.479	H2	0.051	−0.187	−0.070	0.148	0.850	−0.850
C3	−0.099	1.095	4.273	−3.435	0.433	0.716	H3	0.101	−0.276	−3.150	0.673	0.848	−0.850
C4	0.190	0.483	−2.430	−4.827	0.785	0.577	H4	0.061	0.045	0.760	1.568	0.559	0.211
O4	−0.650	−0.544	2.397	2.253	−0.706	−0.764	H5	0.400	0.598	−0.340	−0.583	0.582	0.674
C5	0.190	0.646	−6.136	−0.720	−0.112	0.850	H6	0.061	−0.133	2.029	−0.155	0.248	0.031
O5	−0.650	−0.769	−0.741	0.748	−0.796	−0.483	H7	0.400	0.663	1.503	1.273	0.830	0.366
C6	−0.099	1.099	3.019	2.946	−0.498	−0.388	H8	0.101	−0.270	2.005	−2.289	0.850	−0.850
N7	−0.250	−1.197	1.870	1.275	−0.850	−0.396	H9	0.051	0.000	−0.592	0.107	0.418	−0.850
C7	0.091	0.394	0.746	−1.794	0.236	0.492	H10	0.051	−0.160	0.604	0.595	0.850	−0.850
C20	0.001	−0.541	−0.191	3.592	−0.850	0.844	H11	0.131	−0.016	−1.944	1.854	0.196	−0.850
C21	−0.129	0.242	5.369	−6.820	−0.848	0.823	H12	0.131	0.045	−1.224	−2.102	0.399	−0.179
C22	−0.129	−0.137	0.967	9.451	−0.662	−0.026	H13	0.131	0.203	−0.984	−0.023	0.012	−0.083
C23	0.000	0.272	−2.577	−6.835	0.847	0.654	H14	0.131	0.266	1.707	−1.018	0.075	0.174
C24	−0.129	−0.295	5.402	0.350	−0.137	−0.530	H15	0.131	−0.293	−3.158	1.617	0.256	−0.556
C25	−0.129	−0.675	−6.798	4.066	−0.029	−0.474	H16	0.131	−0.066	0.064	−1.183	−0.157	−0.094
C26	−0.129	0.717	6.173	−6.145	−0.483	0.791	H17	0.131	0.289	0.940	−0.671	0.818	0.747
C27	−0.129	−0.200	0.612	6.736	0.257	−0.705	H18	0.051	0.112	0.165	−0.157	−0.101	0.850
C28	0.000	−0.191	−4.932	−0.662	−0.312	0.081	H19	0.051	0.282	0.539	1.128	0.850	−0.318
C29	−0.129	0.281	2.136	−1.460	−0.288	0.085	H20	0.131	−0.021	2.987	0.254	0.056	0.785
C31	−0.099	−0.622	−3.569	−0.673	−0.744	0.009	H21	0.131	−0.016	0.082	0.683	0.381	−0.253
C32	0.001	−0.077	−1.849	2.175	−0.773	−0.116	H22	0.131	−0.374	−0.647	−0.261	0.363	0.002
C33	−0.129	0.123	−0.799	−0.956	−0.410	−0.258	H23	0.131	−0.025	0.585	−0.896	0.111	0.177
C34	−0.129	−0.173	−1.981	−1.800	−0.407	0.023	H24	0.131	0.183	−0.592	0.657	0.512	−0.107
C35	−0.129	0.445	2.466	1.128	−0.048	−0.587	H25	0.051	0.119	−0.213	−0.839	0.714	0.850
C36	−0.129	−0.484	−4.401	1.563	−0.355	−0.849	H26	0.051	0.256	2.569	1.315	0.727	−0.445
C37	−0.129	0.218	5.025	−1.393	−0.324	0.452	H27	0.131	0.364	−0.857	−0.555	0.648	0.156
C61	−0.099	−0.709	−5.480	−1.806	−0.849	0.474	H28	0.131	−0.207	0.217	−1.438	0.099	−0.013
C62	0.001	0.203	1.399	−0.584	−0.845	0.097	H29	0.131	−0.472	4.420	−0.550	0.381	−0.011
C63	−0.129	−0.304	2.560	1.308	−0.034	0.306	H30	0.131	0.014	−2.298	0.239	0.515	−0.072
C64	−0.129	−0.344	−1.554	2.325	−0.413	−0.845	H31	0.131	−0.048	3.324	−0.220	−0.849	0.107
C65	−0.129	0.643	−6.252	−0.464	−0.580	−0.503	H32	0.131	0.332	1.252	−1.748	0.695	0.499
C66	−0.129	−0.232	6.064	−0.743	0.222	−0.096	H33	0.131	−0.128	−0.384	−1.502	0.241	−0.380
C67	−0.129	0.107	−5.391	1.244	−0.331	0.072	H34	0.131	−0.368	−0.579	−3.557	0.748	0.251
C70	0.001	−0.534	−3.400	0.430	−0.407	0.850	H35	0.131	0.530	−0.427	2.138	0.328	0.456
C71	−0.129	0.315	−1.799	2.497	−0.850	0.727	H36	0.131	0.131	0.190	0.787	−0.005	−0.051
C72	0.000	−0.220	−0.896	−3.430	−0.439	0.490	H37	0.131	0.051	−2.334	0.262	0.074	−0.126
C73	−0.129	−0.186	2.585	3.400	0.168	−0.039	H38	0.131	−0.019	−0.577	−0.455	0.557	−0.650
C74	−0.129	0.999	−2.084	4.949	−0.692	−0.271	ΔG^{rigid}	−15.1	−113.	826.	337.	126.	111.
C75	−0.129	−1.184	2.762	−5.547	−0.377	−0.665	ΔG^{Flex1}	−17.9	−113.	−87.1	176.	−325.	−343.
C76	−0.129	−0.008	−2.024	−0.171	−0.027	−0.062	ΔG^{Flex2}	−14.6	−114.	337.	−87.5	−101	−225.
C77	0.000	0.190	−1.792	1.870	0.338	−0.420							

[a] CHARMM: CHARMM charges of the ligand XK263. Optimized (rigid): charges optimized by solving eq 15 when the ligand is assumed to have the same conformation in free and bound states. Stationary point (Flex 1): charges obtained by solving eq 15 when the ligand is assumed to adopt conformation Flex 1 in the free state. Stationary point (Flex 2): same for the second alternative conformation of the free ligand, Flex 2. Minimized 1 (Flex 1): charges obtained by PRAXIS minimization of $\Delta G_{\text{ll,lp}}$, when the free ligand is in conformation Flex 1; added energy terms restrain net charge of the ligand to 0 and the charge of each atom to $|q_i| \leq 0.85$. Minimized 2 (Flex 1): a second set of charges obtained by the same method but starting from a different initial guess at the charges. $\Delta G_{\text{ll,lp}}^{\text{rigid}}$: energy obtained by substituting each set of charges into eq 17 when the **A** matrix is based upon assumption that the ligand is rigid, so the free conformation is the same as the bound conformation. $\Delta G_{\text{ll,lp}}^{\text{Flex1}}$: same, when the **A** matrix is computed assuming the free ligand is in conformation Flex 1. $\Delta G_{\text{ll,lp}}^{\text{Flex2}}$: same, when the **A** matrix is computed assuming the free ligand is in conformation Flex 2.

this concept by showing how the electrostatic energy of binding $\Delta G_{\text{ll,lp}}$ varies when all atoms of XK263 are changed gradually from their CHARMM charges to their optimal values (blue) and also when the charge of each individual atom is changed to its value in the optimal set of charges (black and red lines), while the other charges are held fixed at their CHARMM values. The parabolic shapes of these graphs are as expected from the theory (see section 2.2). As

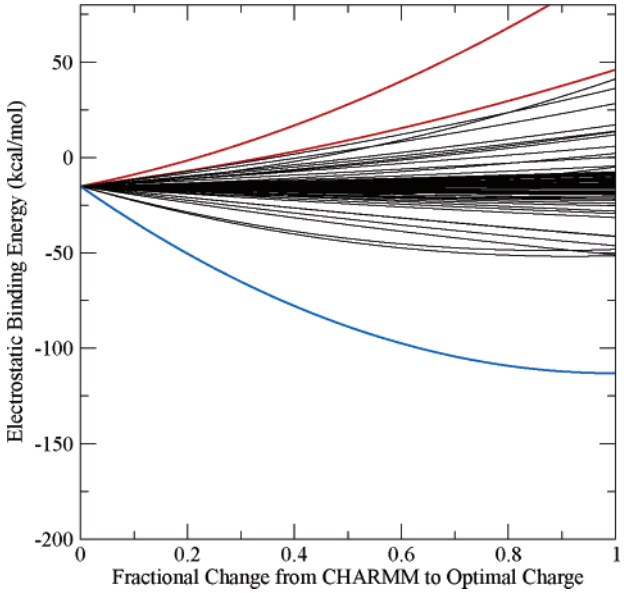

Figure 4. Electrostatic part of binding energy for rigid ligand as a function of the fractional shift from the initial CHARMM charges of XK263 (columns 2 and 9 of Table 2) toward charges optimized for the rigid ligand (columns 3 and 10 of Table 2). **Red**: energy change when each nitrogen atom's charge is varied, with all other atomic charge held fixed. **Black**: same as black graph, for the other 82 atoms of XK263. **Blue**: consequences of changing all ligand charges simultaneously to their optimal values.

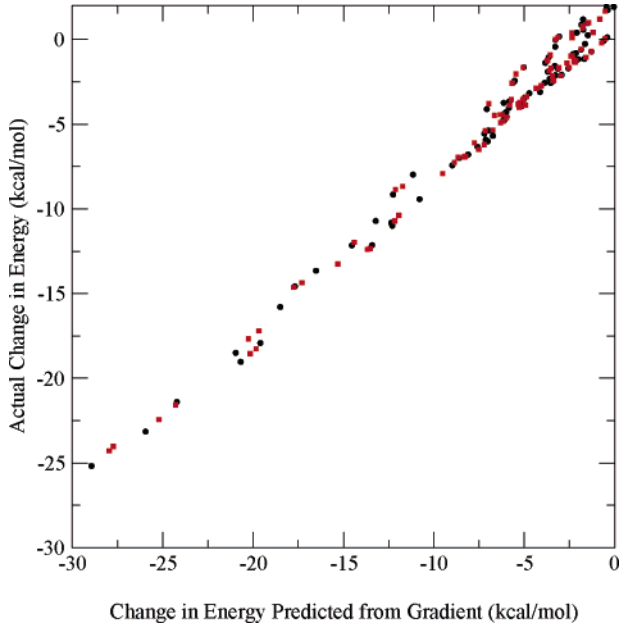

Figure 5. Accuracy of predictions from sensitivity analysis when the ligand is assumed flexible, shown as scatter plots of the change in electrostatic energy computed with the full parabolic energy surface versus the change predicted by sensitivity analysis, for charge changes of 0.2 au. Results are shown in black and red for two alternative free conformations, Flex 1 and Flex 2.

noted in the previous paragraph, changing all the charges to their optimal values produces a very large improvement in the electrostatic energy change upon binding. However, adjusting the charges of an individual atom toward its optimal value does not always improve the electrostatic binding energy. In fact, some changes markedly increase the energy and thus oppose binding.

For example, the two red lines correspond to the nitrogen atoms of XK263, which are situated roughly 5.5 Å from the aspartate groups at the bottom of the active site and roughly 4.2 Å from the amide hydrogens of the flaps at the top of the active site. The CHARMM force field assigns both nitrogens charges of 0.25 au, but the optimal charges are −1.7 and −1.2 au. (See Table 2.) It was initially surprising that optimization directs both atoms to be considerably more negative than their CHARMM values, even though the nearest protein ions are the negative aspartates, and both nitrogens as a consequence have strongly negative energy derivatives $\partial\Delta G/\partial q_i$. (See Table 1.) The explanation appears to be that charge optimization causes atoms closer to the aspartates, and with even more strongly negative derivatives than the nitrogens, to gain substantial positive charge, and these new positive charges effectively shield the nitrogens from the aspartates. In particular, atoms C3 and C6 (Figure 1) change from 0.099 au to 1.1 au, while atoms C4 and C5 also become significantly more positive. (See Table 2.) The resulting large increase in the positive charge situated between the aspartate groups and the nitrogens tends to drive the charges of the nitrogens in the positive direction in the final optimized charge set. This example explains why shifting a subset of the ligand's charges toward their optimal values may not improve the calculated binding affinity, in accord with previous observations.[9]

4.2. Flexible Ligand. Two alternative free conformations of the free ligand XK263 were generated by a simple molecular dynamics approach, as described in Methods, and were used separately as a basis for examining the consequences of ligand flexibility for sensitivity analysis and charge optimization.

4.2.1. Free Energy Derivatives. Using different conformations for the free ligand produces little change in the nature of the free energy derivative results. In particular, the changes in energy due to 0.2 au changes in the charges of single atoms are still well predicted by the derivatives, as shown in Figure 5. Moreover, the energy derivatives themselves are quite similar to those obtained under the assumption of a rigid ligand, as shown in Figure 6.

4.2.2. Charge Optimization. The consequences of ligand flexibility for charge optimization are more complex. First, as shown in Figure 3, the top left $N \times N$ submatrix of **A** matrix now possesses roughly 25 negative eigenvalues when either of the new free conformations is considered (red, green graphs). This implies that the stationary point of the energy surface in the absence of any constraints is a multidimensional saddle point. Thus, there is no optimal set of charges, at least in the case where the atomic charges are completely unconstrained. Instead the electrostatic binding energy can decrease without bound as a function of the charges assigned to the ligand.

Nonetheless, the method of Lagrangian multipliers can still be used to obtain a set of charges which is a stationary point of the energy on the multidimensional hyperplane defined by the constraint $\sum_i^N q_i = Q$. Columns 4, 5, 11, and 12 of

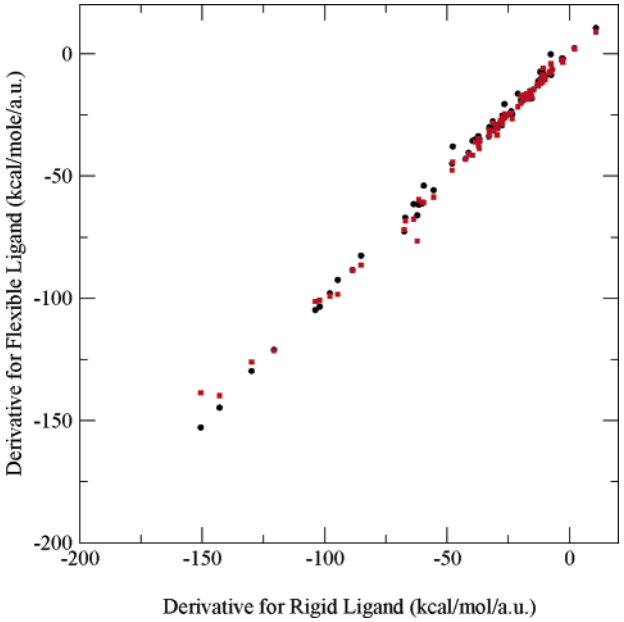

Figure 6. Comparison of free energy derivatives from the sensitivity analysis for two alternative conformations of free ligand, Flex 1 (black) and Flex 2 (red), with derivatives when the ligand is assumed rigid.

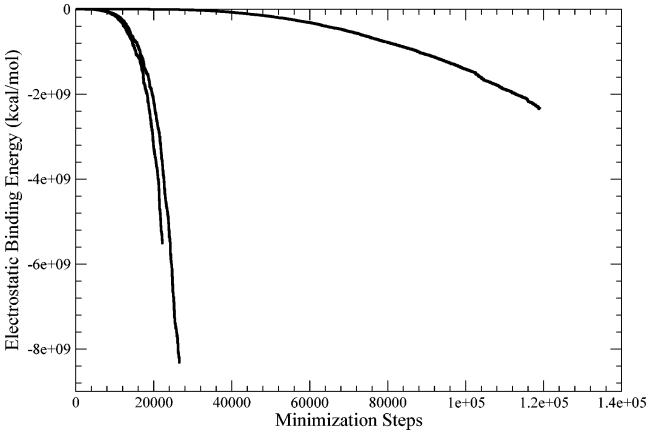

Figure 7. Electrostatic energy as a function of the number of PRAXIS minimization steps, when the free ligand conformation is taken to be different from the bound conformation (Flex 1) and the total ligand charge is restrained to zero. Results are shown for 3 different initial guesses of the atomic charges. The rightmost curve was obtained with a stronger restraining potential on the total ligand charge than the other two curves.

Table 2 present these stationary charges for the two free ligand conformations; they may be compared with those obtained with the assumption of a rigid ligand (columns 3 and 10). The mean absolute values of the stationary charges obtained with the two alternative free conformations are substantially larger, about 2 au, than the optimal charges based upon the assumption of a rigid ligand, about 0.4 au. There is no discernible correlation among the various sets of charges. These results show that changing the free conformation assumed for the ligand can lead to markedly different charge sets when the method of Lagrangian multipliers is used.

The stationary charges for each free ligand conformation can still lead to markedly improved binding energies, as shown at the foot of Table 2. Thus, when the stationary charges for the first flexible conformation are used to compute the binding energy based upon this free conformation, the result is −87.1 kcal/mol; the corresponding energy for the second conformation is similar, at −87.5 kcal/mol. Oddly, however, when these charge sets are used to compute energies under the assumption of alternative free conformations, very poor energies are obtained. For example, when the stationary charges obtained with the first flexible conformation (Table 2, columns 4 and 11) are used with the assumption of a rigid ligand, the energy becomes 826. kcal/mol; and when the same charges are used to compute the binding energy with the free conformation set to the second flexible conformation, the energy becomes 337. kcal/mol.

In contrast, when charges are optimized under the assumption of a rigid ligand (columns 3 and 10), the resulting energies are remarkably stable and favorable at about −113. kcal/mol across all choices of the free conformation. Interestingly, the CHARMM charges also give rather stable energies across all three free conformations, with values of −15 to −18 kcal/mol. In summary, the stationary charges obtained with the two alternate free conformations yield energies that depend strongly upon conformation. In contrast, optimal charges based upon the assumption of a rigid ligand provide a larger improvement in the binding energy, and this improvement is far less sensitive to the choice of free conformation.

This analysis has relied so far on Lagrangian multipliers to find ligand charges that optimize binding energy subject to the constraint on total charge. However, as noted above, it is possible that the binding energy has no lower bound, even when the total charge is constrained, because the top left $N \times N$ submatrix of **A** matrix possesses negative eigenvalues. This possibility is tested here by using a numerical algorithm (PRAXIS;[33] see Methods) to seek charge sets that sum to zero and yield highly favorable binding free energies. Figure 7 shows the results of three such minimizations started with different initial guesses for the atomic charges; one applies a much stronger restraining potential to the net charge of the ligand. The electrostatic binding energies are found to decline precipitously, reaching values as low as -8×10^{-9} kcal/mol with no sign of reaching an asymptote. It is important to note that the total charge was successfully locked near zero: the net charge of the ligand at the end of the three minimizations graphed in Figure 7 are 0.02, 0.0009, and 0.00007 au. However, the atomic charges obtained from these minimizations are completely unrealistic, with values in the hundreds and thousands of atomic units (data not shown). For comparison, Figure 8 shows corresponding convergence plots when the ligand is treated as rigid. All three plots converge toward −113 kcal/mol, the value obtained by the method of Lagrangian multipliers. (See ΔG^{rigid}, columns 3 and 10, in Table 2.) This expected result supports the validity of the numerical methods used in the present study.

The present analysis proves that the stationary point provided by the Lagrangian method is not a true minimum and is instead a saddle point. More importantly, they show

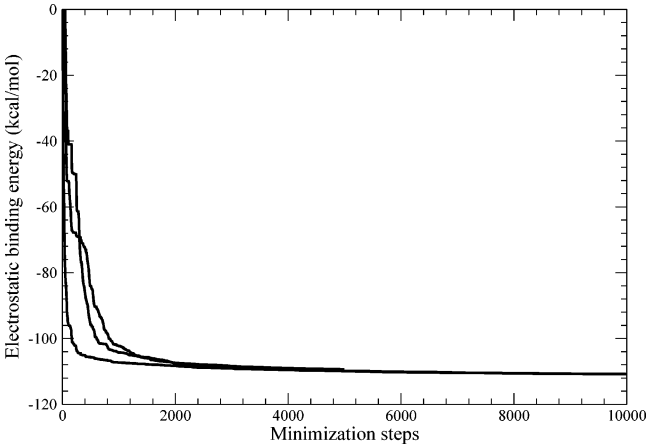

Figure 8. Electrostatic energy as a function of the number of PRAXIS minimization steps when the free ligand conformation is taken to be the same as the bound conformation and the total ligand charge is restrained to zero. Results for 3 different initial guesses of the atomic charges are shown.

that the charge optimization problem need not possess a well-defined solution when the conformation of the ligand is considered to change upon binding.

It is of interest to repeat the minimizations, now applying additional energy restraints that limit the absolute value of the charge of each atom to < 0.85 au, much as done in prior applications of charge optimization that treat the ligand as rigid.[6] These calculations yield convergent energies (data not shown), but the results vary from one minimization to another, depending upon the initial guess for the atomic charges. Two of the resulting charge sets, derived for the same conformation of the free ligand, are shown in columns 6, 7, 13, and 14 of Table 2. Both charge sets yield extremely favorable binding energies, −325 and −343 kcal/mol, when the free ligand is assumed to adopt the conformation for which the charges are optimized. These two charge sets also yield very low energies when the other alternative free conformation is assumed (−101 kcal/mol, −225 kcal/mol), but both are highly unfavorable when the free ligand is assumed to remain in the bound conformation (126 kcal/mol, 111 kcal/mol).

Remarkably, these charge sets can yield such large and negative values of $\Delta G_{ll,lp}$ that the total electrostatic energy of binding is found to be favorable even when the 98.4 kcal/mol penalty for desolvation of the protein is accounted for and when moreover all the electrostatic interactions between the ligand and the protein are artificially neglected. For example, for the second charge set (columns 7 and 14 in the table), the total electrostatic energy of binding is computed to be −31.6 kcal/mol under these assumptions.

This result at first appears necessarily incorrect on physical grounds, since it says that a favorable binding energy is obtained in the absence of any attractive forces between the ligand and the protein. However, the result is correct and is explained by the fact that the charge set in question produces intramolecular Coulombic interactions that powerfully stabilize the bound conformation of the ligand relative to the alternative free conformation, by about −600 kcal/mol, and this stabilization is only partly compensated by a desolvation energy difference of about +150 kcal/mol. As shown in Table 3, similar results apply to the other charge set generated by energy minimization but not to CHARMM charges or the charges generated by Lagrangian optimization when the ligand is assumed to be rigid. Thus, the minimization algorithm finds charges that massively stabilize the bound conformation relative to the alternative free conformation and thus appear to drive binding. It is important to emphasize that a molecule with such charges presumably would not in reality adopt the assumed free conformation but would instead be strongly preorganized into the bound conformation. As a consequence, most or all of the predicted electrostatic contribution to binding would be removed. This type of result is not obtained when when charges are optimized under the usual assumption that the free ligand remains in the bound conformation.

Table 3. Change in Electrostatic Energies (kcal/mol) When the Free Ligand Is Changed from the Flex 1 Conformation to the Bound Conformation (i.e., LigandConf^0 → LigandConf^1), Computed with Various Charge Sets[a]

	charge set			
	CHARMM	Rigid	Flex 1 Min 1	Flex 1 Min 2
Coulombic	−1.7	−12	−603	−607
Solvation (LPB)	−1	12	154	154
Total	−2.7	0	−449	−453

[a] Coulombic energies omit interactions between atoms in 1−2 and 1−3 bonded relationships. Solvation energies are computed with the finite difference, linearized Poisson−Boltzmann method, as described in the text. The total energy is the sum of the Coulombic and solvation terms. CHARMM: unoptimized CHARMM charges (columns 2 and 9 of Table 2). Rigid: charges optimized with the Lagrangian method under the assumption of a rigid ligand (columns 3 and 10 of Table 2). Flex 1 Min 1: the first set of charges adjusted with the PRAXIS algorithm to minimize binding energy when the ligand is assumed to adopt the Flex 1 conformation in the free state, with the total ligand charge restrained to 0 and the absolute values of individual charges restrained ≤ 0.85 au (columns 6 and 13, Table 2). Flex 1 Min 2: the second set of charges adjusted with the PRAXIS algorithm to minimize binding energy when the ligand is assumed to adopt the Flex 1 conformation in the free state, with the total ligand charge restrained to 0 and the absolute values of individual charges restrained ≤ 0.85 au (columns 7 and 14, Table 2).

5. Discussion

Sensitivity analysis is quite accurate for many charge changes large enough to be of interest in lead optimization (±0.5 au), even though it is a linear approximation to an energy function with parabolic curvature. However, its accuracy diminishes significantly for charge changes of ±1 au. Interestingly, accuracy is not degraded when the ligand's conformation is considered to change on binding. Indeed, the energy derivatives are surprisingly insensitive to the assumed conformation of the free ligand. This may result from the fact that the electrostatic potentials at the atoms of the free ligand are dominated by solvation terms which do not depend strongly upon conformation, rather than by conformation-dependent interatomic interactions. Sensitivity analysis is computationally efficient because it requires only one FDPB calculation for each conformation of the ligand. Even greater speed could be be achieved by use of general-

ized Born type models (see, e.g., refs 35−40). Given its efficiency, and the limitations of the linear approximation which forms its basis, sensitivity analysis might be best deployed iteratively. That is, derivatives can be computed and used to guide a first chemical change. Derivatives could then be recomputed for the revised compound to guide a second change and so on. It would furthermore be possible to recompute the conformational preferences of the free ligand after each change and thereby incorporate full-fledged Boltzmann averages of the potentials as derivatives, via eq 7. Finally, it is worth noting that the present generalization of sensitivity analysis is not limited to analyzing the sensitivity of binding free energy to ligand charges but has a much wider range of potential applications. For example, it could be used to examine the sensitivity of protein folding to the strength of an energy term controlling dihedral rotation.

The properties of charge-optimization are more complex, even when the ligand is assumed to be rigid. One important observation is that optimization of *all* ligand charges simultaneously does not appear to be a reliable means of identifying changes in *part* of the ligand that will increase affinity, as previously noted.[9] This issue, illustrated by the case of the urea nitrogens of XK263 in section 4.1.2, could be addressed by applying the optimization formalism only to the part of the ligand that is a candidate for chemical modification, while holding the rest of the charges constant.[9] Mathematically, this would involve merely deleting the rows and columns of the *A* matrix corresponding to the charges which are to be held constant and then applying the method of Lagrangian multipliers as usual. Charge-optimization also has traditionally been rather time-consuming, because it has required at least one FDPB calculation for each atom whose charge is to optimized. Recent methodological advances address this problem, however.[10,11] Charge-optimization, like sensitivity analysis, could be markedly accelerated by replacing FDPB calculations with faster generalized Born calculations.

When the ligand changes conformation upon binding, the Lagrangian charge-optimization formalism may not yield charges that are actually optimal. Instead the stationary point it provides can be a saddle, as suggested by the existence of negative eigenvalues for the upper left $N \times N$ submatrix of the **A** matrix and supported by minimization convergence plots with profoundly negative energies and no evidence of approaching an asymptote. In such cases, charges can be found that lower the nominal binding energy without limit, unless further restraints are applied. Interestingly, charges adjusted to yield a strong nominal binding energy in these circumstances do not driving binding as such but rather a conformational change of the ligand from the free to the bound conformation. There is no guarantee that these charges will actually drive binding. Instead, by strongly stabilizing the bound conformation of the ligand in the free state, these charges violate the assumption of a different free conformation used in generating the charges. More generally, the present analysis highlights the fact that a purely electrostatic analysis cannot yield truly optimal charges, because the best charges depend on the conformational preferences of the ligand which, in turn, are influenced by nonelectrostatic contributions to the energy.

From a practical standpoint, however, if one artificially assumes that the free conformation of the ligand is the same as the bound conformation, as normally done when charge-optimization is employed, the formalism yields charges that robustly improve binding across the different free conformations considered here, as shown at the foot of Table 2. Thus, the current practice of assuming a rigid ligand should often be effective.

In summary, the present results indicate that both sensitivity analysis and charge-optimization can help guide the conversion of a lead compound into a high affinity drug candidate. However, it is important to keep in mind that a predicted improvement in the electrostatic part of the binding energy obtained by applying these methods may not be expressed exactly in the standard free energy of binding because of nonelectrostatic factors. For one thing, it is never possible in reality to change atomic charges without changing other atomic properties, such as atomic radii. In addition, charge changes may produce unforeseen changes in the degree of preorganization of the ligand.

Acknowledgment. This publication was made possible by Grant Number GM61300 from the National Institute of General Medical Sciences (NIGMS) of the NIH. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIGMS. The author thanks the Netlib and LAPACK groups for making the PRAXIS, dsyev, and dgesv, routines available; Mr. Himan Mookherjee for discussions; and Dr. Michael J. Potter and the anonymous referees for their valuable comments.

Supporting Information Available: Three-dimensional coordinates, in PDB format, of the three conformations of XK263 used in the present study: LigandConformation[0] (XK263_docked.pdb), LigandConformation[1] (XK263_md-flex1.pdb), and LigandConformation[2] (XK263_md-flex2.pdb). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Gilson, M. K.; Honig, B. *Proc. Natl. Acad. Sci. U.S.A.* **1989,** *86*, 1524−1528.

(2) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3*, 211−226.

(3) Lee, L. P.; Tidor, B. *J. Chem. Phys.* **1997**, *106*, 8681−8690.

(4) Kangas, E.; Tidor, B. *J. Chem. Phy.* **1998**, *109*, 7522−7545.

(5) Sulea, T.; Purisima, E. O. *J. Phys. Chem. B* **2001**, *105*, 889−899.

(6) Kangas, E.; Tidor, B. *J. Phys. Chem. B* **2001**, *105*, 880−888.

(7) Sulea, T.; Purisima, E. O. *Biophys. J.* **2003**, *84*, 2883−2896.

(8) Mandal, A.; Hilvert, D. *J. Am. Chem. Soc.* **2003**, *125*, 5598−5599.

(9) Sims, P. A.; Wong, C. F.; McCammon, J. A. *J. Comput. Chem.* **2004**, *25*, 1416−1429.

(10) Bardhan, J. P.; Lee, J. H.; Kuo, S. S.; Tidor, M. D. A. B.; White, J. K. Fast methods for biomolecule charge optimization. In *Technical proceedings of the 2003 Nanotechnology Conference and Trade Show*; Nanoscience and Technology Institute: Cambridge, MA, 2003.

(11) Bardhan, J. P.; Lee, J. H.; Altman, M. D.; Leyffer, S.; Benson, S.; Tidor, B.; White, J. K. Biomolecule electrostatic optimization with an implicit Hessian. In *Technical proceedings of the 2004 Nanotechnology Conference and Trade Show*; Nanoscience and Technology Institute: Cambridge, MA, 2004.

(12) Zhang, H.; Wong, C. F.; Thacher, T.; Rabitz, H. *Proteins: Struct., Funct., Genet.* **1995**, *23*, 218−232.

(13) Wong, C. F.; Thacher, T.; Rabitz, H. Sensitivity analysis in biomolecular simulation. In *Rev. Comput. Chem*.; Wiley-VCH: New York, 1998.

(14) Sims, P. A.; Wong, C. F.; McCammon, J. A. *J. Med. Chem.* **2003**, *46*, 3314−3325.

(15) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.

(16) Hill, T. L. *Statistical Mechanics. Principles and selected applications;* McGraw-Hill: New York, 1956.

(17) Cieplak, P.; Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1994**, *101*, 627−633.

(18) Gilson, M. K.; Rashin, A. A.; Fine, R.; Honig, B. *J. Mol. Biol.* **1985**, *183*, 503−516.

(19) Arfken, G. *Mathematical Methods for Physicists*, 3rd ed.; Academic Press: Orlando, FL, 1985.

(20) Lam, P. Y. S.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; Chang, C.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R.; Erickson-Viitanen, S. *Science* **1994**, *263*, 380−384.

(21) Lam, P. Y.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; et al., Y. N. W. *Science* **1994**, *263*, 380−384.

(22) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem*. **1983**, *4*, 187−217.

(23) Accelrys, Inc. San Diego, CA,.

(24) David, L.; Luo, R.; Gilson, M. K. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 157−171.

(25) Kairys, V.; Gilson, M. K. *J. Comput. Chem.* **2002**, *23*, 1656−1670.

(26) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Simul.* **1988**, *1*, 173−185.

(27) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. *Comput. Phys. Commun.* **1991**, *62*, 187−197.

(28) Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151−176.

(29) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *157*, 671−679.

(30) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B. *Proteins: Struct., Funct., Genet.* **1986**, *1*, 47−79.

(31) Gilson, M. K.; Sharp, K. A.; Honig, B. H. *J. Comput. Chem.* **1988**, *9*, 327−335.

(32) Anderson, E.; Bai, Z.; Bischof, C.; Blackford, S.; Demmel, J.; Dongarra, J.; Du Croz, J.; Greenbaum, A.; Hammarling, S.; McKenney, A.; Sorensen, D. *LAPACK Users' Guide;* Society for Industrial and Applied Mathematics: Philadelphia, PA, 3rd ed.; 1999.

(33) http://www.netlib.org/opt/praxis.

(34) http://www.netlib.org.

(35) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem.* **1997**, *101*, 3005−3014.

(36) Ghosh, A.; Rapp, C. S.; Friesner, R. A. *J. Phys. Chem. B* **1998**, *102*, 10983−10990.

(37) Jayaram, B.; Liu, Y.; Beveridge, D. L. *J. Chem. Phys.* **1998**, *109*, 1465−1471.

(38) Dominy, B. N.; Brooks, C. L., III. *J. Phys. Chem. B* **1999**, *103*, 3765−3773.

(39) Lee, M. S.; Salsbury, F. R., Jr.; Charles L.; Brooks, I.

(40) Im, W.; Lee, M. S.; Brooks, C. L., III. *J. Comput. Chem.* **2003**, *24*, 1691−1702.

CT050226Y