Charge Optimization Theory for Induced-Fit Ligands

Citation: Shen, Yang, Michael K. Gilson, and Bruce Tidor 2012Charge Optimization Theory for Induced-Fit Ligands. Journal of Chemical Theory and Computation 8(11): 4580–4592. © 2012 American Chemical Society

As Published: http://dx.doi.org/10.1021/ct200931c

Publisher: American Chemical Society (ACS)

Persistent URL: http://hdl.handle.net/1721.1/79088

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Charge Optimization Theory for Induced-Fit Ligands

Yang Shen,[†,‡] Michael K. Gilson,[§] and Bruce Tidor*[,†,‡,∥]

[†]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

[‡]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

[§]Skaggs School of Pharmacy, University of California San Diego, La Jolla, California 92093, United States

[∥]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

ABSTRACT: The design of ligands with high affinity and specificity remains a fundamental challenge in understanding molecular recognition and developing therapeutic interventions. Charge optimization theory addresses this problem by determining ligand charge distributions that produce the most favorable electrostatic contribution to the binding free energy. The theory has been applied to the design of binding specificity as well. However, the formulations described only treat a rigid ligand—one that does not change conformation upon binding. Here, we extend the theory to treat induced-fit ligands for which the unbound ligand conformation may differ from the bound conformation. We develop a thermodynamic pathway analysis for binding contributions relevant to the theory, and we illustrate application of the theory using HIV-1 protease with our previously designed and validated subnanomolar inhibitor. Direct application of rigid charge optimization approaches to nonrigid cases leads to very favorable intramolecular electrostatic interactions that are physically unreasonable, and analysis shows the ligand charge distribution massively stabilizes the preconformed (bound) conformation over the unbound. After analyzing this case, we provide a treatment for the induced-fit ligand charge optimization problem that produces physically realistic results. The key factor is introducing the constraint that the free energy of the unbound ligand conformation be lower or equal to that of the preconformed ligand structure, which corresponds to the notion that the unbound structure is the ground unbound state. Results not only demonstrate the applicability of this methodology to discovering optimized charge distributions in an induced-fit model, but also provide some insights into the energetic consequences of ligand conformational change on binding. Specifically, the results show that, from an electrostatic perspective, induced-fit binding is not an adaptation designed to enhance binding affinity; at best, it can only achieve the same affinity as optimized rigid binding.

1. INTRODUCTION

Understanding and exploiting the chemical driving forces responsible for tight and specific molecular interactions remains a fundamental challenge in science and engineering. In practical molecular design applications such as structure-based drug design, a lead compound often serves as a starting point for optimization, in which the compound (or *ligand*) is rationally modified in the context of the target protein (or *receptor*), to improve its pharmacological parameters, including potency and selectivity. Geometrical and physicochemical complementarity is often required for tight binding, and the quality of this complementarity is generally reflected in the van der Waals and electrostatic contributions to binding. However, whereas the van der Waals binding contribution is generally strongly favorable, the net electrostatic contribution is often neutral or even somewhat unfavorable.[1,2] This is due to the fact that interfacial protein−ligand electrostatic interactions are acquired through the loss of their individual interactions with solvent in the unbound state, which results in a desolvation penalty. The task of ligand optimization often involves incremental improvement of the shape or electrostatic complementarity, although other sources can also be exploited. This report concerns special considerations for improving electrostatic complementarity.

Charge optimization theory has been developed in the context of linear response theories for the study of electrostatic contributions to binding affinity. The fundamental observation is that the ligand contribution to the binding affinity can be expressed in a simple mathematical form comprising a desolvation penalty that goes as the square of the ligand charge distribution and a generally favorable screened intermolecular interaction term that is linear in the ligand charge distribution. The tradeoff of these terms and their different dependencies on the ligand charge distribution result in an optimum charge distribution corresponding to the most favorable electrostatic contribution to binding. Initially, the geometry of both the free ligand and the bound complex were treated as spherical, and the ligand charge distribution was expressed as a set of multipoles in order to solve the problem.[3,4] Nonspherical geometries and alternative basis sets for ligand charge distributions were then addressed,[5,6] and the method was extended to exact molecular shapes.[7−12] A framework for optimizing electrostatic specificity has also been developed.[13,14] Calculation of the electrostatic potentials is generally done by solving the linearized Poisson−Boltzmann (LPB) equation, but other forms of linear response theory are applicable. A series of

Received: December 26, 2011
Published: June 17, 2012

algorithmic advances has been made to accelerate the calculation[15−18] and produce more-accurate LPB solutions.[19−21]

Charge optimization theory has been broadly applied to probe or design ligands in various molecular systems. Lee and Tidor[7] showed that in the extremely tight-binding barnase−barstar complex, barstar is electrostatically optimized for tight binding to barnase.[7,8] Kangas and Tidor applied charge optimization theory to study the binding between the *Bacillus subtilis* chorismate mutase and an endo-oxabicyclic transition-state analogue.[9] They found that, although the inhibitor showed very good electrostatic complementarity to the enzyme active site, a carboxylate group lost more in desolvation penalty than it gained in interactions with the enzyme; the calculations suggested that substitution with a nitro group would improve the binding affinity. Mandal and Hilvert synthesized the new compound and found, experimentally, a 1.7-kcal/mol improvement in binding affinity in a context corresponding to the calculational study, thus identifying the most potent known inhibitor of this enzyme.[22] Other applications include *E. coli* glutaminyl-tRNA synthetase binding to its cognate substrates,[23] protein inhibitors of HIV-1 cell entry,[24] the interface between protein kinases and their ligands,[25] small-molecule influenza neuraminidase inhibitors,[26] and the celecoxib ligand binding independently to COX2 and CAII.[12] Recently, charge optimization and protein design together identified tighter binding peptides to HIV-1 protease that were studied experimentally.[27] Binding specificity optimization probed ligand binding in the model system of HIV protease, other proteases, and their inhibitors.[14]

The charge optimization approaches described above are based on considering the ligand to remain rigid as it binds to a receptor. A comprehensive theory that includes ligand conformational change on binding has been lacking in the field. Indeed, as one of the authors has shown, a direct replacement of the unbound ligand structure in charge optimization produced charges that strongly stabilized the bound (preconformed) over the unbound conformation, which is physically implausible.[28] In this study, we explore the electrostatic complementarity of an induced-fit ligand to its receptor, and we generalize charge optimization theory to such cases. For our purposes, an induced-fit ligand has one conformation in the unbound state that may or may not be the same as the preconformed structure in the bound state. The origin of the previous unphysical results when the rigid theory was applied to flexible ligands is analyzed, and new methodology is developed and applied to the binding affinity optimization problem for a designed HIV-1 protease inhibitor, MIT-2-KB-98.[29] It generates an optimum partial atomic charge distribution that can be different from that with the rigid-ligand assumption. The key tenet of the new theory is the constraint that the preconformed ligand cannot be lower in free energy than the unbound conformation (otherwise, *it* would be the unbound state).

The remainder of the paper is organized as follows. In Section 2 (Theory), we briefly introduce rigid-ligand charge optimization theory and then describe our treatment of the induced-fit charge optimization problem, in which the ligand adopts an unbound conformation potentially different from the bound. In our formulation, the unbound ligand is either predefined or selected from a set of candidate conformers. The latter framing is useful for future extensions of the theory in which the unbound state is treated as an ensemble representing a population distribution. Whereas nonelectrostatic binding contributions cancel for rigid charge optimization, they can persist in the induced-fit theory developed here. Section 3 (Methods) introduces the test system of HIV-1 protease and its designed inhibitor, as well as methods for calculating continuum electrostatic potentials and nonelectrostatic energies, and approaches to solving the optimization problem. In Section 4 (Results), we analyze the decomposed thermodynamic consequences of the derived results and compare them to those with the rigid-ligand charge optimization approach. Discussion and conclusions are presented in Section 5 (Discussion and Conclusion).

2. THEORY

Molecular binding in an aqueous solvent can be usefully viewed not as an association reaction, in which only new intermolecular interactions are introduced between receptor and ligand, but rather as an exchange reaction in which some receptor−solvent and ligand−solvent interactions present in the unbound state are lost to accommodate the gain of receptor−ligand interactions in the bound complex. The net effect of the exchange nature of binding is potentially interesting and nonintuitive for electrostatics, and leads to charge optimization theory. Here, we first review the theory for the binding of a rigid ligand to form a complex and then extend the theory to treat induced-fit binding, in each case solving for the ligand charge distribution that leads to the most favorable binding free energy.

Rigid-Ligand Charge Optimization. Consider the binding of receptor (R) and ligand (L) to form complex (C) in an aqueous environment. The charge distribution ρ is modeled as a set of point charges q_i at atomic centers $\mathbf{r}_i$. The macromolecules are treated as low dielectric volumes defined by their molecular shapes that are surrounded by high dielectric solvent with ions at physiological concentration. The electrostatic binding free energy of the complex can be written as

$$\Delta G^{\mathrm{es}}_{\mathrm{bind}} = G^{\mathrm{es}}_{\mathrm{C}} - G^{\mathrm{es}}_{\mathrm{R}} - G^{\mathrm{es}}_{\mathrm{L}} \tag{1}$$

where G^{es}_x is the electrostatic free energy of molecule x, which has a particularly convenient form in any linear response theory. Here, we illustrate using solutions to the linearized Poisson−Boltzmann (LPB) equation, without loss of generality:

$$\nabla\cdot(\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})) - \varepsilon(\mathbf{r})\kappa^2(\mathbf{r})\phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}) \tag{2}$$

Equation 2 inter-relates positional maps of electrostatic potential ϕ, charge density distribution ρ, and dielectric constant (or permittivity) ε, where $\kappa(\mathbf{r}) = [8\pi e^2 I(\mathbf{r})/(\varepsilon kT)]^{1/2}$ is the inverse Debye length with I the ionic strength, e the electron charge magnitude, k the Boltzmann constant, and T the absolute temperature.[30,31] The equation is generally solved for the electrostatic potential. In linear response theory, the electrostatic free energy of a molecule x is given by

$$G^{\mathrm{es}}_x = \frac{1}{2}\sum_{j\in x}\phi(\mathbf{r}_j)q_j \tag{3}$$

where the summation runs over the partial atomic charges q_j located at positions $\mathbf{r}_j$. In the LPB model, individual point charges act independently and obey superposition, so their contributions can be calculated separately and summed,

dx.doi.org/10.1021/ct200931c | J. Chem. Theory Comput. 2012, 8, 4580−4592

$$\phi(\mathbf{r}_j) = \sum_{i\in x} \phi_i(\mathbf{r}_j) q_i \tag{4}$$

where $\phi_i(\mathbf{r}_j)$ (the electrostatic potential at position $\mathbf{r}_j$ due to a unit point charge at $\mathbf{r}_i$) can be calculated by solving the LPB equation (eq 2). We note that procedures are generally implemented to avoid the Coulombic singularity for $i = j$, but we will leave the singularity in the equations here. By plugging this equation into eq 3, one can write

$$G_x^{\mathrm{es}} = \frac{1}{2} \sum_{j\in x} \sum_{i\in x} q_i \phi_i(\mathbf{r}_j) q_j \tag{5}$$

After introducing a matrix Φ_x for molecule x ($\Phi_x(i, j) = (1/2)\phi_i(\mathbf{r}_j)$) and splitting it into solvent reaction field and Coulombic terms, as well as expressing the partial atomic charges of x as a vector $\mathbf{q}_x = (q_1, q_2, q_3, \cdots)^{\mathrm{T}}$, one can rewrite eq 5 as

$$G_x^{\mathrm{es}} = \mathbf{q}_x^{\mathrm{T}} \Phi_x^{\mathrm{solv}} \mathbf{q}_x + \mathbf{q}_x^{\mathrm{T}} \Phi_x^{\mathrm{coul}} \mathbf{q}_x \tag{6}$$

Combining eqs 1 and 6, the electrostatic binding free energy of a complex can be expressed as

$$\begin{aligned} \Delta G_{\mathrm{bind}}^{\mathrm{es}} = {} & (\mathbf{q}_{\mathrm{C}}^{\mathrm{T}} \Phi_{\mathrm{C}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{C}} + \mathbf{q}_{\mathrm{C}}^{\mathrm{T}} \Phi_{\mathrm{C}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{C}}) \\ & - (\mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \Phi_{\mathrm{R}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{R}} + \mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \Phi_{\mathrm{R}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{R}}) \\ & - (\mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \Phi_{\mathrm{L}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \Phi_{\mathrm{L}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{L}}) \end{aligned} \tag{7}$$

Denote the matrices $\Phi_{\mathrm{C}}^{\mathrm{solv}}$ and $\Phi_{\mathrm{C}}^{\mathrm{coul}}$ as Φ_{C}^{y} and consider their structure. Each element $\Phi_{\mathrm{C}}^{y}(i, j) = (1/2)\phi_i(\mathbf{r}_j)$ is half the potential (solvent reaction field or Coulombic, depending on y) at point $\mathbf{r}_j$ due to a unit charge at $\mathbf{r}_i$. The locations $\mathbf{r}_i$ and $\mathbf{r}_j$ can both be in the ligand, both be in the receptor, or one in each. It is useful to indicate these three cases by representing Φ_{C}^{y} in block form as

$$\Phi_{\mathrm{C}}^{y} = \begin{pmatrix} \mathbf{R}_{\mathrm{c}}^{y} & \mathbf{C}_{\mathrm{c}}^{y} \\ (\mathbf{C}_{\mathrm{c}}^{y})^{\mathrm{T}} & \mathbf{L}_{\mathrm{c}}^{y} \end{pmatrix} \tag{8}$$

In this format, $\mathbf{R}_{\mathrm{c}}^{y}$ is a square matrix of dimension n_{R} (the number of receptor partial atomic charges) giving half the potential at partial atomic charge locations in the receptor due to the presence of other individual receptor unit charges, and $\mathbf{L}_{\mathrm{c}}^{y}$ is the same for the ligand. $\mathbf{C}_{\mathrm{c}}^{y}$ indicates a half-potential at a ligand atom due to a unit charge on a receptor atom and is not generally square. Because of reciprocity, Φ_{c}^{y}, $\mathbf{R}_{\mathrm{c}}^{y}$, and $\mathbf{L}_{\mathrm{c}}^{y}$ are all symmetric, and $\mathbf{C}_{\mathrm{c}}^{y}$ and $(\mathbf{C}_{\mathrm{c}}^{y})^{\mathrm{T}}$ are indeed transposes of each other. For all cases, the subscript c indicates that the potentials are taken in the protein complex. While this does not affect the Coulombic potentials (taken in uniform low dielectric environment), it is important, because it defines the molecular boundary for the reaction field potentials. The first two terms of eq 7—representing the electrostatic free energy of the complex—then can be rewritten as

$$\begin{aligned} G_{\mathrm{C}}^{\mathrm{es}} = {} & \begin{pmatrix} \mathbf{q}_{\mathrm{R}}^{\mathrm{T}} & \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \mathbf{R}_{\mathrm{c}}^{\mathrm{solv}} + \mathbf{R}_{\mathrm{c}}^{\mathrm{coul}} & \mathbf{C}_{\mathrm{c}}^{\mathrm{solv}} + \mathbf{C}_{\mathrm{c}}^{\mathrm{coul}} \\ (\mathbf{C}_{\mathrm{c}}^{\mathrm{solv}} + \mathbf{C}_{\mathrm{c}}^{\mathrm{coul}})^{\mathrm{T}} & \mathbf{L}_{\mathrm{c}}^{\mathrm{solv}} + \mathbf{L}_{\mathrm{c}}^{\mathrm{coul}} \end{pmatrix} \begin{pmatrix} \mathbf{q}_{\mathrm{R}} \\ \mathbf{q}_{\mathrm{L}} \end{pmatrix} \\ = {} & \mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{R}_{\mathrm{c}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{R}} + \mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{R}_{\mathrm{c}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{R}} + \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \mathbf{L}_{\mathrm{c}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \mathbf{L}_{\mathrm{c}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{L}} \\ & + 2\mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{C}_{\mathrm{c}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{L}} + 2\mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{C}_{\mathrm{c}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{L}} \end{aligned} \tag{9}$$

The factors of 2 in the L−R interaction terms are due to the reciprocity relation $\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}(\mathbf{C}_{\mathrm{c}}^{y})^{\mathrm{T}}\mathbf{q}_{\mathrm{R}} = \mathbf{q}_{\mathrm{R}}^{\mathrm{T}}\mathbf{C}_{\mathrm{c}}^{y}\mathbf{q}_{\mathrm{L}}$. Similarly, for the unbound receptor or ligand, the matrices Φ_{R}^{y} and Φ_{L}^{y} can be represented as $\mathbf{R}_{x}^{y}$ and $\mathbf{L}_{x}^{y}$, respectively, which give the potential at a receptor or ligand atom generated by atomic charges in state x (x = u for the unbound state and x = c for the bound complex). Therefore, the electrostatic contribution to the binding free energy in eq 1 can be rewritten as

$$\begin{aligned} \Delta G_{\mathrm{bind}}^{\mathrm{es}} = {} & (\mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{R}_{\mathrm{c}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{R}} + \mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{R}_{\mathrm{c}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{R}} + \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \mathbf{L}_{\mathrm{c}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \mathbf{L}_{\mathrm{c}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{L}} + 2\mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{C}_{\mathrm{c}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{L}} + 2\mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{C}_{\mathrm{c}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{L}}) \\ & - (\mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{R}_{\mathrm{u}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{R}} + \mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{R}_{\mathrm{u}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{R}}) - (\mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \mathbf{L}_{\mathrm{u}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \mathbf{L}_{\mathrm{u}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{L}}) \\ = {} & \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} (\mathbf{L}_{\mathrm{c}}^{\mathrm{solv}} - \mathbf{L}_{\mathrm{u}}^{\mathrm{solv}}) \mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} (\mathbf{L}_{\mathrm{c}}^{\mathrm{coul}} - \mathbf{L}_{\mathrm{u}}^{\mathrm{coul}}) \mathbf{q}_{\mathrm{L}} + 2\mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{C}_{\mathrm{c}}^{\mathrm{solv}} \mathbf{q}_{\mathrm{L}} + 2\mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \mathbf{C}_{\mathrm{c}}^{\mathrm{coul}} \mathbf{q}_{\mathrm{L}} \\ & + \mathbf{q}_{\mathrm{R}}^{\mathrm{T}} (\mathbf{R}_{\mathrm{c}}^{\mathrm{solv}} - \mathbf{R}_{\mathrm{u}}^{\mathrm{solv}}) \mathbf{q}_{\mathrm{R}} + \mathbf{q}_{\mathrm{R}}^{\mathrm{T}} (\mathbf{R}_{\mathrm{c}}^{\mathrm{coul}} - \mathbf{R}_{\mathrm{u}}^{\mathrm{coul}}) \mathbf{q}_{\mathrm{R}} \end{aligned} \tag{10}$$

When both the receptor and the ligand are regarded as rigid molecules with no conformational change upon binding (u = c for both molecules), the intramolecular Coulombic terms cancel ($\mathbf{L}_{\mathrm{c}}^{\mathrm{coul}} = \mathbf{L}_{\mathrm{u}}^{\mathrm{coul}}$ and $\mathbf{R}_{\mathrm{c}}^{\mathrm{coul}} = \mathbf{R}_{\mathrm{u}}^{\mathrm{coul}}$). Moreover, the difference in solvation matrices can be re-expressed as $\mathbf{L}_{\mathrm{c}}^{\mathrm{solv}} - \mathbf{L}_{\mathrm{u}}^{\mathrm{solv}} = \Delta\mathbf{L}$ and $\mathbf{R}_{\mathrm{c}}^{\mathrm{solv}} - \mathbf{R}_{\mathrm{u}}^{\mathrm{solv}} = \Delta\mathbf{R}$. The matrices $\Delta\mathbf{L}$ and $\Delta\mathbf{R}$ are expected to be positive semidefinite, because desolvation should represent a penalty and symmetric due to reciprocity.[3,5,6] In this rigid "docking" mode with no conformational change, eq 10 can be simplified to

$$\Delta G_{\mathrm{dock}}^{\mathrm{es}} = \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \Delta\mathbf{L} \mathbf{q}_{\mathrm{L}} + \Phi_{\mathrm{R}}^{\mathrm{T}} \mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \Delta\mathbf{R} \mathbf{q}_{\mathrm{R}} \tag{11}$$

where $\Phi_{\mathrm{R}}^{\mathrm{T}} = 2\mathbf{q}_{\mathrm{R}}^{\mathrm{T}}\mathbf{C}_{\mathrm{c}} = 2\mathbf{q}_{\mathrm{R}}^{\mathrm{T}}(\mathbf{C}_{\mathrm{c}}^{\mathrm{solv}} + \mathbf{C}_{\mathrm{c}}^{\mathrm{coul}})$ is the screened electrostatic potential (including both the solvent reaction field and Coulombic terms) at the ligand atoms generated by receptor atomic charges $\mathbf{q}_{\mathrm{R}}$. Also, we have renamed $\Delta G_{\mathrm{bind}}^{\mathrm{es}}$ to $\Delta G_{\mathrm{dock}}^{\mathrm{es}}$ to describe the rigid case for convenience later. Of the three terms in eq 11, ligand desolvation is quadratic in the ligand charge distribution $\mathbf{q}_{\mathrm{L}}$,

$$\Delta G_{\mathrm{L}}^{\mathrm{desolv}} = \mathbf{q}_{\mathrm{L}}^{\mathrm{T}} \Delta\mathbf{L} \mathbf{q}_{\mathrm{L}} \tag{12}$$

the intermolecular screened electrostatic interaction is linear in $\mathbf{q}_{\mathrm{L}}$,

$$\Delta G_{\mathrm{R-L}}^{\mathrm{inter}} = \Phi_{\mathrm{R}}^{\mathrm{T}} \mathbf{q}_{\mathrm{L}} \tag{13}$$

and receptor desolvation is a constant, independent of $\mathbf{q}_{\mathrm{L}}$:

$$\Delta G_{\mathrm{R}}^{\mathrm{desolv}} = \mathbf{q}_{\mathrm{R}}^{\mathrm{T}} \Delta\mathbf{R} \mathbf{q}_{\mathrm{R}} \tag{14}$$

dx.doi.org/10.1021/ct200931c | J. Chem. Theory Comput. 2012, 8, 4580−4592

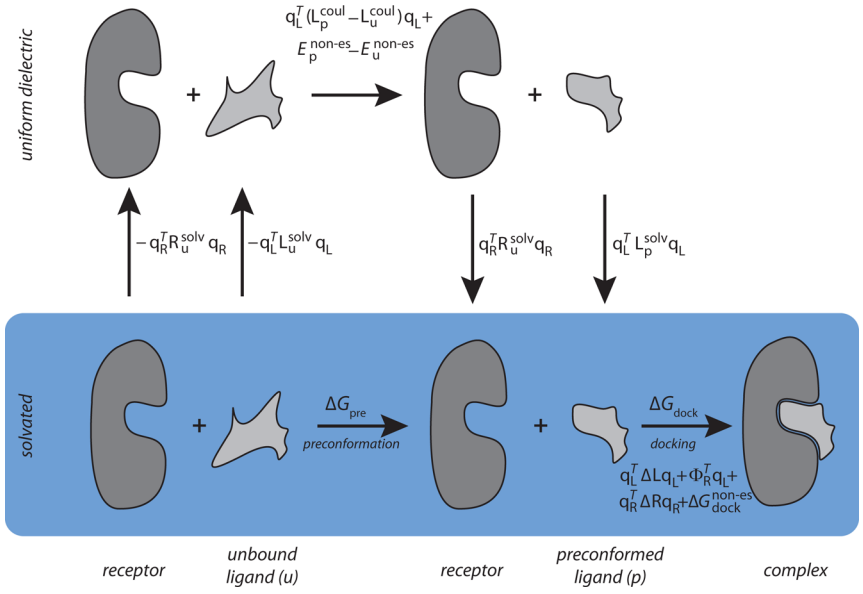

Figure 1. The process of induced-fit binding of a ligand to its receptor. Binding is decomposed into the two steps of preconformation and docking. The ligand has two conformations, unbound (u) and preconformed (p), which can be the same. Calculation of the binding free energy is indicated schematically, using terms from eqs 20 and 21. The blue row indicates solvated molecules and the upper row indicates calculations in uniform dielectric.

Thus, the electrostatic binding free energy for rigid molecules (ΔG_{dock}^{es}) is a paraboloid in the space of $\mathbf{q}_L$.[3] Moreover, there is a unique global optimum $\mathbf{q}_L^{opt} = -\Delta\mathbf{L}^{-1}\mathbf{C}_c^T\mathbf{q}_R$ that balances the unfavorable ligand desolvation penalty and the favorable intermolecular interactions, and gives the minimum electrostatic binding free energy. Kangas and Tidor showed that this minimum value is upper-bounded by zero, and thus the optimized ligand charge distribution leads to favorable electrostatic binding free energy under certain conditions.[6] Natural proteins seem to use this type of optimization to achieve tight binding: the tight-binding inhibitor barstar was found to be electrostatically optimized to its partner barnase, as described by this charge optimization theory.[7]

Induced-Fit Ligand Charge Optimization. In this extension of charge optimization theory, the unbound ligand is treated as a set of individual conformational candidates, one of which is selected by the optimization as a single unbound state conformation. In the current work, the receptor and complex are treated as rigid. Of these two simplifications, the former has no effect on ligand charge optimization, because the unbound receptor contribution to the binding free energy is a constant for all ligand charge distributions considered; multiconformational effects for the bound complex could be significant and will be treated in future work. Also considered in future work are ensemble treatments in which each state is represented as a Boltzmann distribution of conformations rather than a single conformation.

In this framework, the binding free energy can be deconstructed as the sum of two terms as shown in the blue row of the thermodynamic pathway shown in Figure 1. In the first step, the unbound ligand changes conformation to the preconformed ligand structure (i.e., that adopted by the ligand in the bound complex); the free energy change for this step is termed the preconformation contribution ΔG_{pre}. In the second step, the preconformed ligand and receptor dock rigidly to form the bound complex (as modeled in the rigid-ligand charge optimization problem above), and the free-energy change is termed the docking contribution ΔG_{dock}. Thus, the total binding affinity is represented as the sum of two terms corresponding to the two processes:

$$\Delta G_{total} = \Delta G_{pre} + \Delta G_{dock} \tag{15}$$

Each term is expressed as a summation of electrostatic and nonelectrostatic contributions.

$$\Delta G_{pre} = \Delta G_{pre}^{es} + \Delta G_{pre}^{non\text{-}es} \tag{16}$$

$$\Delta G_{dock} = \Delta G_{dock}^{es} + \Delta G_{dock}^{non\text{-}es} \tag{17}$$

Here, the nonelectrostatic terms include covalent energy terms (bonds, angles, torsions, etc.), intramolecular and intermolecular van der Waals interactions, and nonpolar solvation interactions.

In the preconformation step, the ligand is solvated in an aqueous environment alone and changes its conformation from the unbound (u) to the preconformed (p) structure. We stress that even for the preconformed structure, the ligand is free in solution. This is equivalent to a special case of eq 10, in which the receptor is not present. Therefore, the electrostatic contribution to the change in free energy (ΔG_{pre}^{es}) can be expressed as the sum of a solvent reaction field term (ΔG_{pre}^{solv}) and a Coulombic term (ΔG_{pre}^{coul}), given by

$$\begin{aligned}\Delta G_{pre}^{es} &= \Delta G_{pre}^{solv} + \Delta G_{pre}^{coul} \\ &= \mathbf{q}_L^T(\mathbf{L}_p^{solv} - \mathbf{L}_u^{solv})\mathbf{q}_L + \mathbf{q}_L^T(\mathbf{L}_p^{coul} - \mathbf{L}_u^{coul})\mathbf{q}_L\end{aligned} \tag{18}$$

The nonelectrostatic contribution to the change in free energy is

$$\Delta G_{pre}^{non\text{-}es} = E_p^{non\text{-}es} - E_u^{non\text{-}es} \tag{19}$$

where $E_x^{non\text{-}es}$ denotes the nonelectrostatic contribution to the free energy for the ligand in a given conformation (x). In the docking step, the electrostatic binding free energy ΔG_{dock}^{es} is given by eq 11. Therefore,

dx.doi.org/10.1021/ct200931c | J. Chem. Theory Comput. 2012, 8, 4580−4592

$$\Delta G_{\mathrm{pre}} = \mathbf{q}_{\mathrm{L}}^{\mathrm{T}}(\mathbf{L}_{\mathrm{p}}^{\mathrm{solv}} + \mathbf{L}_{\mathrm{p}}^{\mathrm{coul}} - \mathbf{L}_{\mathrm{u}}^{\mathrm{solv}} - \mathbf{L}_{\mathrm{u}}^{\mathrm{coul}})\mathbf{q}_{\mathrm{L}} + (E_{\mathrm{p}}^{\text{non-es}} - E_{\mathrm{u}}^{\text{non-es}}) \tag{20}$$

and

$$\Delta G_{\mathrm{dock}} = \mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\Delta\mathbf{L}\mathbf{q}_{\mathrm{L}} + \Phi_{\mathrm{R}}^{\mathrm{T}}\mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{R}}^{\mathrm{T}}\Delta\mathbf{R}\mathbf{q}_{\mathrm{R}} + \Delta G_{\mathrm{dock}}^{\text{non-es}} \tag{21}$$

Equations 20 and 21 can be interpreted from the thermodynamic pathway of Figure 1. ΔG_{pre} is equal to the difference in free energy of the ligand in two conformations ("u" and "p"). In a given conformation x, the free energy of the ligand includes the cost of assembling the charge distribution $\mathbf{q}_{\mathrm{L}}$ in a uniform low dielectric ($\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\mathbf{L}_x^{\mathrm{coul}}\mathbf{q}_{\mathrm{L}}$) and that of moving it to high-dielectric solvent ($\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\mathbf{L}_x^{\mathrm{solv}}\mathbf{q}_{\mathrm{L}}$), as well as the nonpolar contributions to the chemical potential ($E_x^{\text{non-es}}$).

In the original formulation of charge optimization theory, the binding reaction was treated as the rigid docking of preconformed ligand and receptor;[3] with no ligand conformational change ($\Delta G_{\mathrm{pre}} = 0$ for that case). Here, we remove that limitation and explore the consequences. Specifically, we permit the unbound ligand to adopt one conformation that may be the same or different from the bound (preconformed) ligand conformation. This treatment of ligand flexibility is reminiscent of the induced-fit model.

Charge optimization in the current framing requires optimizing the binding free energy with respect to the ligand charge distribution. In the formulation here, a set of candidate unbound ligand conformations is available, and optimization selects an unbound conformation and charge distribution for the ligand that optimizes binding free energy. One physical constraint that we explore in this study is that the conformation assigned as the ligand unbound state should correspond to the lowest energy conformation in the set of candidate conformations when evaluated with the optimized charge distribution. Said another way, the unbound structure is constrained to be the ground unbound state.

To carry out the desired optimization, we construct a variational binding free energy expression that includes all terms that depend on the ligand charge distribution, either explicitly or implicitly. All terms that depend on the unbound ligand conformation must be kept in the optimization, because they contribute to defining the unbound ligand ground-state energy. Here, those include all terms in ΔG_{pre} except $E_{\mathrm{p}}^{\text{non-es}}$, and, for simplicity, we keep all terms of ΔG_{pre} in our optimization function. The receptor electrostatic desolvation penalty and the nonelectrostatic contribution to ΔG_{dock} are both independent of $\mathbf{q}_{\mathrm{L}}$ and, therefore, have been removed to construct the optimization function $\Delta G'_{\mathrm{total}}$, which can be written as

$$\begin{aligned}\Delta G'_{\mathrm{total}} &= \Delta G_{\mathrm{pre}} + \Delta G'_{\mathrm{dock}} \\ &= (\Delta G_{\mathrm{pre}}^{\mathrm{solv}} + \Delta G_{\mathrm{pre}}^{\mathrm{coul}} + \Delta G_{\mathrm{pre}}^{\text{non-es}}) + (\Delta G_{\mathrm{L}}^{\mathrm{desolv}} + \Delta G_{\mathrm{R-L}}^{\mathrm{inter}}) \\ &= \Big(\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}(\mathbf{L}_{\mathrm{p}}^{\mathrm{solv}} - \mathbf{L}_{\mathrm{u}}^{\mathrm{solv}})\mathbf{q}_{\mathrm{L}} + \mathbf{q}_{\mathrm{L}}^{\mathrm{T}}(\mathbf{L}_{\mathrm{p}}^{\mathrm{coul}} - \mathbf{L}_{\mathrm{u}}^{\mathrm{coul}})\mathbf{q}_{\mathrm{L}} \\ &\quad + (E_{\mathrm{p}}^{\text{non-es}} - E_{\mathrm{u}}^{\text{non-es}})\Big) \\ &\quad + (\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\Delta\mathbf{L}\mathbf{q}_{\mathrm{L}} + \Phi_{\mathrm{R}}^{\mathrm{T}}\mathbf{q}_{\mathrm{L}})\end{aligned} \tag{22}$$

When the unbound ligand conformation is predefined, the optimization task involves the following:

$$\min_{\mathbf{q}_L} \ \Delta G'_{\mathrm{total}} \tag{23}$$

subject to

$$|q_{\mathrm{L}}^k| \leqslant q_{\max} \qquad \forall\ k = 1, \ldots, n_{\mathrm{atom}}$$

$$\sum_{k=1}^{n_{\mathrm{atom}}} q_{\mathrm{L}}^k = q_{\mathrm{total}}$$

where q_{L}^k denotes the kth element of $\mathbf{q}_{\mathrm{L}}$ (i.e., the partial atomic charge for atom k of the ligand). The first constraint limits the magnitude of any single charge to $q_{\max}$ (generally 0.85 units of the electron charge magnitude), and the second fixes the total charge of the ligand at q_{total} (usually an integer).[7] If the unbound and bound ligand conformations are the same, this reduces to the original charge optimization approach;[7] if the two ligand conformations differ, this corresponds to the treatment of Gilson.[28]

When, instead, the unbound ligand conformation is selected from a set of candidates, the optimization problem is somewhat different. The optimized ligand charge distribution $\mathbf{q}_{\mathrm{L}}^{\mathrm{opt}}$ must be selected such that (1) $\Delta G'_{\mathrm{total}}$ is minimized and (2) the chosen unbound ligand conformation has the lowest free energy of all conformations in the unbound set with $\mathbf{q}_{\mathrm{L}}^{\mathrm{opt}}$ applied. The second requirement implements the physical constraint that the conformation claimed for the unbound state be the lowest free energy of those considered, which we show avoids the pathological results observed when rigid theory was applied directly to the nonrigid case.

Denoting the set of unbound candidates by U and indexing each member by i, we write the following mathematical programming problem:

$$(P_0^{\min}) \quad \min_{\mathbf{q}_L} \ \Big(\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}(\mathbf{L}_{\mathrm{p}}^{\mathrm{solv}} + \mathbf{L}_{\mathrm{p}}^{\mathrm{coul}})\mathbf{q}_{\mathrm{L}} + \mathbf{E}_{\mathrm{p}}^{\text{non-es}} - z\Big) + (\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\Delta\mathbf{L}\mathbf{q}_{\mathrm{L}} + \Phi_{\mathrm{R}}^{\mathrm{T}}\mathbf{q}_{\mathrm{L}}) \tag{24}$$

subject to

$$z = \min_{i \in U} \ (\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}(\mathbf{L}_{\mathrm{u},i}^{\mathrm{solv}} + \mathbf{L}_{\mathrm{u},i}^{\mathrm{coul}})\mathbf{q}_{\mathrm{L}} + \mathbf{E}_{\mathrm{u},i}^{\text{non-es}})$$

$$|q_{\mathrm{L}}^k| \leqslant q_{\max} \qquad k = 1, \ldots, n_{\mathrm{atom}}$$

$$\sum_{k=1}^{n_{\mathrm{atom}}} q_{\mathrm{L}}^k = q_{\mathrm{total}}$$

Note that z is not smooth in $\mathbf{q}_{\mathrm{L}}$ (i.e., z has discontinuous derivatives), because of the function $\min(\cdot)$. The lack of smoothness presents a fundamental challenge in solving the optimization problem and is highly undesirable in practice. Thus, we reformulate the problem into the following smooth one with the same optimum solution, using a standard transformation technique that replaces the explicit min function to define z with an implicit one,[32]

$$(P^{\min}) \quad \min_{\mathbf{q}_L} \ \Big(\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}(\mathbf{L}_{\mathrm{p}}^{\mathrm{solv}} + \mathbf{L}_{\mathrm{p}}^{\mathrm{coul}})\mathbf{q}_{\mathrm{L}} + \mathbf{E}_{\mathrm{p}}^{\text{non-es}} - z\Big) + (\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\Delta\mathbf{L}\mathbf{q}_{\mathrm{L}} + \Phi_{\mathrm{R}}^{\mathrm{T}}\mathbf{q}_{\mathrm{L}}) \tag{25}$$

subject to

$$z \leqslant G_{\mathrm{L,i}} = \mathbf{q}_{\mathrm{L}}^{\mathrm{T}}(\mathbf{L}_{\mathrm{u},i}^{\mathrm{solv}} + \mathbf{L}_{\mathrm{u},i}^{\mathrm{coul}})\mathbf{q}_{\mathrm{L}} + \mathbf{E}_{\mathrm{u},i}^{\text{non-es}} \quad \forall\ i \in U$$

$$|q_{\mathrm{L}}^k| \leqslant q_{\max} \qquad k = 1, \ldots, n_{\mathrm{atom}}$$

$$\sum_{k=1}^{n_{\mathrm{atom}}} q_{\mathrm{L}}^k = q_{\mathrm{total}}$$

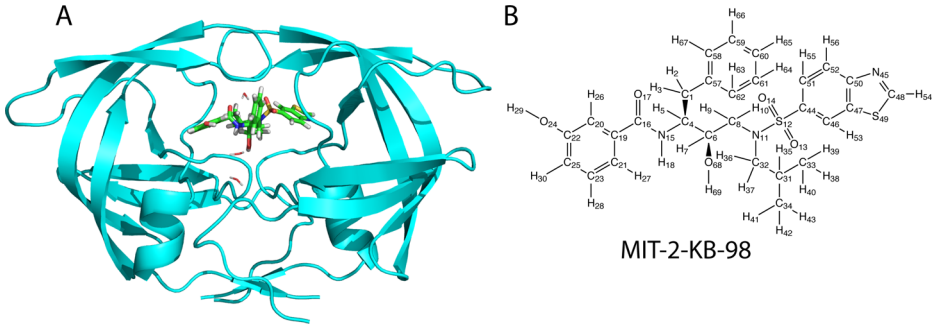

Figure 2. (A) Designed inhibitor MIT-2-KB-98 (stick representation; green) in model complex with HIV-1 protease (cartoon representation; cyan), with the three retained water molecules (line representation). (B) Chemical structure of MIT-2-KB-98.

This formulation is general, in that it includes the special cases of rigid ligand (U containing only the preconformed, bound state)[7] and a two-state induced-fit ligand (U containing only an unbound state different than the preconformed, bound state).[28] The formulation is useful in that it can be extended to treat conformational ensembles representing population distributions rather than unique structures for individual states. Also, note that, although the presentation here is in terms of a basis of point charges, the theory is readily applicable to other bases.

3. METHODS

Test System. The crystal structure of HIV-1 protease alone was extracted from its bound complex with the inhibitor darunavir (PDB ID 1T3R)[33] and prepared as described by Altman et al.[29] with both catalytic residues (Asp 25) deprotonated. Three conserved water molecules were retained, including the so-called "flap water molecule" (residue IDs 1, 2, and 4; Figure 2A). Hydrogen atoms were added to the complex with the HBUILD module[34] of the computer program package CHARMM.[35,36] The parameter set used was CHARMm22.[37]

HIV protease was studied in complex with the inhibitor MIT-2-KB-98.[29] Parameters for MIT-2-KB-98 were assigned from the CHARMm22 set except for the partial atomic charges, which were derived here. Geometry optimization was performed using quantum mechanical calculations at the restricted Hartree−Fock (RHF) 3-21G level as implemented in the program GAUSSIAN 03.[38] Partial atomic charges were calculated for the new geometry by restrained electrostatic potential (RESP) fitting[39,40] to RHF/6-31G* potentials. This charge set will be referred to as "nominal".

A set of 1000 unbound MIT-2-KB-98 conformations was computed from a 30° dihedral grid using dead-end elimination (DEE)[41−44] and A*[45] to select the 1000 lowest energy structures different from the preconformed design structure. This unbound conformer set covered the lowest 1.5 kcal/mol in free energy and over 92% of the Boltzmann distribution of all the enumerated states. The energy function used was the gas-phase energy with a dielectric constant equal to 4 and without nonbond cutoffs. These structures were indexed by rank in energy from 1 (lowest) to 1000 (highest), and, for convenience, the preconformed structure was given an index of 0.

Electrostatic Potentials. Solvation potential matrices were computed with continuum electrostatic calculations implemented in a locally modified version of the program DelPhi.[46−48] Calculations were performed for the preconformed ligand in the free state and complexed with the protein ($\mathbf{L}_{\mathrm{p}}^{\mathrm{solv}} \equiv \mathbf{L}_{\mathrm{u},0}^{\mathrm{solv}}$ and $\Delta\mathbf{L}$), as well as for each of the other 1000 members of the unbound conformer set alone ($\mathbf{L}_{\mathrm{u},i}^{\mathrm{solv}}$, $\forall i = 1, \ldots, 1000$).

Details of the calculations were as given in the work of Altman et al.[29] Partial atomic charges for HIV-1 protease and atomic radii of both the protein and the ligand were obtained from the PARSE parameter set.[49] The dielectric constant was set to 4 for the molecular interior and 80 for the exterior, with the dielectric boundary represented by the molecular surface computed with a 1.4-Å probe. A 2.0-Å ion-excluding Stern layer surrounded all molecules,[50] and a bulk ionic strength of 0.145 M was employed. Linearized Poisson−Boltzmann (LPB) calculations were performed on a 129 × 129 × 129 grid with focusing boundary conditions (successively 23%, 92%, and 184% fill).[51] The final results were averaged from ten translations with respect to the grid.[2,47]

The intramolecular Coulombic electrostatic potential matrices $\mathbf{L}_{\mathrm{u},i}^{\mathrm{coul}}$ were calculated directly from Coulomb's law; 1−2 and 1−3 interactions were excluded and 1−4 interactions were scaled by a factor of 0.50, which are consistent with the CHARMm22 treatment.

Non-Electrostatic Energies. CHARMM was used with the CHARMm22 parameter set to evaluate the nonelectrostatic energies $E_{\mathrm{u},i}^{\text{non-es}}$ for the 1001 members of the set of unbound candidates, including internal energy terms (bond, angle, dihedral, improper, and Urey−Bradley) and van der Waals (excluding the 1−2 and 1−3 interactions, and with special treatment of 1−4 interactions; without nonbonded smoothing or truncation). The nonpolar contribution to the hydration free energy was calculated as a surface-area-dependent term.[49]

Charge Optimization. The induced-fit ligand charge optimization problem formulated in eq 23 is not guaranteed to be convex, which implies possible multiple local minima. In practice, there is no guarantee for a generic solver to determine the global minimum of a nonconvex optimization problem. In this study, we used multistart trajectories with a local minimizer. We chose CONOPT3[52,53] as the local minimizer for its numerical efficiency, as available through GAMS.[54] CONOPT3 is a nonlinear optimization solver based on generic reduced gradient algorithms.[55,56] For each problem, 100 starting ligand charge distributions were randomly generated by uniformly sampling in the feasible space (the intersection of a hypercube defined by $|q_{\mathrm{L}}^{k}| \leqslant q_{\max}$, $\forall k = 1, \ldots, n_{\mathrm{atom}}$ and a hyperplane by $\sum_{k=1}^{n_{\mathrm{atom}}} q_{\mathrm{L}}^{k} = q_{\mathrm{total}}$). For this purpose, we adopted an algorithm decomposing the space into different types of simplexes and sampled within and across these simplexes uniformly.[57] The solution was chosen as that with the lowest value of the goal function from the multistart set.

4. RESULTS

Results are presented for induced-fit charge optimization on a previously designed ligand binding to the enzyme HIV-1 protease. The ligand, MIT-2-KB-98, has an (*R*)-(hydroxyethylamino)sulfonamide scaffold functionalized to make favorable interactions in substrate recognition subsites (see Figure 2B).[33] It has a total of 13 rotatable torsion angles (neglecting the two methyl groups and the amide bond), which were varied to create the set of unbound ligand candidates (see the Methods section). Two different but related schemes were used here. For each, a collection of 1001 different conformations of the ligand MIT-2-KB-98 was used to represent conformational candidates for the unbound ligand. The structure of the bound complex with the receptor HIV-1 protease was taken as the design structure from the study in which this ligand was designed, synthesized, and assayed.[29] Of the 1001 unbound ligand conformers, conformer number 0 corresponds exactly to the preconformed (bound) ligand structure, whereas conformations 1−1000 correspond to ordered low-energy rotameric variants of increasing energy with dihedral angles selected from a 30° grid, identified with the A* algorithm. In one of the approaches, termed "single-conformer unbound state", each of the 1001 unbound ligand conformers was considered in turn to individually be the unbound state. With conformer 0 treated as the unbound state, this corresponded to rigid-ligand charge optimization.[3,5] With conformer 1−1000 as the unbound state, this corresponded to one form of induced-fit charge optimization.[28] In the second approach, termed "dominant-conformer unbound state", an additional constraint was placed on the ligand. Namely, the conformation selected as unbound state was required to be of the lowest free energy in the set of unbound structure candidates, consistent with physical principles. That is, the unbound state is the ground state of the candidate set. In the work presented here, the unbound candidate set was grown by progressively adding members to examine size effects. This second form of induced-fit optimization is the main contribution of the current work.

To facilitate analysis of the results, the overall binding free energy $\Delta G'_{\text{total}}$ is considered as the sum of two processes: a preconformation step, in which the unbound ligand adopts the preconformed structure in the absence of the receptor (ΔG_{pre}), and a docking step, in which preconformed ligand and receptor associate ($\Delta G'_{\text{dock}}$). The prime symbol indicates that receptor desolvation and nonelectrostatic contributions in the docking step, which are independent of ligand unbound conformation and ligand charge distribution, and, thus, are effectively constant, are neglected. Their values are 28.91 and −75.13 kcal/mol, respectively.

Single-Conformer Unbound State. Each member of the collection of 1001 ligand conformations was treated individually as the unbound state, and the rigid-ligand charge optimization approach was used directly to optimize for chemically reasonable partial atomic charge distributions, leading to the most favorable binding free energy. The nonconvex optimization problem was solved using 100 multiple starts with CONOPT3[52,53] with the constraint $q_{\text{total}} = 0$ (neutral ligand). All trajectories reached the same solution, which indicates a wide basin of attraction for $\Delta G'_{\text{total}}$, in terms of the ligand partial atomic charges.

Figure 3A shows the computed optimal binding affinity $\Delta G'_{\text{total}}$ for each of the 1001 individual unbound ligand

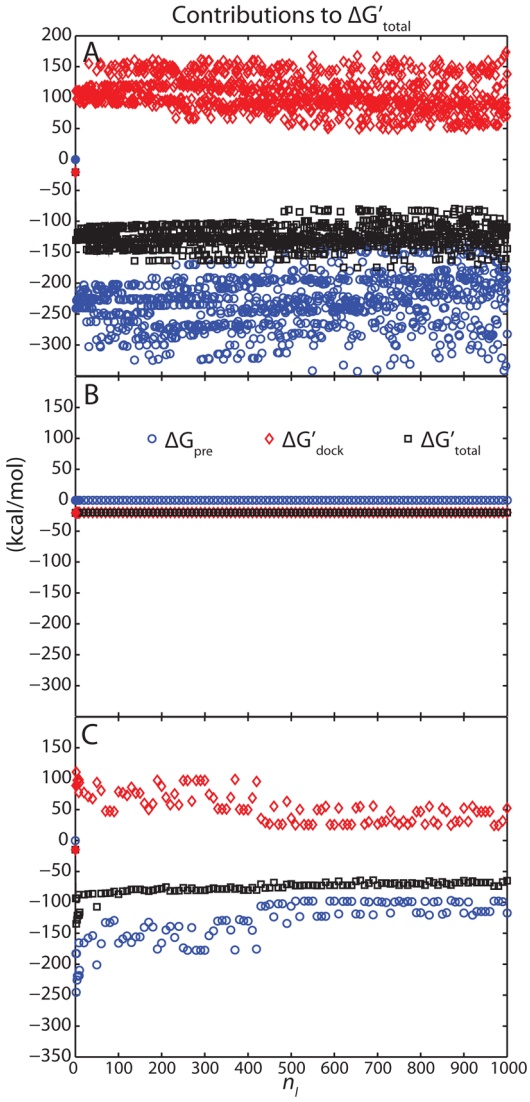

Figure 3. Contributions to $\Delta G'_{\text{total}}$ from three types of optimization: (A) the single-conformer unbound state method, (B) the dominant-conformer unbound state method, and (C) the dominant-conformer unbound state method, with the preconformed structure removed from the unbound candidate set. Filled symbols are used for $n_l = 0$, whereas open symbols are used for all other values. All three graphs are on the same scale.

conformations (black symbols), as well as the preconformation (ΔG_{pre}, blue symbols) and docking ($\Delta G'_{\text{dock}}$, red symbols) contributions. Dramatically different results were observed for rigid binding (structure index $i = 0$; the unbound and bound ligand conformation are identical) and nonrigid binding (structure index $i \in \{1, 2, ..., 1000\}$; the ligand changes conformation on binding). The binding affinity improved from roughly −20 kcal/mol for rigid binding to approximately −150 kcal/mol for nonrigid cases. Rigid binding involved no change in preconformational free energy ($\Delta G_{\text{pre}} = 0$ kcal/mol) and a modestly favorable docking change ($\Delta G'_{\text{dock}} \approx -20$ kcal/mol); nonrigid binding involved a surprisingly favorable preconformational gain (ΔG_{pre} of roughly −300 to −200 kcal/mol) and an unusually large *unfavorable* docking change ($\Delta G'_{\text{dock}}$ of approximately +100 to +150 kcal/mol; see Figure 3A). That is, rigid binding charge optimization resulted in optimized charges that were truly complementary to the binding site;

dx.doi.org/10.1021/ct200931c | J. Chem. Theory Comput. 2012, 8, 4580−4592

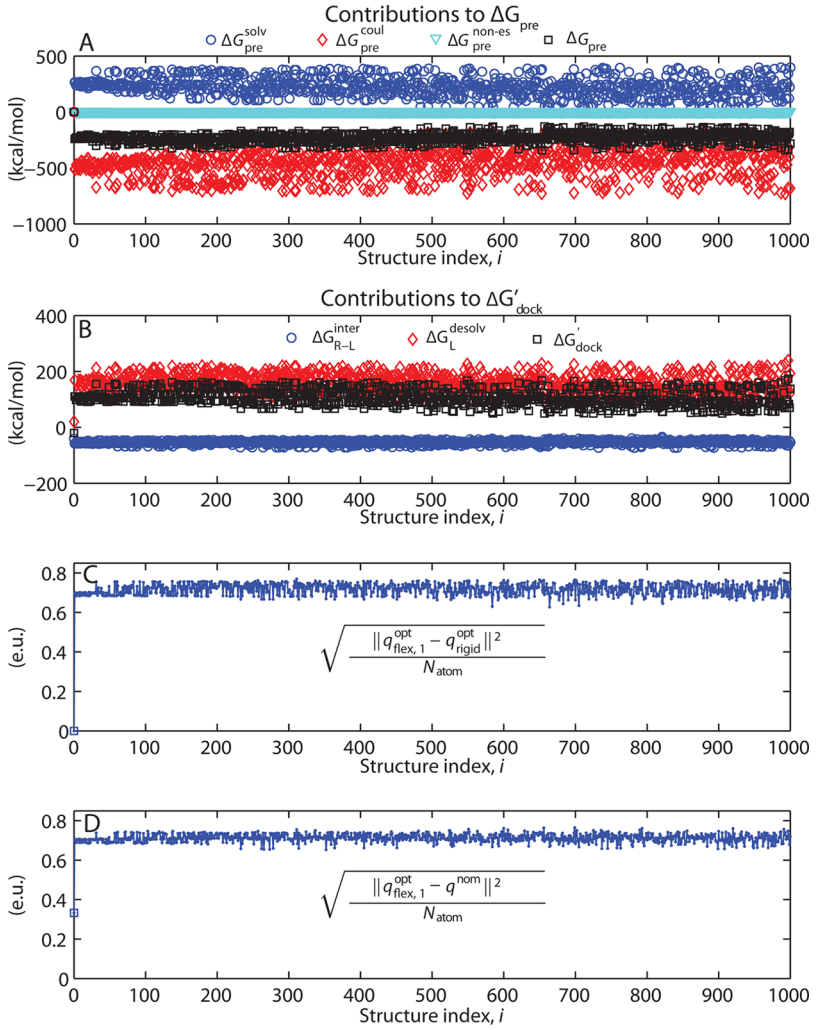

Figure 4. Energetic and atomic-charge analysis of single-conformer studies: (A) contributions to ΔG_{pre}, (B) contributions to $\Delta G'_{dock}$, (C) root-mean-square deviation (rmsd) between single-conformer optimum and rigid optimum partial atomic charge distributions, and (D) rmsd between single-conformer optimum and nominal charge distributions.

flexible charge optimization resulted in charges that were not especially complementary to the binding site (because $\Delta G'_{dock}$ was substantially unfavorable) but that pathologically recovered a large free-energy benefit from changing the conformation from unbound to preconformed free in solution. Further analysis shows the favorable energetics result from a dramatically favorable Coulombic gain ($\Delta G^{coul}_{pre} \approx -500$ kcal/mol) and a more moderate desolvation loss ($\Delta G^{solv}_{pre} \approx 200$ kcal/mol; see Figure 4A). This behavior parallels that seen by one of us in a previous study.[28] This situation presents a physical inconsistency. Although the calculation assumes a given ligand conformation as the unbound state, optimization produces a charge distribution for which another conformation, namely, the preconformed structure, is lower in free energy. That is, the assumed unbound conformation is not the ground unbound state, because, conceptually, the system has a choice of two conformations (p and u). We reasoned that elimination of this unphysical situation might lead to the charge optimization producing useful and meaningful results for flexible binding, which motivated the dominant-conformer unbound state method described in the next subsection.

Before presenting those results, we briefly comment on the mathematical source of the unphysical results. The ligand desolvation matrix (the Hessian of the electrostatic binding free energy) for induced-fit charge optimization is $\Delta\mathbf{L} + \mathbf{L}^{solv}_p + \mathbf{L}^{coul}_p - \mathbf{L}^{solv}_u - \mathbf{L}^{coul}_u$ (eq 22), which is positive semidefinite for rigid binding (where $\mathbf{L}^{solv}_p = \mathbf{L}^{solv}_u$ and $\mathbf{L}^{coul}_p = \mathbf{L}^{coul}_u$ and thus only $\Delta\mathbf{L}$ remains) but need not be for single-conformation unbound state flexible binding. Eigenvectors corresponding to negative eigenvalues present opportunities for improving ligand affinity without limit but are fixed by constraints placed on the ligand charge distribution. These directions produced outsized gains in ΔG_{pre} at the expense of $\Delta G'_{dock}$. Indeed, examination of optimized charges indicates that many terminated at the constraint and differed substantially from both the rigid optimum charge distribution and the nominal one (Figures 4C and 4D). Interestingly, the solvent-screened electrostatic contribution to $\Delta G'_{dock}$ was favorable (about −50 kcal/mol; blue symbols in Figure 4B) but overpowered by the ligand desolvation penalty (~100−150 kcal/mol; red symbols in Figure 4B), which resulted in the unusual large unfavorable $\Delta G'_{dock}$.

These mathematical sources of unphysical behavior correspond to explanations in terms of chemical interaction energetics. Eigenanalysis revealed two mechanisms operating in tandem that were responsible for unphysical behavior. One

dx.doi.org/10.1021/ct200931c | J. Chem. Theory Comput. 2012, 8, 4580−4592

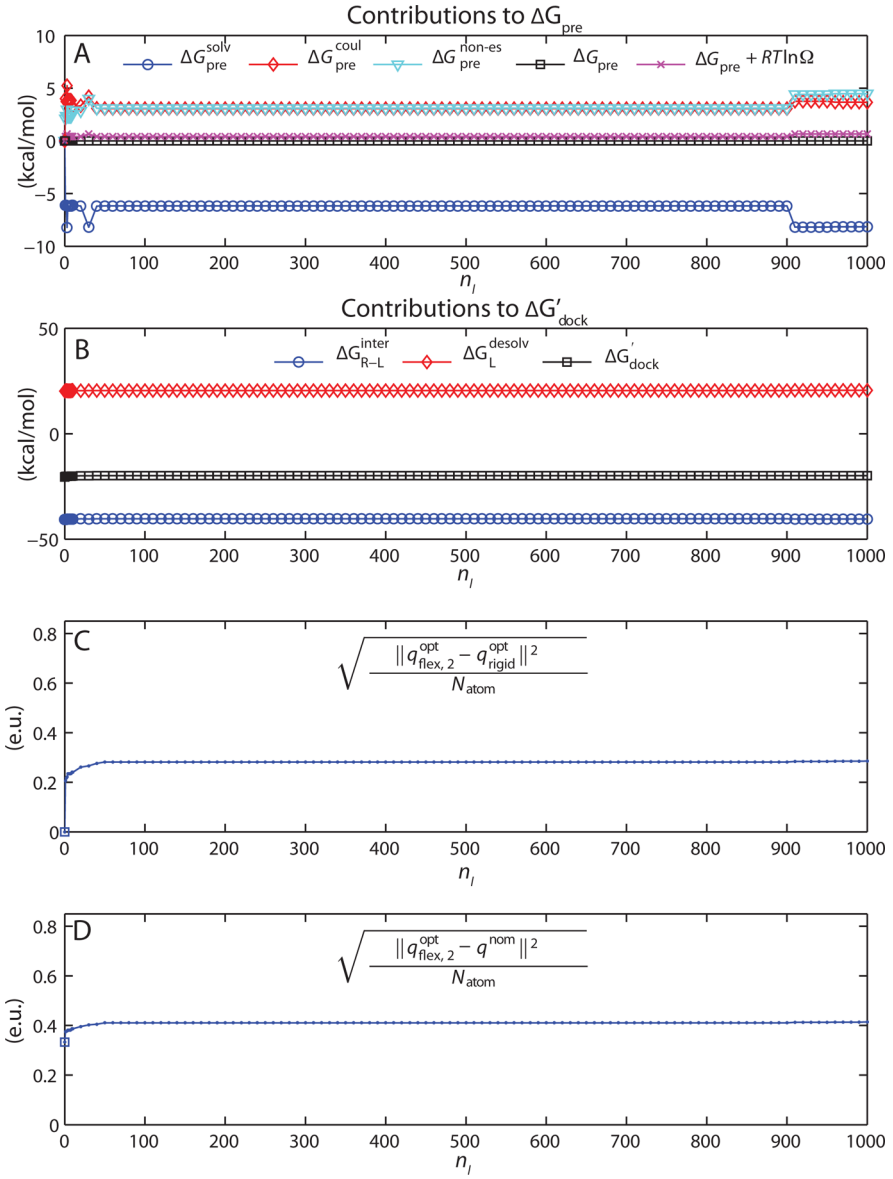

Figure 5. Energetic and atomic-charge analysis of dominant-conformer studies: (A) contributions to ΔG_{pre}, (B) contributions to $\Delta G'_{dock}$, (C) rmsd between dominant-conformer optimum and rigid optimum partial atomic charge distributions, and (D) rmsd between dominant-conformer optimum and nominal charge distributions.

mechanism was that atoms whose solvent exposure increased in going from the actual unbound to preconformed unbound state could improve electrostatic affinity by increasing charge magnitude on these atoms and gaining solvation interactions upon preconformation. The second was that atom pairs whose interatomic distance changed upon preconformation could gain affinity by accumulating like charge on atoms that became more distant (or opposite charge on atoms that became closer) upon preconformation.

Dominant-Conformer Unbound State. We next enforced the notion that the unbound state, if represented as a single conformation, should be the ground state among conformations represented in the set of unbound candidates. The new procedure involved a different formulation in which the optimization simultaneously chose a ligand partial atomic charge distribution and a presumed unbound ligand conformation from a set of candidates. The choice was made self-consistently, such that the presumed unbound ligand conformer did have the lowest free energy in the set when computed with the optimum charge distribution, and the optimum charge distribution was selected to minimize the total binding free energy change for preconformation and docking summed. If multiple unbound conformers had the same unbound state free energy (to within 10^{-4} kcal/mol), they were considered equal contributors to the unbound state, each with the same $\Delta G'_{total}$. In a post-processing step, the energetic cost to ΔG_{pre} of using degenerate conformers for the unbound state was treated by adding an entropic penalty of $RT \ln \Omega$ (Ω is the number of energetically equivalent conformers).

A series of individual optimizations was run using an unbound candidate set (U) of increasing size. With the preconformed ligand structure corresponding to index $i = 0$, the set was of the form $U = \{0, 1, ..., n_l\}$. In successive runs, n_l was incremented from 0 to 1000. The continuous and smooth formulation of the optimization in eq 25 was used.

dx.doi.org/10.1021/ct200931c | J. Chem. Theory Comput. 2012, 8, 4580–4592

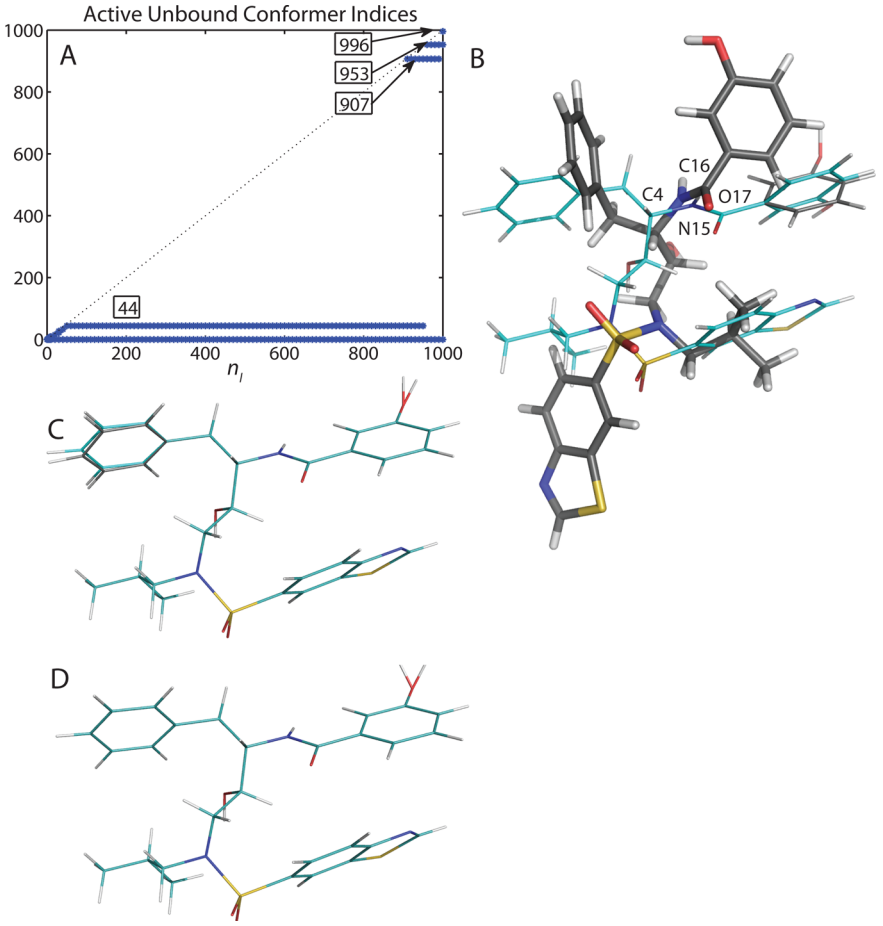

Figure 6. (A) Indices of selected unbound conformations for incrementally increasing unbound conformer candidate set for the dominant-conformer studies. The boxed numbers are index numbers. (B–D) Some selected conformations of unbound MIT-2-KB-98 aligned to each other. (B) Conformer 44 (cyan) and its replacement conformer 953 (gray); conformer 0 (black sticks) was aligned to conformer 44 with all atoms. (C) Conformer 907 (cyan) and 953 (gray) coexisted with conformer 0 as selected unbound conformers. (D) Conformer 907 and its replacement conformer 996 only differed in the position of a hydrogen atom in the hydroxyl group attached to the benzene ring.

Figure 3B shows the overall energetic results. The computed optimal binding affinity $\Delta G'_{\text{total}}$ was approximately −20 kcal/mol and did not vary systematically with n_l; this affinity was due to a zero preconformation energy ΔG_{pre} and dominated by $\Delta G'_{\text{dock}}$ ($\approx$ −20 kcal/mol). These results are dramatically different from those obtained with the single-conformer unbound state method (see Figure 3A). By contrast, the new results produce modest and consistent energetics. The new formulation incorporating the ground-state constraint both eliminated the large favorable ΔG_{pre} and maintained a favorable $\Delta G'_{\text{dock}}$. This latter point is consistent with intuition of electrostatic complementarity.

Further analysis of the results is presented in Figures 5 and 6. The preconformation free-energy change was zero, which suggests the preconformed structure was selected either as the only unbound ligand conformation or as one of several. The Coulombic, solvation, and nonelectrostatic contributions to the preconformation free energy were not always individually zero but summed to zero (different from the single-conformer case), which suggests degenerate unbound conformers were sometimes selected. The docking free-energy contribution was similar to that for the rigid case. The charge distribution was closer to the rigid optimum than to the nominal distribution, but it could be distinct from either. Thus, interestingly, much of the energetics of rigid binding was reproduced for this dominant-conformer case, although differences in energetics and charge distributions did result. The optimized ligand charge distribution $\mathbf{q}^{\text{opt}}_{\text{flex},2}$ (when n_l = 1000) is reported in Table 1, along with that derived with a rigid ligand ($\mathbf{q}^{\text{opt}}_{\text{rigid}}$) and the nominal charge distribution ($\mathbf{q}^{\text{nom}}$).

Figure 6A indicates the indices of conformations selected as the unbound state for each optimization. The preconformed structure (index i = 0) was selected in every optimization, almost always with a small number (up to 2) of other conformations of equivalent free energy. As the size of the unbound candidate set was progressively increased, certain structures were selected until replaced by newer ones. After n_l reached 50, conformer 44 was always chosen, along with conformer 0, until replaced by conformer 953. Conformer 907 joined conformer 44 as one of the selected conformations until being replaced by conformer 996. Conformers 44 and 953 (which replaced conformer 44) only differ in the orientation of the hydroxylated benzene ring (see Figure 6B; upper right). In conformer 953, it was almost in the same plane as the neighboring scaffold atoms N15, C16, N17, and C4 (labeled). Conformer 907 resembled conformer 953, except for a small torsional angle difference for the benzene ring and in the position of the hydroxyl group hydrogen atom (Figure 6C). These two conformers were selected, together with conformer 0, until conformer 907 was replaced by conformer 996, which

dx.doi.org/10.1021/ct200931c | J. Chem. Theory Comput. 2012, 8, 4580–4592

Table 1. Three Sets of Ligand Charge Distributions[a]

atom	$\mathbf{q}^{nom}$	$\mathbf{q}^{opt}_{rigid}$	$\mathbf{q}^{opt}_{flex,2}$	atom	$\mathbf{q}^{nom}$	$\mathbf{q}^{opt}_{rigid}$	$\mathbf{q}^{opt}_{flex,2}$
C1	−0.032	0.813	0.850	H36	0.075	0.300	−0.071
H2	0.049	−0.196	−0.120	H37	0.087	0.248	−0.095
H3	0.065	−0.064	−0.001	H38	0.015	0.030	0.128
C4	0.017	0.233	−0.847	H39	0.016	−0.305	−0.221
H5	0.114	−0.210	0.043	H40	0.040	−0.096	−0.061
C6	0.103	0.169	0.850	H41	0.076	−0.010	0.021
H7	0.119	0.111	0.034	H42	0.066	−0.094	−0.106
C8	−0.008	0.850	0.850	H43	0.051	0.012	−0.017
H9	0.091	0.058	−0.006	C44	0.020	0.342	0.668
H10	0.109	−0.093	−0.211	N45	−0.542	−0.168	−0.164
N11	−0.335	−0.850	−0.850	C46	−0.182	−0.274	−0.525
S12	0.528	−0.206	−0.508	C47	−0.070	0.255	0.360
O13	−0.430	0.013	0.017	C48	0.154	0.135	0.120
O14	−0.419	0.050	0.109	S49	−0.084	−0.392	−0.411
N15	−0.242	−0.850	−0.850	C50	0.461	−0.066	−0.110
C16	0.494	0.561	0.850	C51	−0.032	−0.084	−0.172
O17	−0.545	−0.312	−0.306	C52	−0.279	−0.128	−0.123
H18	0.156	0.292	0.375	H53	0.250	0.159	0.211
C19	−0.118	0.070	−0.334	H54	0.199	−0.072	−0.070
C20	−0.165	0.048	0.357	H55	0.157	0.031	0.032
C21	−0.071	0.113	0.191	H56	0.181	0.169	0.177
C22	0.281	−0.062	−0.188	C57	−0.022	−0.412	−0.850
C23	−0.224	−0.045	−0.049	C58	−0.154	0.164	0.743
O24	−0.607	0.176	0.193	C59	−0.152	−0.005	−0.408
C25	−0.140	−0.713	−0.673	C60	−0.132	0.038	0.293
H26	0.117	−0.043	−0.125	C61	−0.172	−0.031	−0.275
H27	0.138	−0.001	−0.006	C62	−0.072	0.043	0.405
H28	0.169	0.140	0.135	H63	0.118	−0.051	−0.097
H29	0.434	0.015	0.018	H64	0.151	−0.014	0.017
H30	0.166	0.277	0.274	H65	0.134	−0.156	−0.201
C31	0.290	0.355	0.333	H66	0.132	−0.110	−0.052
C32	−0.055	−0.850	0.497	H67	0.134	0.013	−0.101
C33	−0.105	0.056	−0.338	O68	−0.655	−0.099	−0.105
C34	−0.314	0.067	0.007	H69	0.404	0.428	0.372
H35	−0.001	0.227	0.115				

[a]Units of *e*, the magnitude of the electron charge. $\mathbf{q}^{nom}$ is the nominal charge distribution, $\mathbf{q}^{opt}_{rigid}$ is optimized using a rigid ligand, and $\mathbf{q}^{opt}_{flex,2}$ is optimized using a flexible ligand treated with the dominant-state unbound conformer approach using all 1001 candidates.

only differs in the position of the same hydroxyl hydrogen atom.

Thus, the dominant-conformer unbound state method removed the pathological behavior observed in the single-conformer method through elimination of the unphysical situation of the unbound conformer not being the ground state structure. Interestingly, this produced binding energetics extremely similar to rigid-binding charge optimization, and the preconformed ligand structure was consistently chosen as the unbound ground state, sometimes with one to two other degenerate structures. Indeed, an analysis of the Lagrange multipliers from the constrained optimization shows that even when degenerate conformers populated the unbound state, the cost for losing conformer 0 dominated the others substantially. The somewhat varying values for contributions to ΔG_{pre} are due to averaging different multiple members of the unbound ligand state for each optimization. Taken together, these results suggest that there is little or no affinity benefit to changing conformation upon binding, and that adoption of unbound structures corresponding to complementary geometry in the bound state is sufficient to produce optimal affinity. This suggestion is strengthened by similar results obtained when constraining the total ligand charge to other values ($q_{total} \in \{-1, +1\}$), as well as running similar jobs using a crystal structure[29] rather than the design structure as the bound state and updating structure 0 accordingly (data not shown).

As one final control, we reran the dominant-conformer method with the preconformed structure (index 0) removed from the unbound set, so $U = \{1, ..., n_l\}$. This was done to isolate the role of the preconformed ligand structure in the unbound candidate set, because of the many methodological differences between the single-conformer and dominant-conformer methods. The overall results and detailed breakdown (Figure 3C) were generally similar to the single-conformer case, with pathologically favorable affinity arising from unphysically favorable ΔG_{pre} values. These data imply that the preconformed ligand structure should always be considered as a candidate for the unbound conformation, because the alternative not only produces physically unrealistic results but also is physically untrue.

5. DISCUSSION AND CONCLUSION

Here, we have extended charge optimization theory to treat potentially flexible ligands. One additional concept was used to

dx.doi.org/10.1021/ct200931c | J. Chem. Theory Comput. 2012, 8, 4580−4592

ensure physically meaningful results—any conformational change from the unbound state to the preconformed bound conformation must be uphill in free energy (or at least neutral). If it were downhill, then the presumed unbound state would not correspond to the ground state. The addition of this constraint produces physically reasonable optima and eliminates pathologies that appear in its absence.

The flexible binding case was extended from previous work on rigid binding, in which electrostatics could be clearly separated from nonelectrostatic contributions to the binding free energy. The theoretical treatment of the flexible case developed here intimately links nonelectrostatic with electrostatic contributions in the preconformational free-energy contribution in eq 22. The conformational free-energy surface of the unbound ligand can, in principle, affect the optimization of the unbound conformation and its charge distribution. Note that the parameters for calculating nonelectrostatic contributions are assumed independent of ligand partial atomic charges in the treatment presented here. In practice, these parameters are often calibrated to be compatible with these charges in force-field development, and so may be somewhat interdependent. Moreover, to create a perturbed ligand with charge distribution closer to optimal, presumably one would need to alter covalent chemistry, which could further change molecular flexibility and the corresponding internal potential. Nevertheless, an understanding of the charge optimum, in the presence of the current covalent potential, could be quite valuable.

A feature of the problem formulation adopted here is that a set of candidate conformers (including the preconformed ligand structure) is presented as input to the optimization procedure. Optimization selects a conformer (or a few degenerate conformers) for the unbound state and an optimized charge distribution that together minimize the total binding free energy while satisfying all constraints. As the set of unbound conformer candidates is enlarged, the optimization is affected in two ways. The direct effect is that there are more candidates from which to select the unbound conformation. The indirect effect is that if a new conformation is not selected as the unbound ground state, its presence is an additional constraint that limits the charge space available to the eventual ground state, because, to be the ground state, it must be lower in unbound free energy than all other conformers considered. For example, if the previous ground state is chosen as the ground state again, with the optimal charges applied it has to be lower in free energy when unbound than all other conformers including the new conformer added. For the test case used here, the preconformed ligand structure was always selected as the (possibly degenerate) ground unbound state by the optimization, even when the number of candidates grew to 1001. For increasing numbers of candidates, the optimized charge distribution differed from the rigid optimum due to the indirect effect, which diminished the optimum binding affinity by over 0.5 kcal/mol for candidate sets with more than 20 members. For larger sets than those studied here, or for other cases, one imagines that a different conformation might be chosen for the ground unbound state, but this would be a result of needing to satisfy constraints rather than an adaptation to improve affinity.

For the cases studied here, permitting flexibility never improved optimal affinity beyond that achieved by rigid binding optimization, and, frequently, the induced-fit case incurred an affinity penalty, relative to the rigid case. One interpretation of these results is that induced-fit binding, which may have a special role in molecular recognition, may not be an optimization to enhance affinity. This interpretation is based on the theoretical framework developed here, which includes partial atomic charge distributions that remain fixed upon molecular conformational change. It is formally possible that conformational change is designed to create appropriate charge distribution changes that do improve affinity, and we cannot currently rule out this possibility. What we can say, however, is that conformational change upon binding does not appear to enhance affinity by improving the tradeoff between desolvation and interaction energetics, unless it also can somehow simultaneously lead to affinity-enhancing charge distribution changes.

These arguments can be generalized beyond the specific molecular example studied here by noting that induced-fit binding can be conceptualized as a preconformation step, followed by a docking step, as illustrated by the blue row of Figure 1. Constraining the selected unbound conformer to the ground unbound state means that the preconformation step has a free-energy change larger than or equal to zero, with a lower bound of zero. A lower bound on the docking step is the rigid-ligand electrostatic optimum, which is also a lower bound on $\Delta G'_{total}$. This lower bound will be an optimum solution for the induced-fit case unless it violates the ground-state constraint. Thus, the dominant-conformer optimum cannot have higher affinity than the rigid optimum, and any optimum producing an actual ligand conformational change upon binding will have affinity either equivalent to or weaker than the rigid optimization, and be a consequence of having to satisfy the ground-state constraint.

The practical implication for molecular design of high-affinity ligands is that the construction of ligands preconformed for binding and with rigid-ligand optimized partial atomic charges leads to highest affinity in the theory developed here. If the ligand does still change conformation on binding, in many cases, the rigid-ligand optimized partial atomic charges will still be excellent. For the cases studied here, the largest loss in electrostatic affinity from this strategy would be 4.00 kcal/mol (from 8.35 kcal/mol, which is calculated to be the optimal value), and the largest loss in total affinity would be 8.11 kcal/mol (from −66.78 kcal/mol as the optimal value).

The current work adds molecular flexibility to charge optimization theory but still treats each state as a single conformer. Molecular systems actually exist as Boltzmann-weighted distributions of conformations, and the current theory is being extended to treat such ensemble distributions, which could lead to new insights into the sources of conformational changes that accompany binding.

AUTHOR INFORMATION

Corresponding Author

*E-mail: tidor@mit.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Nathaniel Silver, Woody Sherman, Kelly Thayer, and Celia Schiffer for useful discussions and assistance. We also thank an anonymous referee for a number of insightful comments that substantially improved the manuscript. This

work was supported by the National Institutes of Health (GM061300 to M.K.G.; GM065418 and GM082209 to B.T.).

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) Chong, L. T.; Dempster, S. E.; Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1998**, *7*, 206−210.
(5) Kangas, E.; Tidor, B. *J. Chem. Phys.* **1998**, *109*, 7522−7545.
(6) Kangas, E.; Tidor, B. *Phys. Rev. E* **1999**, *59*, 5958−5961.
(7) Lee, L. P.; Tidor, B. *Nat. Struct. Biol.* **2001**, *8*, 73−76.
(8) Lee, L. P.; Tidor, B. *Protein Sci.* **2001**, *10*, 362−377.
(9) Kangas, E.; Tidor, B. *J. Phys. Chem. B* **2001**, *105*, 880−888.
(10) Sulea, T.; Purisima, E. O. *J. Phys. Chem. B* **2001**, *105*, 889−899.
(11) Sulea, T.; Purisima, E. O. *Biophys. J.* **2003**, *84*, 2883−2896.
(12) Bhat, S.; Sulea, T.; Purisima, E. O. *J. Comput. Chem.* **2006**, *27*, 1899−1907.
(13) Kangas, E.; Tidor, B. *J. Chem. Phys.* **2000**, *112*, 9120−9131.
(14) Sherman, W.; Tidor, B. *Chem. Biol. Drug Des.* **2008**, *71*, 387−407.
(15) Kuo, S. S.; Altman, M. D.; Bardhan, J. P.; Tidor, B.; White, J. K. *Proceedings of the 2002 IEEE/ACM International Conference on Computer-Aided Design*; San Jose, CA, Nov. 10−14, 2002; ACM Press: New York, 2002; pp 466−473.
(16) Bardhan, J. P.; Lee, J. H.; Kuo, S. S.; Altman, M. D.; Tidor, B.; White, J. K. *Technical Proceedings of the 2003 Nanotechnology Conference and Trade Show*; Cambridge, MA, Feb. 23−27, 2003; Taylor & Francis: Boca Raton, FL, 2003; pp 508−511.
(17) Bardhan, J. P.; Lee, J. H.; Altman, M. D.; Leyffer, S.; Benson, S.; Tidor, B.; White, J. K. *Technical Proceedings of the 2004 Nanotechnology Conference and Trade Show*; Cambridge, MA, March 7−11, 2004; Taylor & Francis: Boca Raton, FL, 2004; pp 164−167.
(18) Bardhan, J. P.; Altman, M. D.; Tidor, B.; White, J. K. *J. Chem. Theory Comput.* **2009**, *5*, 3260−3278.
(19) Altman, M. D.; Bardhan, J. P.; White, J. K.; Tidor, B. *Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Shanghai, China, Sept. 1−4, 2005; IEEE: Piscataway, NJ, 2006; pp 7591−7595.
(20) Bardhan, J. P.; Altman, M. D.; Willis, D. J.; Lippow, S. M.; Tidor, B.; White, J. K. *J. Chem. Phys.* **2007**, *127*, 014701.
(21) Altman, M. D.; Bardhan, J. P.; White, J. K.; Tidor, B. *J. Comput. Chem.* **2009**, *30*, 132−153.
(22) Mandal, A.; Hilvert, D. *J. Am. Chem. Soc.* **2003**, *125*, 5598−5599.
(23) Green, D. F.; Tidor, B. *J. Mol. Biol.* **2004**, *342*, 435−452.
(24) Green, D. F.; Tidor, B. *Proteins* **2005**, *60*, 644−657.
(25) Sims, P. A.; Wong, C. F.; McCammon, J. A. *J. Comput. Chem.* **2004**, *25*, 1416−1429.
(26) Armstrong, K. A.; Tidor, B.; Cheng, A. C. *J. Med. Chem.* **2006**, *49*, 2470−2477.
(27) Altman, M. D.; Nalivaika, E. A.; Prabu-Jeyabalan, M.; Schiffer, C. A.; Tidor, B. *Proteins* **2008**, *70*, 678−694.
(28) Gilson, M. K. *J. Chem. Theory Comput.* **2006**, *2*, 259−270.
(29) Altman, M. D.; Ali, A.; Reddy, G. S.; Nalam, M. N.; Anjum, S. G.; Cao, H.; Chellappan, S.; Kairys, V.; Fernandes, M. X.; Gilson, M.; Schiffer, C. A.; Rana, T. M.; Tidor, B. *J. Am. Chem. Soc.* **2008**, *130*, 6099−6113.
(30) Sharp, K. A.; Honig, B. H. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301−332.
(31) Sharp, K. A.; Honig, B. H. *J. Phys. Chem.* **1990**, *19*, 7684−7692.
(32) Gill, P. E.; Murray, W.; Wright, M. H. *Practical Optimization*; Academic Press: New York, 1986.
(33) Surleraux, D. L.; Tahri, A.; Verschueren, W. G.; Pille, G. M.; de Kock, H. A.; Jonckers, T. H.; Peeters, A.; De Meyer, S.; Azijn, H.; Pauwels, R.; de Bethune, M. P.; King, N. M.; Prabu-Jeyabalan, M.; Schiffer, C. A.; Wigerinck, P. B. *J. Med. Chem.* **2005**, *48*, 1813−1822.
(34) Brünger, A. T.; Karplus, M. *Proteins* **1988**, *4*, 148−156.
(35) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(36) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(37) Momany, F. A.; Rone, R. *J. Comput. Chem.* **1992**, *13*, 888−900.
(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen,W.;Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision C.02; Gaussian, Inc.; Wallingford, CT, 2004.
(39) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(40) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(41) Desmet, J.; Maeyer, M. D.; Hazes, B.; Lasters, I. *Nature* **1992**, *356*, 539−542.
(42) Goldstein, R. F. *Biophys. J.* **1994**, *66*, 1335−1340.
(43) Pierce, N. A.; Spriet, J. A.; Desmet, J.; Mayo, S. L. *J. Comput. Chem.* **2000**, *21*, 999−1009.
(44) Gordon, D. B.; Hom, G. K.; Mayo, S. L.; Pierce, N. A. *J. Comput. Chem.* **2003**, *24*, 232−243.
(45) Leach, A. R.; Lemon, A. P. *Proteins* **1998**, *33*, 227−239.
(46) Gilson, M. K.; Honig, B. H. *Proteins* **1988**, *4*, 7−18.
(47) Gilson, M. K.; Sharp, K. A.; Honig, B. H. *J. Comput. Chem.* **1988**, *9*, 327−335.
(48) Nicholls, A.; Honig, B. H. *J. Comput. Chem.* **1991**, *12*, 435−445.
(49) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978−1988.
(50) Gilson, M. K.; Honig, B. H. *Nature* **1987**, *330*, 84−86.
(51) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B. *Proteins* **1986**, *1*, 47−59.
(52) Drud, A. *ORSA J. Comput.* **1992**, *6*, 207−216.
(53) Drud, A. *CONOPT*; ARKI Consulting and Development A/S: Bagsvaerd, Denmark, 2003.
(54) Brooke, A.; Kendrick, D.; A., M.; Raman, R.; Rosenthal, R. E. *GAMS: A User's Guide*; GAMS Development Corporation: Washington, DC, 2003.
(55) Abadie, J.; Carpentier, J. In *Optimization*; Fletcher, R., Ed.; Academic Press: London, 1969; pp 33−47.
(56) Drud, A. *Math. Program.* **1985**, *31*, 153−191.
(57) Stafford, R. Random Vectors with Fixed Sum. Available via the Internet at http://www.mathworks.com/matlabcentral/fileexchange/9700, 2006 (accessed Dec. 23, 2008).