

# Multiple drugs and multiple targets: An analysis of the electrostatic determinants of binding between non-nucleoside HIV-1 reverse transcriptase inhibitors and variants of HIV-1 RT

Mona S. Minkara,<sup>1</sup> Pamela H. Davis,<sup>2</sup> and Mala L. Radhakrishnan<sup>1\*</sup>

<sup>1</sup>Department of Chemistry, Wellesley College, Wellesley, Massachusetts 02481

<sup>2</sup>Disabilities Office, Wellesley College, Wellesley, Massachusetts 02481

## ABSTRACT

We present a systematic, computational analysis of the electrostatic component of binding of three HIV-1 RT inhibitors—nevirapine (NVP), efavirenz (EFV), and the recently approved rilpivirine (RPV)—to wild-type (WT) and mutant variants of RT. Electrostatic charge optimization was applied to determine how suited each molecule's charge distribution is for binding WT and individual mutants of HIV-1 RT. Although the charge distributions of NVP and EFV are rather far from being optimal for tight binding, RPVs charge distribution is close to the theoretical, optimal charge distribution for binding WT HIV-1 RT, although slight changes in charge can dramatically impact binding energetics. Moreover, toward the L100I/K103N double mutant, RPVs charge distribution is quite far from optimal. We also determine the contributions of chemical moieties on each molecule toward the electrostatic component of binding and show that different regions of a drug molecule may be used for recognition by different RT variants. The electrostatic contributions of certain RT residues toward drug binding are also computed to highlight critical residues for each interaction. Finally, the charge distribution of RPV is optimized to promiscuously bind to three RT variants rather than to each one in turn, with the resulting charge distribution being a compromise between the optimal charge distributions to each individual variant. Taken together, this work demonstrates that even in a binding site considered quite hydrophobic, electrostatics play a subtle yet varying role that must be considered in designing next-generation molecules that recognize rapidly mutating targets.

Proteins 2012; 80:573–590.  
© 2011 Wiley Periodicals, Inc.

**Key words:** continuum electrostatics; charge optimization; component analysis; HIV-1 reverse transcriptase; promiscuity; nevirapine; efavirenz; rilpivirine; binding.

## INTRODUCTION

HIV-1 reverse transcriptase (RT) has been a key drug target in the treatment of HIV, with several clinically approved inhibitors currently available. However, a major challenge in the effective treatment of HIV is the widespread emergence of drug resistance. A requirement for next-generation RT inhibitors is that they promiscuously bind to many mutant RT variants, thereby creating a higher genetic barrier to resistance. To rationally design inhibitors with broad recognition profiles, one must understand the physicochemical determinants of both tight and promiscuous binding in a particular system. In this work, we analyze the electrostatic determinants of tight and promiscuous binding in the HIV-1 RT system, noting that electrostatics often play a key role in molecular recognition and help mediate either specific or promiscuous binding in many systems.<sup>1–7</sup>

HIV-1 RT is a heterodimer consisting of p66 and p51 subunits, the larger of which provides the binding sites for two classes of drugs.<sup>8</sup> The first class, nucleoside RT inhibitors (NRTIs), are nucleoside analogues that bind in the active site to competitively inhibit enzymatic function of RT.<sup>9</sup> The other class, non-nucleoside RT inhibitors (NNRTIs), is the focus of this work; these inhibitors bind noncompetitively<sup>10,11</sup> in a hydrophobic area several angstroms from the nucleotide binding site.<sup>12,13</sup> Upon binding an NNRTI, HIV-1 RT undergoes a conformational change<sup>14</sup> thought to inactivate the enzyme by altering its kinetics,<sup>15</sup> although the exact mechanism by which this is accomplished is not fully understood and is addressed in several studies and reviews.<sup>16–19</sup> One of the first NNRTIs used in treatment was nevirapine (NVP),<sup>20</sup> a dipyrindodiazepinone with a low barrier to resistance, as it loses affinity and potency

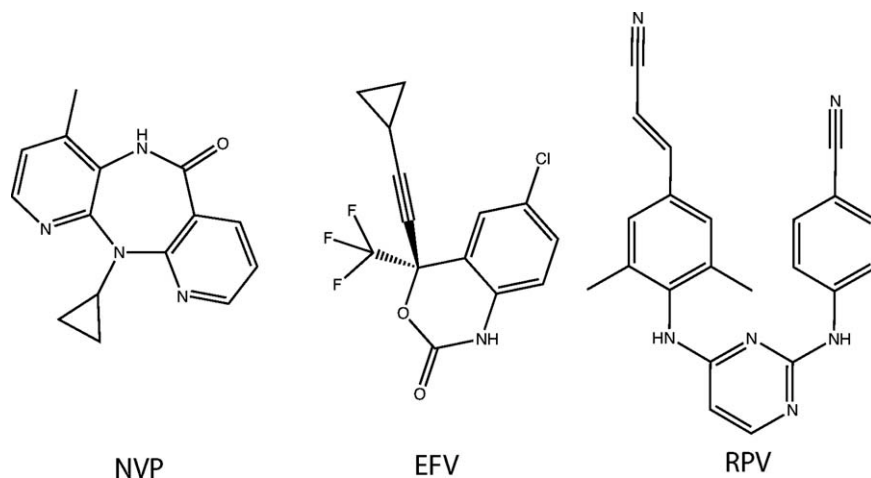
Additional Supporting Information may be found in the online version of this article.  
Grant sponsors: Howard Hughes Medical Institute, Wellesley College.

\*Correspondence to: Mala L. Radhakrishnan, Department of Chemistry, Wellesley College, 106 Central Street, Wellesley, MA 02481. E-mail: mradhakr@wellesley.edu

Received 6 July 2011; Revised 13 September 2011; Accepted 6 October 2011

Published online 18 October 2011 in Wiley Online Library (wileyonlinelibrary.com).

DOI: 10.1002/prot.23221

**Figure 1**

The chemical structure of each drug molecule considered in this study.

upon single amino acid mutations in RT.<sup>21,22</sup> These mutations include Y181C and Y188C, which, from structural analyses, are thought to cause a significant loss of stacking and hydrophobic interactions with NVP.<sup>23</sup> A second-generation inhibitor, efavirenz (EFV),<sup>24</sup> is smaller and makes fewer contacts with Tyr181<sup>23</sup> and Tyr188,<sup>25</sup> thus enabling it to lose less potency against strains with these mutations,<sup>24,25</sup> relative to NVP. However, although still binding more tightly than NVP, EFV loses potency in the presence of K103N, a mutation with significant cross resistance<sup>26–29</sup> that appears to create a kinetic “gate” blocking drug entry into the binding region of RT.<sup>28,30</sup> Recently, a new class of HIV-1 NNRTIs, the diarylpyrimidines (DAPYs), show great promise in maintaining potency against such historically resistant strains. These drugs include etravirine (ETR)<sup>31</sup> as well as the more recently approved rilpivirine (RPV).<sup>32,33</sup> Structural analyses reveal that conformational flexibility and the existence of several possible binding modes play a crucial role in these inhibitors’ ability to bind promiscuously to many variants.<sup>10,33,34</sup> Nevertheless, mutations such as K101P and Y181I, as well as accumulations of other mutations, can confer resistance to ETR<sup>22</sup> and RPV.<sup>35</sup>

There have been many previous computational studies to understand the structural determinants of binding between the NNRTI molecules considered in this work (Fig. 1) and RT variants. Monte Carlo simulations coupled with experimental data<sup>36–38</sup> or QSAR approaches<sup>39–41</sup> have been used to build predictive models for and to highlight physical determinants of inhibitor binding. Other combined approaches involving molecular dynamics have resulted in predictive models that provide insight into physical properties of tight binding.<sup>42,43</sup> The roles of specific amino acid residues in

mediating inhibitor interactions in either WT or mutant complexes have been quantified through multiple methods. Quantum mechanical (QM) approaches<sup>44–49</sup> have in part highlighted the importance of Lys101’s electrostatic interactions with EFV while calling into question the role of stacking interactions between Tyr181 and NVP. QM approaches have also implicated C—H... $\pi$  interactions as being important in the NVP-RT interaction.<sup>50</sup> The effects of specific mutations or residues on binding inhibitors have also been assessed through molecular dynamics simulations, molecular mechanics calculations, or free-energy perturbation methods,<sup>51–56</sup> and several of these studies implicate residues such as Lys101 and Lys103 as important for binding EFV and ETR. The role of water molecules has also been discussed in mediating interactions between RT and drugs such as NVP<sup>57</sup> and ETR and related DAPY analogues.<sup>56</sup> Several studies have also considered the effect of binding NNRTIs on global RT dynamics to elucidate the structural mechanisms of NNRTI inhibition and to understand dynamic structural determinants of binding.<sup>17,18,58,59</sup> Although the information derived from these studies has been substantial, this study is unique in that it provides a comprehensive, comparative, and evaluative analysis of multiple chemically diverse drugs, each in complex with multiple RT variants. Moreover, it focuses on the electrostatic properties of the drug–target interaction, accounting for the crucial effects of desolvation upon binding, which some previous studies—notably many using QM—omitted. Additionally, as there are key hydrogen-bonding interactions between certain drugs and RT, the NNRTI binding pocket is considered quite hydrophobic<sup>10</sup>; consequently, an analysis that focuses on electrostatics in this system could reveal subtle yet interesting

binding determinants that may be “hidden” in an overall energetic analysis and also may be less obvious than those in a highly polar or charged binding site.

In this work, we systematically analyze and compare the electrostatic interactions between each of three drugs and both WT and clinically relevant mutant RT targets to further understand the electrostatic determinants of tight and promiscuous binding. We focus our study on three inhibitors that span both a wide chemical space and the longitudinal history of RT drug development—NVP (first-generation), EFV (second-generation), and RPV (recently approved). Our analysis quantifies the extent to which each drug’s charge distribution is suited for tight binding toward each RT variant, as well as the energetic contribution of various chemical moieties on each drug and certain RT residues toward the electrostatic binding free energy. To quantitatively assess the optimality of each drug’s charge distribution for binding RT, we use the method of electrostatic charge optimization, first developed by Lee, Kangas, and Tidor<sup>60,61</sup> and previously applied to several systems.<sup>62–72</sup> To quantify the contribution of various molecular moieties toward the electrostatic binding free energy, we used electrostatic component analysis techniques similar to those previously applied to various protein–protein and protein–small molecule complexes.<sup>62,72–75</sup> We also optimize the charge distribution of RPV for broad molecular recognition to promiscuously bind to three RT variants. Our approach is similar in aim to previously developed formalisms for multitarget affinity optimization<sup>71,76</sup> that were in part applied to the HIV-1 protease system,<sup>71</sup> but here it is presented as a constrained, convex optimization problem.

Our results demonstrate that while the charge distributions of NVP and EFV are quite dissimilar from their hypothetical, optimal counterparts for binding WT RT, the charge distribution of RPV is quite similar to its WT-optimized counterpart, although slight changes in charge values can greatly improve RPVs binding energetics toward WT RT. We also show that RPVs charge distribution is highly suboptimal for binding the L100I/K103N mutant and that a more muted, hydrophobic charge distribution would allow RPV to more broadly recognize selected RT variants. Additionally, we demonstrate that the drug and target molecular components that contribute most significantly toward or against the binding interaction often depend on the binding partner, although certain components, such as Lys101, Lys103 on RT, are often among the most favorable or unfavorable contributors. Taken together, this work provides a systematic integration of charge optimization and component analysis techniques to provide a comparative study between multiple drug molecules and multiple clinically important target variants within the HIV-1 RT system. Moreover, our systematic analysis of RPV in complex with multiple RT variants demonstrates that electrostatics

**Table 1**

The HIV-1 RT Variant/Drug Complex Crystallographic Structures Studied, Including Protein Data Bank (PDB) Codes (www.pdb.org)

Drug	RT variant	PDB code	Resolution (Å)	References
NVP	WT	1VRT	2.2	77
NVP	Y188C	1JLF	2.6	23
NVP	K103N	1FKP	2.9	25
NVP	Y181C	1JLB	3.0	23
EFV	WT	1FK9	2.5	25
EFV	K103N	1FK0	2.9	25
EFV	Y181C	1JKH	2.5	23
RPV	WT	2ZD1	1.8	33
RPV	L100I/K103N	2ZE2	2.9	33
RPV	K103N/Y181C	3BGR	2.1	33

can play a subtle yet important role even in binding sites that are fairly hydrophobic, and its role should be considered in mediating promiscuous binding toward multiple mutant variants.

## MATERIALS AND METHODS

### Structure preparation

Studies were initiated using four crystalline structures of HIV-1 RT variants bound to NVP, three crystalline structures of RT variants bound to EFV, and three crystalline structures of RT variants bound to RPV (Table 1). For each drug, at least one complex with a resolution  $\leq 2.5$  Å was included to ensure as accurate as possible a model across the drugs given the available experimental data.

In each structure, water molecules with fewer than three potential nonwater hydrogen-bonding interactions were eliminated, as were explicitly resolved solvent ions. The amide groups of asparagine and glutamine residues were flipped as needed using the results from NQ-flipper<sup>78,79</sup> as a guide as well as visual examination of the possible hydrogen bonding contacts as necessary. Imidazole groups of histidines were assigned tautomeric states and flipped if necessary by manual inspection of surrounding contacts. In RT K103N structures, calculations were done using both possible asparagine flip-states at position 103. As a detailed, full-energetic analysis of bound and unbound states for each asparagine conformer is beyond the scope of this work, results using the conformer with the more favorable electrostatic binding free energy using the drugs’ actual charge distributions are shown in each case, with qualitative differences between flipped states discussed in the Results. In a subset of the structures, the sulfinoalanine residue at position 280 on the p66 subunit ( $\sim 30$  Å or more from the drug molecule in all cases) was computationally modified to a cysteine residue.

Hydrogen atoms were modeled onto structures using the hydrogen-building (HBUILD) facility<sup>80</sup> of CHARMM,<sup>81</sup> with the CHARMM22 parameter set and

force field<sup>82</sup> and the TIP3P water model.<sup>83</sup> CHARMM22 atom types were assigned to each atom in NVP, EFV, and RPV; a handful of atoms within drug molecules were assigned alternative CHARMM22 atom types due to the lack of complete parameter availability, but these assignments were expected to have a negligible effect on the structure preparation process as all crystallographically resolved atom positions were fixed during structure preparation. One exception was with the building of hydrogen atoms on the cyclopropyl rings of NVP and EFV. For NVP, the position of the single hydrogen atom on the cyclopropyl carbon proximal to the rest of the molecule was manually adjusted to be similar to its position in the quantum-mechanically geometry-optimized structure. For EFV, this hydrogen was placed similarly to its position in a quantum-mechanically geometry-optimized structure in which angles between the cyclopropyl ring carbons and the alkyne carbons were constrained to the crystallographic values.

Point charge magnitudes for drug molecules were computed by means of the two-stage RESP method.<sup>84</sup> Hydrogen atoms within a methyl group were not constrained to have identical charges, but fitted charges among a group were always within  $0.03e$  (and generally less than  $0.02e$ ) of each other and were therefore assigned arbitrarily to the three hydrogens; a permutation of these charges in one complex altered the binding free energy by less than  $0.04$  kcal/mol. Gaussian03<sup>85</sup> was used to calculate the electrostatic potential used for the RESP fitting procedure, with the geometry of each molecule first optimized at the Hartree-Fock/6-31G level and the electronic wave functions and electrostatic potentials obtained at the Hartree-Fock/6-31G\* level, shown previously to accurately reproduce experimental free energies of solvation via a continuum solvent model.<sup>86</sup> Our goal was to obtain a charge distribution for each drug that reflected its general binding conformation; thus, each geometry optimization used a starting conformation with heavy atoms similar to the crystal conformation of each drug in the wild-type complex. Robustness of RESP-derived charges was evaluated using optimizations starting with each of the exact (WT and non-WT) crystallographic heavy-atom conformations considered in the study, and the resulting RESP-derived charges were extremely similar (RMSD  $\sim 0.01e$  or less between all pairs of complexes involving the same drug, except for PDB ID 1FKO, which had RMSD  $\sim 0.03e$  to the other EFV complexes). The optimized geometries of each drug overlaid with its WT complex crystal coordinates are shown as Supplementary Information (Fig. S1), and it is evident that there are no global conformational changes, providing confidence that the charges reflect the actual binding conformations of each drug.

Areas of missing crystallographic density were adjacent to residues that were at least  $20 \text{ \AA}$  (and generally much farther) from the drug molecule, and the adjacent resi-

dues were patched with capping groups that were subsequently energy minimized. The geometry around nitrogen atoms in each of the two NH linker groups on RPV and nitrogen atoms adjacent to carbonyl groups on NVP and EFV was assumed to be planar. Geometry optimizations on all three molecules in which hydrogen atoms were initially placed out of the C—N—C plane at either the HF/6-31G level or the B3LYP/6-31G\* level resulted in either a planar or a slightly pyramidal (improper dihedral angle  $>160$  degrees) geometry around such atoms. The potential effects of pyramidal geometries around the nitrogen atoms are further addressed in the “Discussion” section.

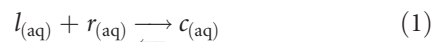
### Structural alignment

ProFit v. 3.1 (A.C.R. Martin and C.T. Porter, <http://www.bioinf.org.uk/software/profit/>)<sup>87</sup> was used to carry out structural alignments of WT complexes used to generate certain figures, using the multiple structure alignment feature. All p66 monomer alpha carbon atoms present in all considered structures were used in the fitting.

### Single target affinity optimization

Electrostatic affinity optimization has been developed, discussed, and applied extensively (see “Introduction” section), and the underlying theory will be only briefly discussed here.

Consider the reversible association of a ligand molecule  $l$  and receptor molecule  $r$  to form a complex  $c$  within solvent:



We model the solvent as a high dielectric continuum and the ligand, receptor, and complex as low-dielectric cavities with point charges located at  $n$  and  $m$  atom centers on the ligand and receptor, respectively. Assuming a linear dielectric response and the rigid binding of ligand and receptor, the electrostatic component of the binding free energy of the above reaction can be written as:

$$\Delta G_{\text{elec}} = q_L^T L q_L + q_R^T R q_R + q_L^T C q_R \quad (2)$$

where  $q_L$  and  $q_R$  are  $n$ - or  $m$ -dimensional vectors of the point charge magnitudes at ligand or receptor atom centers, respectively, and  $L$ ,  $R$ , and  $C$  are  $n \times n$ ,  $m \times m$ , and  $n \times m$  matrices that account for the desolvation free energy paid by the ligand or receptor molecule upon binding and the screened Coulombic interaction between ligand and receptor. Specifically, the  $ij$ th element of  $L$  is the one-half the binding-induced potential difference at ligand atom center  $i$  due to a unit charge at ligand atom center  $j$ . The first term of Eq. (2) represents the ligand desolvation penalty, or the unfavorable free energy

change of replacing high-dielectric solvent with a low-dielectric cavity in the shape of the receptor molecule.  $R$  is defined analogously for the receptor, with the second term of Eq. (2) representing the receptor desolvation penalty. The  $ij$ th element of  $C$  is the bound state potential at ligand atom center  $i$  due to a unit charge at receptor atom center  $j$ ; the last term, therefore, represents the screened Coulombic interaction between the two binding partners.

By differentiating Eq. (2) with respect to  $q_L$  and setting the resulting expression to zero, one obtains the “optimal” charge distribution—the hypothetical ligand charge distribution that minimizes the electrostatic free energy of binding to the target among all possible identically shaped ligands with the same atom centers:

$$q_{L,\text{opt}} = -\frac{1}{2}L^{-1}Cq_R \quad (3)$$

Under the rigid binding assumption,  $L$  is positive semi-definite, and therefore  $q_{L,\text{opt}}$  is strictly a minimum. It is well-known that the binding of NNRTIs induces a conformational change in RT, and therefore, one might question the validity of the rigid binding assumption of the model. However, as we are not varying the charge distribution on RT (we optimize only the drug molecules' charges), the deformation penalty paid by the receptor to assume its bound state conformation is constant for a given drug shape/target combination. Operationally, we can, therefore, assume that the unbound “ $r$ ” molecule on the left side of Eq. (1) is the hypothetical RT molecule predeformed into its bound state conformation for a given drug–RT variant interaction.

To obtain necessary matrix elements needed for Eq. (3) above, a locally built multigrid finite-difference numerical solver of the linearized Poisson-Boltzmann Equation<sup>88</sup> was used to solve for bound and unbound state potentials. The explicit  $L$  matrix, the vector product  $Cq_R$ , and the constant receptor desolvation penalty  $q_R^T R q_R$  were calculated for each binding pair. PARSE charges and radii were used for RT atoms.<sup>89</sup> PARSE radii were assigned as appropriately as possible to drug atoms to determine the dielectric boundary. It had been previously reported that PARSE radii for C and N did not accurately quantify the solvation energies of cyano groups, and for such groups, the previously parameterized radii of 1.85 Å and 1.75 Å for carbon and nitrogen were used here, as were the previously evaluated Parm99 AMBER van der Waals radii for F and Cl.<sup>42</sup> Hydrogens on methyl or cyclopropyl groups of drug molecules retained their RESP-derived charges but had zero radii; these charges fell well within the 2.0 Å-radius sphere encompassed by the central carbon. Retained crystallographic waters were assigned to the target binding partner in binding calculations, based on proximity. Dielectric constants of 4 and 80 were assigned to drug/target molecules and solvent,

respectively. The ionic strength was set to 0.145M with a 2-Å Stern layer. For each structure, the potential was solved on a  $225 \times 225 \times 225$  cubic grid using a three-stage focusing procedure in which the structure occupied 23% of the grid, 92% of the grid, and 184% of the grid concentrating on the drug molecule. The average of three slight translations of the grid was used to account for dependencies of the obtained potentials to the grid placement. The grid resolution at the highest focusing was at least 4 grids/Å in all cases.

Upon computing all matrix elements, Matlab (The Mathworks, Natick, MA) was used to calculate optimal charge distributions and to calculate binding components. The approximate sensitivity of the binding free energy to an atom's charge was taken as proportional to the atom's corresponding diagonal element of the ligand desolvation matrix  $L$ .<sup>62,68</sup>

### Component analyses

Equation (2) can be used to determine the binding free energy for any arbitrary ligand charge distribution  $q_L$ . By setting a subset of atom centers to zero and re-evaluating the electrostatic free energy of binding, one can assess the effect of that charge subset on binding, that is, that molecular component's contribution to binding. Here, component groups on each drug molecule were chosen to roughly correspond to specific chemical moieties within the molecule. Subsets of  $q_L$  corresponding to each chemical moiety were set to zero in turn and  $\Delta G$  was re-evaluated using Eq. (2). The change in  $\Delta G$  can be defined as:

$$\Delta\Delta G_{\text{moiety,mult}} = \Delta G_{\text{orig}} - \Delta G_{\text{moiety}=0} \quad (4)$$

where  $\Delta G_{\text{moiety}=0}$  is the free energy change of binding with the moiety's charges set to zero and  $\Delta G_{\text{orig}}$  is the free energy change of binding using the original ligand charge distribution.  $\Delta\Delta G_{\text{moiety,mult}}$  can be physically interpreted as the change in binding free energy resulting from “charging up” a moiety's hypothetical hydrophobic isostere with its partial atomic charges. With this definition, however, the sum of  $\Delta\Delta G_{\text{moiety,mult}}$  over all moieties in the ligand will not add up to  $(\Delta G_{\text{orig}} - q_R^T R q_R)$  because the sum double counts the interactions between each pair of charges from different moieties. To allow for the contributions of each chemical moiety to be additive, we assigned one-half the interaction between pairs of moieties to each moiety:

$$\Delta\Delta G_{\text{moiety,contrib}} = (\Delta G_{\text{orig}} - \Delta G_{\text{moiety}=0}) - q_z^T L_{zn} q_n \quad (5)$$

Here,  $q_z$  are the original ligand charges at the moiety atoms being set to zero,  $q_n$  are the remaining ligand charges, and  $L_{zn}$  is the corresponding submatrix of  $L$ . The last term in Eq. (5) is one-half the interaction free energy between the relevant chemical moiety and the

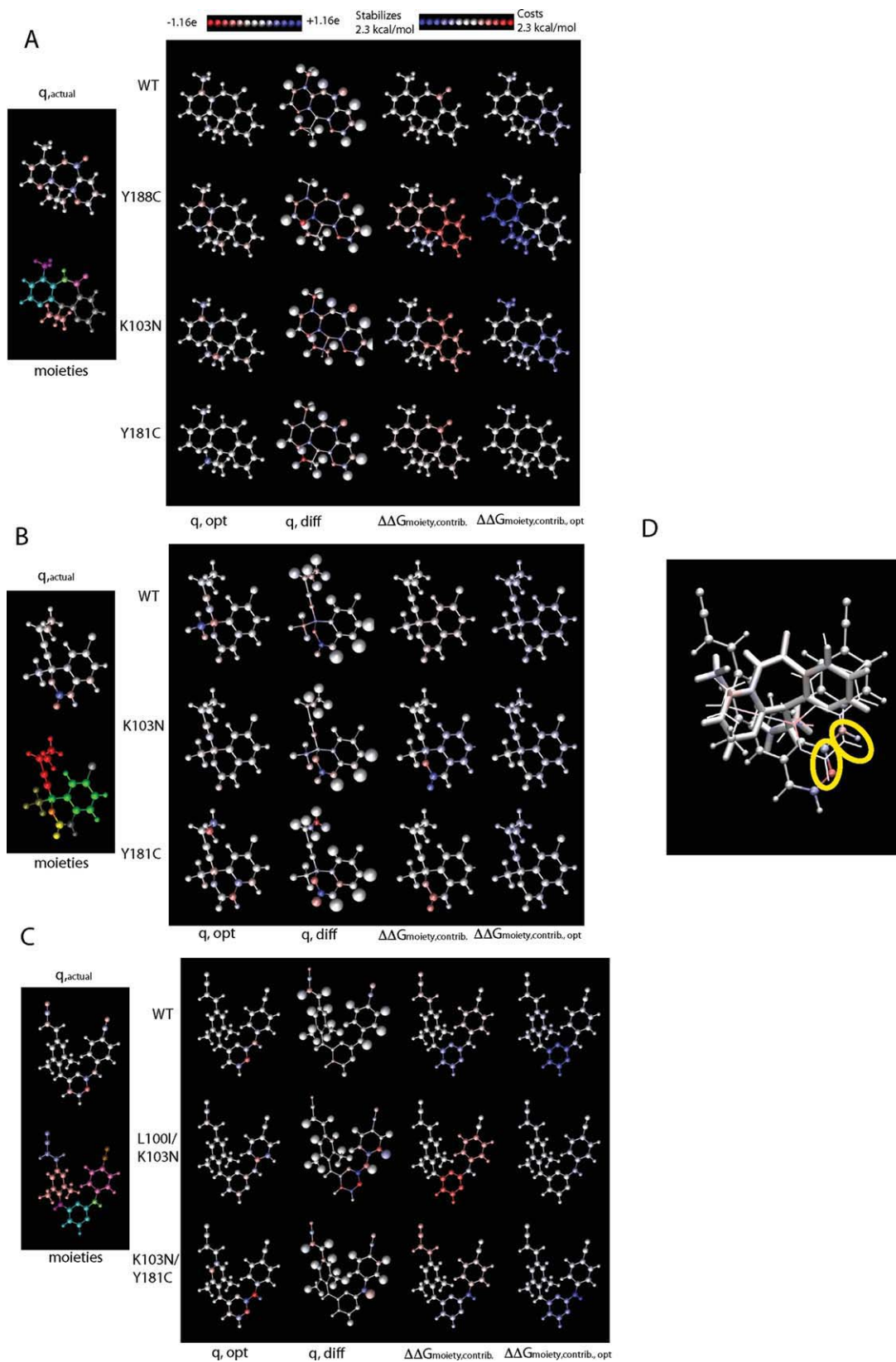


Figure 2

remaining ligand charges. In our analyses of the contributions of drug molecule moieties toward binding, we use the “additive”  $\Delta\Delta G_{\text{moiety, contrib}}$  defined in Eq. (5), as it allows for interpretability when comparing the relative contributions of molecular groups to each other and for assessing each contribution as a fraction of the overall binding free energy change. However, we also quantified the electrostatic contribution of certain amino acid residues on RT toward binding drug molecules. Because the  $R$  matrix was not explicitly calculated, each quantity was obtained by setting the relevant residue’s charges to zero and recalculating the overall binding free energy change between ligand and receptor. In these cases, the “nonadditive”  $\Delta\Delta G_{\text{moiety, mut}}$  [Eq. (4)] was used because only a subset of RT residues were considered and the sum over all considered residues would not have a physically significant interpretation. The definitions used here are similar to the quantities  $\Delta\Delta G^{\text{mut}}$  and  $\Delta\Delta G^{\text{contrib}}$  used in previous work.<sup>73</sup>

### Multitarget affinity optimization

To simultaneously optimize RPV to bind to multiple RT variants, the drug charge distribution  $q_L$  that minimized the sum of the binding free energies toward the set of variants was determined, subject to constraints that the binding free energy toward each individual variant be no greater than a threshold  $t$  above the previously determined global affinity-optimized binding energy toward that variant. The mathematical formulation of the multitarget optimization problem is as follows:

$$\text{Minimize } \sum_i q_L^T L_i q_L + q_L^T C q_{R,i}$$

$$\text{Such that : } \forall i, q_L^T L_i q_L + q_L^T C q_{R,i} \leq W_i + t$$

Here, the variables being optimized are the components of  $q_L$ .  $L_i$  and  $C q_{R,i}$  are the  $L$  matrix and  $C q_R$  vector for the  $i$ th target variant.  $W_i$  is the optimal binding free energy toward the  $i$ th target variant, not

including the constant receptor desolvation penalty. The objective function is convex, as it is a sum of individual convex functions (upward-facing paraboids). The feasible region is also convex, as it is the intersection of multiple convex regions, each of which is a hyperellipse encompassing the feasible charge space in which the affinity toward a given target would be less than  $t$  above its globally optimal value. Therefore, a locally optimal solution found to this optimization problem will be a global optimum.

The threshold value  $t$  was systematically varied until the lowest value (in increments of 0.01 kcal/mol) of  $t$  at which the optimization problem became feasible had been found. This ensured that binding toward each individual variant was sufficiently tight. Optimizations were carried out using GAMS (GAMS Development Corporation, Washington, DC), with the CONOPT solver.<sup>90</sup>

### Figure generation

VMD<sup>91</sup> and Matlab (The Mathworks, Natick, MA) were used to generate all figures and plots shown in this work.

## RESULTS

### Affinity optimization and component analysis of each drug toward WT HIV-1 RT and selected mutants

For each drug-target variant complex, the optimal drug charge distribution was found. Although this optimally tight-binding charge distribution is hypothetical and not reproducible by an actual molecule, a comparison of this charge distribution with the calculated “actual” molecular charge distribution can provide insight into existing determinants of tight binding and potential modifications that could allow the drug’s electrostatic properties to more closely approximate optimality. Additionally, each drug was broken into molecular components, with the partitioning intended to separate

### Figure 2

Affinity optimization and component analysis for the binding of NVP (A), EFV (B), and RPV (C) to WT and mutant RT. For each drug, the actual computed drug charge distribution (“ $q_{\text{actual}}$ ”) is shown in the upper left of its subfigure, with atoms colored according to the legend at top (blue = positive, red = negative). At right, each row contains results for the drug interacting with a particular RT variant, labeled at the row’s left. Column 1 shows the optimal charge distribution ( $q_{\text{opt}}$ ), colored according to the corresponding legend. Column 2 displays “ $q_{\text{diff}}$ ,” the difference between the actual and optimal charge distributions (actual—optimal), colored according to the same legend. Here, a red-colored atom is too negative when compared with its optimal counterpart, and a blue-colored atom is too positive; the size of each sphere is proportional to the approximate sensitivity of the binding free energy to that atom’s charge value (sizes can be compared within a panel but not between panels). Column 3 shows the energetic contribution of various chemical groups on the drug molecule toward the binding free energy as quantified by  $\Delta\Delta G_{\text{moiety, contrib}}$  described in the methods; the groups used to partition the molecule are shown as different colors at left in each panel (“moieties”). The contributions of each group are colored according to the corresponding legend at top—a red-colored group is destabilizing, and a blue-colored group contributes favorably to binding. The fourth column shows the same quantity as the third column, but assuming the hypothetical, optimal drug charge distribution. For ease of visualization, all panels for a given drug are shown in the conformation assumed in its complex with WT. (D) A structural alignment of NVP (licorice), EFV (lines), and RPV (ball-and-stick) in the RT binding pocket, with atoms colored by calculated optimal charge toward WT, according to the same legend. Circled in yellow are chemical moieties on EFV and RPV that make potential hydrogen-bonding interactions with the backbone of Lys101 on RT.

**Table II**

Optimal Monopole ( $Q_{\text{opt}}$ ) and RMSD Between Actual and Optimal Charge Distributions ( $q_{\text{RMSD}}$ ), both in Electron Charge Units ( $e$ ), for Each Drug–Target Pair

NVP			EFV			RPV		
Variant	$Q_{\text{opt}}$	$q_{\text{RMSD}}$	Variant	$Q_{\text{opt}}$	$q_{\text{RMSD}}$	Variant	$Q_{\text{opt}}$	$q_{\text{RMSD}}$
WT	+0.15	0.30	WT	+0.04	0.29	WT	+0.21	0.17
Y188C	+0.69	0.36	K103N	+0.26	0.25	L100I/K103N	+0.22	0.33
K103N	+0.04	0.39	Y181C	+0.06	0.32	K103N/Y181C	+0.34	0.22
Y181C	+0.21	0.32						

out chemically relevant moieties. The electrostatic contribution of each moiety toward binding ( $\Delta\Delta G_{\text{moiety, contrib.}}$ ) was computed, first assuming the drug's "actual" charge distribution and then assuming the drug had the hypothetical optimal charge distribution.

The results from charge optimizations and component analyses are presented for all drug-target variants in Figure 2. In addition, Table II presents the overall monopole ( $Q_{\text{opt}}$ ) of each hypothetical, optimal molecule (not constrained to be integral here) as well as the RMSD between the actual and optimal charge distributions for each drug-target pair considered ( $q_{\text{RMSD}}$ ). All optimal monopoles were slightly positive. With the exception of the EFV—K103N mutant RT complex, the  $q_{\text{RMSD}}$  for a given drug is smaller for its complex with WT RT than toward any mutants, suggesting that each drug is generally most electrostatically optimal for binding WT.

In the following sections, the results presented in Figure 2 are discussed specifically for each drug in turn.

### Nevirapine

In Figure 2(A), the results of charge optimization and component analyses for NVP bound to four RT variants are shown. Toward all variants, the optimal charge distribution is relatively hydrophobic. This result implies that the somewhat polarized NVP molecule is suboptimal for tight electrostatic binding. Indeed, the carbonyl group on NVP contributes unfavorably to binding toward all variants as its red coloring in Figure 2(A), column 3 (all panels) indicates. In each complex, this group is either making no significant interactions with RT or is interacting with an aromatic hydrogen on the side chain of Phe227. Interestingly, no chemical moiety on the actual drug has a significantly favorable electrostatic contribution ( $\Delta\Delta G_{\text{moiety, contrib.}} < -1.0$  kcal/mol) toward binding any variant, and toward the Y188C mutant, multiple molecular components have a significantly unfavorable contribution toward binding. However, for the charge-optimized cases, many contributions are quite favorable, suggesting that charge optimization can significantly alter the role of various moieties in mediating binding. Interestingly, the rightmost column of Figure 2(A) indicates that different chemical moieties on each charge-optimized molecule make the most significant contribution

toward binding each individual variant; for example, in the case of the Y188C mutant, the cyclopropyl ring and the methyl-substituted pyridine ring make the largest contributions, whereas toward the K103N variant, the methyl group and the nonsubstituted pyridine ring do. In the case of the Y181C variant, no chemical moiety makes significantly favorable contributions in the charge optimized structure, suggesting that the binding site geometry and the potential electrostatic contacts on the target are not capable of creating strongly favorable electrostatic interactions, even in the optimal case.

Although the optimal charge distribution is fairly hydrophobic in all cases, there are a few groups, for example, the cyclopropyl ring when binding to the K103N mutant, that obtain a dipole [Fig. 2(A), third row, first panel]. This particular dipole is likely due to the Glu138 residue on the p51 subunit of RT, as a reoptimization upon neutralizing all charges on Glu138 significantly reduced it (data not shown). As with all K103N mutant structures, parallel analyses were carried out assuming each amide flip at Asn103, and Figure 2 shows results for the conformation resulting in the lower binding free energy. The cyclopropyl dipole was even more pronounced in the conformation not shown, due in part to additional effects of Asn103 (data not shown). Nevertheless, this dipole does not contribute significantly to the binding free energy (the cyclopropyl region is nearly white when component analysis is carried out using the optimal charge distribution, Figure 2(A), third row, rightmost panel), perhaps because the binding free energy is relatively insensitive to these charge values [Fig. 2(A), third row, second panel].

Taken together, these data suggest that NVP does not use electrostatics as a significant handle by which to recognize and bind tightly to WT HIV-1 RT or its mutant variants. Nevertheless, with the exception of the Y181C mutant, it is possible in theory for certain groups to contribute favorably to binding were NVPs charge distribution more similar to optimality.

### Efavirenz

Figure 2(B) shows the results of charge optimization and component analyses of EFV bound to three RT variants. The optimal charge distributions toward each variant show more highly charged atoms than those of NVP,



but like NVP, these atoms are at charge locations toward which the binding free energy is relatively insensitive such that these strong dipoles do not significantly contribute to binding. EFV, like NVP, is suboptimal electrostatically in that its polar groups are generally overpolarized. Accordingly, the lactone group contributes unfavorably toward binding WT and the Y181C mutant, as its desolvation penalty is not overcome by its interactions with RT; interestingly, the carbonyl portion of the group contributes favorably toward binding the K103N mutant; in this complex it interacts more closely with the backbone amino group on Lys101 than in the other complexes; the favorability of this group depends on the conformation of Asn103, as it is not significantly favorable with the amide group of Asn103 flipped (data not shown). The other groups on EFV have neither a significant favorable nor unfavorable contribution toward binding any variant.

Unlike NVP, none of the optimal charge distributions contain chemical groups that have significantly favorable contributions toward binding the target. The charge-optimized molecule toward the K103N mutant sacrifices the strongly favorable interaction via the carbonyl group for removing the unfavorable contribution from the ester oxygen, resulting in an overall more favorable binding affinity, but where no group has a significantly favorable contribution toward binding. Overall, the results in Figure 2(B) suggest that like NVP, EFV does not use electrostatics as a major determinant for molecular recognition of WT RT or its variants, although, unlike NVP, the potential for EFV to do so with an altered charge distribution appears to be limited.

### Rilpivirine

The results of charge optimization and component analyses for RPV bound to three RT variants is shown in Figure 2(C). Interestingly, the optimal charge distribution for RPV bound to WT RT is very similar to the actual drug's charge distribution ( $q_{\text{RMSD}} = 0.17$ , lower than the analogous values for NVP and EFV, as shown in Table II), suggesting that RPV's charge values are nearly optimal for tight electrostatic binding to WT RT. The pyrimidine ring contributes favorably to binding WT RT; the same group also is the significant contributor to binding for the optimized drug molecule, demonstrating that RPV is using the same electrostatic binding determinant as its hypothetical, optimal counterpart. The pyrimidine ring and the neighboring NH-linker are involved in two hydrogen-bonding interactions with the amino and carbonyl groups of Lys101 on RT, respectively. This analysis shows that the ligand charge values involved in these hydrogen-bonding interactions are nearly optimal. Nevertheless, the cyanovinyl and cyano groups on RPV are overpolarized and contribute slightly unfavorably toward

binding; these interactions are eradicated in the optimal complex by a neutralization of these groups.

Although the charge distribution is nearly optimal for tight binding to WT RT, it is highly suboptimal for binding the L100I/K103N mutant (Fig. 2, second row). Here, the optimal charge distribution has nearly the opposite polarization as the actual molecule on the pyrimidine ring and the relevant N—H linker atoms. Toward this mutant, no RPV component significantly contributes favorably to binding, and furthermore, even the optimal charge distribution does not have any components that significantly contribute favorably toward binding, although the relevant N—H linker has a slightly favorable contribution. In this structure, the backbone of RT residue Lys101 is greatly shifted relative to the drug molecule when compared with the WT structure, and thus the near-optimal hydrogen bonding interactions observed in the WT structure are suboptimal here. The suboptimality of RPV toward this mutant is qualitatively robust to the flip-state of the Asn103 side chain, although in the conformation not shown, the relevant N—H linker group contributes more favorably in the hypothetical charge optimized structure (data not shown).

The N-H linker on RPV is actually underpolarized toward the Y181C/K103N mutant [third row of Fig. 2(C)]. Toward this mutant, it is the linker group that contributes most significantly toward the binding interaction, but polarizing this group further could lead to an even greater contribution [rightmost panel of third row, Fig. 2(C)]. It is interesting that different molecular components (the pyrimidine ring and the N—H linker) can be the most significant contributor toward different RT variants (WT and Y181C/K103N). The underpolarization of the N—H group toward this mutant is robust to the flip-state of Asn103, although with the alternate conformation, the optimal N—H polarization was somewhat closer to the polarization of the actual molecule, with the group contributing less overall to the actual and optimal binding free energies (data not shown).

These results suggest that, unlike NVP and EFV, certain moieties on RPV not only play a significant role in mediating electrostatic interactions between RPV and RT but also the role greatly varies depending on the RT variant. The pyrimidine ring and N—H linker appear to be optimal for binding WT, overpolarized for the L100I/K103N mutant, and underpolarized for the Y181C/K103N mutant. One might imagine that there may exist a charge distribution that is an ideal “compromise,” such that it is well-suited for binding to all three variants. This idea is further explored below.

### Structural alignment of optimal charge distributions

A structural alignment of the three drugs within the RT non-nucleoside binding pocket is shown in Figure 2(D), with atoms in each drug colored by their optimal

**Table III**

Computed Electrostatic Free Energies of Binding (in kcal/mol) of Drug–RT Variant Complexes

NVP			EFV			RPV		
Variant	$\Delta G_{\text{elec,act}}$	$\Delta G_{\text{elec,opt}}$	Variant	$\Delta G_{\text{elec,act}}$	$\Delta G_{\text{elec,opt}}$	Variant	$\Delta G_{\text{elec,act}}$	$\Delta G_{\text{elec,opt}}$
WT	5.4 (1.5)	2.3	WT	6.4 (0.2)	3.8	WT	7.6 (0.0)	4.1
Y188C	12.7 (3.5)	5.0	K103N	6.0 (−0.3)	4.1	L100I/K103N	11.0 (3.0)	5.5
K103N	6.7 (2.4)	1.2	Y181C	9.6 (1.6)	5.7	K103N/Y181C	7.8 (−0.3)	3.6
Y181C	5.7 (2.7)	2.1						

$\Delta G_{\text{elec,act}}$  is the electrostatic component of binding using the drug's actual charge distribution; the expression in parentheses excludes the receptor desolvation penalty.  $\Delta G_{\text{elec,opt}}$  is the electrostatic component of binding using the calculated optimal drug charge distribution toward each variant.

charge values toward WT RT. The carbonyl oxygen on EFV aligns with the pyrimidine nitrogen on RPV [yellow oval on the left in Fig. 2(D)], and both interact with the backbone amide group of Lys101. N—H groups on both drug molecules are also somewhat aligned [yellow oval on the right in Fig. 2(D)], as these interact with the carbonyl group of Lys101. Nevertheless, the optimal charge distributions of the aligned groups are rather different, and in general, other than most atoms bearing little charge overall, there are no clear commonalities between the optimal charge distributions of the molecules, suggesting the absence of an optimal, molecule-independent electrostatic “signature” within the binding pocket that allows for tight binding in this system. This result may be a consequence of the fact that the NNRTI binding pocket is elastic, changing its structure depending on the drug present.<sup>10</sup>

### Overall electrostatic binding energetics

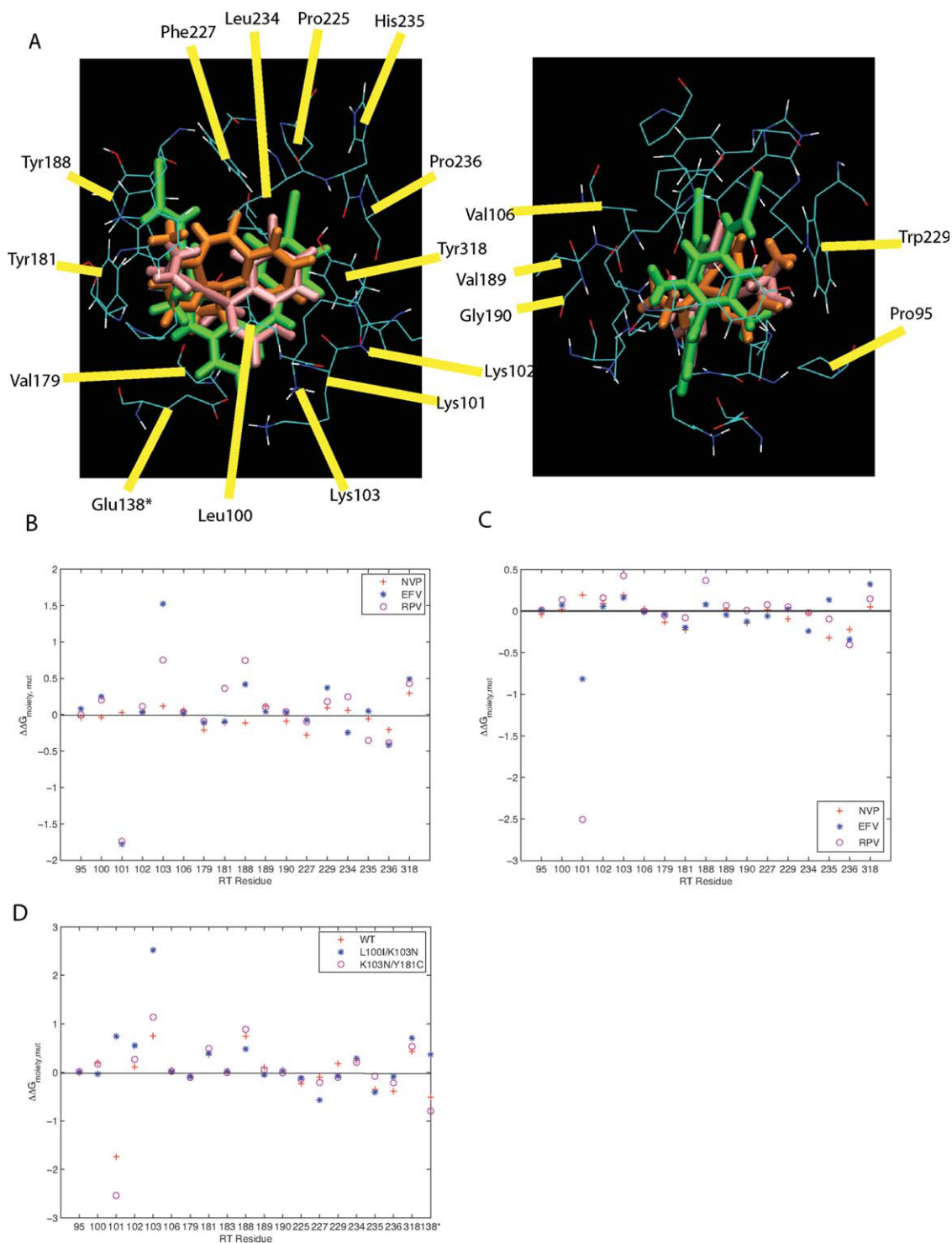
Table III shows the computed energetics of actual and optimal binding for all considered complexes. Both the actual and optimal electrostatic binding free energies are computed to be unfavorable in all complexes. Although the electrostatic component is but one contributor toward the overall binding free energy, it is noteworthy that with the exception of the EFV-K103N mutant complex, the electrostatic binding free energies of all drug–WT complexes are lower than their corresponding drug–mutant complexes. Interestingly, the EFV-K103N mutant is also the only mutant whose  $q_{\text{RMSD}}$  is lower than the corresponding WT  $q_{\text{RMSD}}$  (Table II).

It is surprising that despite the near-optimal charge distribution shown in the above section, RPV has the most unfavorable electrostatic component of binding WT and that it is energetically the furthest from its optimum, with a potential improvement of 3.5 kcal/mol upon optimization compared with 3.1 and 2.6 kcal/mol for NVP and EFV, respectively. The first of these two observations is due to the fact that RPV desolvates RT more than the other drugs. When controlling for receptor desolvation, which is independent of drug charges, one now sees that RPVs charge distribution actually has the least unfavorable electrostatic binding free energy toward WT of the

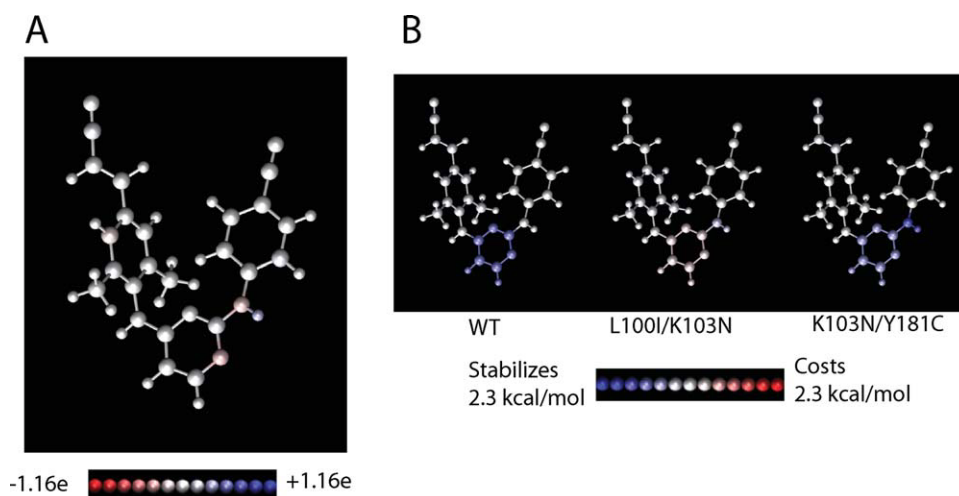
three inhibitors, suggesting its charge distribution is indeed somewhat optimized toward WT (numbers in parentheses in Table III). The large receptor desolvation penalty caused by RPV is likely due to its larger size (28 heavy atoms, compared with 20 and 21 for NVP and EFV). Presumably, its larger size would have a favorable effect on nonpolar interactions to compensate for the unfavorable electrostatic contribution, as RPV experimentally has a very high activity against WT RT.<sup>32,33</sup>

The second of the surprising observations above, that RPV is energetically the furthest from its optimum, highlights the great sensitivity of the energetics to RPVs charge distribution when compared with the other drugs. RPV can gain substantial energetic improvements by relatively small changes in charge values, with its optimal binding free energy toward WT more favorable than those of the other two drugs when controlling for receptor desolvation (−3.5 kcal/mol vs. −1.6 and −2.4 kcal/mol for NVP and EFV). The sensitivity of the binding free energy to RPVs charges highlights the importance of “getting the electrostatics correct” in RPV-RT interaction. Indeed, the largest eigenvalue of the  $L$  matrix for the RPV/WT interaction is higher than those for the other drugs ( $203 \frac{\text{kcal}}{\text{mol} \times e^2}$  vs. 160 and 151  $\frac{\text{kcal}}{\text{mol} \times e^2}$  for NVP and EFV), as are the next several eigenvalues (data not shown). Eigenvalues of the  $L$  matrix are direct measures of the sensitivity of the binding free energy to variations in ligand charge. Additionally, RPVs larger number of atoms may also help explain why, although each atom's charge does not change appreciably to achieve optimality, the overall energetics change appreciably because more charges are involved. Conversely, as discussed previously, electrostatics appear to be less crucial to the interaction between EFV and RT and provide less potential for improvement via optimization; larger changes in charge values to achieve optimality afford relatively modest energetic improvements.

The much higher binding free energy of the Y188C mutant binding to NVP (12.7 kcal/mol) when compared with the other NVP-RT complexes is partly due to conformation of the RT residue Glu138 on the p51 subunit in the complex, which is desolvated by NVP upon binding but does not make significant interactions with the drug; indeed, when the charges on Glu138 are set to

**Figure 3**

The electrostatic contributions of certain RT residues toward binding the studied drug molecules. (A) Two views of the RT binding pocket, showing aligned drug molecules and surrounding residues. For ease of visualization, only the residues from the RPV-WT structure (PDB ID 2ZD1) are shown. NVP is shown in orange, EFV in pink, and RPV in green. (B) The  $\Delta\Delta G_{\text{moiety,mut}}$  defined in the “Methods” section, is shown for surrounding RT residues upon binding to each drug, using wild-type structures for analysis. Panel (C) shows the same quantity as panel (B), although now, each drug bears its hypothetical, optimal charge distribution for binding. (D)  $\Delta\Delta G_{\text{moiety,mut}}$  for selected RT residues, from either WT or mutant structures, toward binding RPV. The starred residue (138) is from the p51 subunit. All energies are in kcal/mol.

**Figure 4**

Optimizing RPVs charge distribution to bind promiscuously to multiple variants. (A) The computed optimally promiscuous charge distribution, colored by charge value according to the legend (blue = positive, red = negative). The computed, optimal charge distributions toward each individual variant are shown in Figure 2(C) for comparison. (B) Groups of atoms are colored according to their  $\Delta\Delta G_{\text{moiety, additive}}$  assuming that the drug had the computed optimally promiscuous charge distribution. Red-colored groups contribute unfavorably toward binding and blue-colored groups contribute favorably. The moieties considered are the same as those shown in Figure 2(C).

zero, the RT desolvation decreases by 4.7 kcal/mol and the overall electrostatic binding free energy becomes 7.02 kcal/mol (data not shown).

### Electrostatic contribution of RT residues toward binding

#### Contribution of WT RT residues toward binding each drug

The electrostatic contribution toward binding was computed for WT RT residues with at least one atom center within 4 Å of all three drug molecules. Selected residues are highlighted in Figure 3(A).  $\Delta\Delta G_{\text{moiety, mut}}$  was used for quantifying the energetic contributions. Physically,  $\Delta\Delta G_{\text{moiety, mut}}$  is the hypothetical change in the binding free energy when a residue's charge distribution replaces its hydrophobic isostere. A negative  $\Delta\Delta G_{\text{moiety, mut}}$  therefore means that the residue's charge distribution contributes favorably to binding. Figure 3(B) is a plot of the  $\Delta\Delta G_{\text{moiety, mut}}$  for the selected residues when interacting with each of the three drugs.

Figure 3(B) indicates that certain residues, such as Phe227, Val179, Pro236, Tyr318, and Trp229 contribute similarly toward binding each drug—slightly favorably in the first 3 cases and slightly unfavorably in the latter two. However, other residues play different electrostatic roles toward different drugs; Tyr188 contributes unfavorably toward binding EFV and RPV, but not NVP, a result that could help explain why NVP is most susceptible to mutations at this position. Lys103 contributes unfavorably toward all drugs, but far more in the case of EFV; in this complex its side chain appears more buried upon binding than in the others and therefore may pay a greater

desolvation penalty. Finally, Lys101 contributes very favorably toward binding EFV and RPV; its backbone makes hydrogen bonding interactions with both drugs, as discussed previously. However, it makes no significant contribution toward binding NVP.

#### Contribution of WT RT residues toward binding each hypothetically optimized drug

Figure 3(C) is a plot of the  $\Delta\Delta G_{\text{moiety, mut}}$  for the same set of residues selected above, although now the charge distribution of the drug in each complex was set to its hypothetical optimum. Most notably, the unfavorable contributions of Lys103 have been somewhat muted, yielding an improved interaction between this residue with the optimized drug charge distributions. Lys101 is still the most favorable contributor for both RPV and EFV, suggesting that the actual drug molecules are using the same RT residues for recognition as their optimal counterparts.

#### Contribution of RT residues toward binding RPV in both WT and mutant complexes

Figure 3(D) is a plot of the  $\Delta\Delta G_{\text{moiety, mut}}$  for all residues with at least one atom within 4 Å of RPV in all three complexes considered here. Across the variants, many residues make similar contributions toward binding RPV, either contributing favorably, unfavorably, or insignificantly in all three cases. Residue 103, which is mutated to Asn in the two mutant variants, contributes unfavorably in all cases, but more so in the L100I/K103N complex. In both mutants, the Asn side chain is buried by the drug upon binding but does not appear to make

**Table IV**

Binding Energetics of a Hypothetical, Optimally Promiscuous Molecule Based on RPV

	WT	L100I/K103N	K103N/Y181C
$\Delta G_{\text{elec,actual}}$	7.6	11.0	7.8
$\Delta G_{\text{elec,opt}}$	4.1	5.5	3.6
$\Delta G_{\text{elec,promiscuous\_opt}}$	5.4	6.8	4.9

$\Delta G_{\text{elec,actual}}$  is the computed electrostatic binding free energy between RPV and an RT variant,  $\Delta G_{\text{elec,opt}}$  is the calculated optimal binding free energy for each individual variant in turn, and  $\Delta G_{\text{elec,promiscuous\_opt}}$  is the calculated binding free energy using the optimally promiscuous charge distribution. All energies are in kcal/mol.

canonical hydrogen-bonding interactions with the drug. Not surprisingly, Lys101 shows great variation in its contribution toward the variants, contributing very favorably in the WT and the K103N/Y181C complexes, but unfavorably in the L100I/K103N complex.

### Multitarget affinity optimization of RPV

While the charge distribution of RPV was shown to be similar to the optimal distribution for binding to WT RT, it was suboptimal toward the mutant variants, especially the L100I/K103N mutant. To understand electrostatic determinants of multitarget recognition, the charge distribution of RPV was optimized to bind as well as possible to three variants (WT, L100I/K103N, and K103N/Y181C) simultaneously. The resulting optimally “promiscuous” charge distribution is shown in Figure 4. The affinities of the optimally promiscuous molecule toward each variant are shown in Table IV; the hypothetical molecule is predicted to have a lower electrostatic binding free energy toward each variant than the actual drug molecule. Additionally, the contribution of the molecular components of the optimized molecule toward binding each RT variant is shown in Figure 4.

The charge distribution of the optimally promiscuous molecule is somewhat of an “average” between the affinity optima toward the individual variants. Although the signs of the charges on the pyrimidine ring and the relevant N—H linker are similar to those on the affinity optima toward WT and the K103N/Y181C mutant, the magnitudes of the charges are much smaller, to account for the opposite polarization required for tight binding to the L100I/K103N variant. The charge distribution is fairly hydrophobic because the optimal charge distributions toward the variants are so dissimilar, and this is the best compromise. The same molecular components on the optimally promiscuous molecule contribute significantly toward binding WT and the K103N/Y181C mutant as they do in the individual affinity optima—the pyrimidine ring in the case of WT and the N—H linker in the case of the K103N/Y181C mutant, although here they do so to a lesser extent. However, the lesser extent of these favorable interactions is accompanied by a loss of unfavorable interactions toward the L100I/K103N variant. In essence, the optimally promiscuous molecule is capable

of having favorably contributing components in two of the three complexes, and no significantly unfavorable contributions in the third, a compromise that the actual charge distribution of RPV is not able to achieve because it is too polarized. The results presented in this section are with the same flip-states of the Asn103 side chains shown in the previous sections. Similar qualitative results (i.e., a more hydrophobic charge distribution for the optimally promiscuous molecule) were obtained assuming the opposite flip-state conformation for the Asn103 side chains in the two double mutants.

## DISCUSSION

In this work, we systematically analyzed the electrostatic interactions between three RT drugs and multiple RT variants. Although our analysis considered only the electrostatic component of binding, it is useful to discuss our results in context with experimental data and other previous studies. First, with only one exception, the electrostatic component of binding was computed to be worse for each drug–mutant complex compared with the corresponding WT complex, suggesting that the worsening of electrostatic interactions is partly responsible for the loss of affinity toward mutant variants. Indeed, our results show that the charge distribution of RPV is highly suboptimal for tight binding to the L100I/K103N RT mutant, likely because of a shift in the orientation of RPV relative to the backbone of Lys101; the corresponding increase in the electrostatic binding free energy computed for this complex can help explain the observed 20-fold increase in the  $EC_{50}$  toward the mutant relative to WT.<sup>33</sup> Nevertheless, some chemical moieties that we show to be electrostatically suboptimal have been shown to be important for other reasons. For example, experimental studies have suggested that the cyano and cyanovinyl groups are important to RPVs broad recognition profile and for the biological activity of a naphthyl-substituted analogue<sup>33,92</sup>; future work toward understanding the tradeoffs between electrostatic and other interactions may help in the rational design of moieties at these positions that are electrostatically well-suited while not sacrificing other crucial properties. Of course, drug properties such as bioavailability and toxicity, while beyond the scope of this discussion, must be considered when evaluating such tradeoffs.

Previous computational studies have quantified the contribution of RT residues toward the drug–target interaction, with results showing some agreement to our findings. For example, Mei *et al.* performed quantum mechanical calculations to compute the interaction energy between RT residues and EFV in WT and mutant complexes.<sup>46</sup> As we also find [Fig. 3(B)], they found that Lys101 was a major contributor toward the interaction in WT. They found that the interaction is less favorable for the K103N mutant because of the movement of the

Lys101 side chain away from the carbonyl moiety on EFV. Interestingly, we show that toward the K103N mutant, the carbonyl moiety on EFV contributes quite favorably toward the binding interaction [Fig. 2(B)], whereas it does not do so toward WT RT. However, as Lys101 contributes quite favorably to binding WT RT [it has a significantly negative  $\Delta\Delta G_{\text{moiety, mut}}$ , Fig. 3(B)], a parallel calculation for Lys101 in the K103N mutant yielded  $\Delta\Delta G_{\text{moiety, mut}} = -0.78$  kcal/mol, a less negative value, which suggests that Lys101 does not contribute as favorably in the mutant structure—agreeing with the previous work. Decha *et al.* also found Lys101 to be the most favorable contributor toward binding EFV,<sup>55</sup> although they also found Lys103 to contribute favorably, a result in conflict with our findings [Fig. 3(B)]. However, their interaction energies also include van der Waals contributions, and while they account for solvation in their total binding free energy calculations, it is unclear whether the effects of desolvation were accounted for in their decomposition into interaction energies; as mentioned in the Results, Lys103 appears more buried in the complex with EFV and may therefore pay a greater desolvation penalty. Finally, Nunriam *et al.*<sup>48</sup> and Weininger *et al.*<sup>53</sup> also found the interaction between Lys101 and EFV to be the most important contributor to binding. Our study provides a more subtle perspective—whereas the contribution from the Lys101 residue toward binding WT RT is quite favorable [Fig. 3(B)], the contributions from the carbonyl and N—H groups on EFV are not [Fig. 2(B), top row]. However, care must be taken when comparing these values; when  $\Delta\Delta G_{\text{moiety, mut}}$  is used, the interpretation is physical, as one is directly calculating the change in binding free energy when charging up a residue from its hydrophobic isostere. When  $\Delta\Delta G_{\text{moiety, contrib}}$  is used, though, there is no direct physical analogue, but the values can be directly compared with values for other moieties on the molecule as true fractions of contribution. As we argue that in this study, it was most appropriate to use  $\Delta\Delta G_{\text{moiety, mut}}$  to quantify RT contributions and  $\Delta\Delta G_{\text{moiety, contrib}}$  to quantify drug contributions, a direct comparison of these values between groups on different binding partners would not be meaningful; indeed, previous work suggested that nonadditive mutational energies may systematically overestimate the favorability of a group because the intergroup interactions tend to be favorable.<sup>73</sup> Nevertheless, it is interesting that, when taking desolvation into account, many groups on either binding partner contribute unfavorably toward binding. This result may not be surprising given that salt bridges—common electrostatic features at binding interfaces—often contribute unfavorably toward stability.<sup>93</sup> Indeed, it is often believed that electrostatic features that do not contribute to stability may be important for determining the specificity of the interaction.<sup>93–97</sup>

Our results provide multiple hypotheses that could be tested experimentally to further understand the studied

interactions as well as to design potentially improved ones. For example, Figure 2(A) shows that the nitrogen atoms in the two pyridine rings of NVP are too negative for optimal binding, suggesting that a replacement of pyridine rings with phenyl rings may improve binding affinity, an experimentally testable hypothesis. Likewise, the cyano groups on RPV are too polarized [Fig. 2(C)], suggesting their replacement with alkyne groups. Of course, such chemical modifications are not isosteric and so nonpolar free energy components will also be affected and could be explicitly modeled in a manner similar to previous work<sup>71</sup>; nevertheless, previous work has also shown that modifications that qualitatively alter a molecule toward its optimal charge distribution often result in increased binding affinity.<sup>69</sup> Additionally, our predictions of the electrostatic contributions of Lys101 and Lys103 could be roughly assessed experimentally by mutating each to methionine and determining the change in binding affinity due to each mutation to an approximate hydrophobic isostere, assuming such mutations do not greatly affect protein stability and native conformation.

Previous structural analyses have highlighted the importance of water molecules in the RT binding pocket that mediate hydrogen-bonding interactions between drug and target.<sup>56,57</sup> Indeed, water plays a crucial role in modulating affinity, promiscuity, and specificity due to its ability to reorient itself and therefore to adjust to various binding interfaces.<sup>98</sup> In our study, we chose to retain only those crystallographic waters whose potential nonwater hydrogen-bonding contacts suggested they would be “locked” in a given position; generally, the retained waters were not near the drug–target interface, with the exception of a water molecule in the NVP-WT RT complex that was shown to play an important role in the drug–target interaction.<sup>57</sup> In several crystal structures, no explicit water molecules were modeled because the crystallographic coordinates did not include any water molecules. Our model does implicitly account for the effect of such water-mediated hydrogen bonds as well as reoriented waters upon binding, as this is the very rationale for modeling the solvent as a dielectric medium. Nevertheless, a future study that quantifies the energetic contribution of certain explicit water molecules can yield additional insight into the electrostatic determinants of binding in this system.

For a given drug, certain observations in this work were fairly robust across RT variants. For example, toward all variants, a relatively hydrophobic optimal charge distribution was found for NVP, with some larger magnitude charges not contributing significantly to overall binding. As future work, one may analyze the overall robustness of the charge distribution to small conformational variations of a given target, analyzing those circumstances in which the optimal charge distribution is sensitive to small interfacial changes and those in which it is not. Indeed, there is a relationship between how

polarized or charged a molecule is and how sensitive its binding free energy is to shape differences amongst targets,<sup>99</sup> suggesting that a sensitivity to conformation might be more apparent in a more charged or polar binding site, but such a hypothesis warrants systematic study. To begin to assess the sensitivity of our results to small conformational changes, we carried out analyses on a different RPV/WT RT structure, PDB ID 3MEE at 2.4-Å resolution.<sup>100</sup> Although the charge distribution of the N—H linker interacting with Lys101 was somewhat less optimal in this structure than in the one presented above (it was underpolarized), the charge of the nitrogen in the pyrimidine ring was again strikingly close to its optimal value, and the cyano and cyanovinyl groups were again similarly suboptimal (data not shown).  $q_{\text{RMSD}}$  for this structure is 0.20, compared with 0.17 for the WT structure primarily considered in this work. Both RMSD values are lower than the analogous values for NVP or EFV, providing additional confidence that the charge distribution of RPV is close to its optimum for binding WT RT.

In this study, we assumed that the geometry around the nitrogen atoms in the NH groups of all drugs was planar, as quantum-mechanical geometry optimizations of each drug resulted in either planar or slightly pyramidal geometries around these atoms. To assess the robustness of our results to the nitrogen geometries, a subset of analyses were performed in which the hydrogen atoms off of certain nitrogen atoms were built to create more significant pyramidal geometries around the nitrogen (improper dihedrals set to 140–160 degrees, depending on the linker considered). Here we briefly describe the qualitative results obtained with the RPV–WT RT complex. Not surprisingly, the optimal charges depended on the choice of inversion at the nitrogen atoms, although if we chose the inversion resulting in the visually superior hydrogen-bonding geometry for the N atom interacting with Lys101, then the RMSD between the optimal charges with the molecules with bent geometry and the planar geometry was only  $0.04e$ . However, while the optimal charges were somewhat robust when assuming the intuitively more favorable inversion, the RESP-derived charges for the actual molecule depended strongly on the geometry of the N atom. The pyramidal nitrogen groups were significantly more polarized than their planar counterparts, resulting in an RMSD value of  $\sim 0.09e$  between the molecule's RESP-derived atomic charges with planar nitrogens and with the inversions considered here, with the charges on N atoms differing by  $\sim 0.4e$  in certain cases. It has been previously shown that RESP-derived charges for amine groups do not accurately reproduce solvation energies,<sup>86</sup> so it is unclear how accurate the RESP-derived charges are as a model for the pyramidal nitrogen atoms within this molecule. As beyond the scope of this work, a study that further probed the dependence of RESP-derived charges on the extent to which a nitrogen atom is pyramidal and attempts to pre-

dict the positions of hydrogens off such “ambiguous” nitrogens through QM calculations would be interesting future work.

Here, we showed that electrostatic interactions were not optimized in the two conformationally rigid molecules, NVP and EFV, and they showed only moderate potential for affinity improvement via charge optimization. However, RPV, a conformationally flexible molecule, has a lower electrostatic binding free energy toward WT than NVP or EFV when controlling for receptor desolvation penalty, with our results suggesting that electrostatics can also be an avenue by which to further improve affinity. One interesting hypothesis that warrants further study is whether flexible molecules—because they may pay a larger entropic penalty upon binding (often estimated as being related to the number of rotatable bonds<sup>101</sup>)—may need to be more electrostatically optimal to compensate for this penalty and may need to use electrostatics more as a determinant for interaction. The small number of molecules studied here does not allow us to test this hypothesis, but a larger scale study comparing electrostatic optimality and conformational flexibility of many diverse molecules could address this hypothesis further.

The optimally “promiscuous” charge distribution of RPV toward three variants of RT was more hydrophobic overall than RPVs actual charge distribution, a result in agreement with the general belief that hydrophobic molecules are often more promiscuous, or “nonspecific,” than polar or charged molecules. In fact, in addition to case-by-case studies describing promiscuous interfaces or binding sites as being hydrophobic,<sup>2,102,103</sup> previous work has demonstrated an overall trend that broadly recognizing molecules are more lipophilic,<sup>104</sup> and a theoretical framework has been developed to demonstrate several specific reasons why this appears to be the case.<sup>99</sup> In the case of rapidly mutating targets, the long-range nature of electrostatic interactions may make the strongly polar or charged molecules unable to fine-tune their interactions to mutant targets, especially if mutations significantly alter the electrostatic potential.

In this work, we have used both electrostatic charge optimization and component analysis techniques to systematically analyze the electrostatic interactions between multiple RT inhibitors and RT variants. We show that the role of electrostatics can vary within the same binding pocket, depending on the drug being bound, and that different molecular components can play key roles in different drug–target complexes. Overall, the subtle yet important role of electrostatics in a hydrophobic binding pocket reveal that it must be considered in the rational design of broadly binding molecules in any system.

## ACKNOWLEDGMENTS

The authors acknowledge Kenneth Merz Jr. and Michael Weaver for assistance, Helena Qi for supplementary

analysis, and Michael Altman, William Coleman, David Green, and Bruce Tidor for helpful software, scripts, and discussions.

## REFERENCES

1. Watkins RE, Wisely GB, Moore LB, Collins JL, Lambert MH, Williams SP, Willson TM, Kliewer SA, Redinbo MR. The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. *Science* 2001;292:2329–2333.
2. Huang K, Kapadia G, Zhu PP, Peterkofsky A, Herzberg O. A promiscuous binding surface: crystal structure of the IIA domain of the glucose-specific permease from *Mycoplasma capricolum*. *Structure* 1998;6:697–710.
3. Loris R, De Greve H, Dao-Thi MH, Messens J, Imberty A, Wyns L. Structural basis of carbohydrate recognition by lectin II from *Ulex europaeus*, a protein with a promiscuous carbohydrate-binding site. *J Mol Biol* 2000;301:987–1002.
4. Friedler A, Vepintsev DB, Rutherford T, von Glos KI, Fersht AR. Binding of Rad51 and other peptide sequences to a promiscuous, highly electrostatic binding site in p53. *J Biol Chem* 2005;280:8051–8059.
5. James LC, Roversi P, Tawfik DS. Antibody multispecificity mediated by conformational diversity. *Science* 2003;299:1362–1367.
6. James LC, Tawfik DS. The specificity of cross-reactivity: promiscuous antibody binding involves specific hydrogen bonds rather than nonspecific hydrophobic stickiness. *Protein Sci* 2003;12:2183–2193.
7. Chen JP, Zhang X, Fernandez A. Molecular basis for specificity in the druggable kinome: sequence-based analysis. *Bioinformatics* 2007;23:563–572.
8. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. Crystal-structure at 3.5 Å resolution of HIV-1 reverse-transcriptase complexed with an inhibitor. *Science* 1992;256:1783–1790.
9. Mitsuya H, Yarchoan R, Broder S. Molecular targets for AIDS therapy. *Science* 1990;249:1533–1544.
10. Das K, Lewi PJ, Hughes SH, Arnold E. Crystallography and the design of anti-AIDS drugs: conformational flexibility and positional adaptability are important in the design of non-nucleoside HIV-1 reverse transcriptase inhibitors. *Prog Biophys Mol Biol* 2005;88:209–231.
11. Prajapati DG, Ramajayam R, Yadav MR, Giridhar R. The search for potent, small molecule NNRTIs: a review. *Bioorg Med Chem* 2009;17:5744–5762.
12. De Clercq E. The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. *Antiviral Res* 1998;38:153–179.
13. Tantillo C, Ding JP, Jacobomolina A, Nanni RG, Boyer PL, Hughes SH, Pauwels R, Andries K, Janssen PAJ, Arnold E. Locations of anti-AIDS drug-binding sites and resistance mutations in the 3-dimensional structure of HIV-1 reverse-transcriptase—implications for mechanisms of drug-inhibition and resistance. *J Mol Biol* 1994;243:369–387.
14. Esnouf R, Ren JS, Ross C, Jones Y, Stammers D, Stuart D. Mechanism of inhibition of HIV-1 reverse-transcriptase by nonnucleoside inhibitors. *Nat Struct Biol* 1995;2:303–308.
15. Spence RA, Kati WM, Anderson KS, Johnson KA. Mechanism of inhibition of HIV-1 reverse-transcriptase by nonnucleoside inhibitors. *Science* 1995;267:988–993.
16. Sluis-Cremer N, Temiz NA, Bahar I. Conformational changes in HIV-1 reverse transcriptase induced by nonnucleoside reverse transcriptase inhibitor binding. *Curr HIV Res* 2004;2:323–332.
17. Ivetac A, McCammon JA. Elucidating the inhibition mechanism of HIV-1 non-nucleoside reverse transcriptase inhibitors through multi-copy molecular dynamics simulations. *J Mol Biol* 2009;388: 644–658.
18. Temiz NA, Bahar I. Inhibitor binding alters the directions of domain motions in HIV-1 reverse transcriptase. *Proteins: Struct Funct Genet* 2002;49:61–70.
19. Ren JS, Stammers DK. Structural basis for drug resistance mechanisms for non-nucleoside inhibitors of HIV reverse transcriptase. *Virus Res* 2008;134:157–170.
20. Merluzzi VJ, Hargrave KD, Labadia M, Grozinger K, Skoog M, Wu JC, Shih CK, Eckner K, Hattox S, Adams J, Rosenthal AS, Faanes R, Eckner RJ, Koup RA, Sullivan JL. Inhibition of HIV-1 replication by a nonnucleoside reverse-transcriptase inhibitor. *Science* 1990;250:1411–1413.
21. Maga G, Amacker M, Ruel N, Hubscher U, Spadari S. Resistance to nevirapine of HIV-1 reverse transcriptase mutants: loss of stabilizing interactions and thermodynamic or steric barriers are induced by different single amino acid substitutions. *J Mol Biol* 1997;274: 738–747.
22. Rhee SY, Gonzales MJ, Kantor R, Betts BJ, Ravela J, Shafer RW. Human immunodeficiency virus reverse transcriptase and protease sequence database. *Nucleic Acids Res* 2003;31:298–303.
23. Ren J, Nichols C, Bird L, Chamberlain P, Weaver K, Short S, Stuart DI, Stammers DK. Structural mechanisms of drug resistance for mutations at codons 181 and 188 in HIV-1 reverse transcriptase and the improved resilience of second generation non-nucleoside inhibitors. *J Mol Biol* 2001;312:795–805.
24. Young SD, Britcher SE, Tran LO, Payne LS, Lumma WC, Lyle TA, Huff JR, Anderson PS, Olsen DB, Carroll SS, Pettibone DJ, Obrien JA, Ball RG, Balani SK, Lin JH, Chen IW, Schleif WA, Sardana VV, Long WJ, Byrnes VW, Emimi EA. L-743,726 (DMP-266)—a novel, highly potent nonnucleoside inhibitor of the human-immunodeficiency-virus type-1 reverse-transcriptase. *Antimicrob Agents Chemother* 1995;39:2602–2605.
25. Ren J, Milton J, Weaver KL, Short SA, Stuart DI, Stammers DK. Structural basis for the resilience of efavirenz (DMP-266) to drug resistance mutations in HIV-1 reverse transcriptase. *Structure* 2000;8:1089–1094.
26. Shulman NS, Zolopa AR, Passaro DJ, Murlidharan U, Israelski DM, Brosgart CL, Miller MD, Van Doren S, Shafer RW, Katzenstein DA. Efavirenz- and adefovir dipivoxil-based salvage therapy in highly treatment-experienced patients: clinical and genotypic predictors of virologic response. *J Acquir Immune Defic Syndr* 2000;23:221–226.
27. Shafer RW. Genotypic testing for human immunodeficiency virus type 1 drug resistance. *Clin Microbiol Rev* 2002;15:247–277.
28. Lindberg J, Sigurdsson S, Lowgren S, Andersson HO, Sahlberg C, Noreen R, Fridborg K, Zhang H, Unge T. Structural basis for the inhibitory efficacy of efavirenz (DMP-266), MSC194 and PNU142721 towards the HIV-1 RT K103N mutant. *Eur J Biochem* 2002;269:1670–1677.
29. Yin PD, Das D, Mitsuya H. Overcoming HIV drug resistance through rational drug design based on molecular, biochemical, and structural profiles of HIV resistance. *Cell Mol Life Sci* 2006;63: 1706–1724.
30. Hsiou Y, Ding JP, Das K, Clark AD, Boyer PL, Lewi P, Janssen PAJ, Kleim JP, Rosner M, Hughes SH, Arnold E. The Lys103Asn mutation of HIV-1 RT: a novel mechanism of drug resistance. *J Mol Biol* 2001;309:437–445.
31. Andries K, Azijn H, Thielemans T, Ludovici D, Kukla M, Heeres J, Janssen P, De Corte B, Vingerhoets J, Pauwels R, de Bethune MP. TMC125, a novel next-generation nonnucleoside reverse transcriptase inhibitor active against nonnucleoside reverse transcriptase inhibitor-resistant human immunodeficiency virus type 1. *Antimicrob Agents Chemother* 2004;48:4680–4686.
32. Janssen PAJ, Lewi PJ, Arnold E, Daeyaert F, de Jonge M, Heeres J, Koymans L, Vinkers M, Guillemont J, Pasquier E, Kukla M, Ludovici D, Andries K, de Bethune MP, Pauwels R, Das K, Clark AD, Frenkel YV, Hughes SH, Medaer B, De Knaep F, Bohets H, De Clerck F, Lampo A, Williams P, Stoffels P. In search of a novel anti-HIV drug: multidisciplinary coordination in the discovery of 4-[[4-[[4-[(1E)-2-cyanoethenyl]-2,6dimethylphenyl] amino]-2-pyrimidinyl]amino]-benzotrile (R278474, rilpivirine). *J Med Chem* 2005;48:1901–1909.



33. Das K, Bauman JD, Clark AD, Frenkel YV, Lewi PJ, Shatkin AJ, Hughes SH, Arnold E. High-resolution structures of HIV-1 reverse transcriptase/TMC278 complexes: strategic flexibility explains potency against resistance mutations. *Proc Natl Acad Sci USA* 2008;105:1466–1471.
34. Das K, Clark AD, Lewi PJ, Heeres J, de Jonge MR, Koymans LMH, Vinkers HM, Daeyaert F, Ludovici DW, Kukla MJ, De Corte B, Kavash RW, Ho CY, Ye H, Lichtenstein MA, Andries K, Pauwels R, de Bethune MP, Boyer PL, Clark P, Hughes SH, Janssen PAJ, Arnold E. Roles of conformational and positional adaptability in structure-based design of TMC125–R165335 (etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants. *J Med Chem* 2004;47:2550–2560.
35. Mascolini M. Key etravirine mutations also render HIV resistant to rilpivirine. In: Conference Reports for NATAP, 18th HIV Drug Resistance Workshop, XVII International Drug Resistance Workshop, Fort Myers, Florida: 2009.
36. Rizzo RC, Tirado-Rives J, Jorgensen WL. Estimation of binding affinities for HEPT and nevirapine analogues with HIV-1 reverse transcriptase via Monte Carlo simulations. *J Med Chem* 2001;44:145–154.
37. Rizzo RC, Udier-Blagovic M, Wang DP, Watkins EK, Smith MBK, Smith RH, Tirado-Rives J, Jorgensen WL. Prediction of activity for nonnucleoside inhibitors with HIV-1 reverse transcriptase based on Monte Carlo simulations. *J Med Chem* 2002;45:2970–2987.
38. Udier-Blagovic M, Watkins EK, Tirado-Rives J, Jorgensen WL. Activity predictions for efavirenz analogues with the K103N mutant of HIV reverse transcriptase. *Bioorg Med Chem Lett* 2003;13: 3337–3340.
39. Gussio R, Pattabiraman N, Zaharevitz DW, Kellogg GE, Topol IA, Rice WG, Schaeffer CA, Erickson JW, Burt SK. All-atom models for the non-nucleoside binding site of HIV-1 reverse transcriptase complexed with inhibitors: a 3D QSAR approach. *J Med Chem* 1996;39:1645–1650.
40. Pungpo P, Hannongbua S. Three-dimensional quantitative structure-activity relationships study on HIV-1 reverse transcriptase inhibitors in the class of dipyrindiazepinone derivatives, using comparative molecular field analysis. *J Mol Graph Model* 2000;18: 581–590,601.
41. Pungpo P, Saparpakorn P, Wolschann P, Hannongbua S. Computer-aided molecular design of highly potent HIV-1 RT inhibitors: 3D QSAR and molecular docking studies of efavirenz derivatives. *SAR QSAR Environ Res* 2006;17:353–370.
42. Wang JM, Morin P, Wang W, Kollman PA. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *J Am Chem Soc* 2001;123:5221–5230.
43. Carlsson J, Boukharta L, Aqvist J. Combining docking, molecular dynamics and the linear interaction energy method to predict binding modes and affinities for non-nucleoside inhibitors to HIV-1 reverse transcriptase. *J Med Chem* 2008;51:2648–2656.
44. Smith MBK, Rouzer CA, Taneyhill LA, Smith NA, Hughes SH, Boyer PL, Janssen PAJ, Moereels H, Koymans L, Arnold E, Ding JP, Das K, Zhang WY, Michejda CJ, Smith RH. Molecular modeling studies of HIV-1 reverse-transcriptase nonnucleoside inhibitors—total-energy of complexation as a predictor of drug placement and activity. *Protein Sci* 1995;4:2203–2222.
45. He X, Mei Y, Xiang Y, Zhang DW, Zhang JZH. Quantum computational analysis for drug resistance of HIV-1 reverse transcriptase to nevirapine through point mutations. *Proteins: Struct Funct Bioinform* 2005;61:423–432.
46. Mei Y, He X, Xiang Y, Zhang DW, Zhang JZH. Quantum study of mutational effect in binding of efavirenz to HIV-1 RT. *Proteins: Struct Funct Bioinform* 2005;59:489–495.
47. Kuno M, Hannongbua S, Morokuma K. Theoretical investigation on nevirapine and HIV-1 reverse transcriptase binding site interaction, based on ONIOM method. *Chem Phys Lett* 2003;380: 456–463.
48. Nunrium P, Kuno M, Saen-Oon S, Hannongbua S. Particular interaction between efavirenz and the HIV-1 reverse transcriptase binding site as explained by the ONIOM2 method. *Chem Phys Lett* 2005;405:198–202.
49. Srivab P, Hannongbua S. A study of the binding energies of efavirenz to wild-type and K103N/Y181C HIV-1 reverse transcriptase based on the ONIOM method. *ChemMedChem* 2008;3:803–811.
50. Raju RK, Burton NA, Hillier IH. Modelling the binding of HIV-reverse transcriptase and nevirapine: an assessment of quantum mechanical and force field approaches and predictions of the effect of mutations on binding. *Phys Chem Chem Phys* 2010;12: 7117–7125.
51. Wang DP, Rizzo RC, Tirado-Rives J, Jorgensen WL. Antiviral drug design: computational analyses of the effects of the L100I mutation for HIV-RT on the binding of NNRTIs. *Bioorg Med Chem Lett* 2001;11:2799–2802.
52. Rodriguez-Barrios F, Balzarini J, Gago F. The molecular basis of resilience to the effect of the Lys103Asn mutation in non-nucleoside HIV-1 reverse transcriptase inhibitors studied by targeted molecular dynamics simulations. *J Am Chem Soc* 2005;127:7570–7578.
53. Weinzing P, Hannongbua S, Wolschann P. Molecular mechanics PBSA ligand binding energy and interaction of efavirenz derivatives with HIV-1 reverse transcriptase. *J Enzym Inhib Med Chem* 2005; 20:129–134.
54. Ceccherini-Silberstein F, Svicher V, Sing T, Artese A, Santoro MM, Forbicci F, Bertoli A, Alcaro S, Palamara G, Monforte AD, Balzarini J, Antinori A, Lengauer T, Perno CF. Characterization and structural analysis of novel mutations in human immunodeficiency virus type 1 reverse transcriptase involved in the regulation of resistance to nonnucleoside inhibitors. *J Virol* 2007;81:11507–11519.
55. Decha P, Intharathep P, Udommaneehanakit T, Sompornpisut P, Hannongbua S, Wolschann P, Parasuk V. Theoretical studies on the molecular basis of HIV-1RT/NNRTIs interactions. *J Enzym Inhib Med Chem* 2011;26:29–36.
56. Ribone SR, Quevedo MA, Madrid M, Brinon MC. Rational approaches for the design of effective human immunodeficiency virus type 1 nonnucleoside reverse transcriptase inhibitors. *J Chem Inf Model* 2011;51:130–138.
57. Treesuwan W, Hannongbua S. Bridge water mediates nevirapine binding to wild type and Y181C HIV-1 reverse transcriptase—evidence from molecular dynamics simulations and MM-PBSA calculations. *J Mol Graph Model* 2009;27:921–929.
58. Paris KA, Haq O, Felts AK, Das K, Arnold E, Levy RM. Conformational landscape of the human immunodeficiency virus type 1 reverse transcriptase non-nucleoside inhibitor binding pocket: lessons for inhibitor design from a cluster analysis of many crystal structures. *J Med Chem* 2009;52:6413–6420.
59. Zhou ZG, Madrid M, Evanseck JD, Madura JD. Effect of a bound non-nucleoside RT inhibitor on the dynamics of wild-type and mutant HIV-1 reverse transcriptase. *J Am Chem Soc* 2005;127: 17253–17260.
60. Lee LP, Tidor B. Optimization of electrostatic binding free energy. *J Chem Phys* 1997;106:8681–8690.
61. Kangas E, Tidor B. Optimizing electrostatic affinity in ligand-receptor binding: theory, computation, and ligand properties. *J Chem Phys* 1998;109:7522–7545.
62. Lee LP, Tidor B. Optimization of binding electrostatics: charge complementarity in the barnase-barstar protein complex. *Protein Sci* 2001;10:362–377.
63. Lee LP, Tidor B. Barstar is electrostatically optimized for tight binding to barnase. *Nat Struct Biol* 2001;8:73–76.
64. Sulea T, Purisima EO. Optimizing ligand charges for maximum binding affinity. A solvated interaction energy approach. *J Phys Chem B* 2001;105:889–899.
65. Kangas E, Tidor B. Electrostatic complementarity at ligand binding sites: application to chorismate mutase. *J Phys Chem B* 2001;105: 880–888.
66. Sims PA, Wong CF, McCammon JA. Charge optimization of the interface between protein kinases and their ligands. *J Comput Chem* 2004;25:1416–1429.

67. Green DF, Tidor B. *Escherichia coli* glutamyl-tRNA synthetase is electrostatically optimized for binding of its cognate substrates. *J Mol Biol* 2004;342:435–452.
68. Ahn JS, Radhakrishnan ML, Mapelli M, Choi S, Tidor B, Cuny GD, Musacchio A, Yeh LA, Kosik KS. Defining Cdk5 ligand chemical space with small molecule inhibitors of Tau phosphorylation. *Chem Biol* 2005;12:811–823.
69. Armstrong KA, Tidor B, Cheng AC. Optimal charges in lead progression: a structure-based neuraminidase case study. *J Med Chem* 2006;49:2470–2477.
70. Altman MD, Nalivaika EA, Prabu-Jeyabalan M, Schiffer CA, Tidor B. Computational design and experimental study of tighter binding peptides to an inactivated mutant of HIV-1 protease. *Proteins: Struct Funct Bioinform* 2008;70:678–694.
71. Sherman W, Tidor B. Novel method for probing the specificity binding profile of ligands: applications to HIV protease. *Chem Biol Drug Des* 2008;71:387–407.
72. Green DF, Tidor B. Design of improved protein inhibitors of HIV-1 cell entry: optimization of electrostatic interactions at the binding interface. *Proteins: Struct Funct Bioinform* 2005;60:644–657.
73. Hendsch ZS, Tidor B. Electrostatic interactions in the GCN4 leucine zipper: substantial contributions arise from intramolecular interactions enhanced on binding. *Protein Sci* 1999;8:1381–1392.
74. Midelfort KS, Hernandez HH, Lippow SM, Tidor B, Drennan CL, Witttrup KD. Substantial energetic improvement with minimal structural perturbation in a high affinity mutant antibody. *J Mol Biol* 2004;343:685–701.
75. Carrascal N, Green DF. Energetic decomposition with the generalized-Born and Poisson-Boltzmann solvent models: lessons from association of G-protein components. *J Phys Chem B* 2010;114:5096–5116.
76. Kangas E, Tidor B. Electrostatic specificity in molecular ligand design. *J Chem Phys* 2000;112:9120–9131.
77. Ren JS, Esnouf R, Garman E, Somers D, Ross C, Kirby I, Keeling J, Darby G, Jones Y, Stuart D, Stammers D. High-resolution structures of HIV-1 RT from 4 RT-inhibitor complexes. *Nat Struct Biol* 1995;2:293–302.
78. Weichenberger CX, Sippl MJ. NQ-Flipper: validation and correction of asparagine/glutamine amide rotamers in protein crystal structures. *Bioinformatics* 2006;22:1397–1398.
79. Weichenberger CX, Sippl MJ. NQ-Flipper: recognition and correction of erroneous asparagine and glutamine side-chain rotamers in protein structures. *Nucleic Acids Res* 2007;35:W403–W406.
80. Brunger AT, Karplus M. Polar hydrogen positions in proteins—empirical energy placement and neutron-diffraction comparison. *Proteins: Struct Funct Genet* 1988;4:148–156.
81. Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM—a program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 1983;4:187–217.
82. Momany FA, Rone R. Validation of the general-purpose QUANTA(R)3.2/CHARMM(R) force-field. *J Comput Chem* 1992;13:888–900.
83. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;79:926–935.
84. Bayly CI, Cieplak P, Cornell WD, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges—the RESP model. *J Phys Chem* 1993;97:10269–10280.
85. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, 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 JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. Gaussian 03, Revision E.01. Wallingford, CT: Gaussian, Inc.; 2004.
86. Green DF, Tidor B. Evaluation of ab initio charge determination methods for use in continuum solvation calculations. *J Phys Chem B* 2003;107:10261–10273.
87. McLachlan AD. Rapid comparison of protein structures. *Acta Crystallogr Sect A* 1982;38:871–873.
88. Altman MD. Computational ligand design and analysis in protein complexes using inverse methods, combinatorial search, and accurate solvation modeling. PhD Thesis. Cambridge: Massachusetts Institute of Technology; 2006.
89. Sitkoff D, Sharp KA, Honig B. Accurate calculation of hydration free-energies using macroscopic solvent models. *J Phys Chem* 1994;98:1978–1988.
90. Drud A. CONOPT—a GRG code for large sparse dynamic nonlinear optimization problems. *Math Program* 1985;31:153–191.
91. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph* 1996;14:33–38.
92. Liang YH, Feng XQ, Zeng ZS, Chen FE, Balzarini J, Pannecouque C, De Clercq E. Design, synthesis, and SAR of naphthyl-substituted diarylpyrimidines as non-nucleoside inhibitors of HIV-1 reverse transcriptase. *Chemmedchem* 2009;4:1537–1545.
93. Hendsch ZS, Tidor B. Do salt bridges stabilize proteins—a continuum electrostatic analysis. *Protein Sci* 1994;3:211–226.
94. Lumb KJ, Kim PS. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled-coil. *Biochemistry* 1995;34:8642–8648.
95. Lumb KJ, Kim PS. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry* 1995;34:8642–8648; Erratum 1998;37:13042–13042.
96. Bosshard HR, Marti DN, Jelesarov I. Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings. *J Mol Recognit* 2004;17:1–16.
97. Hartmann MD, Ridderbusch O, Zeth K, Albrecht R, Testa O, Woolfson DN, Sauer G, Dunin-Horkawicz S, Lupas AN, Alvarez BH. A coiled-coil motif that sequesters ions to the hydrophobic core. *Proc Natl Acad Sci USA* 2009;106:16950–16955.
98. Ladbury JE. Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chem Biol* 1996;3:973–980.
99. Radhakrishnan ML, Tidor B. Specificity in molecular design: a physical framework for probing the determinants of binding specificity and promiscuity in a biological environment. *J Phys Chem B* 2007;111:13419–13435.
100. Lansdon EB, Brendza KM, Hung M, Wang R, Mukund S, Jin DB, Birkus G, Kutty N, Liu XH. Crystal structures of HIV-1 reverse transcriptase with etravirine (TMC125) and rilpivirine (TMC278): implications for drug design. *J Med Chem* 2010;53:4295–4299.
101. Gordon DB, Marshall SA, Mayo SL. Energy functions for protein design. *Curr Opin Struct Biol* 1999;9:509–513.
102. Charnock SJ, Bolam DN, Nurizzo D, Szabo L, McKie VA, Gilbert HJ, Davies GJ. Promiscuity in ligand-binding: the three-dimensional structure of a *Piromyces* carbohydrate-binding module, CBM29-2, in complex with cello-and mannohexaose. *Proc Natl Acad Sci USA* 2002;99:14077–14082.
103. Bencharit S, Morton CL, Hyatt JL, Kuhn P, Danks MK, Potter PM, Redinbo MR. Crystal structure of human carboxylesterase 1 complexed with the Alzheimer's drug tacrine: from binding promiscuity to selective inhibition. *Chem Biol* 2003;10:341–349.
104. Hopkins AL, Mason JS, Overington JP. Can we rationally design promiscuous drugs? *Curr Opin Struct Biol* 2006;16:127–136.