



Published in final edited form as:

*J Chem Inf Model*. 2009 December ; 49(12): 2851–2862. doi:10.1021/ci900320p.

## Prediction of Potency of Protease Inhibitors Using Free Energy Simulations with Polarizable Quantum Mechanics-Based Ligand Charges and a Hybrid Water Model

Debananda Das<sup>1</sup>, Yasuhiro Koh<sup>2</sup>, Yasushi Tojo<sup>2</sup>, Arun K. Ghosh<sup>3</sup>, and Hiroaki Mitsuya<sup>1,2,\*</sup>

<sup>1</sup>Experimental Retrovirology Section, HIV and AIDS Malignancy Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892-1868, USA

<sup>2</sup>Departments of Hematology and Infectious Diseases, Kumamoto University Graduate School of Medical and Pharmaceutical Sciences, Kumamoto 860-8556, Japan

<sup>3</sup>Departments of Chemistry and Medicinal Chemistry, Purdue University, West Lafayette, Indiana 47907, USA

### Abstract

Reliable and robust prediction of binding affinity for drug molecules continues to be a daunting challenge. We have simulated the binding interactions and free energy of binding of nine protease inhibitors (PIs) with wild-type and various mutant proteases by performing GBSA simulations, in which each PI's partial charge was determined by quantum mechanics (QM) and the partial charge accounts for the polarization induced by the protease environment. We employed a hybrid solvation model that retains selected explicit water molecules in the protein with surface generalized Born (SGB) implicit solvent. We examined the correlation of the free energy with antiviral potency of PIs with regard to amino acid substitutions in protease. The GBSA free energy thus simulated showed strong correlations ( $r > 0.75$ ) with antiviral IC<sub>50</sub> values of PIs when amino acid substitutions were present in the protease active site. We also simulated the binding free energy of PIs with P2-*bis*-tetrahydrofuranylurethane (*bis*-THF) or related cores, utilizing a *bis*-THF-containing protease crystal structure as a template. The free energy showed a strong correlation ( $r = 0.93$ ) with experimentally determined anti-HIV-1 potency.

The present data suggest that the presence of selected explicit water in protein, and protein polarization-induced quantum charges for the inhibitor, compared to lack of explicit water and a static force field-based charge model, can serve as an improved lead optimization tool, and warrants further exploration.

### Introduction

Virtual screening has been successful in the discovery of certain novel inhibitors, and a number of these inhibitors have advanced to clinical trials.<sup>1</sup> When the structure of a target protein is available, virtual screening involves docking potential inhibitors against the protein and ranking the inhibitors by their predicted affinity using a scoring function. Molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) or Molecular

\*Correspondence should be addressed to: H.M. Postal address: 10 Center Drive, Room 5A11 - MSC 1868 Bethesda, MD 20892-1868 Phone: 301-496-9238 Facsimile: 301-402-0709 hmitsuya@helix.nih.gov .

**Supporting Information Available:** The simulated GBSA free energies are given in Supplementary Tables S1 and S2. The charges on DRV for OPLS2005 force field (MM level) and protein polarized charges at the B3LYP/6-31G\* level are given in supplementary figures S3 and S4. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Mechanics Generalized Born surface area (MM-GBSA) have been used in some instances in the post-processing and re-ranking of results from molecular docking.<sup>2</sup> Of note, docking and scoring have currently been an integral part of drug discovery efforts and have produced documented successes, however, there is an urgent need for improvement of the accuracy of docking and scoring results.<sup>3</sup> With this regard, Clark has described four areas of improvement, *i.e.*, better scoring functions, treatment of protein flexibility, treatment of water molecules, and improved technology for data analysis of virtual screening results.<sup>1</sup> The scoring functions fail if they do not properly account for solvation, entropy, or polarizability.<sup>1, 4</sup>

Water molecules form polar interactions with both proteins and ligands, fill empty spaces in cavities, and serve as an important component of molecular recognition. Lu *et al.* have analyzed water molecules present at the interfaces of 392 X-ray crystal structures of protein-ligand complexes and have reported high correlations between the polar van der Waals surface area of ligands and the number of ligand-bound water molecules in the crystal structures.<sup>5</sup> In some instances, as many as twenty-one water molecules are bound to a ligand, with the average being 4.6.<sup>5</sup> Despite their importance, the treatment of water molecules in docking calculations have not been widespread because of methodological limitations and poor understanding of how many and which water molecules are to be included in the simulation. By sampling multiple water positions during docking, Huang and Shoichet have recently assessed the ligand enrichment against twenty-four targets.<sup>6</sup> Inclusion of water molecules increased enrichment against twelve targets, while remaining largely unaffected for the others.<sup>6</sup> Fornabaio *et al.* have reported that waters play a significant role in the energetics of binding and have performed a hydrophobic analysis of HIV-1 protease complexes.<sup>7</sup> They have reported a significant improvement of the correlation between their HINT free energy scores and experimentally determined binding constants when appropriate bridging water molecules were taken into account.<sup>7</sup>

Most of the studies measure the accuracy of scoring functions by their ability to correctly rank the activity of a congeneric set of ligands. The prediction of activity of a ligand against mutant proteins is equally important in light of drug resistance in several diseases including acquired immune deficiency syndrome (AIDS) and cancers. In the present study, we focus on the resistance mutations of HIV-1 protease. HIV-1 protease acquires amino acid substitutions under the selection pressure of protease inhibitors (PIs), rendering HIV-1 resistant to such PIs.<sup>8</sup> For example, an Asp30Asn (D30N) substitution causes resistance against nelfinavir. Some amino acid substitutions, while being initially selected under drug pressure against one inhibitor, confer on HIV-1 cross-resistance against other inhibitors.<sup>8</sup> One example of such a substitution is M46I which is a primary indinavir-resistance-associated substitution, but M46I-containing HIV-1 is resistant to other inhibitors such as ritonavir, nelfinavir, and atazanavir.<sup>9</sup> Analysis of the crystal structures of interactions of PIs with mutant proteases have shown that a number of drug resistance-associated mutations, such as G48V, V82A and I84V, occur in the catalytic active site of protease.<sup>10-12</sup> Analyses of crystal structures of mutant proteases has revealed that there are, in general, no major conformational changes to the backbone conformation in such proteases and the changes in binding interactions from the wild-type may involve different polar interactions with a mutant side chain(s) or loss of favorable van der Waals contacts.<sup>13-15</sup> Structural interactions, which are sometimes able to provide rational explanation of the mechanism of resistance, are not able to predict *a priori*, for example, whether V82A causes a higher resistance for ritonavir compared to DRV. More reliable predictions of the potency of inhibitors against protease with drug-resistant mutations would be of use in the design of novel and more potent inhibitors.

The free energy of binding of ligands to proteins can be simulated by methods such as free energy perturbation and linear interaction energy (LIE) approximation.<sup>16, 17</sup> LIE is a semi-empirical method and is based on a linear approximation of polar and non-polar free energy contributions from molecular dynamics simulation averages.<sup>16</sup> The LIE method has recently been used in calculating the binding free energy of N-sulphonyl-glutamic acid inhibitors to MurD ligase, and in probing the DNA replication fidelity.<sup>18, 19</sup>

In the current study we simulated the binding free energies of nine protease inhibitors against wild-type (PRO<sup>WT</sup>) and mutant proteases (PRO<sup>MT</sup>) with standard and hybrid GBSA protocols. While a number of water molecules are present in the X-ray crystal structures of protease-inhibitor complexes, a water molecule that mediates hydrogen bond interactions of the protease inhibitors with Ile50 and Ile50' in the flap is common across several different inhibitor-protease complexes, and is present in the complexes for eight FDA-approved PIs. In this work, we explicitly incorporated water molecule bridging hydrogen bonds with the protease flap. For inhibitors nelfinavir and atazanavir, two additional water molecules that mediate hydrogen bonds between these inhibitors and other protease residues were also explicitly included. We compared the GBSA free energy of binding obtained from simulations with selected explicit water molecules in implicit solvation (a hybrid solvation model) with free energies that did not have the water molecule explicitly present. Furthermore, in the simulations, the inhibitor atoms had either forcefield-derived fixed partial charges or quantum mechanics-based partial charges that accounted for the polarization induced by the surrounding protein environment (a hybrid charge model). We also analyzed the correlation of the GBSA free energies obtained by the simulations with antiviral potency data (IC<sub>50</sub> values). Our current data suggest that selective inclusion of explicit water molecule(s) and protein polarization effects may improve the robustness of GBSA free energy simulations and aid the design of inhibitors that are potent against both wild-type and multi-drug-resistant HIV-1 variants.

## Methods

### Crystal structures used as starting templates

We explored various wild-type protease crystal structures from the protein data bank as starting templates for docking and subsequent free energy simulations. For convenience of protein expression and crystallization, some of the structures deposited in the protein data bank as wild-type structures have several mutations such as Q7K, K14R, R41K, L63P, I64V that are distant from the inhibitor binding site.<sup>15, 20, 21</sup> These non-active site mutations may not drastically alter the conformation of the protease and its interactions with inhibitors compared to a pristine wild-type protease of HIV-1<sub>LAI</sub> or HIV-1<sub>NL4-3</sub>. While one non-active site mutant, depending on the residue and location, may not necessarily affect the binding affinity comparisons, crystal structures with four or five non-active site mutations are unsuitable to be used for free energy simulations, especially when comparing the simulation data with antiviral potency against wild-type HIV-1. We used 2FDE, obtained from the protein data bank, as the starting template for docking against darunavir (DRV), amprenavir (APV), GRL-98065, GRL-02031, and GRL-06579. 2FDE is a co-crystal of brecaonavir and HIV-1<sub>LAI</sub> wild-type protease, and brecaonavir (BCV) has a *bis*-THF ligand as a core.<sup>22</sup> The PDB IDs of the crystal structures used for our simulations of the other inhibitors are as follows: 1HXB<sup>23</sup> for saquinavir (SQV); 2O4P<sup>24</sup> for TPV; 1OHR<sup>25</sup> for nelfinavir (NFV); 1MUI<sup>26</sup> for lopinavir (LPV), and 2AQU<sup>27</sup> for atazanavir (AZV). Waters were not modeled in the crystal structure of LPV,<sup>26</sup> but were present in all other structures.

Our goal was to explore the prediction of free energy of binding once a correct binding mode was obtained. In the present study, we demonstrate that the correct binding mode was reliably obtained when a ligand was docked against a protease structure obtained with a

similar core. To decrease uncertainty arising due to cross docking of ligands to different proteases, we docked ligands against the native protease crystal or against a protease structure obtained with a similar core. It is important to keep in mind that protease side chains may undergo subtle conformational changes to accommodate protease inhibitors of different shapes and sizes (the molecular weights of the PIs in the current study range from 506 to 705) – and these changes might be difficult to capture by simple minimization following ligand docking to non-native crystal structures.

## Docking

The interactions of protease inhibitors with wild-type HIV-1 protease were examined using computational structural modeling and molecular docking. Besides accounting for the conformational flexibility of the inhibitor, the polarization induced in the inhibitor by the protease was taken into consideration by employing polarizable quantum charges in the docking computations. The use of polarizable quantum charges has recently been shown to substantially improve the prediction of protein-ligand complex structures.<sup>28</sup> The QM-polarized ligand docking protocol utilizing Glide version 4.5, QSite version 4.5, Jaguar version 7.0, and Maestro version 8.5 (Schrödinger, LLC, New York, NY 2007) was used as described below. The crystal coordinates described above were obtained from the Protein data bank (<http://www.rcsb.org/>) and used as starting templates. Hydrogens were optimized with constraints on the heavy atoms. The crystal water that mediates the interactions between protease inhibitors and the protease flap was retained, and all other crystal waters were deleted. Close interactions in the protease were annealed, and the docking grid was set-up. Polarizable ligand charges were determined at B3LYP/6-31G\* level. The extra-precision mode of Glide,<sup>29, 30</sup> which has a higher penalty for unphysical interactions, was used. For each docking simulation, up to five final poses were retained and were compared with available X-ray structures to verify that the conformations were reasonable. It was particularly important that the correct ring conformations were obtained during docking. LPV produced ring conformations that were different than the conformations obtained from crystal complexes in some docking solutions, and such conformations were discarded for the subsequent GBSA simulations.

## GBSA scoring simulations

The general principle of a GBSA model has been described before. The free energy of binding,  $\Delta G_{\text{bind}}$  is calculated as,<sup>2</sup>

$$\Delta G_{\text{bind}} = \Delta E + \Delta G_{\text{solv}} + \Delta G_{\text{SA}}$$

$$\Delta E = E_{\text{complex}} - E_{\text{protein}} - E_{\text{ligand}},$$

where  $E_{\text{complex}}$ ,  $E_{\text{protein}}$ , and  $E_{\text{ligand}}$  are the minimized energies of the protease-inhibitor complex, protease, and inhibitor, respectively;

$$\Delta G_{\text{solv}} = G_{\text{solv}(\text{complex})} - G_{\text{solv}(\text{protein})} - G_{\text{solv}(\text{ligand})},$$

where  $G_{\text{solv}(\text{complex})}$ ,  $G_{\text{solv}(\text{protein})}$ , and  $G_{\text{solv}(\text{ligand})}$  are the solvation free energies of the complex, protein, and the inhibitor, respectively.

$$\Delta G_{SA} = G_{SA(\text{complex})} - G_{SA(\text{protein})} - G_{SA(\text{ligand})},$$

where  $G_{SA(\text{complex})}$ ,  $G_{SA(\text{protein})}$ , and  $G_{SA(\text{ligand})}$  are the surface area energies for the complex, protease, and the inhibitor, respectively. The simulations were carried out using the GBSA continuum model<sup>31</sup> in Prime, version 2.0 (Schrödinger, LLC, New York, NY, 2008). Prime uses a surface generalized Born (SGB) model employing a Gaussian surface instead of a van der Waals surface for better representation of the solvent accessible surface area.<sup>31</sup>

GBSA simulations were carried out for the protease-ligand complex structures obtained by molecular docking. The simulations were carried out in four different scenarios: i) no water molecules were retained in the protease, and ligand atoms have fixed charges based on the OPLS force field. This is the standard MM-GBSA simulation carried out in implicit solvation. The change in free energy obtained is denoted by  $\Delta G_{mm}$ , and the correlation coefficients are denoted by  $r_{mm}$ . ii) no water molecules were retained in the protease and the ligand has protein polarized QM charges at B3LYP/6-31G\* level. The protein polarized charge on the ligand is determined from the docked complex, and is used while computing  $E_{\text{complex}}$ ,  $G_{\text{solv}(\text{complex})}$ ,  $G_{SA(\text{complex})}$  as well as for  $E_{\text{ligand}}$ ,  $G_{\text{solv}(\text{ligand})}$ , and  $G_{SA(\text{ligand})}$ . The change in free energy obtained is denoted by  $\Delta G_{qm}$ , and the correlation coefficients are denoted by  $r_{qm}$ . iii) the bridging water molecule mediating the hydrogen bond interactions of inhibitors DRV, GRL-98065, APV, GRL-02031, GRL-06579, NFV, SQV, and AZV with Ile50 and Ile50' in the flap was explicitly retained. This is a hybrid solvation model since implicit GBSA solvation terms for the whole system were used. For tipranavir (TPV), GBSA with the hybrid solvation model was performed by retaining a water molecule that bridges hydrogen bond interactions with Gly48 of one monomer of the protease. NFV and AZV were observed to have two additional bridging water molecules, and additional calculations in the presence of three explicit water molecules were performed for NFV and AZV. In the hybrid solvation model, the inhibitors either had MM charges (change in free energy and correlation coefficient denoted by  $\Delta G_{mm/wat}$  and  $r_{mm/wat}$ , respectively) or iv) protein polarized QM (B3LYP/6-31G\*) charges (change in free energy and correlation coefficient denoted by  $\Delta G_{qm/wat}$  and  $r_{qm/wat}$ , respectively). In all simulations, the protease has OPLS charges. The strain energies of the ligands were taken into account.

### Antiviral agents

DRV, GRL-98065, and GRL-02031 were synthesized as described previously.<sup>32-35</sup> SQV and ritonavir (RTV) were kindly provided by Roche Products Ltd. (Welwyn Garden City, United Kingdom) and Abbott Laboratories (Abbott Park, IL), respectively. APV was a kind gift from Glaxo-Wellcome, Research Triangle Park, NC. NFV and indinavir (IDV) were kindly provided by Japan Energy Inc, Tokyo, Japan. LPV was synthesized by previously published methods.<sup>36</sup> AZV was a kind gift from Bristol-Myers Squibb (New York, NY). TPV was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, National Institutes of Health.

### Generation of recombinant HIV-1 clones

To generate HIV-1 clones carrying desired mutations, site-directed mutagenesis using the QuikChange Site-directed Mutagenesis Kit (Stratagene, La Jolla, CA) was performed, and the mutation-containing genomic fragments were introduced to pHIV-1<sub>NLS $\Delta$ ma</sub> as previously described.<sup>35,37</sup> Determination of the nucleotide sequences of plasmids confirmed that each clone had the desired mutations but no unintended mutations. Each recombinant plasmid was transfected into 293T cells using Lipofectoamine 2000 Transfection Reagent



(Invitrogen, Carlsbad, CA), and thus generated infectious virions were harvested 48 h after transfection and stored at  $-80^{\circ}\text{C}$  until use. HIV-1 carrying D30N substitution (HIV<sup>D30N</sup>) was generated since residue-30 is in the active site (Fig. 1) and the D30N substitution is known to cause primary drug resistance against the FDA-approved protease inhibitor NFV.<sup>38</sup> HIV<sup>I50V</sup> was generated since Ile50 is in the flap region (Fig. 1) and interacts with various PIs through a bridging water molecule.<sup>39</sup> The I50V mutation has been associated with resistance to APV, LPV, and RTV.<sup>40</sup> HIV<sup>V82I/I85V</sup> was also generated since Val82 is located in the active site and its substitution to Ile is associated with HIV-1 resistance to most PIs, presumably due to the expansion of the active site and loss of favorable van der Waals contact.<sup>15, 20, 40, 41</sup> We recently reported the emergence of I85V as a resistant mutation against a PI, GRL-98065 and chose to study the combined effect of V82A/I85V.<sup>32</sup> HIV-1-infected patients who failed to respond to PI-containing regimens often have HIV-1 variants carrying both active site and non-active site mutations in protease, and we chose to explore such clinical HIV-1 isolates, HIV<sup>2840</sup> and HIV<sup>2841</sup>. The former contained L10R, M46I, L63P, V82T and I84V substitutions, while the latter contained M46I, L63P, V82T and I84V substitutions.

### Drug susceptibility assay

To determine the drug susceptibilities of certain laboratory HIV-1 strains, MT-4 cells were employed as target cells, as described previously,<sup>37</sup> with minor modifications. In brief, MT-4 cells ( $10^5/\text{ml}$ ) were exposed to 100 TCID<sub>50</sub>s of drug-resistant HIV-1 strains in the presence or the absence of various concentrations of drugs and were incubated at  $37^{\circ}\text{C}$ . On day 7 of culture, the supernatants were harvested and the amounts of the p24 Gag protein were determined by using a fully automated chemiluminescent enzyme immunoassay system (Lumipulse F; Fujirebio Inc., Tokyo, Japan).<sup>42</sup> The drug concentrations that suppressed the production of p24 Gag protein by 50% (IC<sub>50</sub>) were determined by comparison of the amount of p24 Gag protein produced in drug-treated cell cultures with the level of p24 Gag protein produced in a drug-free control cell culture. All assays were performed in duplicate or triplicate on more than three different occasions and the data are shown as means  $\pm$  1 S.D.

## Results

### Analysis of water molecules in protease crystal structures

Since the number of ligand-bound water molecules in high resolution crystal structures varies widely,<sup>5</sup> we first analyzed the number of water molecules in the five protease-inhibitor crystal structures used in the current study. The total number of water molecules ranged from 51 (protease-NFV complex,<sup>25</sup> PDB ID 1OHR) to 124 (protease-TPV complex,<sup>24</sup> PDB ID 2O4P). Within  $4\text{\AA}$  of the bound ligand, the number of water molecules present in the complexes with PDB ID: 1HXB, 2FDE, 1OHR, 2O4P, and 2AQU were 3, 5, 5, 10, and 10, respectively. We then analyzed the number of water molecules bridging hydrogen bond interactions between the protease and the ligands. The water molecule forming tetra-coordinated hydrogen bond interactions with Ile50 and Ile50' in the flap was the only water molecule interacting with the ligand in the protease complexes of SQV (PDB ID: 1HXB), BCV (PDB ID: 2FDE), and APV (PDB ID: 1HPV<sup>43</sup>). Since this tetra-coordinated water molecule is considered an important pharmacophore for protease-inhibitor interactions,<sup>44</sup> it was explicitly included in our docking and GBSA free energy simulations. Three water molecules formed bridging hydrogen bond interactions in the NFV-protease (PDB ID: 1OHR) and AZV-protease (PDB ID: 2AQU) complexes (Figs. 2c, and q), and simulations were carried out with both one and three bridging water molecules explicitly present for these inhibitors. TPV directly formed hydrogen bonding with Ile50 and Ile50', and a water molecule bridged hydrogen bonding with Gly48 in the flap (Fig. 2g),<sup>24, 45, 46</sup> and was

explicitly included in the simulations. Simulations involving LPV did not include any crystal waters since none was present in the native LPV-protease complex.<sup>26</sup>

### Determination of the susceptibility of recombinant infectious HIV-1 clones carrying amino acid substitutions to PIs in vitro

We determined the susceptibility of six recombinant infectious HIV-1 clones to each of the PIs that were chosen in the present study: 7 clinical available PIs (SQV, NFV, APV, TPV, DRV, and AZV) and 2 experimental PIs (GRL-02031 and GRL-98065) in the HIV-1 p24 Gag production inhibition assay, as previously described.<sup>35, 37</sup> As illustrated in Table 1, most of the recombinant clones showed reduced susceptibility to the PIs examined by up to 16.5-fold. However, it was also noted that HIV<sup>D30N</sup> and HIV<sup>2841</sup> had increased susceptibility to certain PIs. The increased susceptibility of HIV<sup>D30N</sup> to TPV with 33.3-fold was notable, although HIV<sup>D30N</sup> was also less susceptible to SQV, LPV, and NFV (Table 1). Both HIV<sup>I50V</sup> and HIV<sup>2840</sup> were also less susceptible to most of the PIs (Table 1).

### Binding interactions with wild-type and D30N mutant protease

We next determined and analyzed the binding modes of nine different PIs with PRO<sup>WT</sup> and a protease with an amino acid substitution at position 30 from an aspartic acid to asparagine (PRO<sup>D30N</sup>). SQV has four hydrogen bond interactions with Asp29 and Asp30 in the S2 site of the wild-type protease, but has no hydrogen bonds with Asp29', or Asp30' in the S2' site (Fig. 2a). When protease acquires the D30N mutation, SQV loses two hydrogen bonds with Asp29 and Asp30, and does not form any new and compensating hydrogen bonds with other protease residues (Fig. 2b). Comparison of antiviral data of SQV shown in Table 1 indicates that there was a 3.9-fold decrease in antiviral potency with the D30N mutation. It is possible that the decrease of antiviral potency of SQV for D30N mutant is due to the loss of hydrogen bonds with residues 29 and 30 for the mutant. Examining the hydrogen bonds in the S2 site for NFV against PRO<sup>WT</sup> and PRO<sup>D30N</sup> mutant protease (Figs. 2c-d), one observes that NFV has more hydrogen bonds with Asn30 of PRO<sup>D30N</sup> compared to Asp30 of PRO<sup>WT</sup>. An X-ray crystal structure has also demonstrated that NFV has a larger number of hydrogen bonds with PRO<sup>D30N</sup> than with PRO<sup>WT</sup>.<sup>11</sup>

However, D30N is a major amino acid substitution<sup>38</sup> associated with NFV resistance of HIV-1 and HIV<sup>D30N</sup> showed a 5.3-fold increase in IC<sub>50</sub> values in antiviral assays (Table 1). Examining the structural interactions of LPV and TPV against PRO<sup>WT</sup> and PRO<sup>D30N</sup> mutants (Figs. 2e-h), we observed that both the inhibitors have more hydrogen bonds with residues 29 and 30 of PRO<sup>D30N</sup> than they have for those of PRO<sup>WT</sup>. Antiviral data show that while LPV has a 3.1 fold increase in IC<sub>50</sub>, HIV<sup>D30N</sup> was about 30 times more sensitive to TPV (Table 1). Comparison of the structural interactions of DRV, TPV, GRL-02031, and GRL-98065 with PRO<sup>WT</sup> and PRO<sup>D30N</sup> (Figs. 2k-p) revealed that all of them have more hydrogen bond interactions with PRO<sup>D30N</sup>. APV, which is 5-fold more potent against HIV<sup>D30N</sup> (Table 1), has three hydrogen bonds with Asp29, Asp30, and Asp30' for PRO<sup>WT</sup>, and forms three hydrogen bonds with Asp29 and Asn30 for PRO<sup>D30N</sup> in structural models (Figs. 2i-j). Thus, it is clear that while the number of hydrogen bonds may provide an intuitive understanding of binding and/or antiviral potency, it may not always explain why certain PIs show a decrease in antiviral potency with D30N substitution while other PIs show an increase.

### Free energy changes in complexes with the D30N mutation determined by GBSA simulations

Since the number of hydrogen bonds between PIs and protease does not always help predict the potency of PIs as discussed above, we examined the free energies under four different simulation conditions: with and without explicit water(s); and with QM or with MM charges

on the inhibitor (Table 2, Supplementary Tables S1 and S2). It was assumed that an increase in the change of free energy of binding ( $\Delta\Delta G$  more positive) is to be expected for a decrease in antiviral activity, and *vice versa*. With the D30N mutation in PRO<sup>D30N</sup>, SQV showed a reduction in antiviral activity by 3.9-fold. With the bridging water and QM charges on SQV, the free energy change ( $\Delta\Delta G_{\text{qm/wat}}$ ) of the SQV-protease complex increased by 4 kcal/mol for the D30N mutation (Table 2). HIV<sup>D30N</sup> was resistant to NFV by 5.3-fold compared to HIV<sup>WT</sup> and  $\Delta\Delta G_{\text{qm/wat}}$  showed an increase in the free energy of binding by +7 kcal/mol. HIV-1 containing PRO<sup>D30N</sup> was more sensitive to APV and TPV, and  $\Delta\Delta G_{\text{qm/wat}}$  for PRO<sup>D30N</sup> with APV and TPV was -12 and -5 kcal/mol, respectively (Table 2). The  $\Delta\Delta G_{\text{qm/wat}}$  values showed that both TPV and APV had a higher affinity for PRO<sup>D30N</sup> than for PRO<sup>WT</sup> (Table 2) and correlated with the increase in sensitivity to these inhibitors.

We next compared the free energy changes ( $\Delta\Delta G_{\text{mm/wat}}$ ) simulated using the bridging water molecule but with force field based fixed MM charges on the inhibitors. The  $\Delta\Delta G_{\text{mm/wat}}$  values of +4 kcal/mol and +9 kcal/mol for SQV and NFV, respectively, showed that the simulation results in an increase in the free energy of binding correlating with the reduction of antiviral activity. However, the  $\Delta\Delta G_{\text{mm/wat}}$  value showed a wrong trend for APV (+12 kcal/mol) and was alike for TPV (-2 kcal/mol). The  $\Delta\Delta G_{\text{qm}}$  and  $\Delta\Delta G_{\text{mm}}$  values were negative for SQV and NFV, respectively, while it was positive for APV. Thus  $\Delta\Delta G_{\text{qm}}$  and  $\Delta\Delta G_{\text{mm}}$  values, which did not incorporate the bridging water molecule explicitly, simulated inaccurate changes in the free energy of binding for SQV, NFV, and APV. The crystal structure for LPV (PDB ID: 1MUI) did not have any water molecules present, and all simulations involving LPV were carried out with implicit water. The  $\Delta\Delta G_{\text{qm}}$  and  $\Delta\Delta G_{\text{mm}}$  values for LPV were +17 kcal/mol and +21 kcal/mol, respectively. The increase in the change in the free energy of binding of LPV with PRO<sup>D30N</sup> was consistent with its decrease in antiviral potency with D30N substitution. For TPV, the negative  $\Delta\Delta G_{\text{qm}}$  and  $\Delta\Delta G_{\text{mm}}$  values indicated favorable free energy of binding for PRO<sup>D30N</sup> compared to PRO<sup>WT</sup> and correlated with the increase in antiviral potency of TPV with the D30N mutant.

In summary, the  $\Delta\Delta G_{\text{qm/wat}}$  values provided consistent trends of the change in free energy of binding for PRO<sup>D30N</sup>, which is resistant to SQV and NFV, but is more sensitive to APV and TPV.  $\Delta\Delta G_{\text{qm}}$  and  $\Delta\Delta G_{\text{mm}}$  did not always provide the correct trend of change in the free energies of binding.

### Correlation of free energy and antiviral potency for active site mutants

The GBSA free energies were simulated under four conditions: i) with implicit solvation terms and MM charges on both ligand and protein ( $\Delta G_{\text{mm}}$ ); ii) with an explicit water and implicit solvation terms (hybrid solvation model) with MM charges on both ligand and protein ( $\Delta G_{\text{mm/wat}}$ ); iii) with implicit solvation terms and protein polarized QM charges on the ligand and MM charges on the protein ( $\Delta G_{\text{qm}}$ ); and iv) with an explicit water and implicit solvation terms (hybrid solvation model) with QM charges on the ligand and MM charges on the protein ( $\Delta G_{\text{qm/wat}}$ ). All the free energy values determined are shown in supplementary Tables S1 and S2.

We analyzed the correlation of the free energies thus computed with the experimentally-determined antiviral potency data ( $IC_{50}$ ), and the resultant correlation coefficients are shown in Table 3, where Set-1 refers to the correlation coefficient for PRO<sup>WT</sup>, PRO<sup>D30N</sup>, PRO<sup>I50V</sup>; Set-2 refers to the correlation coefficient for PRO<sup>WT</sup>, PRO<sup>D30N</sup>, PRO<sup>I50V</sup> and PRO<sup>V82I/I85V</sup>; and Set-3 refers to the correlation coefficients for PRO<sup>WT</sup>, PRO<sup>D30N</sup>, PRO<sup>I50V</sup>, PRO<sup>V82I/I85V</sup>, PRO<sup>2840</sup>, and PRO<sup>2841</sup>. For Set-1,  $r_{\text{mm}}$  (correlation coefficient of  $\Delta G_{\text{mm}}$  vs  $IC_{50}$ ) showed a strong correlation for only TPV, LPV, and AZV (Table 3). The  $r_{\text{mm}}$  value was poor for the other PIs and indicates the difficulty of obtaining reasonable correlations between free energies and antiviral potency. The  $r_{\text{qm}}$  (correlation coefficient of



$\Delta G_{qm}$  vs  $IC_{50}$ ) values that represented correlation coefficients when the free energies were simulated with polarized QM charges on the ligands showed significant improvement and a strong correlation for DRV, APV, GRL-02031, and LPV. However, both  $r_{mm}$  and  $r_{qm}$  values were poor for GRL-98065 and SQV.

We next determined, for Set-1, the correlation obtained by the hybrid water model that has an explicit bridging water molecule between the inhibitor and the protease flap. The explicit water was treated as a part of the protein, and implicit solvation terms were used. The  $r_{mm/wat}$  value (correlation coefficient of  $\Delta G_{mm/wat}$  vs  $IC_{50}$ ) represented a greater correlation than  $r_{mm}$  for all PIs except TPV and AZV. TPV directly formed hydrogen bonds with Ile50, and Ile50', and the water molecule included in this calculation formed hydrogen bonds with Gly48 of one monomer of the protease dimer. For other PIs, the bridging water molecule formed hydrogen bonds with the flaps from both monomers. The  $r_{qm/wat}$  value (correlation coefficient of  $\Delta G_{qm/wat}$  vs  $IC_{50}$ ) had a high degree of correlation for all PIs except AZV. Thus, the explicit inclusion of the water molecule bridging hydrogen bonds with the flap and protein polarized QM charges for the inhibitors provided strong correlation ( $r > 0.75$ ) for seven out of eight inhibitors. The correlation coefficient  $r_{qm/wat}$  for NFV with three bridging waters was 0.97, a significant improvement over the correlation coefficient of 0.77 obtained with one bridging water molecule. The  $r_{qm/wat}$  value for AZV also improved from 0.16 to 0.64 with the inclusion of three bridging water molecules.

We also determined  $r_{mm}$ ,  $r_{qm}$ ,  $r_{mm/wat}$ , and  $r_{qm/wat}$  values for Set-2, which included PRO<sup>V82I/I85V</sup> as well (Table 3). The  $r_{mm}$  value was poor for all PIs except TPV, while the  $r_{qm}$  values showed good correlations for DRV, LPV and TPV. The  $r_{mm/wat}$  value showed strong correlation for only SQV, and good correlation ( $0.55 < r < 0.75$ ) for NFV. The  $r_{qm/wat}$  value showed strong correlations for DRV, APV, and SQV, and good correlations for GRL-98065 and TPV. The  $r_{qm/wat}$  value for NFV jumped from 0.59 to 0.92 with the incorporation of three bridging water molecules instead of one.

The correlation coefficients with a hybrid water model and with QM polarized ligand charges ( $r_{qm/wat}$ ) on the PIs were higher for most of the PIs compared to the correlation coefficients obtained without any explicit water molecule and MM charges ( $r_{mm}$ ). The only exception was for TPV, which was the only non-peptidomimetic inhibitor among the PIs examined. TPV displaces the tetra-coordinated water molecule and interacts directly with Ile50 and Ile50' in the flap.<sup>46</sup> The hydrogen bond interaction of the bridging water molecule with TPV and Gly48 of one chain might not be an important contributor to its potency. Also, in general, the  $r_{qm/wat}$  values provided better correlations than  $r_{qm}$ .

### Correlation of free energy and antiviral potency for active and non-active site amino acid substitutions

We next analyzed the correlations of the free energies with the antiviral potency ( $IC_{50}$  values) for PRO<sup>WT</sup>, PRO<sup>D30N</sup>, PRO<sup>I50V</sup>, PRO<sup>V82I/I85V</sup>, PRO<sup>2840</sup> that contains L10R, M46I, L63P, V82T, and I84V and PRO<sup>2841</sup> that contains M46I, L63P, V82T, and I84V substitutions (Set-3 in Table 3). The analysis of PRO<sup>2840</sup> and PRO<sup>2841</sup> was substantially complex since both proteases contained non-active site substitutions, but it was worth examining the ability of the GBSA energy function to correlate with antiviral activity when substitutions distant from the inhibitor were present. In general, the correlation coefficients for Set-3 turned out to be low, indicating a lower correlation between the free energies and antiviral potencies when non-active site mutants were present. For DRV, the  $r_{mm}$ ,  $r_{qm}$ , and  $r_{mm/wat}$  values indicated that the corresponding free energies had no correlation with  $IC_{50}$  values. However, the value of  $r_{qm/wat}$  for DRV was 0.83, showing a strong correlation for simulations with polarized QM charges on the inhibitor with a hybrid water model (Table 3, Fig. 3a). For SQV, strong correlations were obtained for both  $\Delta G_{mm}$  with  $IC_{50}$  ( $r_{mm} = 0.78$ )

and  $\Delta G_{qm/wat}$  with  $IC_{50}$  ( $r_{qm/wat} = 0.84$ ). For APV and TPV,  $r_{qm/wat}$  values were 0.60, and 0.50, respectively; but there was no correlation when other GBSA protocols were used for these inhibitors. The  $r_{qm/wat}$  values were greater than 0.75 for two inhibitors, and were greater than 0.55 for five inhibitors. The  $r_{qm}$  and  $r_{mm/wat}$  values were greater than 0.55 for two inhibitors. Thus, the hybrid water model and inclusion of polarization effects, compared to the other protocols, simulated free energies with better correlation with anti-viral  $IC_{50}$  for more inhibitors.

### GBSA with hybrid water model and polarizable quantum charges on PIs as a lead optimization tool

Obtaining a correct relative rank of activity for inhibitors that had potency in the nanomolar range has been a real challenge for scoring methods.<sup>3</sup> Scoring methods providing sufficiently high correlation with potency may serve as a lead optimization tool. In our data set, DRV and GRL-98065 have a *bis*-THF group as the core; GRL-06579 and GRL-02031 have a Cp-THF as the core, and APV has a THF group as the core. All these inhibitors were extremely potent against PRO<sup>WT</sup>, and had a narrow range of activity ranging from 0.3 to 28 nM (Table 1). We computed  $\Delta G_{exp}$  from the antiviral  $IC_{50}$  values (Table-4). Such transformations have recently enhanced the understanding of the binding of N-sulphonyl-glutamic acid inhibitors to MurD ligase, and in understanding the efficiency of DNA catalysis.<sup>19, 47</sup> Substitution of the THF group of APV with the *bis*-THF group (DRV) resulted in a  $-1.3$  kcal/mol improvement in the free energy of binding. GRL-98065 has a 1,3-benzodioxole group as P2' ligand compared to an aniline group in DRV which resulted in  $\Delta G_{exp}$  of GRL-98065 being lower (i.e. better binding affinity) by  $-1.4$  kcal/mol than DRV. Both GRL-06579 and GRL-02031 have a Cp-THF as P2-ligand but have different substituents interacting with the S1' and S2' locations in the protease active site. The  $\Delta G_{exp}$  of GRL-06579 is lower by 1.4 kcal/mol than GRL-02031. Amongst these five PIs, APV and GRL-02031 have the worst  $\Delta G_{exp}$ , and GRL-98065 has the best  $\Delta G_{exp}$ . The absolute magnitudes of our simulated GBSA free energies are larger than those measures experimentally due to the force field parameters, and the form of the energy function, but trends and correlation with experiments can be deduced. Our  $\Delta G_{qm/wat}$  values simulated the correct trend by predicting that the worst binding free energy was for APV and GRL-02031, and the best binding free energy was for GRL-98065 (Supplementary Table S1). None of the other GBSA protocols predicted the trend correctly (Supplementary Table S1 and S2). We further analyzed the correlation of the free energies calculated by our methods against  $IC_{50}$  values for these five PIs. The  $r_{mm}$  and  $r_{mm/wat}$  values were 0.10 and 0.03, respectively indicating that there was no correlation of GBSA free energies simulated by using MM charges, with or without explicit inclusion of water. The difficulty of obtaining good correlation of free energies against antiviral activity is evident from this small but non-trivial data set. The  $r_{qm}$  value was 0.74 indicating that incorporation of protein polarized QM charges on the inhibitors substantially improved the correlation. Including the bridging water molecule with protein polarized QM charges on the PIs resulted in a  $r_{qm/wat}$  value of 0.93 for PRO<sup>WT</sup> (Fig. 3b). Thus a simulation using the hybrid water model and protein polarized QM charges on the ligands resulted in a strong correlation while there was no correlation for simulations with fixed MM charges.

For our docking and subsequent GBSA simulations, we used the crystal coordinates of BCV-protease complex (PDB ID: 2FDE)<sup>22</sup> as our starting template. BCV, DRV and GRL-98065 had a *bis*-THF moiety as the core ligand, and APV, GRL-06579, and GRL-02031 had cores that had a high similarity with *bis*-THF. The high correlation obtained with  $\Delta G_{qm/wat}$ - $IC_{50}$  indicated that free energies obtained with a hybrid water model and polarized QM charges on the ligands would be a promising approach for lead optimization. Currently, we are exploring the utility of GBSA free energy simulations once a correct

binding mode is obtained, and we wanted to use native or native-like crystal structure as a template as a proof-of-principle study. We have not explored docking ligands against non-native protease crystal structures in the current study. A more exhaustive examination that includes poses obtained by cross docking needs to be performed.

## Discussion

Progress has been made in improving the robustness of scoring functions to determine the relative affinity of inhibitors for target proteins, but there are vast scopes where significant improvement can be made.<sup>1, 3</sup> Part of the reason for the inaccuracies in the scoring functions arise because i) the current methodologies largely account for enthalpic changes while completely ignoring entropic changes; ii) they do not properly treat protein flexibility; iii) they do not properly account for solvation and desolvation effects; and iv) they do not account for the polarization induced by the protein and the ligand on each other.<sup>1, 3, 7</sup> Most of the studies on scoring functions deal with rank-ordering the activity of a congeneric set of ligands. The prediction of activity of a ligand against mutant proteins is very important in light of drug resistance mutations that emerge in many therapeutic areas.

We have explored the correlation of GBSA free energies with antiviral potency ( $IC_{50}$  values) of nine different PIs against wild-type and mutant proteases. Correlation with enzymatic  $K_i$  would result in similar conclusions since these PIs have good cell penetration and are highly protease-specific inhibitors. Further more, these PIs do not form aggregates, and do not have problematic properties discussed by Shoichet *et al.*<sup>48, 49</sup> The GBSA free energies were simulated by four different protocols: i) fixed MM charges in implicit solvent ( $\Delta G_{mm}$ ); ii) protein polarized QM charges of PIs in implicit solvent ( $\Delta G_{qm}$ ); iii) fixed MM charges in a hybrid solvent ( $\Delta G_{mm/wat}$ ); and iv) protein polarized QM charges of PIs in a hybrid solvent ( $\Delta G_{qm/wat}$ ). The hybrid solvent protocols have retained an explicit water molecule(s) that bridges hydrogen bonding with the protease in an otherwise implicit solvent environment. The protein polarized ligand charges were determined during ligand docking in the protein environment at B3LYP/6-31G\* level, and these polarized charges were maintained for the ligand for the full GBSA simulation cycle. These enable us to analyze the effect of a different charge model compared to the static charges from a force-field (supplementary figures S3 and S4). Free energies from a successful GBSA protocol should provide a good correlation with the antiviral activity against wild-type and mutant protease. Resistance mutations in the protease active site arise primarily due to loss of favorable binding interactions. Resistance caused by non-active site mutations is more difficult to understand and rationalize although some attempts to elucidate the mechanism have recently been made.<sup>12, 50-52</sup>

We initially compared the correlation coefficients for the PRO<sup>WT</sup>, PRO<sup>D30N</sup>, and PRO<sup>I50V</sup> (Set-1). The antiviral potency of the inhibitors for wild-type and mutant proteases are shown in Table 1 and the inhibitors have potency ( $IC_{50}$  values) in the nanomolar range. The  $r_{qm/wat}$  value was higher than 0.75 for 7 out of 8 PIs for Set-1, with  $r_{qm/wat}$  being more than 0.90 for 4 PIs (Table 3). PRO<sup>D30N</sup> was associated with SQV-, NFV-, and LPV-resistance of HIV-1 and increased susceptibility to APV and TPV (Table 1). It is noteworthy that  $r_{qm/wat}$  showed substantial correlation values of greater than 0.75 even though the fold-change in antiviral activity for PRO<sup>D30N</sup> is non-monotonic. AZV is the only PI that did not show a correlation of  $\Delta G_{qm/wat}$  with the  $IC_{50}$  value with one bridging water molecule. By including two additional bridging water molecules for NFV and AZV,  $r_{qm/wat}$  improved to 0.97 and 0.64 respectively. Selective inclusion of explicit water molecules needs to be explored and validated. For Set-1,  $r_{mm/wat}$ ,  $r_{qm}$  and  $r_{mm}$  had strong correlation for 4, 3 and 3 PIs respectively.

Correlations for Set-2 included the mutant PRO<sup>V82I/I85V</sup>. While V82I is located in the protease active site, and represents one of the major resistance-associated mutations for PIs, I85V was selected as a resistance-associated mutation for GRL-98065 even though it does not form a direct van der Waals contact.<sup>32</sup> For Set-2, the number of PIs that showed strong correlation coefficients for  $r_{qm/wat}$  and  $r_{mm/wat}$  were 4 and 1 respectively, while  $r_{qm}$  and  $r_{mm}$  did not show strong correlations for any PI. Set-3 included Set-2 as well as PRO<sup>2840</sup> and PRO<sup>2841</sup>, which had mutations distant from the active site. We included such mutations because they are seen in drug-resistant HIV-1-harboring patients, and we wanted to test the ability of the GBSA energy function to correlate with the antiviral IC<sub>50</sub> for such protease substitutions. The  $r_{qm/wat}$  values showed strong correlation coefficients for only DRV and SQV for Set-3, and moderate correlation for 3 other PIs. The correlation coefficients,  $r_{mm/wat}$  and  $r_{qm}$  were from 0.55 to 0.75 for two PIs in Set-3, suggesting that those two PIs had a moderate correlation between the free energies and the IC<sub>50</sub> values.

Analysis of correlation coefficients in Table 3 indicated that  $r_{qm/wat}$  had strong ( $r > 0.75$ ) and moderate ( $0.55 < r < 0.75$ ) correlation for more PIs than  $r_{mm/wat}$ ,  $r_{qm}$  or  $r_{mm}$ . This suggested that the GBSA free energies simulated with a hybrid water model with protein polarized QM charges on the inhibitors had a higher correlation with antiviral IC<sub>50</sub> than the other free energy simulation protocols. Others<sup>1, 3</sup> have suggested that improved treatment of solvation and polarizability may improve the robustness of scoring functions, and we have demonstrated that our use of selected explicit water molecule(s) and protein polarized QM partial charges on the inhibitor provided greater correlation with antiviral potency. Further improvement might be achieved by improving upon the hybrid water model, by accounting for the polarization induced by the inhibitors on the protein atoms, and by including changes in entropy.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported in part by the Intramural Research Program of Center for Cancer Research, National Cancer Institute, National Institutes of Health (DD and HM), a grant from the National Institutes of Health (GM 53386 to AKG), a grant from the Kumamoto University Global Center of Excellence Program, Global Education and Research Center Aiming at the control of AIDS (HM) supported by the Ministry of Education, Culture, Sports, Science, and Technology (Monbu-Kagakusho), a Grant-in-Aid for Scientific Research (Priority Areas to HM) from the Ministry of Education, Culture, Sports, Science, and Technology (Monbu-Kagakusho) of Japan (HM), and a Grant for Promotion of AIDS Research from the Ministry of Health, Labor and Welfare (Kosei-Rodosho) of Japan (HM). The work utilized the computational resources of the Biowulf cluster at the NIH.

## REFERENCES AND NOTES

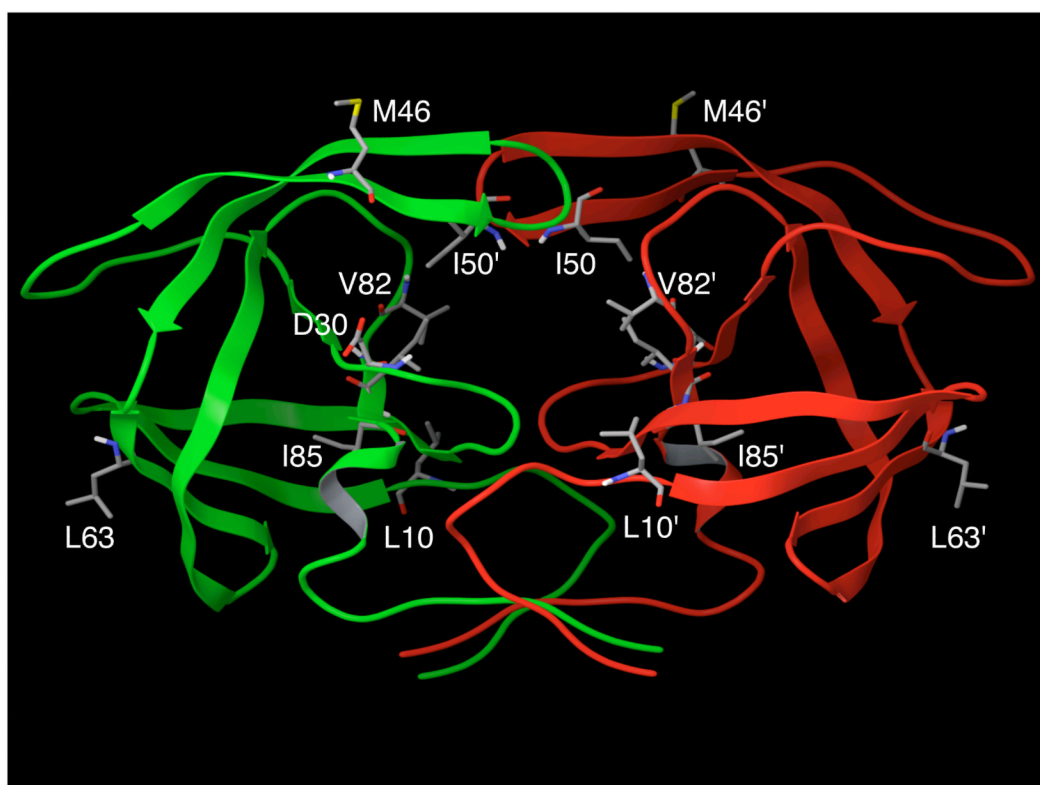
- (1). Clark DE. What has virtual screening ever done for drug discovery? *Expert. Opin. Drug Discov* 2008;3:841–851.
- (2). Lyne PD, Lamb ML, Saeh JC. Accurate prediction of the relative potencies of members of a series of kinase inhibitors using molecular docking and MM-GBSA scoring. *J. Med. Chem* 2006;49:4805–4808. [PubMed: 16884290]
- (3). Leach AR, Shoichet BK, Peishoff CE. Prediction of protein-ligand interactions. Docking and scoring: Successes and gaps. *J. Med. Chem* 2006;49:5851–5855. [PubMed: 17004700]
- (4). Rajamani R, Good AC. Ranking poses in structure-based lead discovery and optimization: Current trends in scoring function development. *Curr. Opin. Drug Discov. Dev* 2007;10:308–315.
- (5). Lu YP, Wang RX, Yang CY, Wang SM. Analysis of ligand-bound water molecules in high-resolution crystal structures of protein-ligand complexes. *J. Chem Inf. Model* 2007;47:668–675. [PubMed: 17266298]

- (6). Huang N, Shoichet BK. Exploiting ordered waters in molecular docking. *J. Med. Chem* 2008;51:4862–4865. [PubMed: 18680357]
- (7). Fornabaio M, Spyraakis F, Mozzarelli A, Cozzini P, Abraham DJ, Kellogg GE. Simple, intuitive calculations of free energy of binding for protein-ligand complexes. 3. The free energy contribution of structural water molecules in HIV-1 protease complexes. *J. Med. Chem* 2004;47:4507–4516. [PubMed: 15317462]
- (8). 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. [PubMed: 16715409]
- (9). Clemente JC, Moose RE, Hemrajani R, Whitford LRS, Govindasamy L, Reutzel R, McKenna R, Agbandje-McKenna M, Goodenow MM, Dunn BM. Comparing the accumulation of active- and nonactive-site mutations in the HIV-1 protease. *Biochemistry* 2004;43:12141–12151. [PubMed: 15379553]
- (10). Hong L, Zhang XJC, Hartsuck JA, Tang J. Crystal structure of an in vivo HIV-1 protease mutant in complex with saquinavir: Insights into the mechanisms of drug resistance. *Protein Sci* 2000;9:1898–1904. [PubMed: 11106162]
- (11). Kozisek M, Bray J, Rezacova P, Saskova K, Brynda J, Pokorna J, Mammano F, Rulisek L, Konvalinka J. Molecular analysis of the HIV-1 resistance development: Enzymatic activities, crystal structures, and thermodynamics of nelfinavir-resistant HIV protease mutants. *J. Mol. Biol* 2007;374:1005–1016. [PubMed: 17977555]
- (12). Mahalingam B, Wang YF, Boross PI, Tozser J, Louis JM, Harrison RW, Weber IT. Crystal structures of HIV protease V82A and L90M mutants reveal changes in the indinavir-binding site. *Eur. J. Biochem* 2004;271:1516–1524. [PubMed: 15066177]
- (13). Ghosh AK, Sridhar PR, Leshchenko S, Hussain AK, Li JF, Kovalevsky AY, Walters DE, Wedekind JE, Grum-Tokars V, Das D, Koh Y, Maeda K, Gatanaga H, Weber IT, Mitsuya H. Structure-based design of novel HIV-1 protease inhibitors to combat drug resistance. *J. Med. Chem* 2006;49:5252–5261. [PubMed: 16913714]
- (14). Kovalevsky AY, Tie YF, Liu FL, Boross PI, Wang YF, Leshchenko S, Ghosh AK, Harrison RW, Weber IT. Effectiveness of nonpeptide clinical inhibitor TMC-114 on HIV-1 protease with highly drug resistant mutations D30N, I50V, and L90M. *J. Med. Chem* 2006;49:1379–1387. [PubMed: 16480273]
- (15). Tie YF, Kovalevsky AY, Boross P, Wang YF, Ghosh AK, Tozser J, Harrison RW, Weber IT. Atomic resolution crystal structures of HIV-1 protease and mutants V82A and I84V with saquinavir. *Proteins* 2007;67:232–242. [PubMed: 17243183]
- (16). Aqvist J, Medina C, Samuelsson JE. New method for predicting binding-affinity in computer-aided drug design. *Protein Eng* 1994;7:385–391. [PubMed: 8177887]
- (17). Straatsma TP, McCammon JA. Computational alchemy. *Annu. Rev. Phys. Chem* 1992;43:407–435.
- (18). Bren U, Martinek V, Florian J. Free energy simulations of uncatalyzed DNA replication fidelity: structure and stability of T.G and dTTP.G Terminal DNA mismatches flanked by a single dangling nucleotide. *J. Phys. Chem. B* 2006;110:10557–10566. [PubMed: 16722767]
- (19). Perdih A, Bren U, Solmajer T. Binding free energy calculations of N-sulphonyl-glutamic acid inhibitors of MurD ligase. *J. Mol. Model* 2009;15:983–996. [PubMed: 19198900]
- (20). Wang YF, Tie YF, Boross PI, Tozser J, Ghosh AK, Harrison RW, Weber IT. Potent new antiviral compound shows similar inhibition and structural interactions with drug resistant mutants and wild type HIV-1 protease. *J. Med. Chem* 2007;50:4509–4515. [PubMed: 17696515]
- (21). Vega S, Kang LW, Velazquez-Campoy A, Kiso Y, Amzel LM, Freire E. A structural and thermodynamic escape mechanism from a drug resistant mutation of the HIV-1 protease. *Proteins* 2004;55:594–602. [PubMed: 15103623]
- (22). Miller JF, Andrews CW, Brieger M, Furfine ES, Hale MR, Hanlon MH, Hazen RJ, Kaldor I, McLean EW, Reynolds D, Sarmond DM, Spaltenstein A, Tung R, Turner EM, Xu RX, Sherrill RG. Ultra-potent P1 modified arylsulfonamide HIV protease inhibitors: The discovery of GW0385. *Bioorg. Med. Chem. Lett* 2006;16:1788–1794. [PubMed: 16458505]

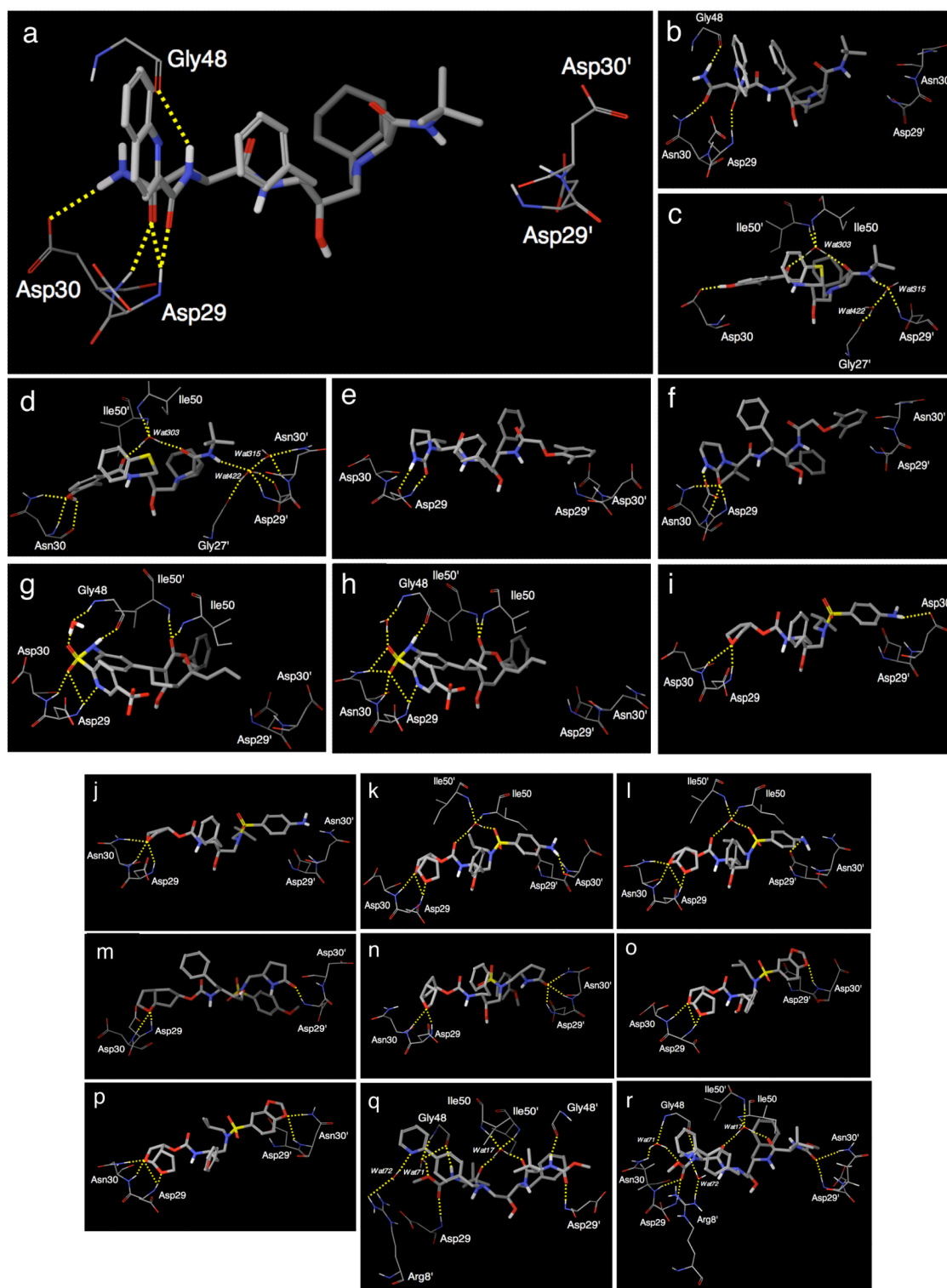


- (23). Krohn A, Redshaw S, Ritchie JC, Graves BJ, Hatada MH. Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere. *J. Med. Chem* 1991;34:3340–3342. [PubMed: 1956054]
- (24). Muzammil S, Armstrong AA, Kang LW, Jakalian A, Bonneau PR, Schmelmer V, Amzel LM, Freire E. Unique thermodynamic response of tipranavir to human immunodeficiency virus type 1 protease drug resistance mutations. *J. Virol* 2007;81:5144–5154. [PubMed: 17360759]
- (25). Kaldor SW, Kalish VJ, Davies JF, Shetty BV, Fritz JE, Appelt K, Burgess JA, Campanale KM, Chirgadze NY, Clawson DK, Dressman BA, Hatch SD, Khalil DA, Kosa MB, Lubbehusen PP, Muesing MA, Patick AK, Reich SH, Su KS, Tatlock JH. Viracept (nelfinavir mesylate, AG1343): A potent, orally bioavailable inhibitor of HIV-1 protease. *J. Med. Chem* 1997;40:3979–3985. [PubMed: 9397180]
- (26). Stoll V, Qin WY, Stewart KD, Jakob C, Park C, Walter K, Simmer RL, Helfrich R, Bussiere D, Kao J, Kempf D, Sham HL, Norbeck DW. X-ray crystallographic structure of ABT-378 (lopinavir) bound to HIV-1 protease. *Bioorg. Med. Chem* 2002;10:2803–2806. [PubMed: 12057670]
- (27). Clemente JC, Coman RM, Thiaville MM, Janka LK, Jeung JA, Nukoolkarn S, Govindasamy L, Agbandje-McKenna M, McKenna R, Leelamanit W, Goodenow MM, Dunn BM. Analysis of HIV-1CRF\_01 A/E protease inhibitor resistance: Structural determinants for maintaining sensitivity and developing resistance to atazanavir. *Biochemistry* 2006;45:5468–5477. [PubMed: 16634628]
- (28). Cho AE, Guallar V, Berne BJ, Friesner R. Importance of accurate charges in molecular docking: Quantum mechanical/molecular mechanical (QM/MM) approach. *J. Comput. Chem* 2005;26:915–931. [PubMed: 15841474]
- (29). Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem* 2004;47:1739–49. [PubMed: 15027865]
- (30). Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem* 2006;49:6177–96. [PubMed: 17034125]
- (31). Yu ZY, Jacobson MP, Friesner RA. What role do surfaces play in GB models? A new-generation of surface-generalized Born model based on a novel Gaussian surface for biomolecules. *J. Comput. Chem* 2006;27:72–89. [PubMed: 16261581]
- (32). Amano M, Koh Y, Das D, Li JF, Leschenko S, Wang YF, Boross PI, Weber IT, Ghosh AK, Mitsuya H. A novel bis-tetrahydrofuranylethane-containing nonpeptidic protease inhibitor (PI), GRL-98065, is potent against multiple-PI-resistant human immunodeficiency virus in vitro. *Antimicrob. Agents Chemother* 2007;51:2143–2155. [PubMed: 17371811]
- (33). Ghosh AK, Leshchenko S, Noetzel M. Stereoselective photochemical 1,3-dioxolane addition to 5-alkoxymethyl-2(5H)-furanone: Synthesis of bis-tetrahydrofuran ligand for HIV protease inhibitor UIC-94017 (TMC-114). *J. Org. Chem* 2004;69:7822–7829. [PubMed: 15527257]
- (34). Ghosh AK, Leshchenko-Yashchuk S, Anderson DD, Baldrige A, Noetzel M, Miller HB, Tie YF, Wang YF, Koh Y, Weber IT, Mitsuya H. Design of HIV-1 Protease Inhibitors with pyrrolidinones and oxazolidinones as novel P1'-Ligands To Enhance Backbone-Binding Interactions with Protease: Synthesis, Biological Evaluation, and Protein-Ligand X-ray Studies. *J. Med. Chem* 2009;52:3902–3914. [PubMed: 19473017]
- (35). Koh Y, Das D, Leschenko S, Nakata H, Ogata-Aoki H, Amano M, Nakayama M, Ghosh AK, Mitsuya H. GRL-02031, a novel nonpeptidic protease inhibitor (PI) containing a stereochemically defined fused cyclopentanyltetrahydrofuran potent against multi-PI-resistant human immunodeficiency virus type 1 In Vitro. *Antimicrob. Agents Chemother* 2009;53:997–1006. [PubMed: 18955518]
- (36). Yoshimura K, Kato R, Yusa K, Kavlick MF, Maroun V, Nguyen A, Mimoto T, Ueno T, Shintani M, Falloon J, Masur H, Hayashi H, Erickson J, Mitsuya H. JE-2147: A dipeptide protease inhibitor (PI) that potently inhibits multi-PI-resistant HIV-1. *Proc. Natl. Acad. Sci. U. S. A* 1999;96:8675–8680. [PubMed: 10411934]

- (37). Koh Y, Nakata H, Maeda K, Ogata H, Bilcer G, Devasamudram T, Kincaid JF, Boross P, Wang YF, Ties YF, Volarath P, Gaddis L, Harrison RW, Weber IT, Ghosh AK, Mitsuya H. Novel bis-tetrahydrofuranylurethane-containing nonpeptidic protease inhibitor (PI) UIC-94017 (TMC114) with potent activity against multi-PI-resistant human immunodeficiency virus in vitro. *Antimicrob. Agents Chemother* 2003;47:3123–3129. [PubMed: 14506019]
- (38). Markowitz, M.; Conant, M.; Hurley, A.; Schluger, R.; Duran, M.; Peterkin, J.; Chapman, S.; Patick, A.; Hendricks, A.; Yuen, GJ.; Hoskins, W.; Clendeninn, N.; Ho, DD. A preliminary evaluation of nelfinavir mesylate, an inhibitor of human immunodeficiency virus (HIV)-1 protease, to treat HIV infection, 1998. Univ Chicago Press; 1998. p. 1533-1540.
- (39). Wlodawer A, Erickson JW. Structure-based inhibitors of HIV-1 protease. *Annu. Rev. Biochem* 1993;62:543–585. [PubMed: 8352596]
- (40). Johnson VA, Brun-Vezinet F, Clotet B, Conway B, Kuritzkes DR, Pillay D, Schapiro J, Telenti A, Richman D. Update of the drug resistance mutations in HIV-1: 2005. *Top HIV Med* 2005;13:51–57. [PubMed: 15849371]
- (41). Baldwin ET, Bhat TN, Liu BS, Pattabiraman N, Erickson JW. Structural basis of drug-resistance for the V82A mutant of HIV-1 proteinase. *Nat. Struct. Biol* 1995;2:244–249. [PubMed: 7773792]
- (42). Maeda K, Yoshimura K, Shibayama S, Habashita H, Tada H, Sagawa K, Miyakawa T, Aoki M, Fukushima D, Mitsuya H. Novel low molecular weight spirodiketopiperazine derivatives potently inhibit R5 HIV-1 infection through their antagonistic effects on CCR5. *J. Biol. Chem* 2001;276:35194–35200. [PubMed: 11454872]
- (43). Kim EE, Baker CT, Dwyer MD, Murcko MA, Rao BG, Tung RD, Navia MA. Crystal-structure of HIV-1 protease in complex with VX-478, A potent and orally bioavailable inhibitor of the enzyme. *J. Am. Chem. Soc* 1995;117:1181–1182.
- (44). Wlodawer A, Vondrasek J. Inhibitors of HIV-1 protease: A major success of structure-assisted drug design. *Annu. Rev. Biophys. Biomol. Struct* 1998;27:249–284. [PubMed: 9646869]
- (45). Thaisrivongs S, Skulnick HI, Turner SR, Strohbach JW, Tommasi RA, Johnson PD, Aristoff PA, Judge TM, Gammill RB, Morris JK, Romines KR, Chrusciel RA, Hinshaw RR, Chong KT, Tarpley WG, Poppe SM, Slade DE, Lynn JC, Horng MM, Tomich PK, Seest EP, Dolak LA, Howe WJ, Howard GM, Schwende FJ, Toth LN, Padbury GE, Wilson GJ, Shiou LH, Zipp GL, Wilkinson KF, Rush BD, Ruwart MJ, Koeplinger KA, Zhao ZY, Cole S, Zaya RM, Kakuk TJ, Janakiraman MN, Watenpaugh KD. Structure-based design of HIV protease inhibitors: Sulfonamide-containing 5,6-dihydro-4-hydroxy-2-pyrones as nonpeptidic inhibitors. *J. Med. Chem* 1996;39:4349–4353. [PubMed: 8893827]
- (46). Turner SR, Strohbach JW, Tommasi RA, Aristoff PA, Johnson PD, Skulnick HI, Dolak LA, Seest EP, Tomich PK, Bohanan MJ, Horng MM, Lynn JC, Chong KT, Hinshaw RR, Watenpaugh KD, Janakiraman MN, Thaisrivongs S. Tipranavir (PNU-140690): A potent, orally bioavailable nonpeptidic HIV protease inhibitor of the 5,6-dihydro-4-hydroxy-2-pyrone sulfonamide class. *J. Med. Chem* 1998;41:3467–3476. [PubMed: 9719600]
- (47). Brown KL, Bren U, Stone MP, Guengerich FP. Inherent stereospecificity in the reaction of aflatoxin B1 8,9-epoxide with deoxyguanosine and efficiency of DNA catalysis. *Chem. Res. Toxicol* 2009;22:913–917. [PubMed: 19301826]
- (48). McGovern SL, Caselli E, Grigorieff N, Shoichet BK. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem* 2002;45:1712–1722. [PubMed: 11931626]
- (49). Shoichet BK. Screening in a spirit haunted world. *Drug Discov. Today* 2006;11:607–615. [PubMed: 16793529]
- (50). Ode H, Matsuyama S, Hata M, Hoshino T, Kakizawa J, Sugiura W. Mechanism of drug resistance due to N88S in CRF01\_AE HIV-1 protease, analyzed by molecular dynamics simulations. *J. Med. Chem* 2007;50:1768–1777. [PubMed: 17367119]
- (51). Ode H, Matsuyama S, Hata M, Neya S, Kakizawa J, Sugiura W, Hoshino T. Computational characterization of structural role of the non-active site mutation M36I of human immunodeficiency virus type 1 protease. *J. Mol. Biol* 2007;370:598–607. [PubMed: 17524421]
- (52). Ode H, Neya S, Hata M, Sugiura W, Hoshino T. Computational simulations of HIV-1 proteases-multi-drug resistance due to nonactive site mutation L90M. *J. Am. Chem. Soc* 2006;128:7887–7895. [PubMed: 16771502]



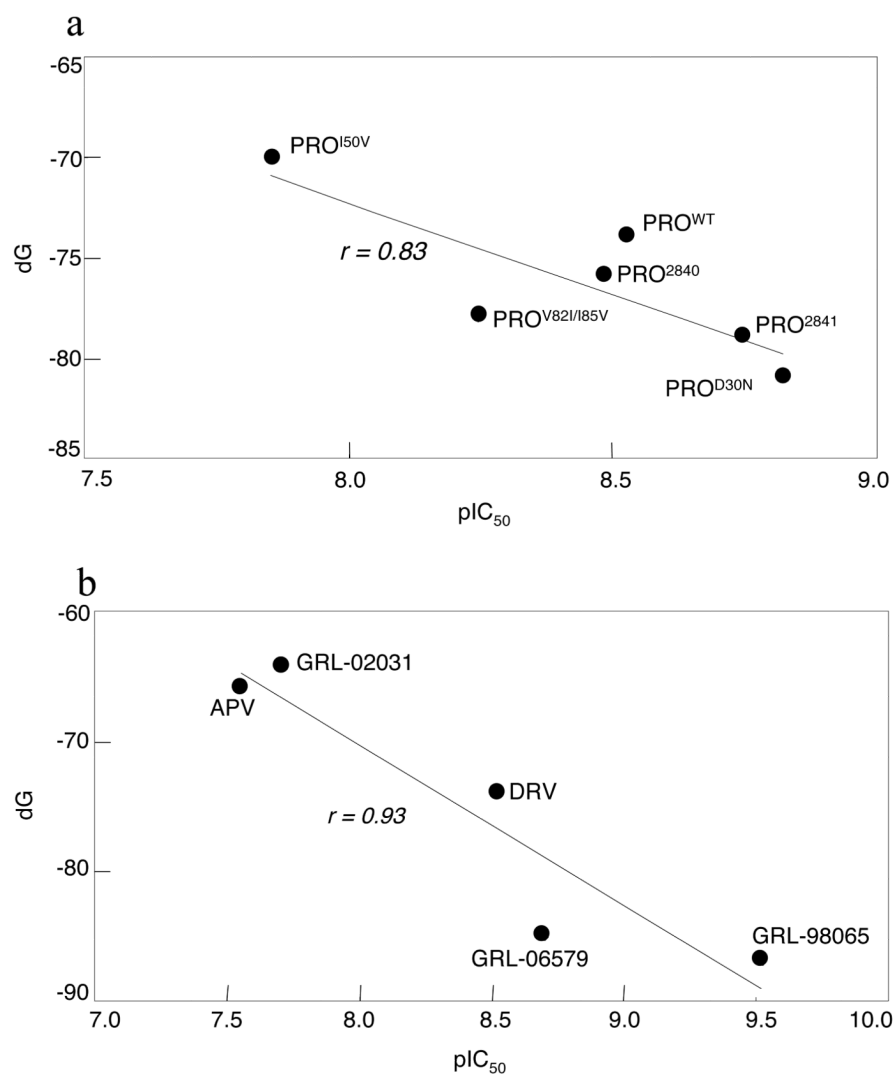
**Figure 1.** Structure of dimerized HIV-1 protease. The monomer subunit is shown in a red or green ribbon. The locations of the mutant residue positions are indicated. Only polar hydrogens are shown and the following atom colors are used in this and all subsequent figures, C: grey; H: white; O: red; N: blue; S: yellow



**Figure 2.** Interactions of protease inhibitors with wild-type and D30N mutant protease. a) The interactions of SQV with wild-type (PRO<sup>WT</sup>) and b) D30N-containing mutant protease (PRO<sup>D30N</sup>) are shown. SQV has hydrogen bond interactions (dotted yellow lines) only in

the S2 site, but no such interactions in the S2' site. The hydrogen bonds of SQV with the mediating water molecule, Asp25, Asp25', and Gly27, are present but not shown. c) The hydrogen bond interactions of NFV with PRO<sup>WT</sup> and d) PRO<sup>D30N</sup>. The bridging water-303 has hydrogen bonds with Ile50, and Ile50', water-315 hydrogen bonds with Asp29', and water-422 hydrogen bonds with Gly27'. Simulations were carried out for NFV with these three waters molecules explicitly present, with one the bridging water molecule, and without any explicit water molecule. Nefinavir has more hydrogen bonds with residues 29 and 30 in PRO<sup>D30N</sup> even though it has a lower antiviral potency against HIV-1<sup>D30N</sup>. e) The hydrogen bonds of LPV with PRO<sup>WT</sup> and f) PRO<sup>D30N</sup>. LPV has more hydrogen bond interactions with the S2 site for the mutant than it has for the PRO<sup>WT</sup>. The crystal structure for LPV 1MUI (PDB ID) does not have waters. g) The hydrogen bond interaction of TPV with PRO<sup>WT</sup> and h) PRO<sup>D30N</sup>. Unlike other inhibitors in this study, TPV directly hydrogen bonds with Ile50, and Ile50', but a bridging water molecule hydrogen bonds with Gly48 in the flap. i) The interactions of APV with PRO<sup>WT</sup> and j) PRO<sup>D30N</sup> are shown. In the S2 site of the protease, the THF ligand of APV has an additional hydrogen bond with Asn30 (PRO<sup>MT</sup>) compared to Asp30 (PRO<sup>WT</sup>). In the S2' site, the aniline nitrogen of APV has a hydrogen bond with the side chain of Asp30 (PRO<sup>WT</sup>) and the hydrogen bond is lost for the Asn30 mutant. k) Structural interaction of DRV with PRO<sup>WT</sup> and l) PRO<sup>D30N</sup> are shown. The *bis*-THF ligand of DRV has more hydrogen bond interactions with PRO<sup>D30N</sup> compared to the PRO<sup>WT</sup>. The water molecule mediating hydrogen bonds between the inhibitor and Ile50 and Ile50' in the flap is shown. The water molecule is present in docking and free energy simulations of APV, GRL-98065, GRL-02031, and SQV, but is not shown. m) The hydrogen bond interactions of GRL-02031 with PRO<sup>WT</sup> and n) PRO<sup>D30N</sup> are shown. o) The hydrogen bond interactions of GRL-98065 with wild-type protease are shown. The *bis*-THF ligand of GRL-98065 has hydrogen bond interactions with backbone atoms of Asp29 and Asp30 in S2 site, and with Asp30' in S2'. p) Besides maintaining the backbone hydrogen bonds with Asp29 and residue-30 and 30' of PRO<sup>D30N</sup>, GRL-98065 forms additional hydrogen bonds with the side chains of Asn30 and Asn30'. q) The hydrogen bond interactions of AZV with PRO<sup>WT</sup> and r) PRO<sup>D30N</sup>.





**Figure 3.** Correlation of free energy and antiviral activity. a) Scatter plot of  $\Delta G_{qm/wat}$  vs  $pIC_{50}$  for DRV for Set-3. b) Scatter plot of  $\Delta G_{qm/wat}$  vs  $pIC_{50}$  (PRO<sup>WT</sup>) for ligands with *bis*-THF related cores. The correlation coefficient is shown in both figures.

**Table 1**  
Anti-viral IC<sub>50</sub> values (μM) of protease inhibitors against wild-type and mutant protease

Inhibitor	HIV <sup>WT</sup> IC <sub>50</sub> (μM)	HIV <sup>D30N</sup> IC <sub>50</sub> (μM)	HIV <sup>150V</sup> IC <sub>50</sub> (μM)	HIV <sup>Y82H/85V</sup> IC <sub>50</sub> (μM)	HIV <sup>284Q</sup> IC <sub>50</sub> (μM)	HIV <sup>284I</sup> IC <sub>50</sub> (μM)
Saquinavir	0.008 ± 0.003 <sup>a</sup>	0.031 ± 0.005 (3.9) <sup>b</sup>	0.034 ± 0.001 (3.9)	0.023 ± 0.01 (2.9)	0.015 ± 0.009 (1.9)	0.004 ± 0.0004 (0.5)
Nelfinavir	0.018 ± 0.002	0.096 ± 0.005 (5.3)	0.11 ± 0.08 (6.1)	0.083 ± 0.071 (4.6)	0.22 ± 0.05 (12.4)	0.031 ± 0.005 (1.7)
Lopinavir	0.018 ± 0.001	0.055 ± 0.011 (3.1)	0.21 ± 0.14 (11.7)	0.037 ± 0.007 (2.1)	0.30 ± 0.009 (16.5)	0.048 ± 0.009 (2.7)
Amprenavir	0.028 ± 0.007	0.0066 ± 0.0049 (0.2)	0.16 ± 0.13 (5.7)	0.058 ± 0.024 (2.1)	0.083 ± 0.018 (3.0)	0.038 ± 0.003 (1.4)
Tipranavir	0.15 ± 0.05	0.0047 ± 0.0002 (0.03)	0.34 ± 0.03 (2.3)	0.28 ± 0.003 (1.9)	0.32 ± 0.01 (2.2)	0.12 ± 0.05 (0.8)
Darunavir	0.003 ± 0.0007	0.0015 ± 0.0013 (0.5)	0.014 ± 0.012 (4.7)	0.0057 ± 0.0024 (1.9)	0.0033 ± 0.0001 (1.1)	0.0018 ± 0.0008 (0.6)
GRL-02031	0.02 ± 0.008	0.0087 ± 0.0067 (0.4)	0.014 ± 0.001 (0.7)	0.036 ± 0.002 (1.8)	0.033 ± 0.0009 (1.7)	0.012 ± 0.001 (0.6)
GRL-98065	0.0003 ± 0.00006	0.00055 ± 0.00035 (1.8)	0.0018 ± 0.001 (6.0)	0.0018 ± 0.0007 (6.0)	0.0004 ± 0.0001 (1.3)	0.00015 ± 0.00007 (0.5)
Atazanavir	0.004 ± 0.0004	0.0016 ± 0.0008 (0.4)	0.0019 ± 0.001 (0.5)	0.0015 ± 0.0001 (0.4)	0.0086 ± 0.0041 (2.2)	0.003 ± 0.001 (0.8)

<sup>a</sup>The IC<sub>50</sub> values shown represent means ± 1 S.D. of assays conducted in duplicate or triplicate on more than three different occasions for the study.

<sup>b</sup>The fold change in activity of the mutants from the wild-type IC<sub>50</sub> are shown within brackets.

**Table 2**

Change in free energy (kcal/mol) associated with D30N mutation<sup>a</sup>

Inhibitor	$\Delta\Delta G_{qm/wat}$ (kcal/mol)	$\Delta\Delta G_{mm/wat}$ (kcal/mol)	$\Delta\Delta G_{qm}$ (kcal/mol)	$\Delta\Delta G_{mm}$ (kcal/mol)	IC <sub>50</sub> fold change <sup>b</sup>
Saquinavir	+4	+4	-1	-2	3.9
Nelfinavir	+7	+9	-11	-9	5.3
Lopinavir	n.d.	n.d.	+17	+21	3.1
Amprenavir	-12	+12	+11	+7	0.2
Tipranavir	-5	-2	-12	-9	0.03

<sup>a</sup> qm subscript refers to the use of protein polarized QM charges on the inhibitor, and mm subscript refer to the use of fixed force-field based charges. The use of a explicit bridging water is shown by the wat subscript. A positive  $\Delta\Delta G$  indicates a decrease in binding affinity while a negative  $\Delta\Delta G$  indicates an increase.

<sup>b</sup> IC<sub>50</sub> fold change greater than 1 indicates a reduction in potency with D30N mutation, while a fold change of less than 1 indicates sensitivity.

Table 3

The correlation coefficient ( $r$ ) of the free energies of binding versus experimental  $IC_{50}$  data<sup>a</sup>

	Set-1 (PRO <sup>WT</sup> , PRO <sup>D30N</sup> , PRO <sup>S0V</sup> )			Set-2 (PRO <sup>WT</sup> , PRO <sup>D30N</sup> , PRO <sup>S0V</sup> , PRO <sup>V82I/I85V</sup> )			Set-3 (PRO <sup>WT</sup> and all mutant proteases)					
	$r_{mm}$	$r_{qm}$	$r_{mm/wat}$	$r_{qm/wat}$	$r_{mm}$	$r_{qm}$	$r_{mm/wat}$	$r_{qm/wat}$	$r_{mm}$	$r_{qm}$	$r_{mm/wat}$	$r_{qm/wat}$
DRV	0.30	<b>0.91</b>	<b>0.93</b>	<b>0.93</b>	0.14	<b>0.65</b>	<b>0.53</b>	<b>0.81</b>	0.29	0.23	0.46	<b>0.83</b>
98065	0.35	0.18	<b>0.99</b>	<b>0.94</b>	0.31	0.40	0.32	<b>0.66</b>	<b>0.53</b>	0.17	0.46	0.48
APV	0.26	<b>0.66</b>	<b>0.57</b>	<b>0.99</b>	0.34	0.23	0.47	<b>0.88</b>	0.31	0.19	0.29	<b>0.60</b>
02031	0.16	<b>0.78</b>	0.24	<b>0.78</b>	0.46	0.48	0.21	0.10	0.61	<b>0.62</b>	0.02	0.19
TPV	<b>0.96</b>	<b>0.70</b>	0.49	<b>0.83</b>	<b>0.91</b>	<b>0.71</b>	0.01	<b>0.68</b>	0.14	0.04	0.04	<b>0.50</b>
NFV	0.21	<b>0.51</b>	<b>0.83</b>	<b>0.77</b>	0.26	0.35	<b>0.71</b>	<b>0.59</b>	0.11	0.31	0.41	0.39
NFV	0.21	<b>0.51</b>	<b>0.62</b> <sup>b</sup>	<b>0.97</b> <sup>b</sup>	0.26	0.35	0.48 <sup>b</sup>	<b>0.92</b> <sup>b</sup>	0.11	0.31	0.18 <sup>b</sup>	<b>0.58</b> <sup>b</sup>
LPV	<b>0.76</b>	<b>0.97</b>	n.d.	n.d.	0.42	<b>0.59</b>	n.d.	n.d.	0.16	<b>0.66</b>	n.d.	n.d.
SQV	0.06	0.45	<b>0.96</b>	<b>0.99</b>	0.11	0.43	<b>0.94</b>	<b>0.87</b>	<b>0.78</b>	<b>0.22</b>	<b>0.57</b>	<b>0.84</b>
AZV	<b>0.76</b>	<b>0.54</b>	0.06	0.16	0.13	0.14	0.35	0.16	0.24	0.17	<b>0.66</b>	<b>0.59</b>
AZV	<b>0.76</b>	<b>0.54</b>	<b>0.91</b> <sup>b</sup>	<b>0.64</b> <sup>b</sup>	0.13	0.14	0.34 <sup>b</sup>	<b>0.53</b> <sup>b</sup>	0.24	0.17	<b>0.62</b> <sup>b</sup>	<b>0.58</b> <sup>b</sup>

<sup>a</sup> qm subscript refers to the use of protein polarized QM charges on the inhibitor, and mm subscript refers to the use of fixed force-field based charges. The use of a explicit bridging water is shown by the wat subscript. A bold font indicates correlation ( $r > 0.5$ ).

<sup>b</sup> calculated with three water molecules that mediate hydrogen bonds between the protease and the inhibitor.

**Table 4**

Experimental free energies<sup>a</sup> of protease inhibitors against PRO<sup>WT</sup>

	DRV	APV	GRL-98065	GRL-02031	GRL-06579
IC <sub>50</sub> (μM)	0.003	0.028	0.0003	0.02	0.002
ΔG <sub>exp</sub> (kcal/mol)	-11.6	-10.3	-13.0	-10.5	-11.9

<sup>a</sup>Experimental free energies determined as  $\Delta G_{\text{exp}} = RT \ln(\text{IC}_{50})$ , where R and T represent the universal gas constant, and temperature, respectively.