ELSEVIER

Can an optimization/scoring procedure in ligand–protein docking be employed to probe drug-resistant mutations in proteins?

Y.Z. Chen, X.L. Gu, and Z.W. Cao

Department of Computational Science, National University of Singapore Singapore

A simple ligand–protein structural optimization and binding evaluation procedure has been routinely used in high-speed ligand–protein docking studies. In this work, we examine whether such an optimization/scoring procedure is useful in indicating possible drug-resistant mutations in proteins. Crystal structures of three wild-type enzymes (HIV-1 protease, HIV-1 reverse transcriptase, and Mycobacterium tuberculosis H37Rv enoyl-ACP reductase) complexed to a variety of inhibitors are studied. Mutations are introduced into these structures by using the molecular modeling software, SYBYL. Structural optimization and scoring of a mutant complex is conducted by a procedure similar to that used in a recent docking study (Wang et al., 1999). The computed results are compared with observed drug resistance data and the profile of nonresistant mutations. Most mutations studied show an energy change in the same direction as those indicated by observed resistance data. 50% of the polar to polar or nonpolar to nonpolar mutations are found to correlate qualitatively with observed drug resistance data. Van der Waals interactions account for most of these changes, which is in agreement with conclusions from structural studies. Substantially larger deviations are found between computed results and observed data for most polar to nonpolar or nonpolar to polar mutations, which result from deficiency in modelling and scoring ligand–protein interactions in our procedure. Our results suggest that an optimization/docking scoring procedure is useful for qualitatively probing polar to polar or nonpolar to nonpolar resistant mutations in addition to its application in screening active compounds. More accurate description of ligand–protein interactions and the use of methods such as free energy perturbation and Poisson-Boltzmann may be needed to further improve the quality of prediction. © 2001 by Elsevier Science Inc.

Keywords: drug–receptor interaction, ligand–protein interaction, molecular mechanics of drug resistant mutation, molecular modeling of resistant mutation

INTRODUCTION

Computer ligand–protein docking has developed into a useful tool in facilitating new drug discover.[1,2] To achieve high-speed docking that involves screening of a database containing a large number of chemical compounds, a simple optimization/scoring procedure has been routinely applied to optimization and scoring of docked ligand–protein structures. Such a procedure is capable of finding ligands and binding conformations at a receptor site close to experimentally determined structures.[3–7]

This capability of distinguishing between binding compounds/conformations and nonbinding compounds/conformations raises an interesting question as to whether such a procedure can be extended to indicate possible drug resistant mutations in proteins. If currently available ligand–protein docking tools can be shown to be useful in indicating drug resistant mutations, then they may be readily extended to facilitate prediction of drug resistance in addition to rational drug design. This can potentially enhance the capability of computer-aided drug design tools.

Methods such as free energy perturbation[8,9] have been used to model and evaluate drug resistant mutations. While more accurate in modeling resistant mutations than a simple optimization/scoring procedure commonly used in ligand-protein docking, these methods have yet to be employed in a high-speed ligand–protein docking study partly due to their CPU-intensive nature. Instead of studying a more accurate method for drug-resistant mutations, this work examines whether an

Corresponding author; Y. Z. Chen, Department of Computational Science, National University of Singapore, Lower Kent Ridge Road, Singapore 119260.

E-mail address: yzchen@cz3.nus.edu.sg (Y.Z. Chen)

Journal of Molecular Graphics and Modelling 19, 560–570, 2001

655 Avenue of the Americas, New York, NY 10010

1093-3263/01/$–see front matter
PII S1093-3263(01)00091-2

optimization/scoring procedure is useful in indicating drug resistant mutations.

A number of drug-resistant mutations occur at receptor sites in direct contact with the binding drug.[10–14] Analysis of crystallographic structures of several receptors[11,12] indicated that a large portion of these mutations alter the tight packing between the binding drug and receptor without substantial change in overall conformation in the complex. Structural comparison between mutant and wild-type complexes showed that, in a majority of these cases, the only significant change is in the pattern of local van der Waals and hydrogen bonding interactions at a mutation site.[12,13] While the hydrophobic effect is important in some cases,[15] variation in local packing interactions is expected to be a major factor in the reduced drug binding affinity leading to resistance[11] for a number of mutations at sites in contact with the binding drug.

These studies suggest that the enthalpic factor likely plays a major role in a number of contact-region-resistant mutations. Therefore, currently used optimization/scoring procedures, which are based on molecular mechanics, are expected to be applicable to this kind of mutation. In this work, we use a standard method to generate mutant structures from wild-type structures and employ an optimization/scoring procedure similar to that of Wang et al.[5] to optimize and evaluate mutant structures. This procedure is selected because it is based on molecular mechanics as opposed to more simplified energy functions used in some of the other docking studies. The methods used in this procedure are similar to some of the methods used in predicting ligand-binding affinity.[16,17] It is noted that some important factors are not well represented in the force fields used in this work as well as in most docking studies. These include interactions with water, hydrophobic effect, and electrostatic interactions involving aromatic rings. This deficiency likely leads to certain degree of error in optimization and scoring. The intention of this work is to evaluate a procedure similar to that used in a docking study, therefore no attempt is made to improve the force fields.

The mutation-induced changes in ligand–protein interaction energy for each of these mutants are compared to observed resistance data to examine whether there is a consistent correlation between them. The results are also compared to that of nonresistant mutations to evaluate whether the modeling procedure is capable of distinguishing between resistant and nonresistant mutations.

Based on the availability of drug-resistance data and X-ray crystallographic structure, several enzymes complexed to a variety of inhibitors were studied. Various single and multiple mutations at the drug-binding site were introduced into these structures by using the molecular-mechanics-based mutation modelling and structure optimization tools provided in the molecular modelling software, SYBYL. A recent study has shown that a molecular-mechanics-based modeling method is useful in generating reasonably good mutant structures from wild-type structures.[18]

The enzymes investigated in this work are HIV-1 protease, HIV-1 reverse transcriptase, and *Mycobacterium tuberculosis* enoyl-ACP reductase. Each of these enzymes is a major therapeutic target of a host of clinical drugs for the treatment of AIDS and tuberculosis. Resistant mutations against these drugs have been observed.[10,19,20] The structure of each of these wild-type enzymes complexed to different inhibitors has been determined by X-ray crystallography.[13,21–27] The resistance profile and structural features of the mutants of these enzymes have been subjects of extensive investigations.[10,11] Therefore these enzymes, with available inhibitor-bound 3D structures, are excellent systems for testing the applicability of an optimization/scoring procedure in indicating drug-resistant mutations in viral and bacterial proteins.

METHODS

Structures

Crystal structures of eight wild-type enzyme-inhibitor complexes are used in this study. These include HIV-1 protease complexed to MK 639 (indinavir), Ro31-8959 (saquinavir), SB 203386, U89360e, and VX478[21–25]; HIV-1 reverse transcriptase complexed to nevirapine and TIBO R82913[26,27]; and *Mycobacterium tuberculosis* H37Rv enoyl-ACP reductase complexed to NADH.[13]

These complexes along with the corresponding PDB identification code, literature and introduced mutations are given in Table 1. The mutations in Table 1 are all located within the ligand-binding site and they are of two main types. The first type is polar to polar or nonpolar to nonpolar substitutions at the binding sites involving relatively small changes in hydrophobic packing. It is expected that these mutations can be modelled and evaluated with relatively fair accuracy by a molecular mechanics-based mutation modelling and scoring procedure. The second type of mutations considered in this work are polar to nonpolar or nonpolar to polar substitutions that might involve substantial changes in hydrophobic packing. No further conformation optimization was conducted on these crystal structures before the introduction of mutations. Only nonhydrogen atoms are considered in our study. The mass and partial charge of hydrogen is added to its donor atom.

The structures of three of the studied ligands MK639, U89360E, and NADH are given in Figure 1, 2, and 3 respectively. Both the ligand and receptor residues at the binding region (within 6 Å of the ligand) are shown. In these structures, hydrogens are added by SYBYL add-hydrogen module to facilitate the examination of protonation state of each ligand after checking atom type and bond type. Only U89360E and NADH are found to have protonated or deprotonated atoms. The protonated or deprotonated atoms in these structures are marked by a + or − sign, respectively.

Modeling of Mutation

A mutation is introduced into the crystal structure of a wild-type ligand–protein complex by the use of biopolymer mutation modelling tool provided in the molecular modelling software SYBYL. The sidechain of a selected amino acid residue is removed and then replaced by that of a new amino acid whose Cα-R bond matches that of the original residue.[28] The conformation of this sidechain is then optimized through torsion angle search of relevant rotatable bonds by the use of the SYBYL Fix Side-chain program to release bad contact between this residue and its neighboring residues and the binding ligand. In this SYBYL-based procedure, standard Tripos force fields are used in which steric, electrostatic, and hydrogen-bond interactions are considered.[28] However, effects associated with aromatic interaction, solvation, and hydrophobic packing are neglected.

Table 1. Enzyme–inhibitor complexes and mutations

Enzyme–inhibitor complex	PDB Id	Reference	Mutation introduced	Mutation type
HIV-1 protease + MK 639	1HSG	23	V82A, V82F, V82I, I84V, V82F/I84V, M46I/L63P, V82L	Nonpolar–Nonpolar
			V82T/I84V, M46I/L63P/V82T/I84V	Nonpolar–Polar
HIV-1 protease + Saquinavir	1HXB	21	V82F, V82I, I84V, G48V, V82F/I84V	Polar–Polar or Nonpolar–Nonpolar
			V82T/I84V	Nonpolar–Polar
HIV-1 protease + SB 203386	1SBG	22	I32V/V47I/I84V	Nonpolar–Nonpolar
HIV-1 protease + U89360E	1GNO	25	V82D, V82N, V82Q, D30F	Nonpolar–Polar or Polar–Nonpolar
HIV-1 protease + VX 478	1HPV	24	V82T/I84V, M46I/L63P/V82T/I84V	Nonpolar–Polar
HIV-1 RT + Nevirapine	1VRT	26	L100I, K103N, V106A, G138L, E138K, Y181C, Y188H	Nonpolar–Nonpolar or Polar–Polar
HIV-1 RT + TIBO R82913	1TVR	27	L100I, K103N, V106A, E138K, Y181C, Y188H	Nonpolar–Nonpolar or Polar–Polar
Mycobacterium tuberculosis H37Rv enoyl-ACP reductase + NADH	1ZID	13	D30F	Polar–Nonpolar

Structure Optimization and Computation of Ligand–Protein Interaction Energy

In a recent molecular docking study,[7] docked structures were subjected to structure refinement including molecular mechanics minimization, conformational sampling at the binding site and a short run of molecular dynamics-based stimulated annealing. Here, a similar scheme is adopted for the refinement of mutant ligand–protein structures. The local conformation of a mutant complex is first optimized by 200 iterations of energy minimization for all residues with at least one atom located within 6 Å of the mutated residue. This is then followed by 10 ps molecular dynamics (MD) simulation at 300 K for the same residues. The Powell algorithm is used for energy minimization. The number of steps of energy minimization used in recent docking studies was generally below 300.[7] In this work, we use a medium value of 200 iterations. This number appears to be sufficient in minimization of local conformation in the mutant complexes studied. Our analysis shows that an increase of the number to 300 does not result in substantial change in the minimized structure. The AMBER force field is used in both energy minimization and molecular dynamics. The MMFF94 charges are assigned to the ligand atoms. The computation was conducted without explicit solvent. A distance-dependent dielectric constant ($\varepsilon = 4r$)[29] is used to mimic solvent effect. A 6 Å cutoff is chosen as that used in a recent docking study[5] that allowed that such a cutoff speeds up the computation significantly while giving reasonably docked structure for a number of ligand–protein complexes. The CPU time for running MD is around 190 s per structural refinement on an SGIR10000 Octane workstation.

The interaction energy between a protein and its binding ligand is computed by the following empirical potential energy function that includes hydrogen bond, non-bonded van der Waals and electrostatic interaction, and atomic solvation free energy terms:

$$V = \sum_{\text{H bonds}}[V_0\,(1 - e^{-a(r-r0)})^2 - V_0] + \sum_{\text{nonbonded}}[A_{ij}/r_{ij}^{12} - B_{ij}/r_{ij}^{6} + q_i q_j/\varepsilon_r r_{ij}] + \sum_{\text{atoms i}} \Delta\sigma_i A_i$$

where r is hydrogen-bond donor-acceptor distance, and V_0, a, and r_0 are hydrogen bond potential parameters; A_{ij} and B_{ij} are nonbonded van der Waals parameters; ε_r is the dielectric constant, q_i and q_j are the partial charges of the i-th and j-th atoms, and r_{ij} is the distance between them; $\Delta\sigma_i$ is atomic solvation parameter and A_i is the solvent-accessible surface area of the i-th atom.

The nonbonded van der Waals and electrostatic terms and their parameters are the same as those used in structure optimization in this work. Many crystal structures contain only nonhydrogen atoms. To avoid the difficulty in modeling the dynamics of hydrogens, the Morse potential,[30] which is a function of donor–acceptor distance, is used to represent hydrogen bond terms. This potential has been shown to give reasonable description of hydrogen-bond energy and dynamics in biomolecules.[31,32] The published Morse potential parameters[31] are used in this work. To partially estimate solvent effect on ligand-protein binding, atomic solvation energy term developed by Eisenberg et al.[33,34] are added to the scoring function. It is noted that this solvation term is only used in the computation of the ligand-protein interaction energy. No solvation effect is considered in the SYBYL-based modeling of mutation and structural optimization.

RESULTS AND DISCUSSION

Comparison Between Modelled and Crystallographic Mutant Structure

To evaluate the quality of our procedure for generating mutant structures, we have compared our modelled structure of HIV-1 protease mutant V82D complexed to inhibitor U89360E with that obtained from crystallographic data.[25] The SYBYL Match

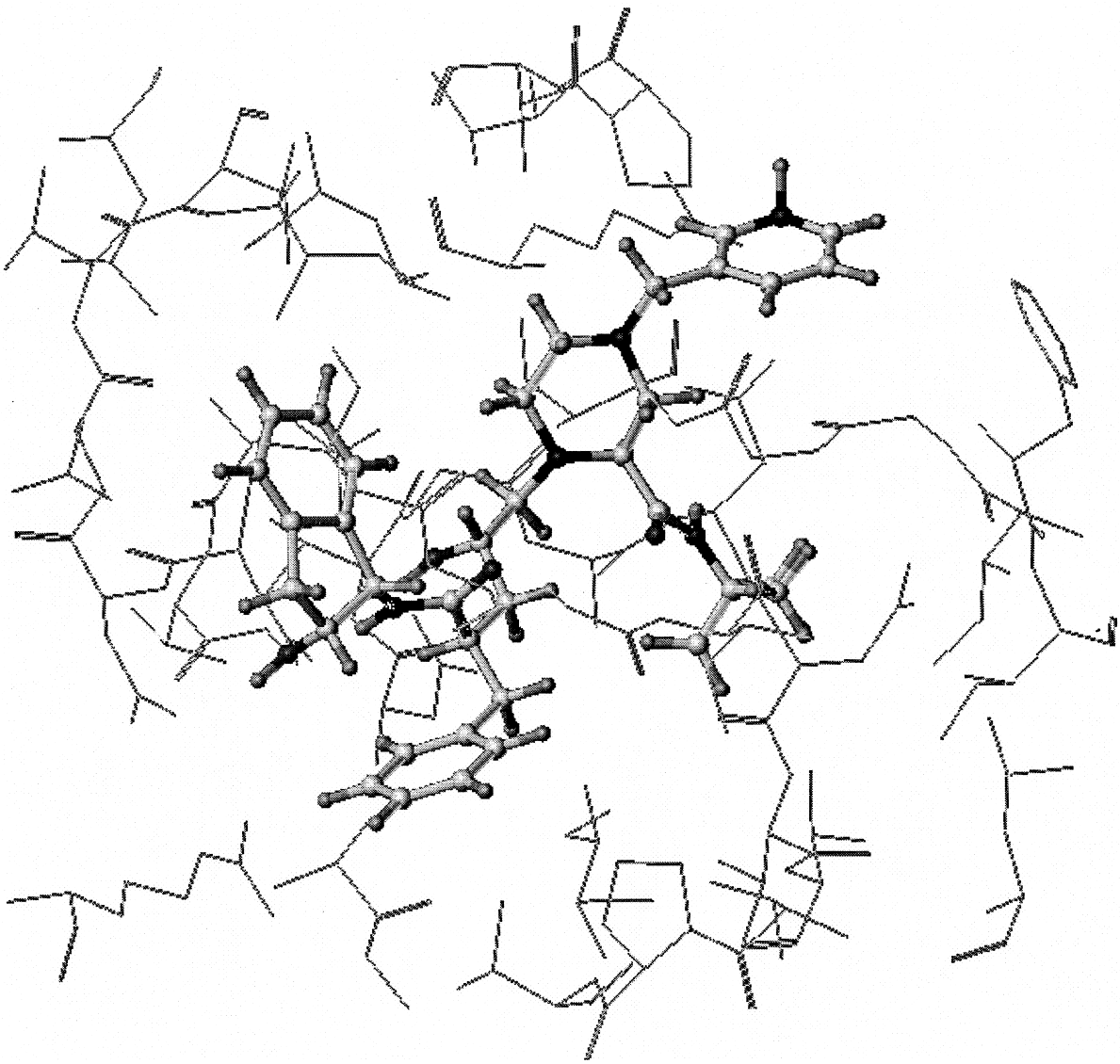

Figure 1. Structure of MK639 (ball-and-stick) and receptor residues (line drawing) within 6 Å of the ligand (PDB identification code 1HSG). Hydrogens are added by using SYBYL add-hydrogen module to facilitate the examination of protonation state. No protonated or deprotonated atom is found in this structure. No histidine residue is found at the binding site.

program, which performs a least squares fit for the two structures, is used to conduct the comparison. Our results show that the modelled and crystallographic structures fit each other fairly well. The RMSD between the two structures is 1.18 Å. For comparison, the RMSD between the wild-type and mutant crystallographic structure is 1.96 Å. These RMSDs are computed for all ligand and protein atoms in the respective ligand–protein structures. Structural deviation mostly occurs within a region covering several residues around the mutation site. The relatively smaller RMSD value seems to indicate that the proposed modelling procedure is capable of generating reasonable mutant structures at least for mutations that primarily induce localized conformation changes.

An analysis shows that the largest structural deviation occurs at both ends of the ligand and the nearby amino-acid residues Pro81 and Asp82. The deviation is 1.25 Å and 1.32 Å, respectively, which is substantially smaller than the RMSD value of 1.96 Å for the crystal structure. In our modelled mutant structure, the aromatic ring end of the ligand appears to be closer to the nearby polar Asp82 residue in chain B of the enzyme than that in the crystal mutant structure. On the other hand, the polar end of the ligand is closer to the nearby pyrrolidine ring of residue Pro81 in chain A of the enzyme. The smaller than observed distances likely result from the deficiency in modelling ligand–protein interactions in the SYBYL mutation tool. In this tool, the electrostatic interactions are modelled by a simplified charge distribution. Interaction with structured waters around protein is also ignored. Moreover, no account is given to the induced electrostatic interactions between aromatic rings and other atoms. These interactions have been shown to affect ligand–protein binding and thus cannot be ignored.[35] In general, it is difficult to say which particular factor contributes most to the discrepancy between computation and observed structure. It is likely that multiple factors are involved.

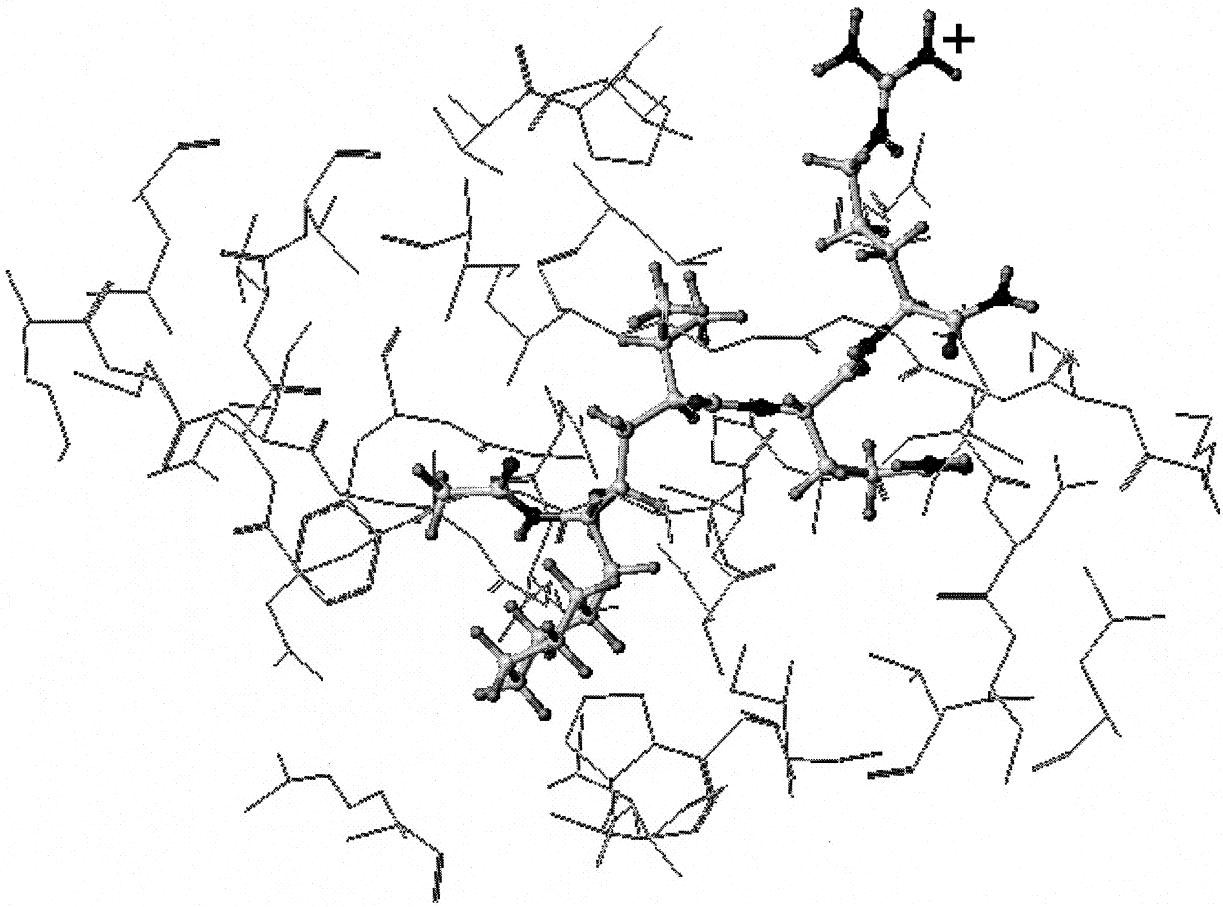

Figure 2. Structure of U89360E (ball-and-stick) and receptor residues (line drawing) within 6 Å of the ligand (PDB identification code 1GNO). The protonated atom N8 is marked by a + sign. No histidine residue is found at the binding site.

HIV-1 Protease

A variety of single and double mutations are introduced at the respective binding site of HIV-1 protease complexed to each of three different inhibitors. These inhibitors and mutations are selected because of the availability of crystallographic structure for the wild-type inhibitor-enzyme complexes[21–25] and the corresponding drug resistance data.[14,19,36] All these inhibitors bind to approximately the same site (the substrate-binding pocket) of the enzyme and they inhibit the viral protease competitively against the substrates.

Table 2 gives the computed change in inhibitor–enzyme interaction energy ΔE between mutant and wild-type complex for different inhibitors and mutations. Only the polar to polar and nonpolar to nonpolar mutations are included. The polar to nonpolar and nonpolar to polar mutations will be discussed separately in this article. The leading interactions that contribute to ΔE are also given in Table 2. For comparison, the logarithm of the ratio of observed inhibitor dissociation constant between mutant and-wild-type enzyme $RT\ln(K'_i/K_i)$ for all the mutants studied are also included. Here K'_i and K_i are observed dissociation constant for a mutant and the corresponding wild-type enzyme respectively.[14,19,36] R is the gas constant and T is the temperature (taken as 300 K in this study). It is noted that $RT\ln(K'_i/K_i)$ is associated with mutation-induced free energy change, ΔG. Because of the predominantly enthalpy-driven nature of the majority of binding-site-mutations studied in this work, our computed ΔE can be roughly compared with $RT\ln(K'_i/K_i)$.

Our computed ΔEs show some degree of correlation with observed resistance data. It can be seen from Table 2 that, for the majority of the mutants studied, their respective ΔE changes in the same direction as $RT\ln(K'_i/K_i)$. All the ΔEs are found to be positive, consistent with the experimental finding that nearly all of these mutants exhibit reduced binding affinity for their inhibitor.[14,19,34] As shown in Table 2, the reduced ΔE primarily results from reduced van der Waals contacts. This is consistent with results from earlier structural analysis.[12–14]

Figure 4 gives the plot of ΔE as a function of $RT\ln(K'_i/K_i)$ for the three enzyme–inhibitor complexes investigated in our study. With the exception of a few mutations, ΔEs for majority of mutants are found to change in the same direction as observed $RT\ln(K'_i/K_i)$s. 50% of them show qualitative correlation. Also included in Figure 4 are results for nonresistant mutations along with experimental data. The computed ΔE of two out of three mutants are-in-the same direction as observed $RT\ln(K'_i/K_i)$s. Overall, the ΔE versus $RT\ln(K'_i/K_i)$ curve modelled as a linear equation has a Pearson coefficient of 0.4, which shows that the adopted optimization/scoring procedure only captures some of the predominant features of mutation-induced change in ligand–protein binding. It is noted that by

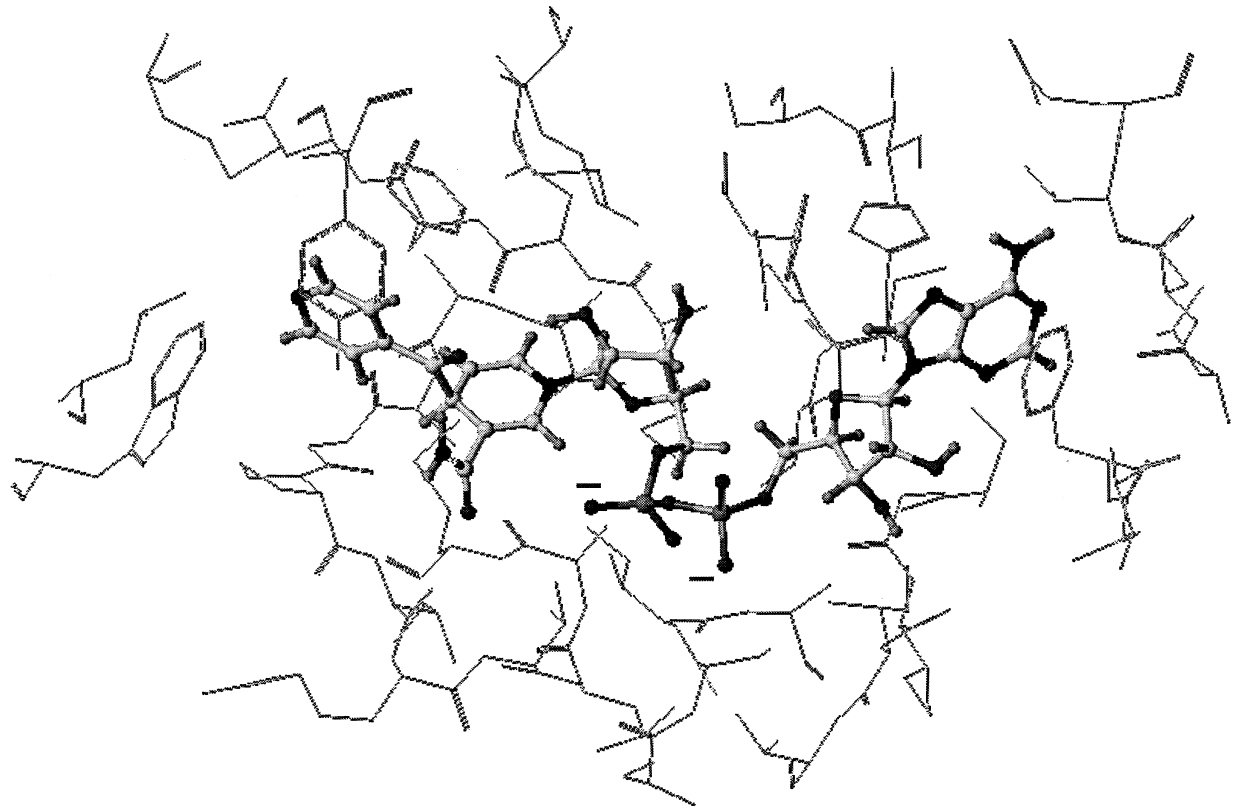

Figure 3. Structure of NADH (ball-and-stick) and receptor residues (line drawing) within 6 Å of the ligand (PDB identification code 1ZID). The two deprotonated atoms O2 are marked by a − sign. No histidine residue is found at the binding site.

eliminating two extreme outliers, V82F of MK 639 and I32V/V47I/I82V of SB 203386, the Pearson coefficient is brought to 0.7. Elimination of all the outliers leads to a coefficient of 0.88. However, 1/3 of the data is eliminated. Therefore, some aspects of the modelling procedure such as force fields and modeling of solvation need to be further improved so as to get a more quantitative agreement with experiments.

While the computed energy profile for the majority of mutants roughly shows the same trend as that from observations, there are a few exceptions, which all seem to be related to the

Table 2. Comparison of calculated change in ligand–protein interaction energy ΔE with the observed logarithm of the ratio of inhibitor dissociation constant $RT\ln(K'_i/K_i)$ between mutant and wild-type HIV-1 protease complexed with three different inhibitors. K'_i and K_i are the observed inhibitor dissociation constant for mutant and wild-type complexes respectively.[12,14,30] R is the gas constant and T the temperature chosen as 300 K. Only polar to polar or nonpolar to nonpolar mutations are included here

Inhibitor	Mutation	ΔE (kcal/mol)	$RT\ln(K'_i/K_i)$ (kcal/mol)	Leading contribution to ΔE
MK 639	V82A	3.04	1.11	van der Waals
	V82F	5.87	−0.46	van der Waals
	V82I	2.11	1.01	van der Waals
	I84V	5.12	0.56	van der Waals
	V82L	−0.10	−0.21	van der Waals
	V82F/I84V	5.84	2.62	van der Waals
Saquinavir	V82F	4.62	0.72	van der Waals
	V82I	3.01	1.19	van der Waals
	I84V	1.71	1.42	van der Waals
	G48V	−2.59	−2.80	van der Waals
	V82F/I84V	3.75	1.76	van der Waals
	V82T/I84V	5.15	2.22	van der Waals
SB 203386	I32V/V47I/I84V	7.13	1.10	solvation

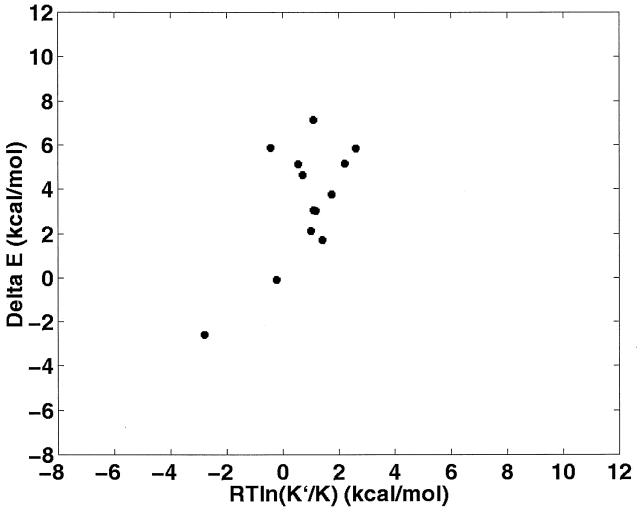

Figure 4. Computed binding energy change ΔE as a function of observed RTln(K_i'/K_i) between wild-type and mutant HIV protease and inhibitors. Experimental data are from other studies.[12,15,30] R is the gas constant and T the temperature chosen as 300 K.

deficiency in modelling and scoring ligand–protein interactions in our SYBYL-based procedure. For instance, in the MK 639 complex, ΔE for the mutant V82F structure is found to be 5.87 Kcal/mol, which indicates a reduced binding affinity upon the mutation. In contrast, the experimental RTln(K'_i/K_i) value of −0.46 kcal/mol implies a slightly increased binding affinity upon the mutation. It has been reported that, in proteins, aromatic rings in close contact tend to form stacking interactions.[37] Visual inspection of our modelled mutant structure in Figure 5 suggests that the aromatic benzyl ring of the mutated residue Phe82 should have a more favored stacking interaction with the nearby benzyl ring of MK 639 than that of the wild-type residue Va182. However, the SYBYL mutation and conformation optimization tools adopted in this work give a larger separation between the benzyl rings in our modelled structure. This also explains why, in the V82F mutant of HIV-1 protease inhibitor complexed to DMP323,[12] the Phe82 benzyl ring in the crystal structure is closer to the hydrophobic P1 group of DMP323 than that in the modelled structure.

There are also cases that likely relate to the modeling deficiencies other than aromatic interactions. It is noted that the computed ΔE for the double mutation V82F/I84V (5.84 kcal/mol) is similar to that of the single mutation V82F (5.87 kcal/mol). In contrast, the observed RTln(K'_i/K_i) for the double mutation is substantially higher than that of the single mutation. In our modelled mutant structure, Va184 is further away from the binding ligand than Phe82 and its sidechain is constrained by the neighboring Phe82. Hence the contribution of the I84V mutation to ΔE is substantially smaller than that of the V82F mutation. As a result, ΔE of the double mutation is primarily determined by V82F and it is approximately the same as that of the single mutation. Therefore, the discrepancy between computation and experiment indicates problems other that aromatic interactions in the modelling procedure.

HIV-1 Reverse Transcriptase

Compared with HIV-1 protease, HIV-1 reverse transcriptase is a substantially larger protein with 967 residues, whose inhibitor binding sites are not necessarily at the active site. All the inhibitors considered in this work bind to a region outside of the active sites, and all the mutations are located in the binding region. These inhibitors and mutations are selected because of the availability of wild-type crystal structures.[26,27] and drug resistance data.[39,40] For HIV-1 reverse transcriptase, the published resistance data is given in terms of IC_{50}, which measures antiviral inhibitory activity of the studied drugs. It has been pointed out that, under certain conditions, IC_{50} can be considered to correlate with K_i and thus free energy change.[40] However, strictly speaking, changes in IC_{50} values may not be directly compared with the calculated ΔEs for the reverse transcriptase complexes discussed here. The mutations are in the active site, which means that substrate binding might be affected. While such effects are included implicitly in measured IC_{50}s, no provision has been made for them in the interaction energetics in our optimization/scoring procedure. Therefore, the comparison of computed ΔEs with observed IC_{50} values here should be regarded as a very rough evaluation of the computed results.

Table 3 gives the computed change in inhibitor−enzyme interaction energy ΔE between mutant and wild-type complex for two inhibitors, nevirapine and TIBO R82913, and various mutants. Figure 6 shows ΔE as a function of RTln(IC_{50}'/IC_{50}) value. The computed ΔEs for most mutants show qualitative correlation with observed RTln(IC_{50}'/IC_{50}) values. The degree of correlation is somehow less apparent as that in the HIV-1 protease-inhibitor complexes as shown in Figure 4. Overall, the ΔE versus RTln(IC_{50}'/IC_{50}) curve modeled as a linear equation has a Pearson coefficient of 0.3. No qualitative discrepancy between computed ΔE and observed RTln(IC_{50}'/IC_{50}) is found.

Mycobacterium tuberculosis H37Rv Enoyl Reductase

The computed ΔE for the S94A mutant in *Mycobacterium tuberculosis* H37Rv enoyl reductase-complexed to NADH is 20.96 kcal/mol, which is compared with observed RTln(K_i'/K_i) value of 3.0 ∼ 4.8 kcal/mol.[10] The substantially larger energy change arises because of sizable change in the computed conformation of both ligand and its receptor at the binding site. This conformation change substantially alters van der Waals and hydrogen-bond interactions between the ligand and its receptor. The location of the S94A mutation is within the ligand-binding region. An analysis of X-ray crystallographic structure indicates that the S94A mutant displays a reduced hydrogen-bond pattern between NADH and the enzyme, resulting in an increased Ki for NADH.[13] This is consistent with our computation that shows an increase of 10.3 kcal/mol in hydrogen-bond energy. However, the substantially large discrepancy between computed and observed energetics seems to indicate that some of the important interactions may not be properly accounted for in the optimization/scoring procedure. With two protonated oxygen atoms, NADH is a highly charged and heavily hydrated molecule in which structured water is almost certainly involved in binding. These and other factors such as hydrophobic packing need to be taken into consideration in the modelling procedure.

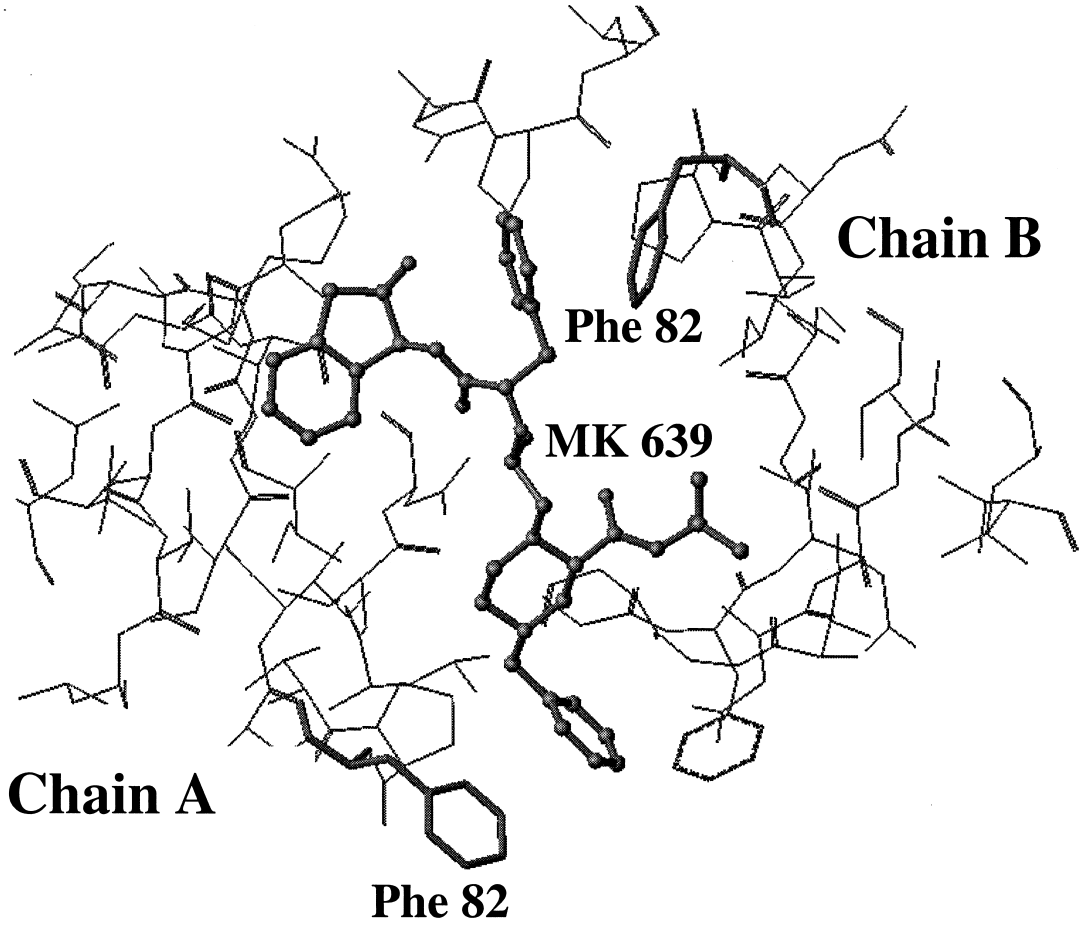

Figure 5. Modeled V82F mutant structure of HIV-1 protease complexed with MK 639. The ligand is represented by ball-and-stick, the mutant residue Phe82 by stick, and the rest of the protein is represented by line drawing.

Leading Factors for Mutation-Induced Change in Ligand–Protein Interaction Energy

Four energy terms are included in computation of mutation-induced change in ligand–protein interaction energy ΔE. These are van der Waals interaction, electrostatic interaction, hydrogen bonding, and solvation free energy terms. As shown in Table 1 to Table 3, van der Waals interaction is found to be the leading factor in computed ΔE for majority of polar to polar and nonpolar to nonpolar mutations studied. This is consistent with a recent study of binding stability for a number of TIBO inhibitors of HIV-1 reverse transcriptase using Monte Carlo simulations in a linear response method.[40]

However, there are cases where other factors appear to be dominant in computed ΔE. For instance, solvation free energy change is the leading term in the computed ΔE of the I32V/V47I/I84V mutation in HIV-1 protease complexed with SB 203386. This arises because the mutation involves changes in the size of hydrophobic residues that likely affects solvation behavior of the molecule. In the case of HIV-1 reverse transcriptase, solvation free energy term is also significant in mutant K103N for both nevirapine and TIBO R82913. For the S94A mutation in tuberculosis H37Rv enoyl-reductase-NADH complex, electrostatic interaction is the predominant factor for reduced ligand–protein interaction.

Comparison with Nonresistant Mutations

To examine whether our modelling procedure can distinguish between resistant and nonresistant mutations, we have constructed three HIV-1 protease models with nonresistant mutations: V82F and V82L mutant complexed to inhibitor MK639, and G48A mutant complexed to saquinavir. The computed energy change along with observed data is shown in Figure 4. Results for V82F mutant differs from experimental data and it has been discussed earlier in this paper. The V82L and G48V mutants have not been found to confer resistance to the protease.[14,19,36] Both leucine and valine are nonpolar amino acids with similar sidechains, except that the sidechain of leucine contains one more carbon atom than valine. Similarly, alanine has an additional methyl group than glycine. The additional methyl group in the sidechain of these two mutant residues likely incurs a slightly favorable change in binding stability compared with wild-type protease because of greater van der Waals contact. Our calculation shows that the ΔE of V82L with MK639 and G48A with saquinavir are −0.10 and −2.59 kcal/mol, respectively, which is consistent with the structural characteristics of these two nonresistant mutations.

Two additional nonresistant mutants are constructed for the nevirapine- HIV-1 reverse transcriptase complex. These are G138L and E138K. As shown in Figure 6, fair agreement

Table 3. Comparison of calculated change in ligand–protein interaction energy ΔE with observed logarithm of inhibition concentration ratio RTln(IC'_{50}/IC_{50}) between mutant and wild-type HIV-1 reverse transcriptase complexed with two different inhibitors. IC'_{50} and IC_{50} are observed mutant and wild-type inhibition concentration respectively.[32,33] R is the gas constant and T the temperature chosen as 300 K. Only polar to polar or nonpolar to nonpolar mutations are included here

Inhibitor	Mutation	ΔE (kcal/mol)	RTln(IC'_{50}/IC_{50}) (kcal/mol)	Leading contribution to ΔE
Nevirapine	L100I	1.57	1.29	van der Waals
	K103N	1.36	2.56	solvation
	V106A	0.26	2.74	van der Waals and solvation
	E138K	0.44	0.00	solvation
	G138L	−0.14	0.00	van der Waals
	Y181C	10.4	3.21	van der Waals
	Y188H	0.83	3.15	van der Waals
TIBO R82913	L100I	4.73	2.88	van der Waals and solvation
	K103N	1.90	2.49	solvation
	V106A	6.25	1.16	van der Waals
	E138K	6.59	1.63	van der Waals
	Y181C	2.02	1.59	van der Waals
	Y188H	5.71	1.94	van der Waals

between computation and experiment is found for the G138L mutant. However, the E138K mutant shows slight discrepancy between computed ΔE and observed resistant data. Computed ΔE indicates a slight decrease in binding interaction. On the other hand, the observed resistant data shows essentially no change due to mutation. This mutation involves changes from a negatively charged residue to a positively charged residue. Better modelling of solvent, such as explicit solvent, might give rise to a better result. This study is only focused on the method commonly used in molecular docking studies. Therefore, consideration of explicit solvent will be left to future investigation.

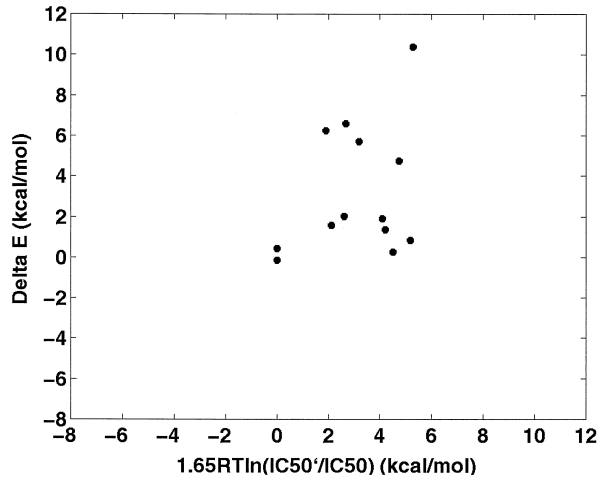

Figure 6. Computed binding energy change ΔE as a function of observed RTln(IC'_{50}/IC_{50}) of mutant HIV-1 Reverse Transcriptase and inhibitors. Experimental data are from other studies.[32,33] R is the gas constant and T the temperature chosen as 300 K.

Effect of Hydrophobic Packing on Polar to Nonpolar and Nonpolar to Polar Mutations

Simulations have shown that proteins fold so that the maximum number of nonpolar contacts are formed.[41] Such hydrophobic packing can be affected by a polar to nonpolar or a nonpolar to polar mutation. In the SYBYL mutation modelling procedure adopted in this work, the effect associated with hydrophobic packing is neglected in generating a mutant structure,[28] which might result in a discrepancy between computation and observations in some cases. We have carried out a study of mutations involving polar to nonpolar or nonpolar to polar substitutions in HIV-1 protease complexed to inhibitors U89360E, MK 639, saquinavir, and VX 478, respectively. Table 4 gives corresponding ΔE and observed RTln(K'_i/K_i)[36,38,42] for these mutants. Our modelling results-show substantial deviations between computed ΔE and the corresponding RTln(K'_i/K_i) value, even though they are generally in the same direction. The average difference between ΔE and RTln(K'_i/K_i) for these mutations is 7.70 kcal/mol, which is considerably larger than that of 2.57kcal/mol for HIV-1 protease mutants not involving polar to nonpolar or nonpolar to polar mutations. This seems to suggest that hydrophobic packing may need to be considered in the modelling of resistant mutations involving changes from polar to nonpolar or nonpolar to polar residues.

CONCLUSIONS

The certain degree of correlation between computed energy change and observed drug resistance data for most polar to polar and nonpolar to nonpolar mutants studied in this work indicates the potential of a ligand–protein docking optimization/scoring procedure in predicting possible drug resistance for this kind of mutation. Several improvements are needed to achieve a more accurate prediction. Proper modelling of aromatic interactions, interaction with waters, hydrophobic packing, and further refinement of currently used optimization/

Table 4. Comparison of computed energy change ΔE with observed logarithm of the ratio of inhibitor dissociation constant RTln(K′i/Ki) between mutant and wild-type HIV-1 protease for mutations involving polar to non-polar or non-polar to polar substitutions. The experimental inhibition constants are from other studies.[30,31,34] R is the gas constant and T the temperature chosen as 300 K

Inhibitor	Mutation	ΔE (kcal/mol)	$RT\ln(K'_i/K_i)$ (kcal/mol)
U89360E	V82D	15.3	2.00
	V82N	13.6	2.79
	V82Q	13.0	0.97
	D30F	5.96	3.73
MK 639	V82T/I84V	9.08	2.49
	M46I/L63P, V82T/I84V	8.16	2.39
Saquinavir	V82T/I84V	5.15	2.22
VX 478	V82T/I84V	12.9	1.72
	M46I/L63P, V82T/I84V	10.4	1.39

scoring force field parameters may help improve the quality of modelled mutant structure and the accuracy of ligand–protein interaction energy. Moreover, given that computer speed is increasing rapidly, more sophisticated methods, such as free energy perturbation and Poisson-Boltzmann, may be introduced into a docking optimization/scoring procedure to better model docked structures and mutation-induced changes.

REFERENCES

1 Kuntz, I.D. Structure-based strategies for drug design and discovery. *Science* 1992, **257,** 1078–1082

2 Blundell, T. L. Structure based drug design. *Nature* 1996, **384** suppl., 23–26

3 Rarey M., Kramer B., Lengauer T., and Klebe G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol*. 1996, **261**, 470–489

4 Jones G., Willet P., Glen R.C., Leach A.R., and Taylor R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol*. 1997, **267**, 727–748

5 Wang J., Kollmann P.A., and Kuntz, I. D. Flexible ligand-docking: A multistep strategy approach. *Proteins* 1999, **36**, 1–19

6 Charifson P.S, Corkery J.J, Murcko M.A., and Walters W.P. *J. Med. Chem.* 1999, **42**, 5100–5109

7 Muegge I., and Martin Y.C. A general and fast scoring function for protein–ligand interactions: A simplified potential approach. *J. Med. Chem.* 1999, **42**, 791–804

8 Wade, R.C., and McCammon J. A. Binding of an antiviral agent to a sensitive and a resistant human rhinovirus. Computer simulation studies with sampling of amino acid side-chain conformations. II. Calculation of free-energy differences by thermodynamic integration. *J. Mol. Biol.* 1992, **225**, 697–712

9 Frecer V., Miertus S., Tossi A., and Romeo, D. Rational design of inhibitors for drug-resistant HIV-1 aspartic protease mutants. *Drug Des. Discov.* 1998, **15**, 211–231

10 Blanchard, J.S. Molecular mechanisms of drug resistance in mycobacterium tuberculosis. *Annu. Rev. Biochem.* 1996, **65**, 215–239.

11 Erickson, J.W., and Burt, S.K. Structural mechanisms of HIV drug resistance. *Annu. Rev. Pharmacol. Toxicol.* 1996, **36,** 545–571

12 Ala, P.J., Huston, E.E., Klade, R.M., McCabe, D.D., Duke, J.L., Rizzo, C.J., Korant, B.D., Deloskey, R.J., Lam, P.Y.S., Hodge, C.N., and Chang, C.-H. Molecular basis of HIV-1 protease drug resistance: Structural analysis of mutant proteases complexed with cyclic urea inhibitors. *Biochemistry* 1997, **36**, 1573–1580

13 Rozwarski, D.A., Grant, G.A., Barton, D.H.R., Jacobs Jr., W.R., and Sacchettini, J.C. Modification of the NADH of the isoniazid target (InhA) from mycobacterium tuberculosis. *Science* 1998, **279**, 98–102

14 Klabe, R.M., Bacheler, L.T., Ala, P.J., Erickson-Vitanen, S., and Meek, J.L. Resistance to HIV protease inhibitors: A comparison of enzyme inhibition and antiviral potency. *Biochemistry* 1998, **37**, 8735–8742

15 Sussman F., Villaverde M.C., Davis A. Solvation effects are responsible for the reduced inhibitor affinity of some HIV-1 PR mutants. *Protein Sci.* 1997, **6**, 1024–1030

16 Aqvist J., Medina C., and Samuelsson J.E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* 1994, **7**, 385–391.

17 Jones-Hertzog D.K., and Jorgensen W.L. Binding affinities for sulfonamide inhibitors with human thrombin using Monte Carlo simulations with a linear response method. *J. Med. Chem.* 1997, **40**, 1539–1549

18 Simpson, M.A., and Bernlohr, D.A. Analysis of a series of phenylalanine 57 mutants of the adipocyte lipid-binding protein. *Biochemistry* 1998, **37**, 10980–10986

19 Roberts, N.A., Craig, J.C., and Sheldon, J. Resistance and cross-resistance with saquinavir and other HIV protease inhibitors: theory and practice. *AIDS* 1998, **12**, 453–460

20 De Clercq, E. HIV resistance to reverse transcriptase inhibitors. *Biochem. Pharmacol.* 1994, **47**, 155–169

21 Krohn, A., Redshaw, S., Ritchie, J.C., Graves, B.J., and Hatada, M.H. Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere. *J. Med. Chem.* 1991, **34**, 3340–3342

22 Abdel-Meguid, S.S., Metcalf, B.W., Carr, T.J., Demarsh, P., DesJarlais, R.L., Fisher, S., Green, D.W., Ivanoff, L., Lambert, D.M., and Murthy, K.H. An orally bioavailable HIV-1 protease inhibitor containing an imidazole-derived peptide bond replacement: Crystallographic and pharmacokinetic analysis. *Biochemistry* 1994, **33**, 11671–11677

23 Chen, Z., Li, Y., Chen, E., Hall, D.L., Darke, P.L., Culberson, C., Shafer, J.A., and Kuo, L.C. Crystal structure at 1.9-A resolution of human immunodeficiency virus (HIV) II protease complexed with L-735,524, an orally bioavailable inhibitor of the HIV proteases. *J. Biol. Chem.* 1994, **271**, 26344–26359
24 Kim, E.E., Baker, C.T., Dwyer, M.D., Murcko, M.A., Rao, B.G., Tung, R.D., and Navia, M.A. 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
25 Hong, L., Treharne, A., Hartsuck, J. A., Foundling, S., Tang, J. Crystal structures of complexes of a peptidic inhibitor with wild-type and two mutant HIV-1 proteases. *Biochemistry* 1996, **35**, 10627–10633
26 Ren, J., Esnouf, R., Garman, E., Somers, D., Kirby, C.R.I., Keeling, J., Darby, G., Jones, Y., Stuart, D.I., and Stammers, D. High resolution structures of HIV-1 RT from four RT-inhibitor complexes. *Nat. Struct. Biol.* 1995, **2**, 293–302
27 Das, K., Ding, J., Hsiou, Y., Clark, A.D. Jr., Moereels, H., Koymans, L., Andries, K., Pauwels, R., Janssen, P.A., Boyer, P.L., Clark, P., Smith, R.H. Jr., Kroeger Smith, M.B., Michejda, C.J., Hughes, S.H., and Arnold, E. Crystal structures of 8-C1 and 9-C1 TIBO complexed with wild-type HIV-1 RT and 8-CI TIBO complexed with the Tyr181Cys HIV-1 RT drug-resistant mutant. *J. Mol. Biol.* 1996, **264**, 1085–1100
28 SYBYL Biopolymer modeling manual. Version 6.5. 1998. Tripos Inc. http://www.tripos.com
29 McCammon, J.A., and Karplus, M. Internal motions of antibody molecules. *Nature* 1977, **268**, 765–766
30 Baird, N.C. Simulation of hydrogen bonding in biological systems: Ab initio calculations for NH_3-NH_3 and NH_3-NH_4^+. *Int. J. Quantum Chem. Symp.* 1974, **1**, 49–53
31 Chen, Y.Z., Zhuang, W., and Prohofsky, E. W. Premelting thermal fluctuational interbase hydrogen-bond disrupted states of a B-DNA guanine-cytosine base pair: significance for amino and imino proton exchange. *Biopolymers* 1991, **31**, 1273–1281
32 Chen, Y.Z., and Prohofsky, E.W. The role of a minor groove spine of hydration in stabilizing poly-(dA).poly(dT) against fluctuational interbase H-bond disruption in the premelting temperature regime. *Nucleic. Acids. Res.* 1992, **20**, 415–419
33 Eisenberg, D., and McLachlan, D. Solvation energy in protein folding and binding. *Nature* 1986, **319**, 199–203
34 Wesson, L., and Eisenberg, D. Atomic solvation parameters applied to molecular dynamics of proteins in solution *Protein Sci.* 1992, **1**, 227–235
35 Dougherty D.A. Cation-π interactions in chemistry and biology: A new view of benzene, Phe, Tyr, and Trp. *Science* 1996, **271**, 163–168
36 Schock, H.B., Garsky, V.M., and Kuo, L.C. Mutational anatomy of an HIV-1 protease variant conferring cross-resistance to protease inhibitors in clinical trials. *J. Bio. Chem.* 1996, **271**, 31947–31963
37 McGaughey, G. B., Gagne, M., and Rappe, A.K. π-stacking interactions. Alive and well in proteins. *J. Biol. Chem.* 1998, **273**, 15458–15463
38 Bardi J. S., Luque I., and Freire E. Structure-based thermodynamic analysis of HIV-1 protease inhibitors. *Biochemistry* 1997, **36**, 6588–6596
39 Smith R.H. Jr, Jorgensen W.L., Tirado-Rives J., Lamb M.L., Janssen P.A., Michejda C.J., and Kroeger Smith, M.B. Prediction of binding affinities for TIBO inhibitors of HIV-1 reverse transcriptase using Monte Carlo simulations in a linear response method. *J. Med. Chem.* 1998, **41**, 5272–5286
40 Tute, M.S. Drug design: The present and the future. *Advances Drug Res.* 1995, **26**, 48–50
41 Dill, K.A. Dominant forces in protein folding. *Biochemistry* 1990, **31**, 7133–7155
42 Lin Y.Z., Lin X.L., Hong L., Foundling S., Heinrikson R.L., Thaisrivongs S., Leelamanit W., Raterman D., Shah M., Dunn B.M., and Tang J. Effect of point mutations on the kinetics and the inhibition of human immunodeficiency virus type 1 protease: Relationship to drug resistance. *Biochemistry* 1995, **34** 1143–1152