Rationale for More Diverse Inhibitors in Competition with Substrates in HIV-1 Protease

Nevra Ozer,[†‡] Celia A. Schiffer,[§*] and Turkan Haliloglu[†‡*]
[†]Polymer Research Center and [‡]Chemical Engineering Department, Bogazici University, Istanbul, Turkey; and [§]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts

ABSTRACT The structural fluctuations of HIV-1 protease in interaction with its substrates versus inhibitors were analyzed using the anisotropic network model. The directions of fluctuations in the most cooperative functional modes differ mainly around the dynamically key regions, i.e., the hinge axes, which appear to be more flexible in substrate complexes. The flexibility of HIV-1 protease is likely optimized for the substrates' turnover, resulting in substrate complexes being dynamic. In contrast, in an inhibitor complex, the inhibitor should bind and lock down to inactivate the active site. Protease and ligands are not independent. Substrates are also more flexible than inhibitors and have the potential to meet the dynamic distributions that are inherent in the protease. This may suggest a rationale and guidelines for designing inhibitors that can better fit the ensemble of binding sites that are dynamically accessible to the protease.

INTRODUCTION

One of the most important factors in elucidating the pathogenesis of HIV-1 is viral resistance; thus, it is important to understand the development of this drug resistance to improve the therapeutic management of AIDS (1). The homodimeric HIV-1 protease is an effective therapeutic target of the most effective antiviral drugs for the treatment of HIV-1 infection. The protease sequentially cleaves at least 10 asymmetric and nonhomologous sequences in the Gag and Gag-Pol polyproteins, and allows for maturation of the immature virion that facilitates the spread of the virus (2). These peptidomimetic drugs are the result of structure-based drug design efforts on the part of both academia and the pharmaceutical industry. Indeed, protease inhibitors are considered the most potent drugs currently available for the treatment of AIDS (1).

Protease inhibitors are all competitive inhibitors that bind at the active site and compete directly with the enzyme's ability to recognize substrates (1,3). They all have large, generally hydrophobic moieties that interact with the mainly hydrophobic pockets in the active site (1). Unfortunately, the medical efficacy of the current inhibitors is proving to be short-lived, as viable mutant variants of HIV-1 protease confer drug resistance. Drug resistance results from a subtle change in the balance of recognition events between the relative affinity of the enzyme to bind inhibitors and its ability to bind and cleave substrates. Since HIV-1 protease binds substrates and inhibitors at the same active site, the change that alters inhibitor binding also alters substrate binding. However, the substrate recognition does not seem to be greatly altered when inhibitors contact the residues that are not contacted extensively by the substrates (4). This may not be the case for residues that are important for both substrate and inhibitor binding. Although they are chemically different, the three-dimensional shape and electrostatic character of the protease inhibitors are fairly similar. A small set of mutations can thus result in a protease variant with multidrug resistance. This evolution of drug resistance in HIV-1 protease presents a new challenge to future structure-based drug design efforts (1).

The HIV-1 protease functions as a homodimer with a single active site (residues 25–27 of each chain) that is formed by the dimer interface and capped by two flexible flaps (5). Despite the symmetry conferred on its active site by being a homodimer, the enzyme recognizes a series of nonhomologous asymmetric octomeric substrate sites within the Gag and GagPol polyproteins. Yet, despite the fact that the substrate sites are asymmetric, the currently prescribed inhibitors are relatively symmetric around the cleavage site. This allows a single mutation to impact the inhibitor binding twice, while possibly impacting substrate binding to a lesser extent. Two solvent-accessible loops of the protease (residues 33–43 of each chain) followed by the two flexible flaps (residues 44–62 of each chain) are important for ligand-binding interactions (6). The terminal residues 1–4 and 95–99 of each chain play a role in dimerization and stabilization of the active protease (6). A large conformational change occurs during ligand binding, which consists of the opening and closing of the flaps over the binding site.

Molecular recognition in ligand binding is dependent on the intrinsic dynamics of the protein (7). Although structural changes have been observed experimentally with ligand binding, the intrinsic dynamics of the protein, which is likely evolutionarily optimized, is not well described. An induced fit in ligand recognition is favored by long-range interactions, whereas conformational selection in binding

Submitted November 13, 2009, and accepted for publication June 28, 2010.

*Correspondence: celia.schiffer@umassmed.edu or halilogt@boun.edu.tr

Nevra Ozer's present address is Department of Bioengineering, Marmara University, Istanbul, Turkey.

Editor: Ruth Nussinov.

0006-3495/10/09/1650/10 $2.00

doi: 10.1016/j.bpj.2010.06.064

is favored by short-range interactions (7). The diversity of conformations and the insufficient data on the energetics of protein-ligand interactions make it very difficult to incorporate the intrinsic dynamics into drug discovery efforts.

Fluctuations of biomolecular complexes around their native states are important for functional analysis in molecular biophysics. Several features, such as entropy changes upon binding, possible drug-binding sites, and the overall stability, flexibility, and function, can be deduced from detailed analyses of these fluctuations (8,9). There is a significant correlation between cooperative motions of the structure and its biological function (7). There are several computational methods that can be used to identify these dominant correlated motions. The common approach is to decompose the dynamics into a collection of modes of motion focusing on a few low-frequency/large-amplitude modes that are expected to be relevant to function (10,11). The process of extracting the dominant collective modes from fluctuations in molecular-dynamics (MD) trajectories, also called principal component analysis (PCA), is now an established computational method for studying protein dynamics. The major disadvantage of this method is the sampling inefficiency of MD simulations, especially in large molecular systems (12,13). Alternatively, the cooperative motions can be studied by normal mode analysis, in which the concerted motions of a protein are expressed in terms of a set of collective variables (normal modes) (14–20). There are also some geometry-based methods, such as CONCOORD/tCONCOORD (21,22), and methods based on constraint theory, such as FIRST (23), that can be used to tackle protein flexibility and sample the conformational space to predict functional behaviors. The similarities between these approaches and elastic network models were previously presented in detail (22,24).

Recently, elastic network models have gained considerable attention for studying the large-scale motion of protein structures that are relevant to function (25–31). This suggests that the structures evolved in such a way that the intrinsic elastic low-frequency modes are the most efficient way for the structures to function. Elastic network models originated from the work of Tirion (32), who developed a model in which a single uniform harmonic potential reproduces the complex vibrational properties of macromolecular systems. The fluctuations predicted by the isotropic Gaussian network model (GNM) (33,34) applied to coarse-grained proteins with one point mass per residue agree significantly with experimental crystallographic B-factors for many proteins. The model has been extended to the anisotropic network model (ANM) (35), in which the directions as well as the magnitudes of motions are predicted. GNM and ANM applied to the HIV-1 protease system have also produced results that are highly in accord with those of both experimental and MD simulations, despite their simplicity (6,10,36–40). Computational studies of the dynamics of HIV-1 protease suggest that the structural fluctuations in low-frequency modes can be utilized for intrinsic protein flexibility and motion to maintain function.

In protein-ligand interactions, the ligand prefers the conformations that best match its structural and dynamic properties among those intrinsically accessible to the unbound protein (7). Here, to search for the determinants of structural changes that accompany ligand binding in HIV-1 protease, we performed a comparative analysis of the conformational space spanned by the protease bound to different ligands and its intrinsic dynamics. We analyzed the crystal structures of both substrate- (1,3,4) and inhibitor-bound (41–49) proteases by means of a simple physics-based ANM (35) constructed by incorporating all of the atoms (except hydrogen atoms) of the structure. We elaborated the magnitude and orientation of motion of protease and peptide residues in the low-frequency modes, emphasizing specific regions (i.e., the dimerization, active site, flap, and substrate cleft regions), by performing a comparative analysis between different natural substrate and inhibitor complex structures. The results of this analysis may help elucidate both the binding and drug-resistance mechanisms of HIV-1 protease.

MATERIALS AND METHODS

Structures

In this study, the crystal structures of both substrate (1f7a, ca-p2 (3); 1kj4, ma-ca (1); 1tsu, nc-p1 (4); 1kjf, p1-p6 (1); 1kj7, p2-nc (1); 1kjh, rh-in (1); 1kjg, rt-rh (1)) and inhibitor (1hpv, amprenavir-apv (41); 2fxe, atazanavir-atv (42); 1t3r, darunavir-drv (43); 1hsg, indinavir-idv (44); 1mui, lopinavir-lpv (45); 1ohr, nelfinavir-nfv (46); 1hxw, ritonavir-rtv (47); 1hxb, saquinavir-sqv (48); 1d4y, tipranavir-tpv (49)) liganded protease were analyzed by ANM (35). All atoms of the structure, except hydrogens, were incorporated into the calculations.

Anisotropic network model

ANM (35) performs a harmonic vibrational analysis of protein structures around their equilibrium states and predicts the directionalities of the collective motions, as well as their magnitudes. The elastic network is formed by connecting all neighboring atoms except hydrogen atoms, and the conformations that describe the fluctuations of residues from the average in the principal directions of motion are generated. The total potential energy for a system of N nodes is the summation over all harmonic interactions of close-neighboring (i, j) pairs, calculated as:

$$V = (\gamma/2)\left[\sum_{i,j}^{N} h\left(r_c - R_{ij}\right)\left(R_{ij} - R_{ij}^0\right)^2\right], \qquad (1)$$

where γ is the harmonic force constant; R_{ij} is the instantaneous distance; R_{ij}^0 is the equilibrium distance between sites i and j in the native structure; $h(r_c - R_{ij})$ is the Heaviside step function (which is one if $(r_c - R_{ij}) \geq 0$, and zero otherwise); and r_c is the cutoff distance (here taken as 9 Å, which leads to a higher correlation of the predicted fluctuations with the experimental B-factors compared to lower cutoff values tried). A 9 Å cutoff distance has successfully been used to account for interresidue interactions in all-atom structure models (50). Higher cutoff distances are preferred in the ANM compared to the GNM calculations, whereas lower cutoff distances are

appropriate in all-atom models compared to a reduced description of residues with only α carbons (51,52).

The Hessian matrix $\mathbf{H}$ is a $3N \times 3N$ symmetric matrix composed of $N \times N$ super elements $\mathbf{H}_{ij}$, each of size 3×3, given by the second derivatives of the potential V. An orthogonal transformation of the real symmetric Hessian matrix gives the normal modes of the elastic network with $3N - 6$ nonzero eigenvalues λ_i and corresponding eigenvectors $\mathbf{u}_i$:

$$\mathbf{H}^{-1} = \sum_{i=1}^{3N-6} \frac{1}{\lambda_i} \mathbf{u}_i \mathbf{u}_i^T. \quad (2)$$

The cross-correlation between the fluctuations of sites i and j is calculated as:

$$\langle \Delta\mathbf{R}_i . \Delta\mathbf{R}_j \rangle = (3k_B T/\gamma) tr[\mathbf{H}^{-1}]_{ij}, \quad (3)$$

where $tr[\mathbf{H}^{-1}]_{ij}$ is the trace of the ij^{th} submatrix $[\mathbf{H}^{-1}]_{ij}$ of $\mathbf{H}^{-1}$. When $i = j$, the self-correlations between the components $\Delta\mathbf{R}_i$ are obtained. Here, the knowledge of fluctuation vectors permits us to construct and explicitly view pairs of alternative conformations sampled by the individual modes simply by adding the fluctuation vectors $\pm\Delta\mathbf{R}_i$ to the equilibrium position vectors in the respective modes.

RESULTS AND DISCUSSION

Motion in principal directions

The fluctuations in the low-frequency modes (principal directions) refer to the main functional motion of the structures, and thus all HIV-1 protease complex structures that are functional should display similar modes of motion. These collective modes of motion should be highly robust against sequence and structure variations.

Indeed, the motion in these most cooperative modes is similar for all substrate- and inhibitor-bound structures. In the first mode (Fig. 1 *A*), both monomers of the protease rotate around two axes parallel to the z direction, and the peptide fluctuates from the C-terminus to the N-terminus in the negative y direction. In the second mode (Fig. 1 *B*), there are two axes around which the monomers rotate parallel to the x (longest axis along which the protease lies) and z directions. The monomers rotate around the x axis in opposite directions, and the motion in the substrate is significant in the edges in the second mode. In the third-slowest mode, the protease monomers rotate around two axes parallel to the y direction (see Fig. S1 in the Supporting Material). The contribution of each of these modes to the overall protease-ligand dynamics on the average of all structures is 11%, 10%, and 6%, respectively, with the contribution of the next-slowest modes decaying very rapidly.

Diversity in orientational correlations

The fluctuations in the low-frequency modes are expected to describe the functional motion of the ligand-bound HIV-1 protease structure. Therefore, it is of interest to elucidate whether there is a difference between substrate- and inhibitor-bound HIV-1 protease structures in these functional motions, and what such a difference might imply. The inner products of fluctuation vectors for the two cooperative modes are calculated to observe the orientational correlations between the fluctuations of the same residues in different substrate- and inhibitor-bound complex structures. The orientational correlations in the slowest two modes are given in Figs. 2 and 3, where the normalized correlation values range between 1 and −1. The correlations for all substrate- and inhibitor-bound complex structures are included in the plots, which reflect the dynamic conformational ensemble of the protease complex structures. The peaks with negative correlation values in the charts (the troughs) indicate the residues that fluctuate more diversely, i.e., the residues that display the maximum variations in the direction of their fluctuations. The lower correlation values in the substrate- compared to inhibitor-bound complexes imply that the protease residues are able to depict fluctuations in more diverse directions in interaction with the substrates than with the inhibitors. The orientational freedom of the protease residues is more restricted when the protease is bound to the inhibitors. The standard deviation from the average orientational correlation for the substrate-bound structures is higher than that for the inhibitor-bound structures, which clearly indicates the richer conformational space spanned by the former (Fig. S2).

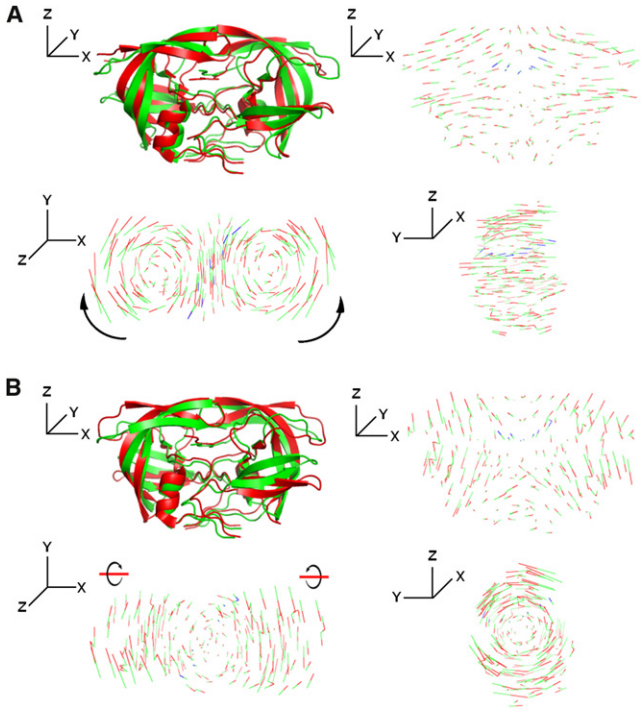

FIGURE 1 Motion of HIV-1 protease complex structures in the (*A*) slowest and (*B*) second-slowest modes. The fluctuation of the structure in each mode is represented as the protease moving between the conformations shown in green and red, and the peptide moving between the conformations shown in green and blue.

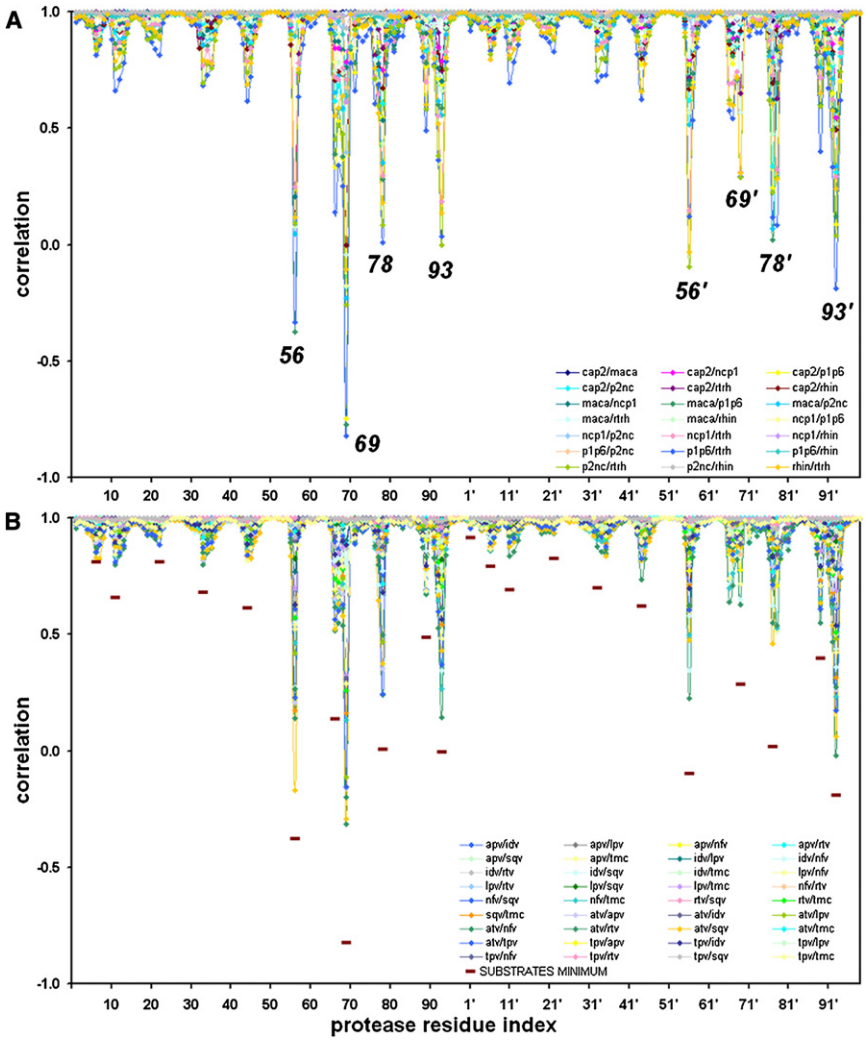

FIGURE 2 Orientational correlation of protease residues in the slowest mode in (*A*) substrate- and (*B*) inhibitor-bound complex structures. The minimum value of orientational correlation among the substrate-bound complexes is shown in red bars in panel *B*.

Also, the asymmetry in the fluctuations between the two monomers of the substrate-bound protease structures is higher than that of the inhibitor-bound protease structures.

The relatively restricted structural motion of inhibitor-bound protease complex structures, and thus the limited conformational space of the protease, may suggest tighter binding as well as less diversity for dynamic freedom of the protease in the bound state. Binding to different substrates gives the protease sample a larger conformational space, apparently due to the flexible nature of the substrate structures, which in turn implies more flexibility of the protease in its interactions with the substrates. In other words, the substrates are better able to match the conformational states that are intrinsically accessible to the protease structure. This is in agreement with recent work (7) in which an analysis of a larger set of structures in inhibitor-bound complex structures with the unbound structures revealed that drugs select the conformation that best matches its energetic and dynamic properties among those that are intrinsically accessible to the unbound protein structure.

The protease residues that display the maximum variations in the direction of their fluctuations between different structures in the slowest mode correspond to residues 56, 69, 78, and 93 (Fig. 2). Residues 56 and 78 are close in space, 56 is the hinge point connecting the flap loop (residues 45–55) to the solvent-exposed upper arm of the flap (residues 36–44), and 78 is the hinge point connecting the same flap loop (residues 45–55) to the lower arm of the flap (residues 57–77). Residues 69 and 93 are also close in space; 69 is the tip of the lower arm of the flap (residues 61–73) and 93 is the tip of another loop (residues 85–96) (Fig. 4 *A*). The other minima observed, reflecting the diversity in the fluctuations, yet less celebrated, are at residues 6, 11–13, 22, 33–35, 44, 45, 57, 66–68, and 71. Of interest, these residues are observed to lie along the rotational axes (Fig. 4 *A*) around which the protease monomers rotate in the slowest mode (Fig. 1 *A*). In the second-slowest mode, on the other hand, the residues that display maximum variations in the direction of their fluctuations correspond to residues 25–27, 49–51, 84, and 97 (Fig. 3). Residues 49–51 are at the flap tips. In addition, residues 25–27 are active-site residues at the tip of the

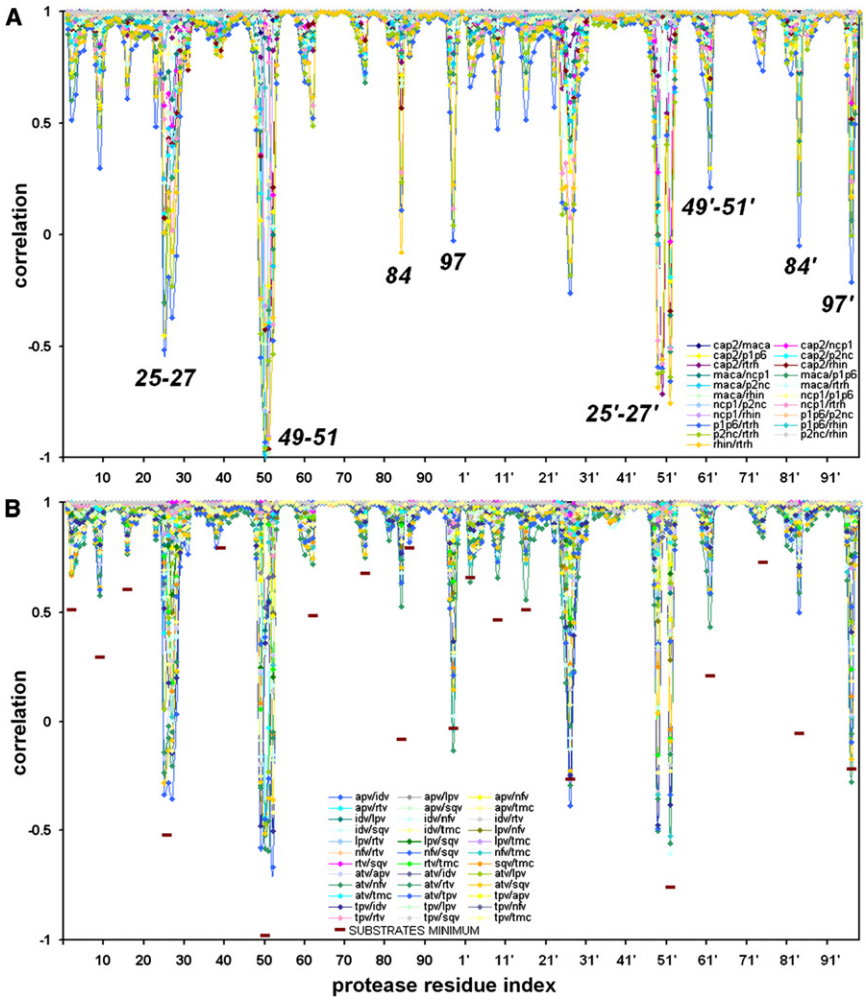

FIGURE 3 Orientational correlation of protease residues in the second-slowest mode in (*A*) substrate- and (*B*) inhibitor-bound complex structures. The minimum value of orientational correlation among the substrate-bound complexes is shown in red bars in panel *B*.

loop right in the middle of the substrate cleft that is connected to the 12–22 loop. Residues 84 and 97 are at the edges of the 85–96 loop (Fig. 4 *B*). Residues 2, 9, 16, 23–25, 28, 29, 39, 48, 52, 53, 61, 62, 74, and 75 also display diversity, albeit to a lesser extent, in the fluctuation direction of the second mode. These residues are also observed to lie along the rotational axes (Fig. 4 *B*) around which the monomers rotate in this mode (Fig. 1 *B*). The same applies for the third-slowest mode, which has less significant motion compared to the slowest two modes. The diversity in the fluctuation directions is observed at residues 13, 19, 20, 33, 64, 73, and 81, which lie along the hinge axes that mediate the behavior in this mode (Fig. S3 and Fig. S4). Apparently, the orientational difference in fluctuations, and the asymmetry of fluctuations between the monomers of the substrate- and inhibitor-bound protease are caused mainly by these residues, which are part of the hinge axes that are responsible for the most cooperative motions of the complex structure.

The conformational freedom of the protease residues is restricted when the protease is bound to the ligand. This is expected because of the additional interaction between protease residues and the ligand, which appears with

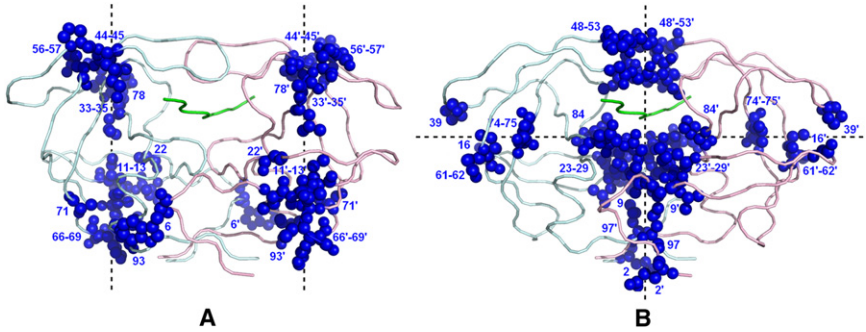

FIGURE 4 Regions that cause an orientational difference in the (*A*) slowest and (*B*) second-slowest modes. The protease monomers are displayed in light blue and light pink, and the ligand is displayed in green. The residues that display maximal variations in the direction of their fluctuations between different complex structures are displayed as blue spheres. The dashed lines indicate the rotational axes around which the monomers rotate, i.e., the hinge axes along which the residues with maximum orientational difference lie.

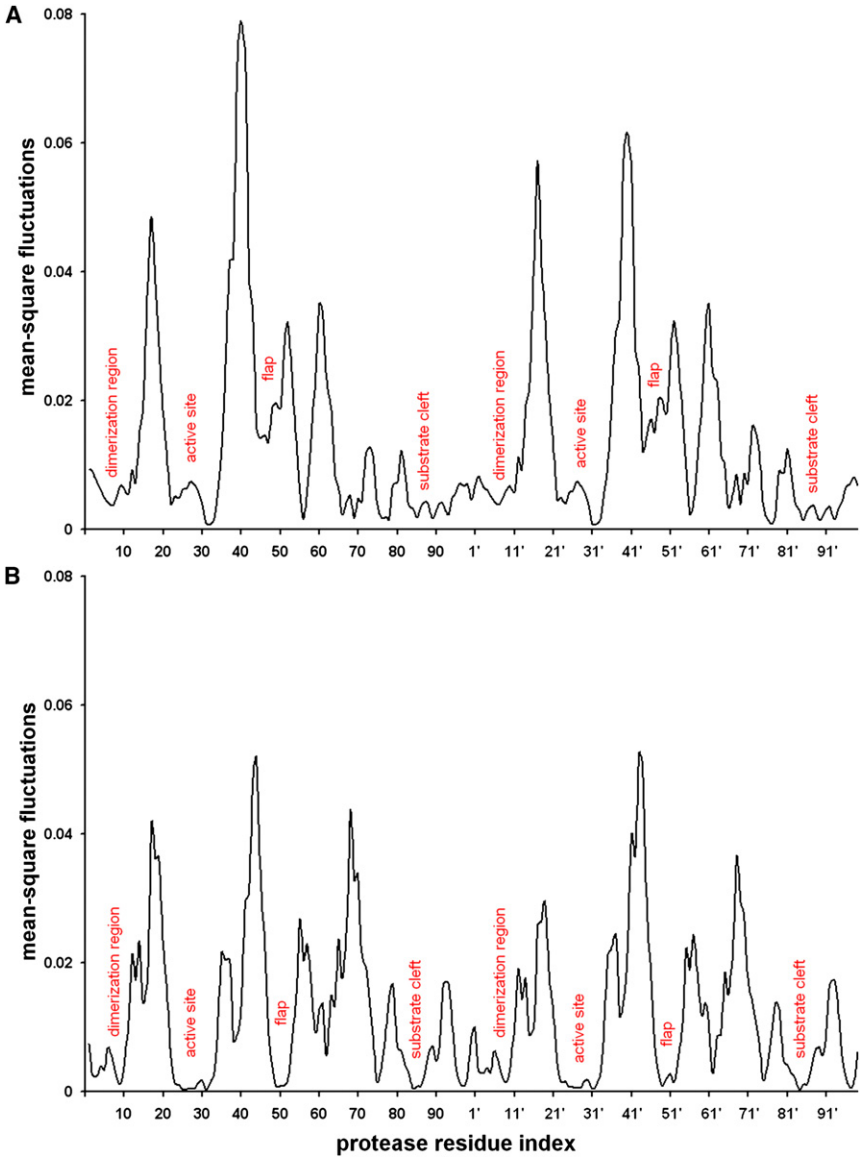

FIGURE 5 Mean-square fluctuations averaged over all the substrate-bound protease complex structures in the (*A*) slowest and (*B*) second-slowest modes. The fluctuations are calculated for all the atoms in the structure, but the fluctuations of Cα atoms are plotted for clearer representation, and corresponding residue numbers are indicated on the *x* axis.

additional springs in the elastic network description. Nevertheless, it should be noted that the same cutoff distance value is used for both the substrate- and inhibitor-bound structures studied, and the relative behavior of the protease in interaction with the substrates versus inhibitors is of interest here. Moreover, on average, the substrates comprise a higher number of atoms than the inhibitors, indicating relatively more additional springs for the substrate-bound conformations; that is, although the number of springs added in the ANM is larger for the substrates than for the inhibitors, the proteases display more flexibility in interaction with the substrates than the inhibitors.

Extent of mobility in the slowest modes of motion

The distribution of mobilities among residues driven by different frequency modes can be represented by mode shapes, such that the minima correspond to the hinge regions (53). The residues that display the maximum differences in the direction of their fluctuations between different structures in the respective modes have low mobility, i.e., low mean-square fluctuations, in the mode shape of the bound HIV-1 protease structure. Hence, the distribution of the mean-square fluctuations in the two low-frequency modes shows that the residues with maximal orientational differences between different structures in the first mode (i.e., 56, 69, 78, and 93) and those in the second mode (i.e., 25–27, 49–51, 84, and 97) are in the minima of the corresponding mode shapes presented in Fig. 5, where the values are averaged over all of the substrate-bound protease structures. A similar average mode shape is also revealed by the inhibitor-bound protease structures. The residues causing the orientational difference in the third-slowest mode can also be seen in the fluctuation profile in the

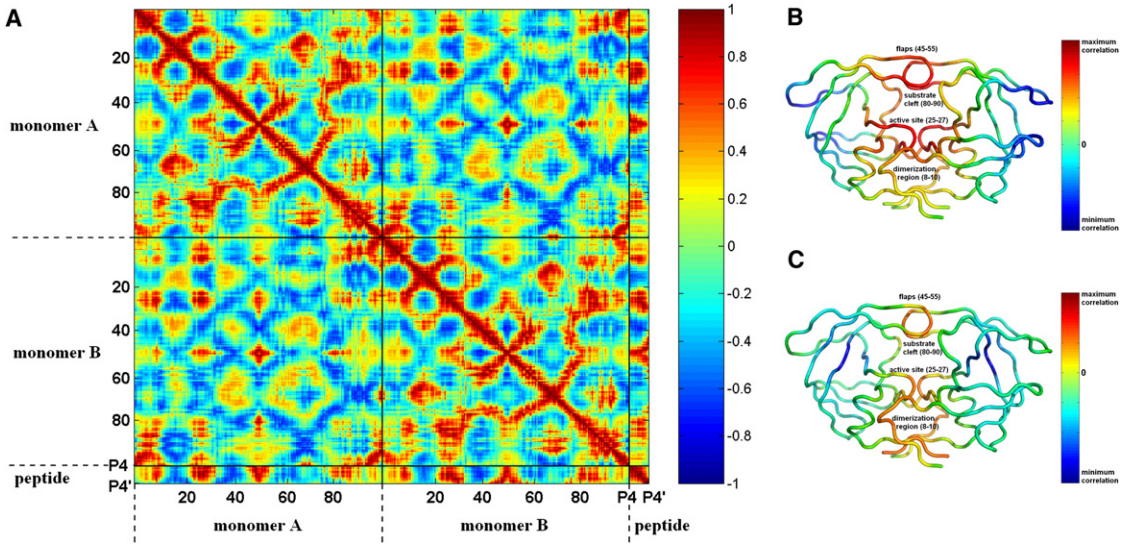

FIGURE 6 Cross-correlations between the fluctuations of residues. (*A*) Correlations shown for the representative ca-p2 complex structure of HIV-1 protease in the 10 slowest ANM modes. (*B*) Protease regions significant for binding are color-coded by the correlations of protease residues to peptide residues. (*C*) Protease regions significant for dimerization are color-coded by the correlations between the residues of the two protease monomers. For clarity, the peptide is not included.

respective mode (Fig. S5). The crucial sites in the protease, such as the dimerization region (residues 5–10), the active site (residues 25–27), the flap (residues 45–55), and the substrate cleft (residues 80–90), are observed to display relatively low fluctuations in all three slowest mode shapes. Yet, the slow fluctuations of the active site and flap residues are emphasized more in the second mode (Fig. 5), as confirmed by the hinge axis along which they lie in this mode (Fig. 4). The similar positions for the hinge regions are also observed when conformation generation is performed with geometric constraints using tCONCOORD (22).

Network of cooperative fluctuations

The motion coordinated by the hinge axes results in a network of cooperative fluctuations within a structure that could have a plausible link to its function. When the inter- and intrachain correlations in the protease dimer for the average of the first 10 modes, which account for 48% of the overall dynamics, are analyzed in both substrate- and inhibitor-bound protease structures (Fig. 6), the dimerization region (residues 5–10), the active sites (residues 25–27), the flaps (residues 45–55), and the substrate cleft (residues 80–90) are observed to be highly positively correlated with each other. The positive correlations of these hinge regions across the interface account for the association of the two chains. As for the interaction between the protease and the peptide, the residues of the protease that display positively correlated fluctuations with the fluctuations of the peptide's residues also correspond to these hinge regions. On the other hand, the highly mobile regions corresponding to the loops of residues 12–22 and 36–44 in the two slowest modes display negative correlations with both the hinge regions and the peptide in the average of the first 10 modes (Fig. 6). The mobile regions and residues that are important for protein stability or that take part in the key native contacts have been addressed for HIV-1 protease in previous studies (13,37,38). Although the flap region 45–55 is part of the relatively mobile flap, it has reduced mobility in the bound state due to its low tolerance to mutations (37,38). Nevertheless, region 36–44 is located at the solvent-exposed parts of the flap and has the highest mobility. It is possible that these regions surround and anchor the peptide in the cleft between the two monomers (38). The coupled fluctuations of the flaps, the active site, the dimerization interface, and the substrate cleft together with the ligand, as emphasized in the cross-correlation map, constitute a dynamic domain, suggesting that the enzymatic activity is an entire property of the protease structure, rather than localized in or near the active site (36).

The hinge regions composing the minima in the slowest mode shapes (Fig. 5) that also cause diversity in orientation (Fig. 4) serve as the mechanistically crucial sites that mainly coordinate the intrinsic dynamics of the structure, as indicated by the three-dimensional structural motion described above (Fig. 1). The slowest two modes contribute to the overall dynamics in similar amounts (11% and 10%, respectively) and suggest the following when looked individually: The hinges suggested by the slowest mode mostly coordinate the intrachain cooperative motions, whereas the hinges of the second-slowest mode are mostly responsible for the correlations across the dimerization interface.

The correlated fluctuations (mainly referred to as positively correlated fluctuations here) between the two protease

monomers, and those between each monomer and peptide positions have some differences when studied separately. The number of peptide atoms that are positively correlated to each protease atom is higher for inhibitor-bound than substrate-bound complex structures, even though the substrates comprise a higher number of atoms. The significant residues for binding in all complex structures are mainly located in four regions (residues 8–10 (dimerization region), 25–27 (active site), 45–55 (flap), and 80–90 (substrate cleft)), yet the active sites and the flaps are more emphasized compared to the other two regions. The well-conserved residues (e.g., residues 25, 27–29, and 49 (54)) in these hinge regions have been demonstrated to be critical for substrate binding (55,56), and thus resistance-evading potent drugs should interact strongly with these residues. In dimerization, on the other hand, the active sites of the protease monomers interact strongly in the inhibitor-bound complex structures, yet the interactions between the flaps and between the N- and C-termini of the monomers are weakened compared to the substrate-bound complex structures. The interactions between the two monomers in all complex structures, i.e., the critical residues in dimerization, are also mainly located in the same four specific regions as in binding to the peptide. However, the dimerization region, the active site, and the flaps are more emphasized in the cooperative fluctuations compared to the substrate cleft. In addition, the N- and C-termini of the monomers are highly correlated with each other. The importance of a dimer interface for drug targeting was noted in previous works (57,58), which demonstrated that inhibitors that act as allosteric inhibitors in binding at the dimer interface and alter the conformation of the protease can indirectly reduce the binding affinity of substrates.

The network of cooperative fluctuations can be analyzed in terms of the positions of drug resistant mutations. In general, drug resistance occurs when mutations in the target protein enable it to retain function while no longer being effectively inhibited by the drug (59). The drug-resistant mutations in the HIV-1 protease render the variant protease resistant to the inhibitor while allowing the enzyme to cleave substrates (59). The invariant positions include the active site (residues 25–27), some positions in close contact with the active site or near the substrate cleft (residues 28, 29, 31, 80, 81), most of the N- and C-terminal sites on the dimerization interface, and other positions associated with maintaining the conformational flexibility, such as the conserved glycines and the flap tips (residues 49, 51, and 52) (54). Drug resistance that alters the balance between substrate recognition and inhibitor binding often occurs by combinations of mutations both inside and outside the active site. The most frequent mutations found in drug-resistant isolates of protease involve positions 10, 36, 46, 54, 71, 77, 82, 84, and 90 (60). Of interest, positions 10, 36, 46, 71, and 77 correspond to either the residues or the first neighbors of the residues lying along the hinge axes in the slowest ANM mode, whereas positions 10, 54, 82, 84, and 90 are associated in a similar way with the second-slowest mode (see Fig. S6, *A* and *B*). These hinge axes appear to mostly coordinate the intrachain cooperative fluctuations in the slowest mode, and the interchain cooperative fluctuations together with ligand interactions in the second-slowest mode. Color-coding of the cooperative fluctuations of protease relevant to intramolecular interactions, dimerization, and substrate binding can be viewed with respect to these frequent mutations in Fig. S6, *C–E*. Residues 10, 36, 46, 54, 71, 77, 82, 84, and 90 are highly positively correlated in the intramolecular interactions. Residues 10, 46, 54, 82, 84, and 90 are associated with the highly positively correlated regions specific to binding. Residues 10, 82, and 84, which are in contact with the ligand, are positively correlated at the dimerization interface. Residue positions 10, 36, 46, 54, 71, 77, and 90 are the common mutation sites outside the active site and the substrate cleft. These may not only impact inhibitor binding but also compensate for the viability and fitness of the enzyme and thus increase the growth rate of the mutant virus.

CONCLUSIONS

As presented in this study, the stronger binding of the inhibitors (55,56,58,61) results in restricted motion of the complex structures compared to the substrates. By contrast, because of their higher flexibility, the substrates are more adaptable to the protease's backbone rearrangements or conformational changes. The similarity in orientation of the fluctuations of inhibitor complexes, together with their similarity in three-dimensional shape and electrostatic character, may also have implications for multidrug resistance. In conclusion, our analysis of the fluctuations of the ligand-bound HIV-1 protease structures using ANM identifies a functionally plausible dynamic motion between substrate- and inhibitor-bound complex structures. The results presented here may help elucidate the plasticity of the ensemble of ligand-bound HIV-1 protease conformations and aid in drug design.

SUPPORTING MATERIAL

Figures corresponding to the third-slowest mode of ANM of the HIV-1 protease complex structures, the standard deviations from the average orientational correlations in the slowest and second-slowest modes, and the location of most frequent drug-resistant mutations on the protease in relation to the orientation and network of cooperative fluctuations are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(10)00834-9.

This research was supported by the National Institutes of Health AIDS-FIRCA RO3 TW006875-01. T.H. acknowledges Turkish State Planning Organization grant 2009K120520 and Betil Fund. C.A.S. acknowledges grant R01 GM65347.

REFERENCES

1. Prabu-Jeyabalan, M., E. Nalivaika, and C. A. Schiffer. 2002. Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure.* 10:369–381.
2. Chou, K. C. 1996. Prediction of human immunodeficiency virus protease cleavage sites in proteins. *Anal. Biochem.* 233:1–14.
3. Prabu-Jeyabalan, M., E. Nalivaika, and C. A. Schiffer. 2000. How does a symmetric dimer recognize an asymmetric substrate? A substrate complex of HIV-1 protease. *J. Mol. Biol.* 301:1207–1220.
4. Prabu-Jeyabalan, M., E. A. Nalivaika, …, C. A. Schiffer. 2004. Structural basis for coevolution of a human immunodeficiency virus type 1 nucleocapsid-p1 cleavage site with a V82A drug-resistant mutation in viral protease. *J. Virol.* 78:12446–12454.
5. Wlodawer, A., and J. W. Erickson. 1993. Structure-based inhibitors of HIV-1 protease. *Annu. Rev. Biochem.* 62:543–585.
6. Yang, L., G. Song, …, R. L. Jernigan. 2008. Close correspondence between the motions from principal component analysis of multiple HIV-1 protease structures and elastic network modes. *Structure.* 16:321–330.
7. Bakan, A., and I. Bahar. 2009. The intrinsic dynamics of enzymes plays a dominant role in determining the structural changes induced upon inhibitor binding. *Proc. Natl. Acad. Sci. USA.* 106:14349–14354.
8. Micheletti, C., G. Lattanzi, and A. Maritan. 2002. Elastic properties of proteins: insight on the folding process and evolutionary selection of native structures. *J. Mol. Biol.* 321:909–921.
9. Keskin, O., S. R. Durell, …, D. G. Covell. 2002. Relating molecular flexibility to function: a case study of tubulin. *Biophys. J.* 83:663–680.
10. Bahar, I., A. R. Atilgan, …, B. Erman. 1998. Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability. *Phys. Rev. Lett.* 80:2733–2736.
11. Bahar, I. 1999. Dynamics of proteins and biomolecular complexes: inferring functional motions from structure. *Rev. Chem. Eng.* 15: 319–349.
12. Doruker, P., A. R. Atilgan, and I. Bahar. 2000. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to α-amylase inhibitor. *Proteins.* 40:512–524.
13. Micheletti, C., F. Cecconi, …, A. Maritan. 2002. Crucial stages of protein folding through a solvable model: predicting target sites for enzyme-inhibiting drugs. *Protein Sci.* 11:1878–1887.
14. Go, N., T. Noguti, and T. Nishikawa. 1983. Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc. Natl. Acad. Sci. USA.* 80:3696–3700.
15. Brooks, B., and M. Karplus. 1983. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA.* 80:6571–6575.
16. Case, D. A. 1994. Normal mode analysis of protein dynamics. *Curr. Opin. Struct. Biol.* 4:285–290.
17. Hinsen, K. 1998. Analysis of domain motions by approximate normal mode calculations. *Proteins.* 33:417–429.
18. Tama, F., and Y. H. Sanejouand. 2001. Conformational change of proteins arising from normal mode calculations. *Protein Eng.* 14:1–6.
19. Ma, J. 2005. Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure.* 13:373–380.
20. Cui, Q., and I. Bahar. 2005. Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems. CRC Press, Boca Raton, FL.
21. de Groot, B. L., D. M. F. van Aalten, …, H. J. Berendsen. 1997. Prediction of protein conformational freedom from distance constraints. *Proteins.* 29:240–251.
22. Seeliger, D., J. Haas, and B. L. de Groot. 2007. Geometry-based sampling of conformational transitions in proteins. *Structure.* 15:1482–1492.
23. Thorpe, M. F., M. Lei, …, L. A. Kuhn. 2001. Protein flexibility and dynamics using constraint theory. *J. Mol. Graph. Model.* 19:60–69.
24. Rader, A. J., and I. Bahar. 2004. Folding core predictions from networks models of proteins. *Polymer (Guildf.).* 45:659–668.
25. Bahar, I., and A. J. Rader. 2005. Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* 15:586–592.
26. Schuyler, A. D., and G. S. Chirikjian. 2005. Efficient determination of low-frequency normal modes of large protein structures by cluster-NMA. *J. Mol. Graph. Model.* 24:46–58.
27. Nicolay, S., and Y. H. Sanejouand. 2006. Functional modes of proteins are among the most robust. *Phys. Rev. Lett.* 96:078104.
28. Kim, M. K., Y. Jang, and J. I. Jeong. 2006. Using harmonic analysis and optimization to study macromolecular dynamics. *Int. J. Control Autom. Syst.* 4:382–393.
29. Leherte, L., and D. P. Vercauteren. 2008. Collective motions in protein structures: applications of elastic network models built from electron density distributions. *Comput. Phys. Commun.* 179:171–180.
30. Csermely, P. 2008. Creative elements: network-based predictions of active centres in proteins and cellular and social networks. *Trends Biochem. Sci.* 33:569–576.
31. Zheng, W. 2008. A unification of the elastic network model and the Gaussian network model for optimal description of protein conformational motions and fluctuations. *Biophys. J.* 94:3853–3857.
32. Tirion, M. M. 1996. Large amplitude elastic motions in protein from a single-parameter atomic analysis. *Phys. Rev. Lett.* 77:1905–1908.
33. Bahar, I., A. R. Atilgan, and B. Erman. 1997. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.* 2:173–181.
34. Haliloglu, T., I. Bahar, and B. Erman. 1997. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* 79:3090–3093.
35. Atilgan, A. R., S. R. Durell, …, I. Bahar. 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80:505–515.
36. Zoete, V., O. Michielin, and M. Karplus. 2002. Relation between sequence and structure of HIV-1 protease inhibitor complexes: a model system for the analysis of protein flexibility. *J. Mol. Biol.* 315:21–52.
37. Kurt, N., W. R. Scott, …, T. Haliloglu. 2003. Cooperative fluctuations of unliganded and substrate-bound HIV-1 protease: a structure-based analysis on a variety of conformations from crystallography and molecular dynamics simulations. *Proteins.* 51:409–422.
38. Micheletti, C., P. Carloni, and A. Maritan. 2004. Accurate and efficient description of protein vibrational dynamics: comparing molecular dynamics and Gaussian models. *Proteins.* 55:635–645.
39. Hamacher, K., and J. A. McCammon. 2006. Computing the amino acid specificity of fluctuations in biomolecular systems. *J. Chem. Theory Comput.* 2:873–878.
40. Hamacher, K. 2009. Temperature dependence of fluctuations in HIV-1 protease. *Eur. Biophys. J.* 39:1051–1056.
41. Kim, E. E., C. T. Baker, …, M. A. Navia. 1995. Crystal structure of HIV-1 protease in complex with Vx-478, a potent and orally bioavailable inhibitor of the enzyme. *J. Am. Chem. Soc.* 117:1181–1182.
42. Klei, H. E., K. Kish, …, S. Sheriff. 2007. X-ray crystal structures of human immunodeficiency virus type 1 protease mutants complexed with atazanavir. *J. Virol.* 81:9525–9535.
43. Surleraux, D. L., A. Tahri, …, P. B. Wigerinck. 2005. Discovery and selection of TMC114, a next generation HIV-1 protease inhibitor. *J. Med. Chem.* 48:1813–1822.
44. Chen, Z., Y. Li, …, L. C. Kuo. 1994. 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.* 269:26344–26348.
45. Stoll, V., W. Qin, …, D. W. Norbeck. 2002. X-ray crystallographic structure of ABT-378 (lopinavir) bound to HIV-1 protease. *Bioorg. Med. Chem.* 10:2803–2806.

46. Kaldor, S. W., V. J. Kalish, …, J. H. Tatlock. 1997. Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease. *J. Med. Chem.* 40:3979–3985.

47. Kempf, D. J., K. C. Marsh, J. F. Denissen, E. McDonald, S. Vasavanonda, C. A. Flentge, B. E. Green, L. Fino, C. H. Park, X. P. Kong…, 1995. ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. *Proc. Natl. Acad. Sci. USA*. 92:2484–2488.

48. Krohn, A., S. Redshaw, …, M. H. Hatada. 1991. Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere. *J. Med. Chem.* 34:3340–3342.

49. Thaisrivongs, S., …, H. I. Skulnick, K. D. Watenpaugh. 1996. Structure-based design of HIV protease inhibitors: sulfonamide-containing 5,6-dihydro-4-hydroxy-2-pyrones as non-peptidic inhibitors. *J. Med. Chem.* 39:4349–4353.

50. Kurkcuoglu, O., R. L. Jernigan, and P. Doruker. 2006. Loop motions of triosephosphate isomerase observed with elastic networks. *Biochemistry.* 45:1173–1182.

51. Pandey, B. P., C. Zhang, …, Y. Zhou. 2005. Protein flexibility prediction by an all-atom mean-field statistical theory. *Protein Sci.* 14:1772–1777.

52. Sen, T. Z., Y. Feng, …, R. L. Jernigan. 2006. The extent of cooperativity of protein motions observed with elastic network models is similar for atomic and coarser-grained models. *J. Chem. Theory Comput.* 2:696–704.

53. Liu, X., H. A. Karimi, …, I. Bahar. 2004. Protein functional motion query and visualization. *Proc. Annu. Int. COMPSAC, 28th.* 86–89.

54. Wu, T. D., C. A. Schiffer, …, R. W. Shafer. 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. *J. Virol.* 77:4836–4847.

55. Wang, W., and P. A. Kollman. 2001. Computational study of protein specificity: the molecular basis of HIV-1 protease drug resistance. *Proc. Natl. Acad. Sci. USA*. 98:14937–14942.

56. Hou, T., W. A. McLaughlin, and W. Wang. 2008. Evaluating the potency of HIV-1 protease drugs to combat resistance. *Proteins.* 71:1163–1174.

57. Bowman, M. J., S. Byrne, and J. Chmielewski. 2005. Switching between allosteric and dimerization inhibition of HIV-1 protease. *Chem. Biol.* 12:439–444.

58. Hornak, V., and C. Simmerling. 2007. Targeting structural flexibility in HIV-1 protease inhibitor binding. *Drug Discov. Today.* 12:132–138.

59. King, N. M., M. Prabu-Jeyabalan, …, C. A. Schiffer. 2004. Combating susceptibility to drug resistance: lessons from HIV-1 protease. *Chem. Biol.* 11:1333–1338.

60. Hertogs, K., S. Bloor, …, B. A. Larder. 2000. Phenotypic and genotypic analysis of clinical HIV-1 isolates reveals extensive protease inhibitor cross-resistance: a survey of over 6000 samples. *AIDS.* 14:1203–1210.

61. Luque, I., M. J. Todd, …, E. Freire. 1998. Molecular basis of resistance to HIV-1 protease inhibition: a plausible hypothesis. *Biochemistry.* 37:5791–5797.