PROTEINS: Structure, Function, and Genetics 49:61–70 (2002)

Inhibitor Binding Alters the Directions of Domain Motions in HIV-1 Reverse Transcriptase

N. Alpay Temiz and Ivet Bahar*

Center for Computational Biology and Bioinformatics, and Department of Molecular Genetics and Biochemistry, School of Medicine, University of Pittsburgh, Pennsylvania

ABSTRACT **Understanding the molecular mechanisms of HIV-1 reverse transcriptase (RT) action and drug inhibition is essential for designing effective antiretroviral therapies. Although comparisons of the different crystal forms of RT give insights into the flexibility of different domains, a direct computational assessment of the effect of inhibitor binding on the collective dynamics of RT is lacking. A structure-based approach is used here for exploring the dynamics of RT in unliganded and inhibitor-bound forms. Non-nucleoside RT inhibitors (NNRTI) are shown to interfere directly with the global hinge-bending mechanism that controls the cooperative motions of the p66 fingers and thumb subdomains. The net effect of nevirapine binding is to change the direction of domain movements rather than suppress their mobilities. The second generation NNRTI, efavirenz, on the other hand, shows the stronger effect of simultaneously reorienting domain motions and obstructing the p66 thumb fluctuations. A second hinge site controlling the global rotational reorientations of the RNase H domain is identified, which could serve as a target for potential inhibitors of RNase H activity. Proteins 2002;49:61–70.**

Key words: molecular mechanisms; domain motions; hinge bending; collective dynamics; ligand binding; HIV-1 reverse transcriptase

INTRODUCTION

HIV-1 reverse transcriptase (RT) continues to be a major target for anti-AIDS chemotherapy since the early design of pharmaceutical compounds against HIV. Two groups of RT inhibitors have been extensively investigated: nucleoside-analogue inhibitors and non-nucleoside RT inhibitors (NNRTI)s. NNRTIs are chemically diverse, generally hydrophobic and relatively noncytotoxic compounds, including a number of drugs licensed for clinical use, such as nevirapine, delavirdine, and efavirenz. Despite the advances in RT control, the effective inhibition of RT remains as a challenge, because of the continual emergence of drug-resistant mutants with reduced susceptibility to inhibitors.

Our understanding of the molecular basis of RT function significantly advanced after the X-ray elucidation of RT structures in multiple forms, complexed with NNRTI[1,2] or nucleic acid template primers,[3,4] or in the absence of ligands,[5,6] and the early comparative studies of these structures.[7] RT is a heterodimer. Its larger (66 kDa) subunit, p66, is composed of two domains, DNA polymerase (pol) and RNase H, and the pol domain is, in turn, composed of four subdomains, referred to as the p66 fingers, thumb, palm, and connection, in analogy to the pol I enzyme family. The smaller subunit, p51, is composed of the same subdomains, fingers, palm, thumb, and connection, but assembled in a different geometry. The NNRTI-binding pocket is located in a large cleft in the p66 palm and opens toward the interface of the p66/p51 heterodimer. Three aspartates located in the palm subdomain of p66, Asp110, Asp185, and Asp186, form the catalytic carboxylate triad of the pol domain. The latter two belong to the highly conserved YXDD motif common to all retroviral RTs.

Usually, the structural changes induced by point mutations near the inhibitor-binding pocket and the reduction in ligand-binding affinities are examined as determinants of drug efficiency. Several recent studies have focused on the effects of ligand binding[8–10] as well as on the binding affinities of different ligands.[11–13] Although the effect of ligand binding on structure and internal interactions can be deduced from these studies, several questions concerning the dynamic implications of ligand binding remain unanswered: (a) what is the effect of inhibitor binding on the mechanisms of molecular motions, or how does inhibitor binding alter the collective dynamics of the molecule, (b) how do local effects induce a global change in the mechanism of action of the enzyme? On a broader basis, one might ask (c) which sites, other than those located near the catalytic triad at the p66 palm, play a significant role in mediating the collective dynamics and could be looked on as suitable targets for novel inhibitor binding.

We use an analytical approach based on a model that originates from the statistical mechanics of elastic networks[14,15] and graph theory. The earlier version of this model, known as Gaussian Network Model (GNM),[16,17] proved useful in predicting X-ray crystallographic B factors,[16] H/D exchange free energies near native state

*Correspondence to: I. Bahar, Center for Computational Biology and Bioinformatics, School of Medicine, University of Pittsburgh, Suite 601, L. Kaufmann Building, 3471 Fifth Avenue, Pittsburgh, PA 15213. E-mail: bahar@pitt.edu

Received 16 April 2002; Accepted 25 April 2002

TABLE I. PDB Codes of Crystal Structures Used

	PDB accession code	Resolution (Å)	Reference
RT (wild type)	1DLO	2.7	6
RT-nevirapine	1RTH	2.2	71
RT-efavirenz	1FK9	2.5	10

conditions,[18] and ^{15}N-NMR order parameters.[19] Notably, GNM can identify the "key" residues important for function and stability,[20] as shown in several applications[21–23] including HIV-1 RT.[24] GNM was recently extended to predict the mechanisms or directionalities of collective motions, in addition to their size.[25] This new version, referred to as the anisotropic network model (ANM), is exploited here for comparing the dynamics of RT in unliganded and liganded forms. The structures are shown in Table I. A change in domain reorientation, rather than suppression in domain flexibility, is detected with nevirapine binding. Efavirenz, on the other hand, is shown to both dampen down the amplitudes of the p66 thumb fluctuations and alter the directions of domain movements, which could be consistent with the greater resilience of efavirenz to drug-resistance mutations.

The two underlying features of the approach are the validity of a structure-based assessment of equilibrium dynamics and the adoption of a purely mechanical description.

Structure-Based Studies of Dynamics: a Direct Method of Inferring Functional Mechanisms From Structure

A common method of inferring dynamics from structure has been to compare different crystal structures for a given enzyme and determine after suitable superimposition the regions that show the largest movements. For example, the swiveling motion that allows for the rotation of the RT pol domain relative to the rest of the molecule has been identified by such a comparative study.[13] Here we predict the collective dynamics for a given structure, based on its overall architecture, or topology of contacts. Two major premises in such studies of "structure-based dynamics" are as follows: (a) the overall three-dimensional (3D) structure of the molecule uniquely determines the types (or modes) of collective conformational motions that are undergone at equilibrium and (b) the most cooperative (global) modes among these motions have functional implications, on the premise that each particular 3D structure has evolved to perform a well-defined function. So, the dynamics is assumed to be "encoded" by structure, and biological function is ensured by dynamics.

Each protein has a well-defined 3D architecture. Each architecture, in turn, has access to a well-defined distribution of relaxation motions, or collective modes.[26] These modes allow for functional flexibility while maintaining the overall stability or structural integrity. The lowest frequency modes among these, also called global modes, usually embody the entire molecule. Their cooperative nature and uniqueness for a given architecture lend support to their direct involvement in functional mechanisms. A wealth of theoretical[27,28] and experimental[29,30] studies provide evidence for the close link between dynamics and function.[31,32]

A variety of methods based on principal component analysis (PCA) have been developed for extracting information on global dynamics from known structures and their molecular dynamics (MD) trajectories. The two most widely used computational methods have been normal mode analysis (NMA)[33-36] and essential dynamics analysis (EDA).[37] Domain movements relevant to function have been determined for several proteins by using NMA[38–43] or MD coupled with EDA or PCA.[27,44–50] The ANM approach presently adopted can also be viewed as a simplified form of NMA, as explained in more details in Model and Methods, with the major advantages of being much simpler, both conceptually and computationally, and thereby particularly useful for analyzing the equilibrium dynamics of large multimeric proteins or complexes, as shown in two recent applications.[51,52]

Mechanical Versus Chemical Effects in Determining Structure and Dynamics

Several independent studies suggest that the global dynamics of large systems are insensitive to the details of the model and parameters.[26,53–56] In line with these studies, a simple mechanical model that uses the fundamental concepts of statistical mechanics and solid state physics is adopted here. Thus, the results of this study refer to this simple mechanical model. It is clear that chemical interactions or specific energetics also contribute, in addition to mechanical drives, to shape the global modes. However, our understanding is that the specificities play a dominant role in selecting the structure, rather than the dynamics. And once a given structure is selected, the dynamics can be described as a mechanical process to a good approximation. The (bonded and nonbonded) connectivity of the 3D structure, or the contact topology of the network of interactions stabilizing the folded state, is the major determinant of dynamics in this case.

From another point of view, specific interactions, or energetics, determine the global energy well; on the other hand, the equilibrium motions are determined by the curvature of the energy surface near the well. The depth of the energy well is important for selecting the most stable (native) structure, whereas the curvature is important for defining the fluctuations near the global minimum. The curvature is the major ingredient and the only adjustable parameter of the GNM[16]; it is represented by the spring constant γ associated with inter-residue potentials. In parallel with the approximation originally proposed by Tirion[57] for atomic interactions in folded structures, we adopt a uniform harmonic potential (or identical γ values) for all pairs of interacting residues. We note that Hinsen and Kneller[58] also used a uniform (distance-dependent) force constant for all residue pairs (represented by their α-carbons) to show that the domain movements found by such coarse-grained models and uniform potentials can

closely reproduce those predicted by detailed, atomic NMA.[38,54,58]

MODEL AND METHOD

In the GNM, the dynamics of the interactions is controlled by the connectivity (or Kirchhoff) matrix $\mathbf{\Gamma}$, by analogy with the statistical mechanical theory of elasticity originally developed by Flory and coworkers for polymer networks.[14,15,59] The elements of $\mathbf{\Gamma}$ are defined as[16]

$$\Gamma_{ij} = \begin{cases} -1 & \text{if } i \neq j \text{ and } R_{ij} \leq r_c \\ 0 & \text{if } i \neq j \text{ and } R_{ij} \geq r_c \\ -\sum_{i,i\neq j} \Gamma_{ij} & \text{if } i = j \end{cases} \quad (1)$$

Here r_c is the cutoff distance defining the range of interaction of residues, each residue being represented by its α-carbon, and $\mathbf{R_{ij}}$ is the distance between ith and jth residues. The inhibitor is also included in $\mathbf{\Gamma}$, by taking an additional (N+1)th interaction site (or network node) located at the centroid of the inhibitor. The value of $r_c = 7$ Å includes the neighboring residues located in the first coordination shell near a central residue.[60] This value has been adopted in previous GNM studies.[16,21,22,24] Multiplied by γ, $\mathbf{\Gamma}$ may be viewed as the matrix of second derivatives of inter-residue interactions, for the special case of all pairs interacting via a uniform harmonic potential. Comparison with experimental B-factors showed that a force constant γ of the order of 1.0 kcal/(mol · Å^2) is needed to quantitatively reproduce the size of residue fluctuations. The same force constant is found with a cutoff distance of 12–15 Å in the ANM.[25] $\mathbf{\Gamma}$ gives a complete description of the connectivity of the network. The ith diagonal element of $\mathbf{\Gamma}$ characterizes the local packing density or the coordination number of residue i, $1 \leq i \leq n$, for a protein of n residues.

The equilibrium correlations between the fluctuations $\Delta\mathbf{R}_i$ and $\Delta\mathbf{R}_j$ of residues i and j are given by[14–16]

$$\langle \Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j \rangle = (k_B T/\gamma)[\Gamma^{-1}]_{ij} \quad (2)$$

where k_B is the Boltzmann constant, T is the absolute temperature, and $[\Gamma^{-1}]_{ij}$ denotes the ijth element of the inverse of $\mathbf{\Gamma}$. The mean-square fluctuations of residues in the global (lowest frequency) GNM mode are calculated by eigenvalue decomposition of $\mathbf{\Gamma}^{-1}$ as[17] as

$$(\Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_i)_1 = (3k_B T/\gamma)\,[\lambda_1^{-1}\mathbf{u}_1\mathbf{u}_1^T]_{ii} \quad (3)$$

where λ_1 is the smallest nonzero eigenvalue of $\mathbf{\Gamma}$ and $\mathbf{u}_1$ is the corresponding eigenvector reflecting the global mode frequency and shape, respectively, the subscript 1 refers to mode 1 (global mode), and the subscript ii refers to the ith diagonal element of the N × N matrix enclosed in square brackets. We note that the columns (or rows) of $\mathbf{\Gamma}$ are interdependent (all sum up to zero), and thus $\mathbf{\Gamma}$ cannot be inverted; instead, it is reconstructed after removal of its zero eigenvalue and corresponding eigenvector.

In the ANM, $\mathbf{\Gamma}$ is replaced by its 3N × 3N counterpart $(1/\gamma)\mathcal{H}$ where $\mathcal{H}$ is the Hessian matrix of the second derivatives of the intramolecular potential $V = (\gamma/2)\,\Delta\mathbf{R}^T\,\mathbf{\Gamma}\,\Delta\mathbf{R}$.[25] $\mathcal{H}$ is a 3N × 3N symmetric matrix composed of N × N superelements $\mathcal{H}_{ij}$ each of size 3 × 3, given by the second derivatives of V with respect to position vectors $\mathbf{r}_i$ and $\mathbf{r}_j$. For more details and explicit expressions of the elements of $\mathcal{H}$, the reader is referred to our recent article.[25] As discussed in detail therein, $\mathcal{H}$ is equivalent to the second derivative matrix $\mathcal{H}$ calculated in NMA near a local energy minimum, with the major simplification that a uniform harmonic potential is adopted for all pairs, as originally proposed by Tirion.[57] NMA usually considers the mass-weighted second derivative (i.e., the product $\mathbf{M}^{-1/2}\,\mathcal{H}\,\mathbf{M}^{1/2}$ where $\mathbf{M}$ is the diagonal 3N × 3N matrix of atomic masses, for extracting normal modes. ANM differs from the classical NMA of proteins in that all interacting pairs are assumed to be identical; this removes the need for introducing masses or specific potentials. Furthermore, the analysis is performed at the residue level, and there is no need to prior energy minimization, the PDB structure being directly adopted as the energy-minimized conformation, in conformity with the GNM approach.[16] With these approximations there is actually no need to invoke energetics, and the elements of $\mathcal{H}$ can be readily found by geometrical and mechanical considerations. See for example the force balance presented in Ref. 25 for deriving $\mathcal{H}$, and the proof of its equivalence to the matrix of second derivatives commonly used in NMA.

It is clear that the collective motions of HIV-1 RT result from the superimposition of 3N-6 modes in this model, the contribution of the individual modes scaling with their inverse frequencies (or eigenvalues). The low-frequency modes are the most robust, insensitive to the details of the adopted model, and usually a subset of slow modes account for the dominant mechanisms of motion. The molecular dynamics of RT resulting from the collective contribution of all modes has been explored in other detailed simulations.[61,62] We focus here on the slowest (global) mode among these, which has a dominant effect on the longtime dynamics of the molecule. We have also examined a series of low-frequency mode shapes for HIV-1 RT (not shown), and for other proteins,[51,52] we chose to report here the results for the global mode of HIV-1 RT because these directly relate to the interference of inhibitor binding with functional motions.

RESULTS AND DISCUSSION

Ligand Binding Has Minor Effect on the Size of Residue Motions

Figure 1 displays the global mode shapes of three RT structures: unliganded (solid curve), liganded with nevirapine (dashed), and with efavirenz (dotted) for the subunits p66 (left) and p51 (right). The global mode shape represents the distribution of the square displacements, $[(\Delta R_i)^2]_1$ for $1 \leq i \leq N$, as driven by the lowest frequency mode of motion. It is a unique function, or a fingerprint, of the particular 3D structure. The main structural difference between unliganded and liganded structures is the closed conformation of p66 thumb and finger domains in unliganded RT. The p66 thumb and fingers of efavirenz-bound RT has a more open conformation than those of nevirapine-

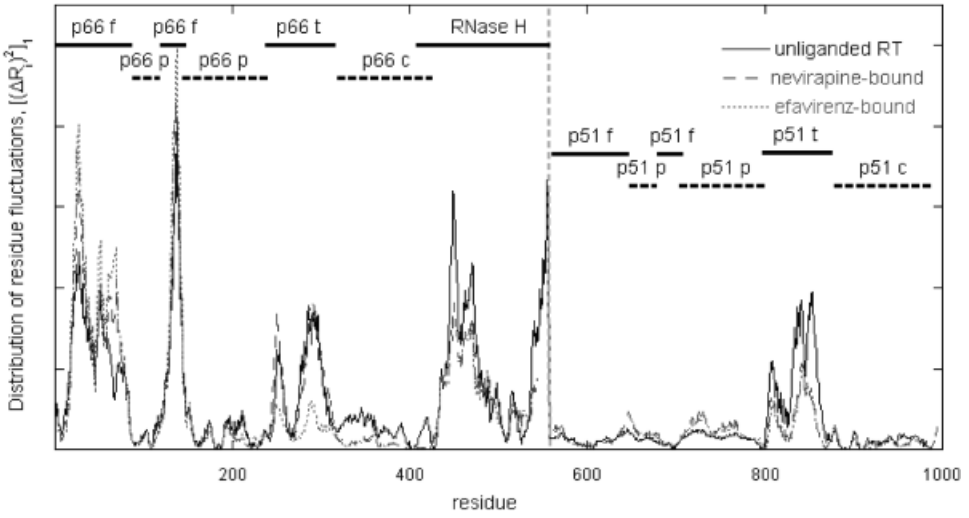

Fig. 1. Distribution of mean-square fluctuations of HIV-1 RT residues in the global motions of the unliganded form of RT (solid), and the forms liganded with the NNRTIs nevirapine (dashed) and efavirenz (dotted). The different structural domains of subunits p66 and p51 are indicated by the upper bars. Note the close superposition of the curves, except at the p66 thumb region.

bound RT. The fingers, thumb, and RNase H subdomains of p66 and the thumb of p51 are found to enjoy high flexibilities, as evidenced by the peaks in the global mode shape. These results are consistent with our earlier calculations,[24] molecular dynamics simulations,[61,62] and previous comparisons of crystal structures.[63–65] The largest size motions occur at the p66 fingers in all three structures. The palm and connection subdomains of p66, on the other hand, are severely constrained. Minima at the palm and connection indicate the regions acting as hinge-bending sites in the global mode. The p51 thumb is the only subdomain of subunit p51, which moves, all other subdomains being practically frozen in the global mode. It is of interest that the p51 thumb has been pointed out to be essential for the activation of HIV-1 RT.[66]

An intriguing feature in Figure 1 is the rather small difference in the fluctuation behaviors of the wild-type RT and the RTs complexed with NNRTIs. The RNase H and p51 thumb flexibilities appear to be slightly reduced on inhibitor binding, whereas that of the p66 fingers shows a slight increase. The most strongly affected region is the p66 thumb in the RT complexed with efavirenz.

The weak effect of ligand binding on the collective dynamics of RT is contrary to expectations because these ligands are drugs that inhibit the enzyme activity and should be expected to impede the dynamics of the molecule. How does the NNRTI interfere with the dynamics of the enzyme, if the flexibilities of the individual residues are unaffected? The answer lies in the directionalities of motions.

The Directions of Domain Motions Are Changed by Ligand Binding

Figure 2 displays the global conformational changes computed for nevirapine-bound (A) and unliganded RTs (B). The left and right diagrams, labeled (I) and (II), are the conformations between which the molecule fluctuates at equilibrium under the action of the global mode. They are generated by adding (I) or subtracting (II) the ANM-derived fluctuation vectors $\Delta\mathbf{R}_i$ of individual residues to/from their X-ray crystallographic position vectors $\mathbf{R}_i$. Residue positions are identified with the coordinates of the Cα atoms. The p66 subdomains are colored cyan (fingers), yellow (palm), red (thumb), green (connection), and pink (RNase H) for clarity. The displayed structures between which RT fluctuates are reminiscent of the open and closed conformations of p66.[67]

The most striking aspect of Figure 2 is the drastic difference in the mechanisms of global fluctuations for the liganded and unliganded RT. This difference is significant given that the sizes or distributions of fluctuations were practically unaffected by ligand binding (see Fig. 1).

In the form complexed with nevirapine [Fig. 2(A)], the p66 fingers and RNase H fluctuate in opposite directions (anticorrelated motions), giving rise to alternating open (I) and closed (II) conformations. The palm and connection serve as a rigid support for these flexible regions. The p66 thumb, on the other hand, is subject to orthogonal, but cooperative, motions with respect to the p66 fingers and RNase H. It projects away from the p66 palm and connection when the fingers and RNase H close down (II) and is folded back onto the connection (I) when the fingers and RNase H open up. The alternating opening/closing motions of the fingers and RNase H are thus accompanied by the concerted folding/dissociation of the thumb onto/from the p66 connection. We note that Sarafianos et al.[65] drew attention to the closing down of the fingers on the dNTP at the polymerase active site in conformity with present predictions.

The p66 thumb and fingers of the unliganded RT, on the other hand, form a highly unified block that undergoes in-phase oscillations about the p66 palm, the latter serving

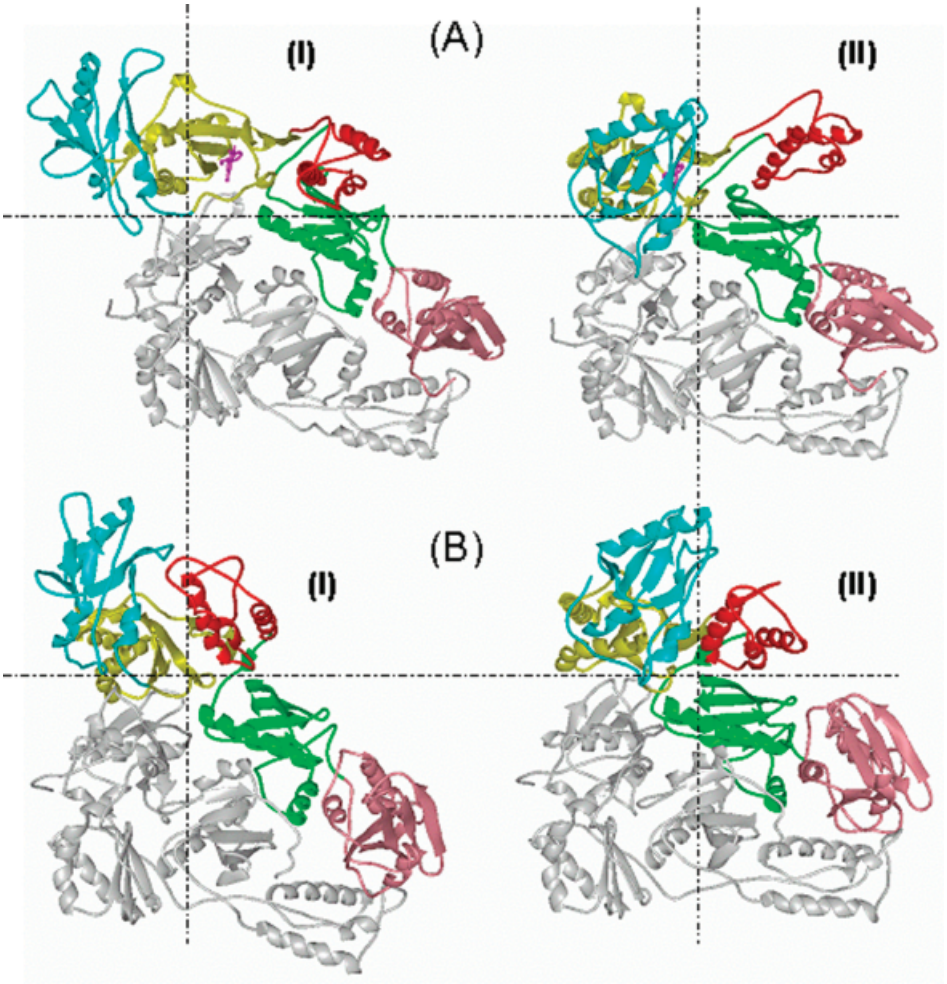

Fig. 2. Fluctuating conformations of the nevirapine-bound (**A**) and unliganded (**B**) forms of RT. The left (I) and right (II) diagrams are the conformations between which the molecule fluctuates, as derived by the ANM computation[25] of global dynamics. The p66 subdomains are colored cyan (fingers), yellow (palm), red (thumb), green (connection), and pink (RNase H). Note the significant difference in the orientations of domain motions in structures (A) and (B), despite the unaffected flexibility of the individual domains (Fig. 1).

as an anchor. The RNase H domain and the p51 thumb also perform en bloc motions. Their motions are negatively correlated with those of the p66 fingers and thumb. In contrast to the liganded RT, the motions of the RNase H, the p66 finger and the p51 thumb are not confined to the plane of the paper, but they enjoy an additional degree of freedom along the perpendicular direction (see below). This enhanced flexibility, as well as the concerted swinging motions of the p66 fingers and thumb about the palm, leave a relatively larger accessible surface between the p66 thumb-finger and RNase H domain, which should facilitate the optimal translocation and processing of the polynucleotide.

Drug Binding Selectively Hinders Certain Domain's Reorientations

Figures 1 and 2 indicate that nevirapine binding changes the directionality of the domain movements, without restricting their amplitudes. This effect can be seen more clearly in Figure 3. The figure shows the components ΔX_i, ΔY_i, and ΔZ_i of ΔR_i for unliganded (solid curve) and liganded (dashed curve) RT. See the ribbon diagrams in Figure 3 for the definitions of the *X*, *Y*, and *Z* axes. The p66 fingers, thumb, RNase H, and p51 thumb are colored blue, red, pink, and magenta, respectively. The *X* axis coincides with the out-of-plane direction with respect to the view presented in Figure 2; the *Y* and *Z* axes are in-plane.

Figure 3 shows that the out-of-plane fluctuations of all domains, except for p66 thumb, are suppressed in the ligand-bound form (upper panel); on the other hand, in the unliganded form, the p66 fingers and thumb fluctuate in opposite direction with respect to the RNase H and p51 thumb, allowing for a cooperative enlargement or contraction of the nucleotide-processing cavity between these two regions. The RNase H and the p51 thumb remain strongly coupled irrespective of NNRTI binding (upper and lower panels). Finally, the p66 fingers tend to be decoupled from the p66 thumb in the presence of nevirapine. See the sharp extrema at the p66 fingers in Figure 3 (middle panel), and the inversion in the direction of fluctuations in the lower

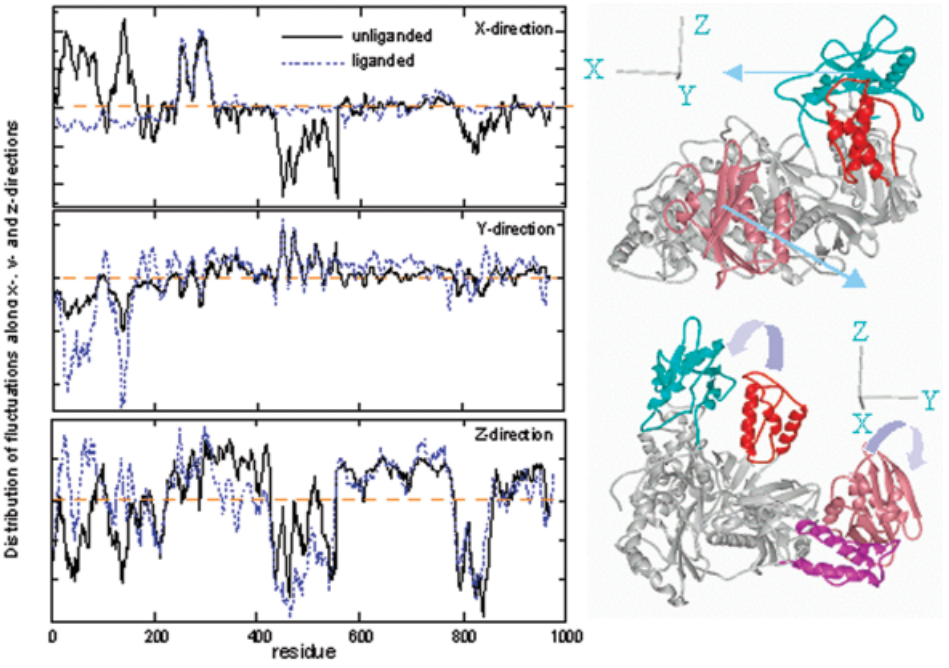

Fig. 3. Residue fluctuations along the X-, Y-, and Z-directions (ΔX_i, ΔY_i, and ΔZ_i) for unliganded (solid curve) and nevirapine bound (dashed curve) RT. The *X* axis coincides with the out-of-plane direction with respect to the view shown in Fig. 2; the *Y* and *Z* axes lie along the in-plane directions. See the ribbon diagrams on the right. The p66 fingers, thumb, RNase H, and the p51 thumb are colored blue, red, pink, and magenta, respectively.

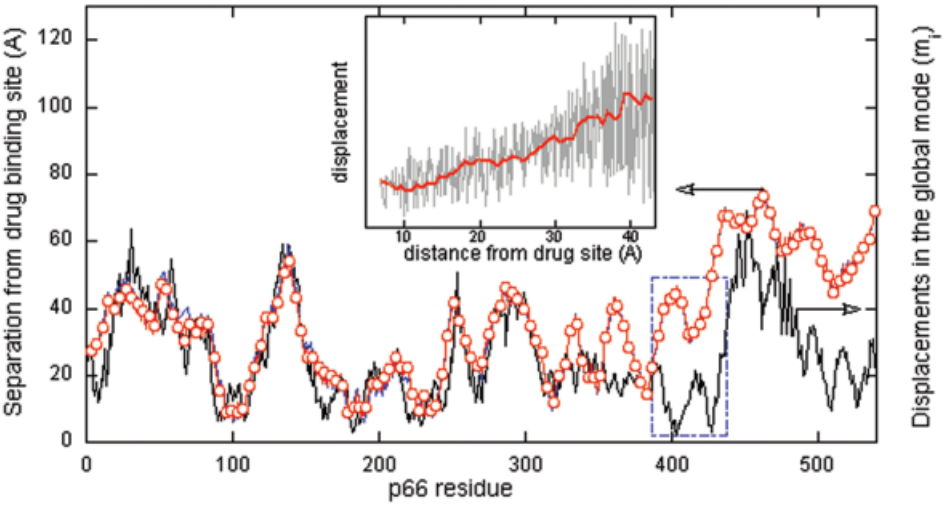

Fig. 4. Mobilities $m_i = [[(\Delta R_i)^2]_1/(\langle z \rangle / z_i)]^{1/2}$ of p66 residues ($1 \leq i \leq 560$) in the first (most cooperative) mode (continuous curve; right ordinate) compared with the separation (l_i) of individual residues from drug-binding site (curve with open circles, left ordinate). z_i is the coordination number of residue i, $\langle z \rangle$ is the average over all z_i. The superposition of the curves in $1 \leq i \leq 311$ (all p66 subdomains, excluding the connection and RNase H domain) evidences the binding of the drug to the global hinge center that coordinates the movements of the p66 fingers and thumb. The inset shows the correlation between m_i and l_i.

panel. The coupling of the p66 fingers and thumb to cooperatively sequester the nucleotide-binding cleft appears thus to be hampered by NNRTI binding.

Comparison of the Two Different NNRTIs-Bound Forms of RT

The efavirenz-bound form exhibits approximately the same mechanism of fluctuations as the nevirapine-bound structure. The only, but important, difference is the suppression of the mobility of the p66 thumb; the out-of-plane fluctuations of the thumb, which were shown above to persist even in the nevirapine-bound form, are hindered in the efavirenz-bound form. This more effective hindering of the thumb motion in the global mode of RT could be a reason for the stronger efficiency of efavirenz, compared to nevirapine, in inhibiting RT.

The reader is referred to the site http://www.ccbb/pitt.edu/hiv for the movies of the global fluctuations of the three (unliganded, nevirapine-bound, and efavirenz-bound) structures of RT.

Inhibitor Efficiency Correlates With the Susceptibility of the Binding Site

In the zeroth approximation, the GNM-predicted mean-square (ms) fluctuations of residues are inversely proportional to their local packing density. The local packing density near each residue is represented by the number of residues, z_i, within its first coordination shell. This approximation follows from the expansion of the inverse Kirchhoff matrix of contacts as a series of matrices, the first term of which is a diagonal matrix of elements $1/z_i$. Clearly, more exposed residues enjoy a greater mobility, irrespective of the topology of inter-residue contacts.

The fluctuations predicted by the GNM include both the effect of local packing density and the higher order effects imparted by contact topology. It is possible to extract the net effect of contact topology by normalizing the GNM-predicted ms fluctuations with respect to $1/z_i$ values. To this aim, we define a new measure of mobility, $m_i = [[(\Delta R_i)^2]_1/(\langle z \rangle / z_i)]^{1/2}$, where $\langle z \rangle$ is the average over all z_i values. m_i reflects the linear displacements of individual residues in the global mode, after eliminating the changes induced by local packing density fluctuations.

The resulting mobilities of the p66 residues in the global mode are shown by the solid curve in Figure 4. This curve is qualitatively similar to the distribution shown in Figure 1, which indicates that the local packing density changes do not have a significant influence on the molecular machinery driven by mode 1. The topology of contacts is the major determinant of global dynamics. Local minima occur at residues 89–95, 107–110, 161–165, 180–188, 219–231, 397–408, and 425–429. See also Figure 5(A). These residues occupy critical loci (minima) in the global mode and could all be viewed as susceptible sites for controlling the molecular machinery. For example, most of these residues interact with DNA:DNA[63] and polypurine tract RNA:DNA[68] complexes. It is not surprising that the catalytic triad of Asp residues and the drug-resistant mutations M184[13] coincide with these sites. However, not all of these sites are easily accessible because of the specific topology of contacts. Some other drug-resistant mutants, such as K103N,[64] select more accessible residues that closely interact with these susceptible sites. Efficient recognition is indeed a prerequisite for binding and inhibition, and an efficient binding, in turn, requires a certain conformational freedom to accommodate or optimize the interaction with the substrate.

Figure 5 also shows the distance l_i between the ith residue and the inhibitor (left ordinate; curve with open dots). l_i is found from the average distance between the ith α-carbon and all inhibitor atoms. The l_i values for the nevirapine-bound and efavirenz-bound forms are almost indistinguishable in this coarse-grained, residue-based calculation.

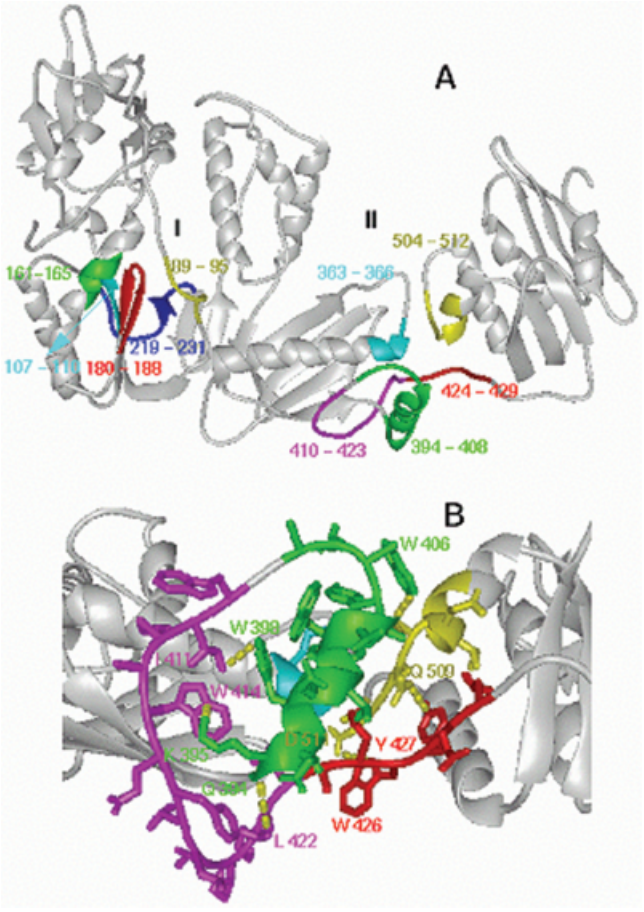

Fig. 5. **A:** Two hinge-bending centers on RT forming minima in Figure 1 (or 4): (I) near the NNRTI binding site, involving residues 107–110 (cyan), 161–165 (green), 180–188 (red) and 219–231 (blue), and (II) near the p66 connection and RNase H interface, comprising residues 363–366 (cyan), 394–408 (green), 410–423 (loop, magenta), 424–429 (interdomain linker, red), and 504–512 (yellow). **B:** A closer view of region II, showing explicitly the side-chains near the hinge site. Close tertiary contacts are indicated by the yellow dots.

A striking observation is the close superposition of the m_i and l_i curves in Figure 4, at the thumb, palm, and fingers of subunit p66. Thus, the mobilities of these three subdomains scale linearly with their separation from the drug-binding site, after eliminating the perturbations introduced by local density fluctuations. The physical meaning of this—at first unexpected—observation is that the drug precisely binds to the hinge site that coordinates the movements of these subdomains. In fact, the mobilities simply increase with the moment arm with respect to the hinge center.

Figure 4 provides a firm evidence for the binding of the inhibitor to a central hinge site about which the p66 fingers and thumb undergo global rotations. The hinge site is the most susceptible region of the molecule from the dynamic point of view in the sense that a local perturbation, or interference, at this site can give rise to allosteric effects or even disrupt the entire cooperativity of the functional motions. And the binding near the hinge site

may be the essence of the efficiency of inhibition of RT by the NNRTIs, in the absence of drug-resistant mutations.

Targeting Hinge-Bending Sites Is an Effective Means of Inhibiting Function

It is noteworthy that (a) drug resistance is usually mediated by a single amino acid substitution and (b) the structural changes induced by the mutations that confer resistance to inhibitors are usually highly localized. Yet the effects of such local changes on the dynamics of the molecule are far stronger than expected. This is due to their localization on critical sites controlling the global mechanics of the molecule. The juxtaposition of hinge sites and catalytic sites in many enzymes is not probably a coincidence, but a functional requirement. In several instances, substrate binding does indeed lock the structure in a closed conformation with respect to the binding site and eradicates one or more degrees of freedom.[69] The hinge site of RT also neighbors the catalytic Asp residues. Substitutions at the adjacent residues Tyr183 and Met184 impair the efficiency of reverse transcription.[70] This and several other examples suggest that the sites acting as hinges in the global domain movements can be conceived as potential targets for inhibitors.

A Hinge Site That Can Impede RNase H Machinery if Targeted by an Inhibitor

The proportionality between the m_i and l_i curves in Figure 4 vanishes starting approximately from the p66 connection subdomain. This is consistent with the fact that the NNRTIs target the hinge site near the interface of the p66 palm with the p66 fingers and thumb. The movement controlled by NNRTI binding does not involve the connection subdomain and the RNase H domain of the p66 subunit nor the entire p51 unit. One might then ask if it is possible to identify another target site for inhibitor binding that could interfere with the RNase H activity.

The mobility curve in Figure 4 invites attention to another susceptible region that serves as a hinge site for the global movements of the RNase H domain with respect to the pol domain of RT. This site refers to the minima enclosed in a dotted square in Figure 4. It is comprised of a stretch of about 20 residues at the interface between the p66 connection and RNase H domain, illustrated in Figure 5(B). The deepest minima occur at the helical residues 394–408 (green) and the residues 424–429 (red) on the interdomain linker. These two segments are connected with a relatively more flexible region prone to substrate recognition, the loop 410–423 (magenta).

Figure 5(A) shows the entire p66 subunit to illustrate the two regions of interest, which form the minima in Figure 1 or 4: (I) the hinge region near the NNRTI binding site and (II) the newly proposed target site controlling the global reorientation of the RNase H domain. Two noncontiguous sequences, 363–366 and 504–512, participate in the stabilization of region II (colored cyan and yellow in Fig. 5). Tertiary contacts at region II include the amino-aromatic interactions Lys395-Trp414, Gln509-Tyr427, and Asp511-Trp426, and the hydrophobic pairs Trp398- Ile411, Thr400-Leu425, and Trp406-Ala508. Exposed residues that could potentially serve as targets for inhibitor recognition are Glu415 and Asn418 on the loop connecting the hinge sites, and more importantly, the polar and highly flexible residue Gln428 directly located at the linker between the p66 connection and RNase H domain.

CONCLUSION

A new feature presently elucidated is the net effect of inhibitor binding on the orientation of global motions. It is striking that nevirapine binding imparts a significant change in the orientation of global motions, as illustrated in Figures 2 and 3, whereas the sizes of motions are marginally affected (Fig. 1). The changes in orientation, particularly those at the p66 thumb, could be correlated with those observed by site-directed spin labeling mutagenesis.[67] Efavirenz, the second generation NNRTI known to inhibit RT more effectively, has the additional effect of effectively suppressing the amplitude of motion at the p66 thumb, in addition to affecting the orientations of domain movements.

Thus, two mechanisms appear to underlie drug inhibition: (a) changing the directions of domain fluctuations and (b) obstructing them all together. The former is operative on nevirapine binding, whereas efavirenz has the stronger effects of involving changes both in the size and orientation of motions. The efficiency of these inhibitors originates from their interference with the global hinge-bending site at the connection of the p66 fingers and thumb with the palm subdomain of p66. Our detailed analysis of the molecular mechanisms of domain movements reveals another hinge site controlling the movements of the RNase H domain. It remains to be seen how effectively RT function can be hampered if this second hinge-bending region is targeted by an inhibitor.

REFERENCES

1. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science 1992;256:1783–1790.
2. Das K, Ding J, Hsiou Y, Clark AD Jr, Moereels H, Koymans L, Andries K, Pauwels R, Janssen PA, Boyer PL, Clark P, Smith RH Jr, Kroeger Smith MB, Michejda CJ, Hughes SH, Arnold E. Crystal structures of 8-Cl and 9-Cl TIBO complexed with wild-type HIV-1 RT and 8-Cl TIBO complexed with the Tyr181Cys HIV-1 RT drug-resistant mutant. J Mol Biol 1996;264:1085–1100.
3. Arnold E, Jacobo-Molina A, Nanni RG, Williams RL, Lu X, Ding J, Clark AD Jr, Zhang A, Ferris AL, Clark P. Structure of HIV-1 reverse transcriptase/DNA complex at 7 A resolution showing active site locations. Nature 1992;357:85–89.
4. Jacobo-Molina A, Ding J, Nanni RG, Clark AD Jr, Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. Proc Natl Acad Sci USA 1993;90:6320–6324.
5. Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolf DJ, Debouck C, Harrison SC. The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. Proc Natl Acad Sci USA 1995;92:1222–1226.
6. Hsiou Y, Ding J, Das K, Clark AD Jr, Hughes SH, Arnold E. Structure of unliganded HIV-1 reverse transcriptase at 2.7 A resolution: implications of conformational changes for polymerization and inhibition mechanisms. Structure 1996;4:853–860.
7. Patel PH, Jacobo-Molina A, Ding J, Tantillo C, Clark AD Jr, Raag R, Nanni RG, Hughes SH, Arnold E. Insights into DNA polymer-

ization mechanisms from structure and function analysis of HIV-1 reverse transcriptase. Biochemistry 1995;34:5351–5363.
8. Sluis-Cremer N, Dmitrienko GI, Balzarini J, Camarasa MJ, Parniak MA. Human immunodeficiency virus type 1 reverse transcriptase dimer destabilization by 1-[Spiro[4″-amino-2″,2″ -dioxo-1″,2″ -oxathiole- 5″,3′-[2′, 5′-bis-O-(tert-butyldimethylsilyl)-beta-D-ribofuranosyl]]]-3- ethylthy mine. Biochemistry 2000;39: 1427–1433.
9. Alber F, Carloni P. Ab initio molecular dynamics studies on HIV-1 reverse transcriptase triphosphate binding site: implications for nucleoside-analog drug resistance. Protein Sci 2000;9:2535–2546.
10. Ren J, Milton J, Weaver KL, Short SA, Stuart DI, Stammers DK. Structural basis for the resilience of efavirenz (DMP-266) to drug resistance mutations in HIV-1 reverse transcriptase. Struct Fold Des 2000;8:1089–1094.
11. Eriksson MA, Pitera J, Kollman PA. Prediction of the binding free energies of new TIBO-like HIV-1 reverse transcriptase inhibitors using a combination of PROFEC, PB/SA, CMC/MD, and free energy calculations. J Med Chem 1999;42:868–881.
12. Wang J, Morin P, Wang W, Kollman PA. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. J Am Chem Soc 2001;123: 5221–5230.
13. Gao HQ, Boyer PL, Sarafianos SG, Arnold E, Hughes SH. The role of steric hindrance in 3TC resistance of human immunodeficiency virus type-1 reverse transcriptase. J Mol Biol 2000;300:403–418.
14. Flory PJ. Statistical thermodynamics of random networks. Proc R Soc Lond A 1976;351:351–380.
15. Kloczkowski A, Mark JE, Erman B. Chain dimensions and fluctuations in random elastomeric networks. 1. Phantom Gaussian networks in the undeformed state. Macromolecules 1989;22: 1423–1432.
16. Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in proteins using a single- parameter harmonic potential. Fold Des 1997;2:173–181.
17. Haliloglu T, Bahar I, Erman B. Gaussian dynamics of folded proteins. Phys Rev Lett 1997;79:3090–3093.
18. Bahar I, Wallqvist A, Covell DG, Jernigan RL. Correlation between native-state hydrogen exchange and cooperative residue fluctuations from a simple model. Biochemistry 1998;37:1067–1075.
19. Haliloglu T, Bahar I. Structure-based analysis of protein dynamics: comparison of theoretical results for hen lysozyme with X-ray diffraction and NMR relaxation data. Proteins 1999;37:654–667.
20. Bahar I, Atilgan AR, Demirel MC, Erman B. Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability. Phys Rev Lett 1998;80:2733–2736.
21. Bahar I, Jernigan RL. Vibrational dynamics of transfer RNAs: comparison of the free and synthetase-bound forms. J Mol Biol 1998;281:871–884.
22. Bahar I, Jernigan RL. Cooperative fluctuations and subunit communication in tryptophan synthase. Biochemistry 1999;38: 3478–3490.
23. Demirel MC, Atilgan AR, Jernigan RL, Erman B, Bahar I. Identification of kinetically hot residues in proteins. Protein Sci 1998;7:2522–2532.
24. Bahar I, Erman B, Jernigan RL, Atilgan AR, Covell DG. Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. J Mol Biol 1999;285:1023–1037.
25. Atilgan AR, Durrell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophys J 2001;80:505–515.
26. Kitao A, Go N. Investigating protein dynamics in collective coordinate space. Curr Opin Struct Biol 1999;9:164–169.
27. Abseher R, Nilges M. Efficient sampling in collective coordinate space. Proteins 2000;39:82–88.
28. Berendsen HJ, Hayward S. Collective protein dynamics in relation to function. Curr Opin Struct Biol 2000;10:165–169.
29. Goodman JL, Pagel MD, Stone MJ. Relationships between protein structure and dynamics from a database of NMR-derived backbone order parameters. J Mol Biol 2000;295:963–978.
30. Wand AJ. Dynamic activation of protein function: a view emerging from NMR spectroscopy. Nat Struct Biol 2001;8:926–930.
31. Frauenfelder H, McMahon B. Dynamics and function of proteins: the search for general concepts. Proc Natl Acad Sci USA 1998;95: 4795–4797.
32. Ma J, Sigler PB, Xu Z, Karplus M. A dynamic model for the allosteric mechanism of GroEL. J Mol Biol 2000;302:303–313.
33. Brooks B, Karplus M. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. Proc Natl Acad Sci USA 1983;80:6571–6575.
34. Case DA. Normal mode analysis of protein dynamics. Curr Opin Struct Biol 1994;4:285–290.
35. Go N, Noguti T, Nishikawa T. Dynamics of a small globular protein in terms of low-frequency vibrational modes. Proc Nat Acad Sci USA 1983;80:3696–3700.
36. Levitt M, Sander C, Stern PS. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. J Mol Biol 1985;181:423–447.
37. Amadei A, Linssen AB, Berendsen HJ. Essential dynamics of proteins. Proteins 1993;17:412–425.
38. Hinsen K, Thomas A, Field MJ. Analysis of domain motions in large proteins. Proteins 1999;34:369–382.
39. Miller DW, Agard DA. Enzyme specificity under dynamic control: a normal mode analysis of alpha-lytic protease. J Mol Biol 1999;286:267–278.
40. Sanejouand YH. Normal-mode analysis suggests important flexibility between the two N- terminal domains of CD4 and supports the hypothesis of a conformational change in CD4 upon HIV binding. Protein Eng 1996;9:671–677.
41. Tama F, Sanejouand YH. Conformational change of proteins arising from normal mode calculations. Protein Eng 2001;14:1–6.
42. Thomas A, Field MJ, Mouawad L, Perahia D. Analysis of the low frequency normal modes of the T-state of aspartate transcarbamylase. J Mol Biol 1996;257:1070–1087.
43. Thomas A, Hinsen K, Field MJ, Perahia D. Tertiary and quaternary conformational changes in aspartate transcarbamylase: a normal mode study. Proteins 1999;34:96–112.
44. Aalten DMF, Amadei A, Linssen ABM, Eijsink VGH, Vriend G, Berendsen HJC. The essential dynamics of thermolysin: confirmation of the hinge-bending motion and comparison of simulations in vacuum and water. Proteins 1995;22:45–54.
45. de Groot BL, Hayward S, van Aalten DM, Amadei A, Berendsen HJ. Domain motions in bacteriophage T4 lysozyme: a comparison between molecular dynamics and crystallographic data. Proteins 1998;31:116–127.
46. Garcia AE, Hummer G. Conformational dynamics of cytochrome c: correlation to hydrogen exchange. Proteins 1999;36:175–191.
47. Hayward S, Berendsen HJ. Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. Proteins 1998;30:144–154.
48. Kitao A, Hayward S, Go N. Energy landscape of a native protein: jumping-among-minima model. Proteins 1998;33:496–517.
49. Nolde SB, Arseniev AS, Orekhov VY, Billeter M. Essential domain motions in barnase revealed by MD simulations. Proteins 2002;46: 250–258.
50. Roccatano D, Mark AE, Hayward S. Investigation of the mechanism of domain closure in citrate synthase by molecular dynamics simulation. J Mol Biol 2001;310:1039–1053.
51. Isin B, Doruker P, Bahar I. Functional motions of influenza virus hemagglutinin: a structure-based analytical approach. Biophys J 2002;82:569–581.
52. Keskin O, Bahar I, Flatow D, Covell DG, Jernigan RL. Molecular mechanisms of chaperonin GroEL-GroES function. Biochemistry 2002;41:491–501.
53. Doruker P, Atilgan AR, Bahar I. Dynamics of proteins predicted by molecular dynamics simulations and analytical methods: application to α-amylase inhibitor. Proteins 2000;40:512–524.
54. Hinsen K. Analysis of domain motions by approximate normal mode calculations. Proteins 1998;33:417–429.
55. Sinha N, Nussinov R. Point mutations and sequence variability in proteins: redistributions of preexisting populations. Proc Natl Acad Sci USA 2001;98:3139–3144.
56. Tama F, Gadea FX, Marques O, Sanejouand YH. Building-block approach for determining low-frequency normal modes of macromolecules. Proteins 2000;41:1–7.
57. Tirion MM. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Phys Rev Lett 1996;77:1905–1908.
58. Hinsen K, Kneller GR. A simplified force field for describing the vibrational protein dynamics over the whole frequency range. J Chem Phys 1999;111:10766–10769.

59. Eichinger BE. Elasticity theory. I. Distribution functions for perfect phantom networks. Macromolecules 1972;5:496
60. Bahar I, Jernigan RL. Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. J Mol Biol 1997;266:195–214.
61. Madrid M, Lukin JA, Madura JD, Ding J, Arnold E. Molecular dynamics of HIV-1 reverse transcriptase indicates increased flexibility upon DNA binding. Proteins 2001;45:176–182.
62. Madrid M, Jacobo-Molina A, Ding J, Arnold E. Major subdomain rearrangement in HIV-1 reverse transcriptase simulated by molecular dynamics. Proteins 1999;35:332–337.
63. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science 1998;282:1669–1675.
64. Hsiou Y, Ding J, Das K, Clark AD Jr, Boyer PL, Lewi P, Janssen PA, Kleim JP, Rosner M, Hughes SH, Arnold E. The Lys103Asn mutation of HIV-1 RT: a novel mechanism of drug resistance. J Mol Biol 2001;309:437–445.
65. Sarafianos SG, Das K, Ding J, Boyer PL, Hughes SH, Arnold E. Touching the heart of HIV-1 drug resistance: the fingers close down on the dNTP at the polymerase active site. Chem Biol 1999;6:R137–R146.
66. Morris MC, Berducou C, Mery J, Heitz F, Divita G. The thumb domain of the P51-subunit is essential for activation of HIV reverse transcriptase. Biochemistry 1999;38:15097–15103.
67. Kensch O, Restle T, Wohrl BM, Goody RS, Steinhoff HJ. Temperature-dependent equilibrium between the open and closed conformation of the p66 subunit of HIV-1 reverse transcriptase revealed by site-directed spin labelling. J Mol Biol 2000;301:1029–1039.
68. Mao C, Sudbeck EA, Venkatachalam TK, Uckun FM. Structure-based design of non-nucleoside reverse transcriptase inhibitors of drug-resistant human immunodeficiency virus. Antivir Chem Chemother 1999;10:233–240.
69. Keskin O, Jernigan RL, Bahar I. Proteins with similar architecture exhibit similar large-scale dynamic behavior. Biophys J 2000;78:2093–2106.
70. Harris D, Yadav PN, Pandey VN. Loss of polymerase activity due to Tyr to Phe substitution in the YMDD motif of human immunodeficiency virus type-1 reverse transcriptase is compensated by Met to Val substitution within the same motif. Biochemistry 1998;37:9630–9640.
71. Ren J, Esnouf R, Garman E, Somers D, Ross C, Kirby I, Keeling J, Darby G, Jones Y, Stuart D. High resolution structures of HIV-1 RT from four RT-inhibitor complexes. Nat Struct Biol 1995;2:293–302.