


RESEARCH ARTICLE

The effect of dimerization and ligand binding on the dynamics of Kaposi's sarcoma-associated herpesvirus protease

David Bern¹ | Dror Tobi^{1,2} ¹Department of Molecular Biology, Ariel University, Ariel, Israel²Department of Computer Sciences, Ariel University, Ariel, Israel**Correspondence**Dror Tobi, Department of Computer Sciences, Ariel University, Ariel 40700, Israel.
Email: drorto@ariel.ac.il**Abstract**

The Kaposi's sarcoma-associated herpesvirus protease is essential for virus maturation. This protease functions under allosteric regulation that establishes its enzymatic activity upon dimerization. It exists in equilibrium between an inactive monomeric state and an active, weakly associating, dimeric state that is stabilized upon ligand binding. The dynamics of the protease dimer and its monomer were studied using the Gaussian network model and the anisotropic network model, and its role in mediating the allosteric regulation is demonstrated. We show that the dimer is composed of five dynamical domains. The central domain is formed upon dimerization and composed of helix five of each monomer, in addition to proximal and distal domains of each monomer. Dimerization reduces the mobility of the central domains and increases the mobility of the distal domains, in particular the binding site within them. The three slowest ANM modes of the dimer assist the protease in ligand binding, motion of the conserved Arg142 and Arg143 toward the oxyanion, and reducing the activation barrier for the tetrahedral transition state by stretching the bond that is cleaved by the protease. In addition, we show that ligand binding reduces the motion of helices $\alpha 1$ and $\alpha 5$ at the interface and explain how ligand binding can stabilize the dimer.

KEYWORDS

allostery, anisotropic network model, comparative dynamics, Gaussian network model, herpesvirus protease, Kaposi sarcoma, normal mode analysis, normal modes alignment, protein dynamics, varicella-zoster virus

1 | INTRODUCTION

Herpesviruses are one of the most common viral families, including the human herpesviruses that consist of nine herpes viruses and are known to cause illness in humans.¹ Kaposi's sarcoma-associated herpesvirus (KSHV), also known as HHV-8, is an oncogenic virus that causes Kaposi's sarcoma (KS) cancer in untreated AIDS patients. Infection with KSHV is necessary for KS to develop but is not sufficient,

and additional cofactors are needed.² It is now known that a combination of KSHV infection and host immune deficiency, either by HIV infection or iatrogenic immunodeficiency that is seen in organ transplant recipients, can cause KS cancer.³ Currently, most therapies approved against herpesviruses are based on inhibition of viral DNA replication, but these therapies suffer from poor efficiency, high toxicity, and more. Therefore, there is a growing interest in discovering new alternative therapies against herpesvirus infection.⁴

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Proteins: Structure, Function, and Bioinformatics* published by Wiley Periodicals LLC.

All herpesviruses express unique and conserved serine protease that is critical to the maturation of the virus. In KSHV, the role of the protease is to cleave the assembly protein near its C-terminus to enable maturation of KSHV procapsids.⁵ Mutants expressing an inactive form of protease in herpes simplex virus (HSV) were shown to form capsids lacking viral DNA, hence affirming the importance of the proteases.⁶ The herpesvirus proteases are an example in which allosteric regulation of an enzyme activity is achieved through the formation of quaternary structure.⁷ Regulation of the protease is believed to be concentration-dependent; when the protease level is high enough it undergoes dimerization that activates a binding site pocket located approximately 15–20 Å from the dimer interface.⁸ The KSHV Pr exist in equilibrium between an inactive monomeric state and an active, weakly associating dimeric state. The relatively weak micromolar dimerization affinity of herpesviral proteases⁹ is thought to act as a regulatory mechanism that can limit protease activity to the interior of the maturing viral particle where the concentration is sufficient to support dimerization. Thus, understanding the effect of dimerization on the dynamics of the associating monomers is fundamental for understanding its regulation.

Many studies of protein dynamics are based on principal components analysis (PCA) of biomolecular structures and dynamic simulation. These studies have proven useful in unraveling the collective modes, and, in particular, those at the low-frequency end of the mode spectrum that underlie protein equilibrium dynamics.¹⁰ Elastic network model (ENM) is a simplified normal mode analysis (NMA) of equilibrium structures using a single parameter harmonic potential,^{11,12} while essential dynamics analysis (EDA) of the covariance matrices retrieved from molecular dynamics (MD) runs¹³ and singular value decomposition (of MD or Monte Carlo [MC]) trajectories^{14–16} all fall into this category of PCA-based methods. We note that technically “dynamics” should be reserved to time-dependent processes, while NMA “motions” describe the equilibrium fluctuations of a system (near its native state). However, nowadays, the phrase “dynamics” is used rather inclusively to also describe equilibrium fluctuations. Two common ENM methods are the Gaussian network model (GNM)¹⁷ and the anisotropic network model (ANM).^{18,19} The fluctuations predicted by GNM are isotropic and, hence, describe the “size” of the motion but do not provide information on the “directions” of motions in different modes, just their sizes and sign. The fluctuations predicted by ANM provide information of the “size” and “direction” of the motions. The fluctuation predicted by the former model shows higher agreement with experiments while the latter model is more informative.²⁰

Looking at the allosteric mechanism of KSHV Pr, Nomura et al.²¹ showed that folding of helices five and six upon dimerization positions a loop that contains key components of the catalytic machinery required for transition-state stabilization during substrate hydrolysis; that is, Arg142 and Arg143. In addition, several works have shown relations between allostery and protein dynamics.^{22–26} The purpose of this work is to explore the effect of dimerization on the dynamics of the dimer and the implications for its function, effects that are beyond structural change and complementary to it. Furthermore, a

comparison between KSHV Pr and varicella-zoster virus (VZV or HHV-3), another member of HHV Pr, was carried out to assess the implications of the results to the HHV Pr family.

2 | METHODS

2.1 | Protein structures

Protein data bank (PDB) entry 2PBK⁸ representing the dimeric structure of KSHV Pr was used in the calculations. The structure shows the protein in complex with a hexapeptide transition state analog. The transition state analog inhibitors covalently phosphorylate the active site serine, freezing the enzyme structure during catalysis and stabilizing the dimeric state of the enzyme. The monomeric structures of KSHV Pr do not contain the C-terminal end (helices five and six), since they are unstructured in the monomeric state, and hence do not enable us to accurately compare the dynamics of the monomeric and dimeric states. The dynamics of the monomeric subunit was done by simply removing the other subunit from the dimeric structure and calculating its dynamics. These calculations enable us to deduce the effect of dimerization on the dynamics of the subunits per se, beyond the structural rearrangement of the C-terminal. Complementary analysis was carried out on the VZV Pr protein structure entry 1VZV²⁷ to compare the dynamics of different HHV Pr members. This structure shows the native state of the protein without an inhibitor.

2.2 | ANM and GNM calculation

2.2.1 | Gaussian network model

In the GNM calculations, each residue is represented by a single node positioned at its C^α atom.^{17,28,29} The peptidic ligands (or inhibitors) were also represented by their C^α atoms, in addition to the phosphorus atom of the phosphonate group that is covalently bound to S114. Nodes within a cutoff distance of $r_c = 7.3$ Å are connected by springs of uniform force constant γ , which leads to Gaussian fluctuations in the node positions and inter-residue distances. We define the equilibrium position of a node i by vector R_i^0 and its instantaneous position by R_i . The fluctuation, or deformation, from this mean position is defined by the vector $\Delta R_i = R_i - R_i^0$. The deformation in the distance vector R_{ij} that extends from residue i to j is defined as: $\Delta R_{ij} = R_{ij} - R_{ij}^0 = \Delta R_j - \Delta R_i$. Using the components of the deformation vector $\Delta R_{ij} = [\Delta X_{ij} \Delta Y_{ij} \Delta Z_{ij}]$, the potential V_{GNM} of two interacting residues is defined as: $V_{\text{GNM}}(i,j) = \frac{\gamma}{2} \{ (\Delta X_{ij})^2 + (\Delta Y_{ij})^2 + (\Delta Z_{ij})^2 \}$.

The topology of the network of N nodes (residues) is fully defined by the Kirchhoff matrix Γ , the elements of which are

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

In terms of individual fluctuations, the Kirchhoff matrix simplifies the expression for the potential of the entire network of N residues to

$$V_{\text{GNM}} = \frac{\gamma}{2} [\Delta X^T \Gamma \Delta X + \Delta Y^T \Gamma \Delta Y + \Delta Z^T \Gamma \Delta Z] \quad (2)$$

where the fluctuation vectors ΔX^T , ΔY^T , and ΔZ^T have the components $[\Delta X_1, \Delta X_2 \dots \Delta X_N]$, $[\Delta Y_1, \Delta Y_2 \dots \Delta Y_N]$, and $[\Delta Z_1, \Delta Z_2 \dots \Delta Z_N]$, respectively. The isotropic assumption implies that $\langle \Delta X \Delta X^T \rangle = \langle \Delta Y \Delta Y^T \rangle = \langle \Delta Z \Delta Z^T \rangle = \frac{1}{3} \langle \Delta \vec{R} \Delta \vec{R}^T \rangle$. Thus, the potential function in Equation (2) can be rewritten as: $V_{\text{GNM}} = \frac{\gamma}{2} [\Delta R^T \Gamma \Delta R]$.

The mean-square fluctuations (MSFs) of a particular residue i or the cross-correlations $\langle \Delta \vec{R}_i \cdot \Delta \vec{R}_j \rangle$ between residue fluctuations is obtained by the statistical mechanical average over all fluctuations and leads to³⁰

$$\langle \Delta \vec{R}_i \cdot \Delta \vec{R}_j \rangle = \frac{3k_B T}{\gamma} [\Gamma^{-1}]_{ij} \quad (3)$$

where $[\Gamma^{-1}]_{ij}$ denotes the ij th element of the inverse of Γ . The MSF is simply found by replacing j with i in Equation 3, and B-factor calculation of residue i is defined as: $B_i = \frac{8\pi^2}{3} \langle \Delta \vec{R}_i \cdot \Delta \vec{R}_i \rangle$.

2.2.2 | Mode analysis

The motions along different GNM modes are found by eigenvalue decomposition $\Gamma = \mathbf{U} \Lambda \mathbf{U}^{-1}$, where \mathbf{U} is the orthogonal matrix of eigenvectors u_k of Γ and Λ is the diagonal matrix of the eigenvalues (λ_k) , $1 \leq k \leq N$. The eigenvalues represent the frequencies of the $N - 1$ nonzero GNM modes and are organized in ascending order such that $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{N-1}$ and $\lambda_N = 0$. The i th element $(u_k)_i$ of the k th eigenvector describes the fluctuation (deformation) of residue i from its equilibrium position along the k th principal coordinate. The MSF of residue i can be rewritten as a weighted sum of the square fluctuations driven by all modes as:

$$\langle \Delta \vec{R}_i^2 \rangle = \sum_k [\Delta \vec{R}_i^2]_k = 3k_B T / \gamma \sum_k [\lambda_k^{-1} (\vec{u}_k)_i]^2 \quad (4)$$

GNM enables us to predict the relative sizes of motions accessed by different modes, not their directions, the GNM fluctuations being isotropic by definition. High fluctuations in absolute values correspond to protein structural regions with high mobility and vice versa. Regions with an opposite GNM sign move in anticorrelated manner.^{17,28,29} The directions of collective motions can be characterized by the ANM.¹⁸

2.2.3 | Anisotropic network model

The ANM^{18,31} Hessian matrix \mathbf{H} is based on a harmonic potential of the form

$$V_{\text{ANM}} = \frac{\gamma}{2} \sum_{ij}^N (R_{ij} - R_{ij}^0)^2 \Gamma_{ij} \quad (5)$$

where R_{ij}^0 and R_{ij} are the original (native state) and deformed (by ANM modes) distances between residues i and j . A relatively large cutoff value ($r_c = 15 \text{ \AA}$) is adopted in the ANM to ensure the coherence of the network. Eigenvalue decomposition of \mathbf{H} yields $3N - 6$ eigenvectors $(\vec{u}_k^{\text{ANM}})^T = [u_k^{X1}, u_k^{Y1}, u_k^{Z1}, \dots, u_k^{ZN}]$, the components of which describe the motions of the N residues in the X-, Y-, and Z- directions according to the k th mode. The components of $3N$ -dimensional vector \vec{u}_k^{ANM} are conveniently organized into N superelements of size 3×1 , designated as $[\vec{u}_k^{\text{ANM}}]_i$, each corresponding to a given residue.

The first or slowest mode (of GNM or ANM) is defined as the mode with the smallest nonzero eigenvalue and so on. The higher the mode number, the higher the energetic cost of the protein to move along this mode, and the motions become more local and less global.

2.3 | Global alignment of anisotropic network model modes

The Needleman-Wunsch dynamic programming algorithm³² was modified for aligning ANM modes. ANM mode analysis results in a set of vectors $\{\vec{U}\}$ describing the deformation of residues from their equilibrium position (native structure) in the Cartesian space. Let \vec{U}_i^k be the deformation vector of residue i in mode k of one protein and \vec{V}_j^l the deformation vector of residue j in mode l of another protein. The score for residues i and j upon alignment of modes k and l is defined as:

$$S_{ij} = \frac{\vec{U}_i^k \cdot \vec{V}_j^l}{|\vec{U}_i^k| |\vec{V}_j^l|} - C, \text{ where } 0 \leq C \leq 1 \quad (6)$$

where S_{ij} will be positive if their cosine value is greater than C that is the two deformation vectors pointing in the same direction and negative if their cosine is smaller than C . Here, we used $C = 0.7$ radians ($\sim 40^\circ$) to define the threshold for vector similarity. In the case of alignment of homologous or identical proteins, it is possible to guide the algorithm to prefer the matching of spatially close residues by applying distance constraints. Distance constraints were applied in the present work by modifying the alignment score S_{ij} as follows:

$$S_{ij} = \begin{cases} \frac{\vec{U}_i^k \cdot \vec{V}_j^l}{|\vec{U}_i^k| |\vec{V}_j^l|} - C, & r_{ij} \leq R_c \\ -1, & r_{ij} > R_c \end{cases} \quad (7)$$

where r_{ij} is the C^α distance between residues i and j and R_c is the cutoff distance set here to 10 \AA . The algorithm enables alignment of two modes even if their length is different as it creates insertions/

deletions along the alignment to obtain the optimal alignment. Gap opening and gap extension penalties were set to 1.0 and 0.1, respectively. Increasing or decreasing the penalties by a factor of two did not change the results significantly. After the alignment process, to obtain a unified alignment score that is not dependent on the alignment length, an average alignment score (AAS) was calculated. The AAS is calculated as the average S_{ij} along the alignment with the exception of gap position and distance constraint regions that received a score of zero. This algorithm is an extension of our previous GNM modes alignment³³ and was previously reported.³⁴

2.4 | Clustering calculations

Clustering calculations were performed in R environment for statistical computing and graphics.³⁵ Clustering was performed using the hierarchical clustering function “hclust” with default parameters after converting the GNM cross-correlation matrix into a distance matrix.

3 | RESULTS

The KSHV Pr is active as a weakly associating dimer. Upon dimerization, helix five ($\alpha 5$) and six ($\alpha 6$), which are unstructured in the monomeric state, refold with the former being at the newly formed interface. Over 80% of the dimer interface of KSHV Pr is formed by helices $\alpha 1$ and $\alpha 5$, where $\alpha 5$ of each monomer is packed between helices one and five of the other monomer.⁵ The homodimeric structure of KSHV Pr is presented in Figure 1. Each monomer contains a

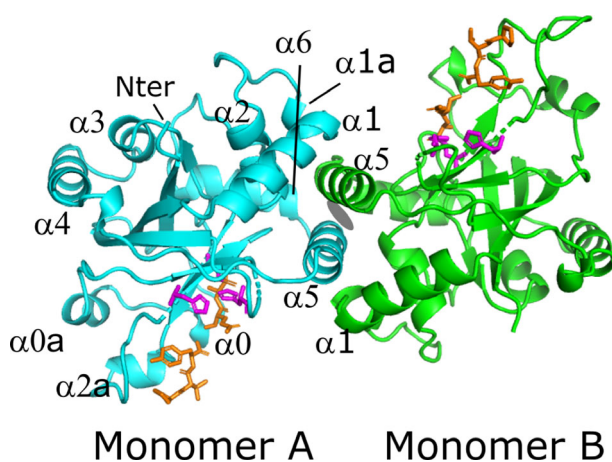


FIGURE 1 The dimeric structure of KSHV Pr. The protein is depicted using cartoon representation with monomers A and B colored cyan and green, respectively. The Ser-His-His catalytic triad of each monomer and the inhibitory peptides are shown using stick representation and colored purple and orange, respectively. All helices of monomer A and helices one ($\alpha 1$) and five ($\alpha 5$) of monomer B are marked. Helices $\alpha 1$ and $\alpha 5$ of each monomer form the interface. The dimer axes are shown using gray ellipse. All protein figures were done using PyMOL software.³⁶ KSHV, Kaposi's sarcoma-associated herpesvirus

noncanonical Ser-His-His catalytic triad (represented using purple sticks) and a substrate (orange sticks) within a substrate binding pocket located approximately 15–20 Å from the dimer interface.⁸

Theoretical B-factors were calculated for the dimer and each of its monomers separately using GNM analysis. The correlation coefficient between theoretical and experimental B-factors is 0.73 indicating good agreement between the two methods. Figure 2 presents the calculated B-factors of monomer A (A) and B (B) in the dimeric (blue line) and the monomeric (red line) states. The B-factor values of $\alpha 5$ (black bar) are lower in the dimer state by 50%, on average, compared to the monomeric state. The opposite effect is seen for amino acids at the binding site (gray bar). The mobility of these amino acids is increased by 36%, on average, compared to the monomeric state. Figure 2C depicts the KSHV Pr dimeric structure colored according to the ratio between the dimer and monomer state B-factors. Regions with lower mobility (or B-factor values) in the dimeric state, compared to the monomeric state, are colored in blue while regions with higher mobility are colored red. In addition to the binding site, the β -sheet underneath it also shows higher mobility in the dimeric state. The above calculations were done without including inhibitory peptides. However, inclusion of the inhibitory ligands did not significantly change the results. Thus, the dimerization forms a global hinge region at the interface, reduces the mobility of the interfacial region especially that of $\alpha 5$, and increases the mobility of the binding site.

Amino acids within the dimer can move together in a correlated manner, or in opposite directions in an anticorrelated manner. The cross-correlation map obtained from the GNM analysis is presented in Figure 3A. A negative value (red color) refers to an anticorrelation between residue fluctuations, whereas a positive value indicates concerted motion in the same direction (blue). In general, the amino acids of each monomer (or chain) move in a correlated manner while the amino acids of each monomer move in an anticorrelated manner with respect to the opposite monomer. The exception is helix $\alpha 5$ of each monomer (marked with a green rectangle) that move in an anticorrelated manner to the other residues of its own monomer while showing correlated motions with residues of the opposite monomer. The average correlation of helix $\alpha 5$ residues with each residue of the dimer was calculated from the cross-correlation matrix and is presented in Figure 3B for monomers A (blue) and B (red). The bottom bar presents the following regions: $\alpha 1$ (gray), $\alpha 2$ (black), and $\alpha 5$ (orange). Helix $\alpha 5$ of monomer A (blue line) shows high average positive correlation with helices $\alpha 1$, $\alpha 2$, and $\alpha 5$ of monomer B but does not show high positive correlation with any residues of monomer A besides itself. A similar picture is observed for helix $\alpha 5$ of monomer B; it correlates with itself and helices $\alpha 1$, $\alpha 2$, and $\alpha 5$ of monomer A. Thus, dynamically, helix $\alpha 5$ of each monomer correlates and affects the mobility and motion of helices $\alpha 1$, $\alpha 2$, and $\alpha 5$ of the opposite monomer. A similar analysis was performed on the VZV serine-protease to assess the implication or generality of the results to the HHV Pr family. The VZV protease has a similar structure to KSHV with helices A6 ($\alpha 5$ of KSHV) and A1 and A3 ($\alpha 1$ and $\alpha 2$ of KSHV) of the opposite monomer forming the interface. The cross-correlation map is presented in Figure S1. The average correlation of helix A6 residues with

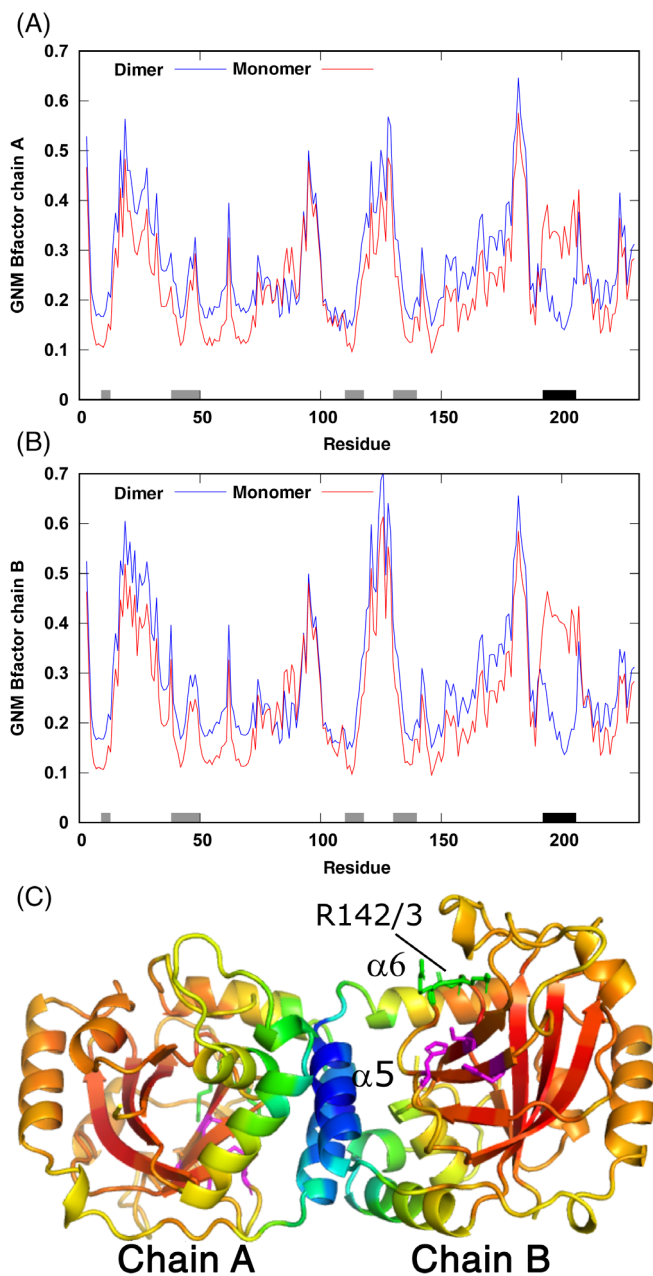


FIGURE 2 B-factor comparison between the monomeric and dimeric forms of KSHV Pr. GNM-based B-factor profile of (A) monomer A and (B) B is presented for monomeric (red line) and dimeric (blue line) forms of KSHV Pr. The bottom bar indicates amino acids at the binding site (gray) and $\alpha 5$ residues (black). (C) The KSHV Pr dimeric structure colored according to the ratio between the dimer and monomer state B-factors. Regions with lower mobility or B-factor values in the dimeric state (compared to the monomeric state) are colored in blue while regions with higher mobility are colored red. GNM, Gaussian Network Model; KSHV, Kaposi's sarcoma-associated herpesvirus

each residue of the dimer was calculated from the cross-correlation matrix and is presented in Figure 3C. Helix A6 shows high average positive correlation with helices A1 and A6 of the opposite monomer. It also shows a positive correlation with helix A1 and strands B2 and

B3 of its own monomer unlike KSHV Pr. The dimer interface of KSHV and VZV is depicted in Figure 3D. KSHV monomers A and B are colored green and yellow while VZV monomers A and B are colored blue and cyan, respectively. The $\alpha 5$ of KSHV is located at the center of the interface while A6 helix is tilted toward helix A1 and is in closer contact with it compared to $\alpha 1$ and $\alpha 5$ contacts of KSHV. Helix A1 in turn forms contacts with strands B1 and B2. The loop between strands B2 and B3 include the catalytic residue H52, with the other two catalytic residues S120 and H139 located on strands B5 and B6, respectively. The higher cross-correlation of helix A6 with strands B2 and B3 of its own monomer and with helices A1 and A6 of the opposite monomer enables more of a direct effect of one monomer and the dynamics of the active site of the opposite monomer compared to KSHV.

Clustering of the KSHV dimer residues was done based on the above cross-correlation map. The obtained dendrogram is presented in Figure S2. The cutting level of the dendrogram determines the number of clusters and the resolution of the clustering, when each cluster represents a dynamical domain within the dimer. We note that the term dynamical domains refer to domains that are based on the cross-correlation map and do not necessarily represent the classical structure-based domains. The dimer can be separated into three main clusters: The two $\alpha 5$ helices of monomers A and B, monomer A without its $\alpha 5$, and monomer B without its $\alpha 5$. The two $\alpha 5$ helices that are associated in a process that recalls the term "domain swapping"³⁷ became a single dynamical domain. A more refined cut further separates each of the two latter clusters into two to get a total of five dynamical domains: The central or interfacial domain composed of helices $\alpha 5$ of both the monomers; the proximal domain composed of helices $\alpha 1$, $\alpha 1a$, and $\alpha 6$ (or part of it for monomer A); and the distal domain mostly composed of the β -sheet that includes the ligand binding site. Figure 4 presents the dimer colored according to the five dynamical domains.

The average GNM cross-correlation values between the five domains are listed in Table S3. The average cross-correlation values for residues within the proximal and distal domains of each monomer range between 0.1 and 0.13, whereas they are approximately half for residues across these domains. The average correlation of the interfacial domain with the proximal domain is low since each helix $\alpha 5$ does not show positive cross-correlation with residues of its own monomer. Since each monomer is divided into proximal and distal domains, the low mobility of the interfacial domain reduces the mobility of the proximal domains which, in turn, increases the mobility of the distal domains that includes the binding sites of the dimer.

The GNM modes are isotropic, hence give information on the magnitude of the fluctuations of each mode but not their direction. Alternatively, we can use ANM analysis which gives information on the magnitude and directions of the motions. ANM analysis was carried out to obtain the modes of motions of the KSHV dimer that included the ligands. The first two or three slowest modes usually describe global functional motions that are of great biological interests.³⁸ The three slowest ANM modes of KSHV Pr, in agreement with the GNM results, describe anticorrelated rotation motions of one monomer relative to the other. Figure 5 presents the motions using

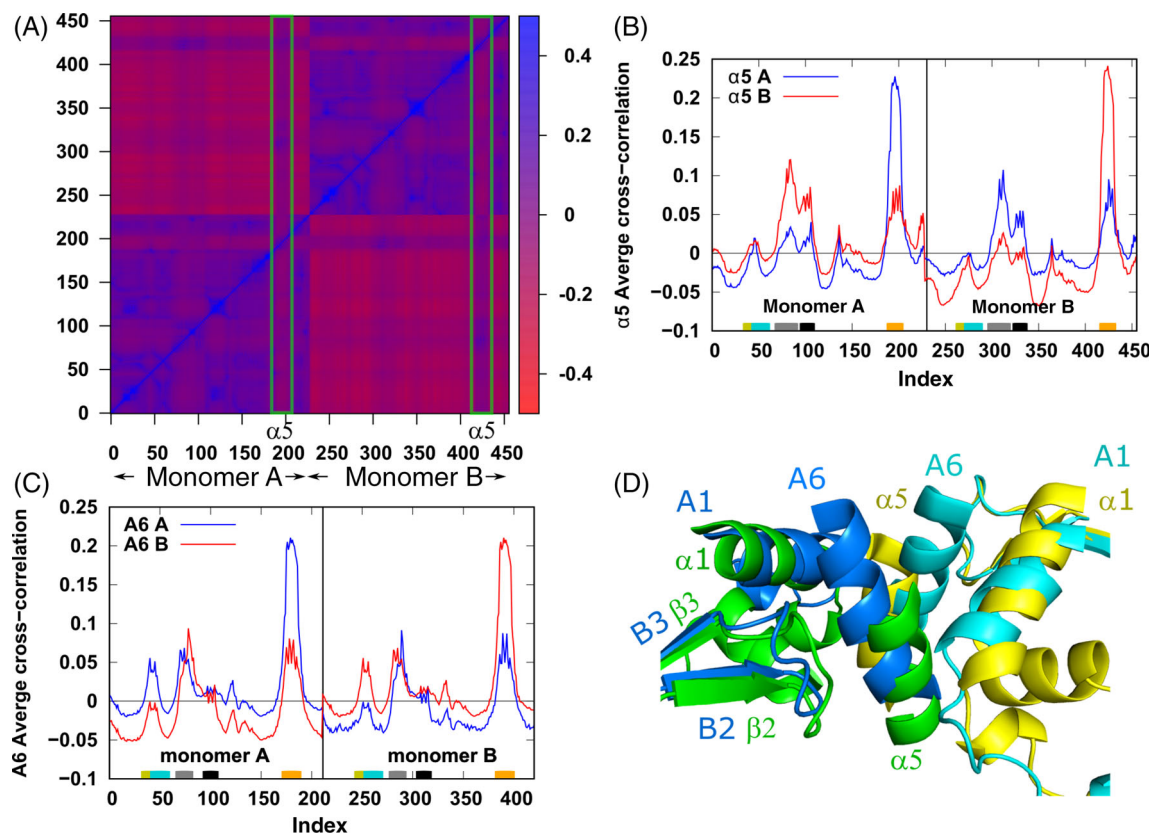


FIGURE 3 KSHV and VZV Pr cross-correlation analysis. (A) Cross correlation map of the KSHV dimer. A positive value indicates concerted motion in the same direction (blue), whereas a negative value refers to an anticorrelation between residue fluctuations (red). The green rectangles indicate the $\alpha 5$ region of each monomer. (B) The average correlation of each KSHV Pr residue with $\alpha 5$ residues of monomer A (blue line) and monomer B (red line). Indexes 0–227 and 228–455 indicate monomers A and B, respectively. The bottom bars indicate the following regions: $\beta 2$ (yellow), $\beta 3$ (turquoise), $\alpha 1$ (gray), $\alpha 2$ (black), and $\alpha 5$ (orange). (C) The average correlation of each VZV Pr residue with A6 residues of monomer A (blue line) and monomer B (red line). Indexes 0–210 and 211–421 indicate monomers A and B, respectively. The bottom bars indicate the following regions: B2 (yellow), B3 (turquoise) A1 (gray), A3 (black), and A6 (orange). (D) The dimer interface of KSHV and VZV Pr. KSHV Pr monomers A and B are colored green and yellow, while VZV Pr monomers A and B are colored blue and cyan, respectively. KSHV, Kaposi's sarcoma-associated herpesvirus; VZV, varicella-zoster virus

porcupine plots, with the red arrows showing the direction and magnitude of motion of the residues. For clarification, black arrows pointing to the general rotation direction and a black line indicating the rotation axis were added in Figure 5B,C. The first or slowest mode of motion (Figure 5A) describes a rotation motion of each monomer around itself with the rotation axis at the center of each monomer and perpendicular to the plane of the figure. The second mode of motion (Figure 5B) describes a rotation motion of mainly the distal domain of each monomer with a single rotation axis that crosses the two domains and is perpendicular to the $\alpha 5$ helices at the interface. The third mode describes a rotation motion with a separate rotation axis for each monomer (Figure 5C). The functional role of these motions is of great interest. Stretching of the ligand (or inhibitor) is observed in all three modes and especially in the first mode for monomer A and the second mode for monomer B. In addition, we observed stretching of the bond between S114 and the phosphorus atom of the inhibitor, a bond that is cleaved upon completion of the proteolytic reaction of a natural ligand, in all modes and especially in mode three. Arginines 142 and 143 provide hydrogen bonds shown

previously to be essential for catalysis.³⁹ The motion of R142 and R143 is observed in modes two and three. The guanidino group of R142 and R143 is moving toward the oxyanion in modes two and three, respectively. A movie showing the motion of the first mode is available in Video S1. A similar analysis was performed on the VZV Pr structure to see if the ANM slow modes are similar among other family members. The three slowest modes of the protease are presented using porcupine plots in Figure S5. These modes are similar to the KSHV modes where the first and second modes switched their order. This order is the same as the ANM modes calculated to the KSHV Pr without the ligand.

Since the three slowest modes have a functional role, it is interesting to check whether they exist in the monomers without dimerization. ANM modes were calculated for the dimer and each of the monomers separately without the inhibitory ligand. The modes part of the dimer that correspond to each monomer were aligned with the modes that were calculated for the monomer alone. The AAS for the alignment of the 10 slowest modes between the dimer and monomer states are presented in Figure 6 for monomers A (top) and B (bottom).

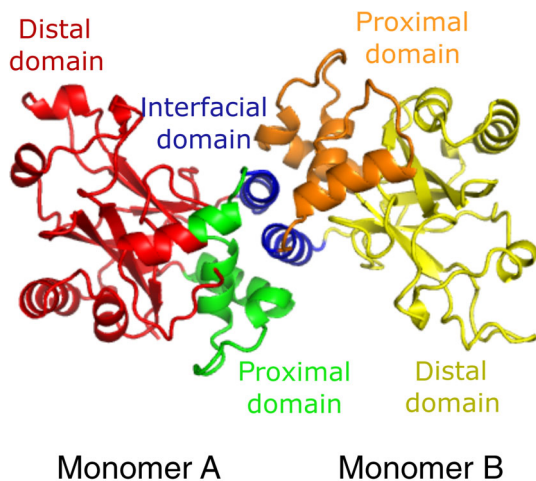


FIGURE 4 KSHV Pr dynamical domains. The KSHV dimer structure is colored according to five dynamical domains. The central or interfacial domain is composed of helix $\alpha 5$ of the two monomers and colored blue. The proximal domains are composed of helices $\alpha 1$, $\alpha 1a$, and $\alpha 6$ (or part of it for monomer A) are colored green and orange for monomers A and B, respectively. The distal domains mostly composed of the β -sheet and include the ligand binding site are colored red and yellow for monomers A and B, respectively. KSHV, Kaposi's sarcoma-associated herpesvirus

The AAS value ranges from one (full match) to zero (no match). The alignments show that the first and third modes of monomer A in the dimeric state have moderate similarity to mode seven (AAS = 0.56) and ten (AAS = 0.63) in the monomeric state, respectively. No matching mode was found for the second mode in the dimeric state among the 10 slowest modes of the monomer. In the case of monomer B, the first and fourth modes of this monomer in the dimeric state show partial similarity to mode seven of the monomer state (AAS of 0.63 and 0.46, respectively) and the second mode of the dimer state shows moderate similarity to mode six of the monomeric state (AAS = 0.53). No matching mode was found for the third mode in the dimeric state among the 10 slowest modes of the monomer. The higher the mode number is, the higher is the energetic cost for execution of motion. Thus, the three slowest modes of motion that are observed for the monomers in their active dimeric state are less accessible prior to dimerization.

The differences in the alignment scoring matrices in the above results imply that there is some asymmetry between the two monomers that was further studied. A reverse dimer was created by overlapping helix $\alpha 5$ of monomer A with the one of monomer B. Figure 7A depicts the original dimer with monomers A and B colored cyan and green and the reverse dimer with monomers B and A colored gray. The overall root-mean-square deviation (RMSD) between the two structures is 0.93 Å, indicating a close structural similarity of the two monomers. The exemptions are residues H119-P130 (binding site loop) that show structural difference between the two monomers as indicated by the black arrow heads. These residues connect the two β -strands that form most of the binding site of the inhibitory ligands and part of these residues are in direct contact with the ligands. ANM

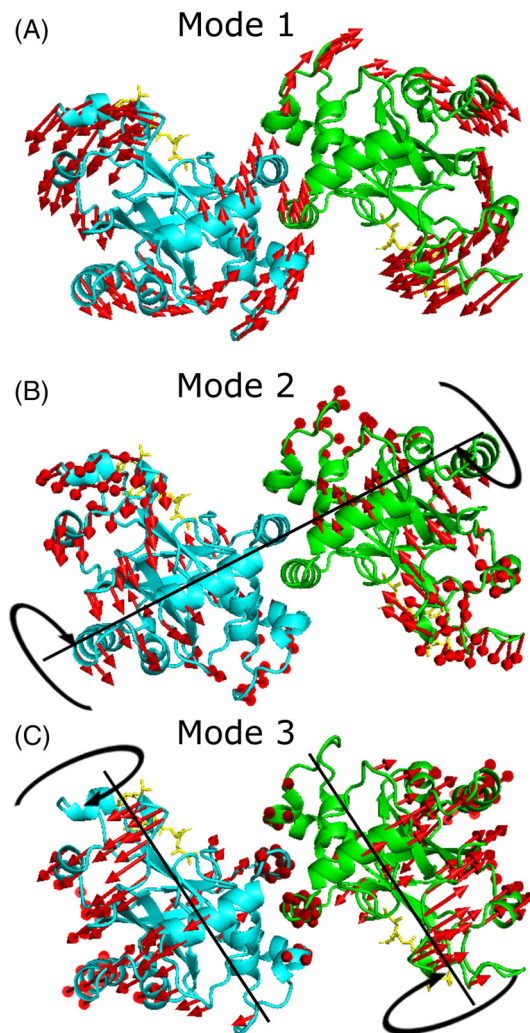


FIGURE 5 Porcupine plot of the three slowest modes of motion of KSHV Pr. The (A) first, (B) second, and (C) third ANM modes of motion of the KSHV dimer. The protein is depicted using cartoon representation with monomer A (left) and B (right) colored as in Figure 1 and the ligands are shown using yellow stick representation. Red arrows display the direction and magnitude of motion of the residues in each mode, black arrows (B and C) point to the general rotation direction of each monomer along the rotation axis shown as a black line. ANM, anisotropic network model; KSHV, Kaposi's sarcoma-associated herpesvirus

modes of motion were calculated (without the presence of the ligands) for the reverse dimer and modes alignment were done between the ten slowest modes of the dimer and its reverse. The resulting AASs are presented in Figure 7B as a heat map. The modes of motion of the dimer and its reverse show good similarity besides mode six. Thus, the structural changes at the binding site loop affect the sixth mode of motion. Similar results were obtained when the ligands were included in the ANM calculation of the dimer and its reverse.

The results indicate that changes in the ligand binding site may change modes of motion. Therefore, we investigated how ligand binding affects the modes of motion of the dimer. The ten slowest ANM modes of motion of the dimer that were calculated with and without

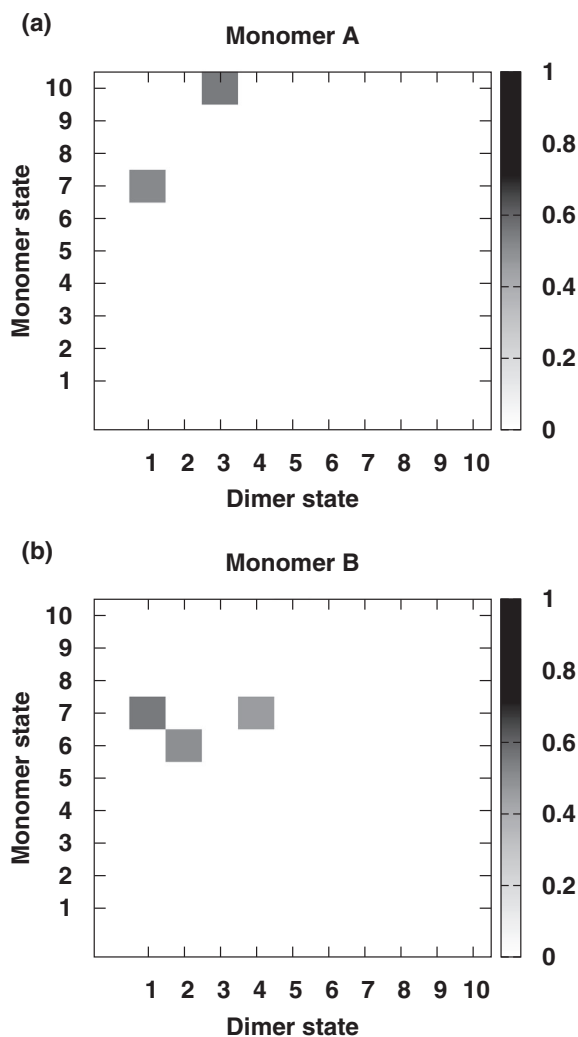


FIGURE 6 Global ANM AASs matrix between the dimer and monomer states of KSHV Pr. Global mode alignment was carried out between the ten slowest ANM modes of each monomer in the monomeric and dimeric states. The AAS values range from one (full match) to zero (no match) and the scoring matrix is depicted as a heat map for (A) monomer A and (B) monomer B. AAS, average alignment score; ANM, anisotropic network model; KSHV, Kaposi's sarcoma-associated herpesvirus

the presence of the ligand were aligned. The AASs of the alignments are presented in Table S6. The average AAS between the first five slowest modes is 0.94 and drops down to 0.72 in the alignment of the sixth mode of the dimer with and without the ligand. The sixth mode is depicted in Figure 8A, and describes the motion of the ligand binding loop of both monomers, the interfacial region, and many residues of monomer B. The red arrows describe the motion of the dimer in this mode without the ligand while the black arrows describe the motion with the presence of the ligand. A close-up of the interface is presented in Figure 8B, showing that without the ligand helix $\alpha 1$ of monomer A and helix $\alpha 5$ of monomer B are moving apart from each other by 45° . However, in the presence of the ligand, this motion is suppressed. The mobility of the residues in the sixth mode is presented in Figure 8C,D for monomers A and B, respectively. The black,

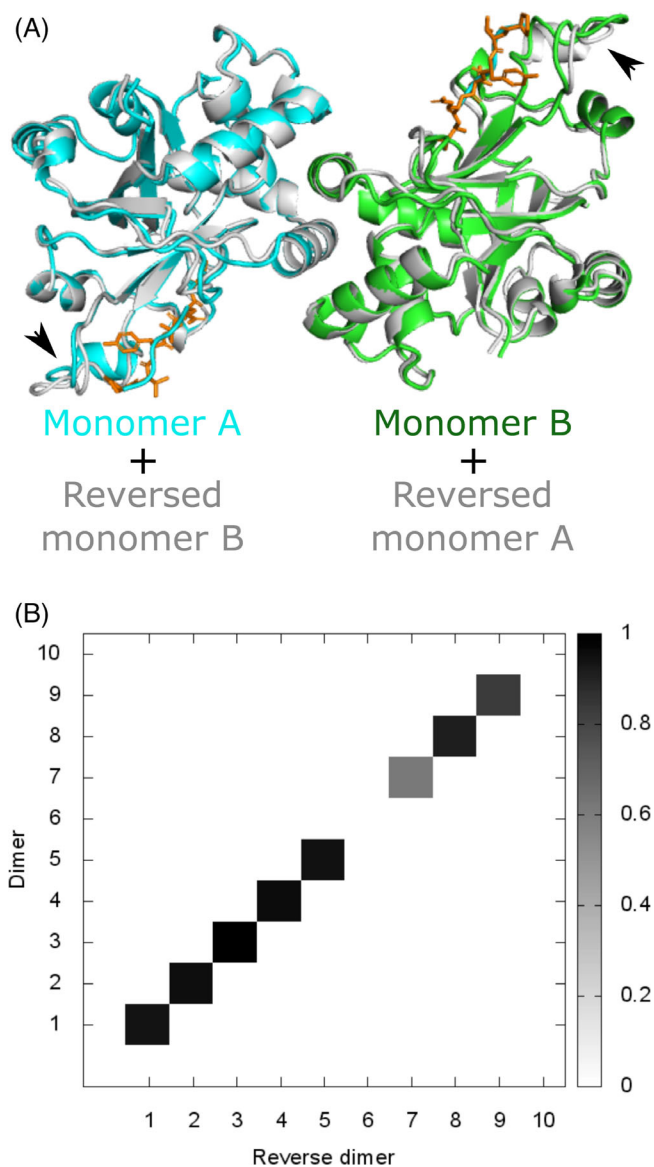


FIGURE 7 Structural changes in the binding site affects mode six. (A) Superimposition of the dimer and the reverse dimer. The dimer is shown using cartoon representation with monomers A and B colored cyan and green, respectively. The ligands are shown using stick representation and colored orange. A reverse dimer was created by overlapping helix $\alpha 5$ of monomer A with the one of monomer B and is colored gray. Arrow heads indicate regions with structural differences between the monomers. (B) Global ANM modes alignment was carried out between the dimer and the reverse dimer. The ANM modes were calculated without the presence of the ligand in both cases. AASs matrix is depicted as a heat map. AAS, average alignment score; ANM, anisotropic network model

orange, and red bars designate helices $\alpha 1$, $\alpha 5$, and $\alpha 6$. The motion of helices $\alpha 1$ and $\alpha 5$ that form the interface is suppressed for monomers A and B in the presence of the ligand. However, the motion of $\alpha 6$ and the oxyanion stabilizing loop that includes R142 and R143 is increased in the presence of the ligands for monomer B. Thus, the sixth mode of motion connects the ligand binding region with the interface and ligand binding modulates the dynamics of the interface.

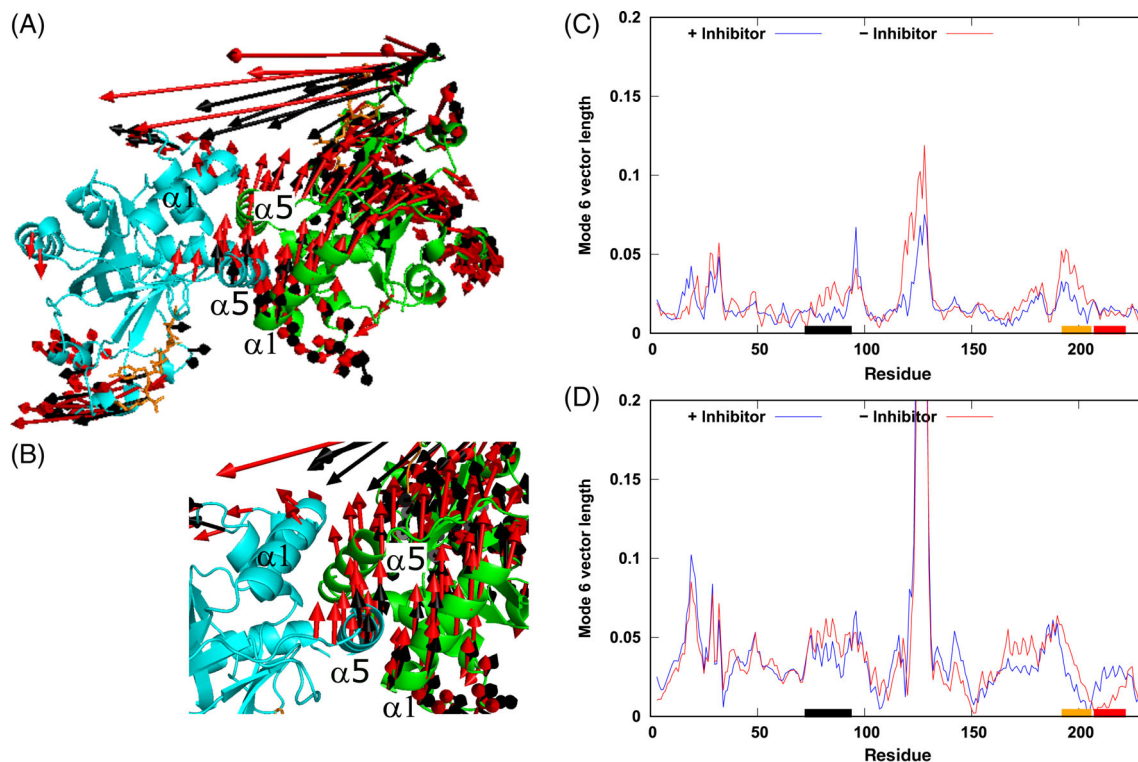


FIGURE 8 KSHV Pr dynamics changes in presence of ligand in mode six. (A) ANM sixth mode of motion in KSHV dimer is represented using porcupine plot. Monomers A (cyan) and B (green) are shown using cartoon representation and the ligand (orange) is shown using sticks representation. Black and red arrows display the direction and magnitude of motion of the dimer residues with and without ligand, respectively. (B) Close-up of the interfacial area of KSHV Pr showing the changes. Without the ligand, helix $\alpha 1$ of monomer A and helix $\alpha 5$ of monomer B are moving apart from each other by 45° . Mode six mobility plot for (C) monomer A and (D) monomer B of the dimer with ligand (blue line) and without ligand (red line). The bars indicate helices $\alpha 1$ (black), $\alpha 5$ (orange), and $\alpha 6$ (red). ANM, anisotropic network model; KSHV, Kaposi's sarcoma-associated herpesvirus

4 | DISCUSSION

This work shows the effect of KSHV Pr dimerization on the dynamics of the subunits and the implications for the dimer activation and function. The GNM analysis shows that the dimerization process creates a dimer with five dynamical domains. The central domain is composed of the two $\alpha 5$ helices, one from each monomer, with each helix affecting the dynamics of the proximal domain of the opposite monomer. The central domain becomes a hinge region and its mobility is reduced, which, in turn, partially tethers the proximal domain, leading to the increased mobility of the distal domain and mainly the binding site within it. These sequences of events result in dynamical changes between the monomeric and dimeric states of the monomers. These changes are intertwined with structural changes that establish the proper conformation of the substrate recognition/oxyanion hole to favor the transition state of catalysis.⁴⁰ Based on the functional role of the slow modes and their ability to stretch the ligand and the ligand-Serine114 cleavage bond, we conclude that the increased mobility of the binding site enables the dimer to perform its activity more efficiently.

The way by which one monomer affects the dynamics of the other monomer can be adjusted by how the interfacial helices are packed. In the case of KSHV Pr, the $\alpha 5$ helix is loosely connected to

its own monomer and moves in an anticorrelated manner to it, while moving in correlation with helices $\alpha 1$, $\alpha 2$, and $\alpha 5$ of the opposite monomer. In the case of VZV Pr, the A6 helix is more tightly connected to its own monomer and moves in a correlated manner with helix A1 and strands B2 and B3 that form part of the binding site as well as with helix A6 of the opposite monomer. Hence, it enables each monomer to exert a stronger effect on the dynamics of the opposite monomer. These two examples show that the level of dynamical coupling between the two monomers can vary between different members of the HHV Prs.

The GNM analysis is more accurate than the ANM analysis in mobility calculation but does not give information on the direction of motion of the different modes.²⁰ Therefore, a complementary ANM analysis was performed to explore in depth the dynamics of the dimer and its monomers. This method provides better biochemically related data. The three slowest ANM modes support specific conformation of the ligand/inhibitor, which mimics a major step in the catalytic cycle, and show functional importance. These modes stretch and bend the inhibitory ligand and hence can assist in its binding and release. The PDB structure we used for the analysis captures the KSHV Pr in its intermediate state with the catalytic S114 covalently bound to the phosphorus atom of the inhibitor. This bond is cleaved upon completion of the proteolytic processing of the natural ligand. The dynamics

in these modes stretch the bond and exemplify the adaptation of the dynamics to the function of the protease. In addition, we see that the two conserved arginines R142 and R143 (also Arg 165 and 166 for other human cytomegaloviruses [CMV]) are moving toward the oxyanion in modes two and three, respectively. Based on the crystal structure of the CMV protease, these arginines were proposed to stabilize the oxyanion intermediate formed during the reaction.^{41,42} Arg166 was shown previously to hold a water molecule that can act as a hydrogen-bond donor to the oxyanion and was thus proposed to stabilize the oxyanion intermediate. Liang et al.³⁹ argue that Arg165 only utilizes the backbone NH for catalysis as suggested by the crystal structure of the Arg165 side chain and is not involved in the stabilization of the oxyanion, leaving the question as to why then the side chain of Arg165 (142) is highly conserved. We show in our calculations that the side chains of these arginines can move toward the oxyanion and hence can interact with it. These modes of motion that have a functional role are the slowest modes of motion in the active dimeric state and are energetically favored. While, in the monomeric state, they become modes of a higher number and less favored energetically. Complementary ANM analysis to the VZV Pr dimer showed similarity to KSHV Pr of the first three slow modes, indicating that these modes are conserved across the HHV Prs.

Two comparisons of the modes of motion: (i) the dimer and its reverse and (ii) the dimer with and without the ligand, indicate that the sixth mode of motion connects the ligand binding region with the interface. Ligand binding and structural changes in the ligand region can affect its dynamics and the dynamics of $\alpha 1$ and $\alpha 5$ that form the interface. Marnett et al.⁴³ showed that ligand binding stabilizes the dimer. Phosphorylation of the active site serine by the inhibitory peptide yields a tetrahedral phosphonate adduct. The transition state places a strong negative charge in the oxyanion hole that forms hydrogen bonds with R142 and R143 and stabilizes the dimer. We suggest that this allosteric effect is also mediated through modulation of the sixth mode of motion of the dimer. Thus, dynamically, the KSHV Pr has evolved to function as a dimer and dynamics play an important role in its function and regulation.

The HHV Prs are of great interest as targets for therapeutic intervention. Recently, Kaynak et al.⁴⁴ developed the Essential Site Scanning Analysis (ESSA), an ENM-based method to identify allosteric and orthosteric ligand-binding residues as well as hinge residues. Our work shows the utility of ENM in studying HHV Prs dynamics and especially the relation between binding site and interface dynamics. Thus, this work calls for use of methods like the ESSA in developing inhibitors to these targets.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Dror Tobin  <https://orcid.org/0000-0002-1125-2751>

REFERENCES

- Gable JE, Lee GM, Jaishankar P, et al. Broad-spectrum allosteric inhibition of herpesvirus proteases. *Biochemistry*. 2014;53(28):4648-4660.
- Mesri EA, Cesarman E, Boshoff C. Kaposi's sarcoma and its associated herpesvirus. *Nat Rev Cancer*. 2010;10(10):707-719.
- Cesarman E, Damania B, Krown SE, Martin J, Bower M, Whitby D. Kaposi sarcoma. *Nat Rev Dis Primers*. 2019;5(1):9.
- Coen DM, Schaffer PA. Antihherpesvirus drugs: a promising spectrum of new drugs and drug targets. *Nat Rev Drug Discov*. 2003;2(4):278-288.
- Reiling KK, Pray TR, Craik CS, Stroud RM. Functional consequences of the Kaposi's sarcoma-associated herpesvirus protease structure: regulation of activity and dimerization by conserved structural elements. *Biochemistry*. 2000;39(42):12796-12803.
- Unal A, Pray TR, Lagunoff M, Pennington MW, Ganem D, Craik CS. The protease and the assembly protein of Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8). *J Virol*. 1997;71(9):7030-7038.
- Zuhlsdorf M, Hinrichs W. Assemblins as maturational proteases in herpesviruses. *J Gen Virol*. 2017;98(8):1969-1984.
- Lazic A, Goetz DH, Nomura AM, Marnett AB, Craik CS. Substrate modulation of enzyme activity in the herpesvirus protease family. *J Mol Biol*. 2007;373(4):913-923.
- Pray TR, Nomura AM, Pennington MW, Craik CS. Auto-inactivation by cleavage within the dimer interface of Kaposi's sarcoma-associated herpesvirus protease. *J Mol Biol*. 1999;289(2):197-203.
- Kitao A, Go N. Investigating protein dynamics in collective coordinate space. *Curr Opin Struct Biol*. 1999;9(2):164-169.
- Cui Q, Bahar I. *Normal Mode Analysis. Theory and Applications to Biological and Chemical Systems*. Taylor & Francis Group; 2006.
- Bahar I, Rader AJ. Coarse-grained normal mode analysis in structural biology. *Curr Opin Struct Biol*. 2005;15(5):586-592.
- Amadei A, Linssen AB, Berendsen HJ. Essential dynamics of proteins. *Proteins*. 1993;17(4):412-425.
- Bahar I, Erman B, Haliloglu T, Jernigan RL. Efficient characterization of collective motions and interresidue correlations in proteins by low-resolution simulations. *Biochemistry*. 1997;36(44):13512-13523.
- Romo TD, Clarage JB, Sorensen DC, Phillips GN Jr. Automatic identification of discrete substates in proteins: singular value decomposition analysis of time-averaged crystallographic refinements. *Proteins*. 1995;22(4):311-321.
- Garcia AE, Harman JG. Simulations of CRP:(cAMP)₂ in noncrystalline environments show a subunit transition from the open to the closed conformation. *Protein Sci*. 1996;5(1):62-71.
- Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold Des*. 1997;2(3):173-181.
- Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J*. 2001;80(1):505-515.
- Eyal E, Chennubhotla C, Yang LW, Bahar I. Anisotropic fluctuations of amino acids in protein structures: insights from X-ray crystallography and elastic network models. *Bioinformatics*. 2007;23(13):i175-i184.
- Eyal E, Yang LW, Bahar I. Anisotropic network model: systematic evaluation and a new web interface. *Bioinformatics*. 2006;22(21):2619-2627.
- Nomura AM, Marnett AB, Shimba N, Dotsch V, Craik CS. Induced structure of a helical switch as a mechanism to regulate enzymatic activity. *Nat Struct Mol Biol*. 2005;12(11):1019-1020.

22. Bahar I, Chennubhotla C, Tobi D. Intrinsic dynamics of enzymes in the unbound state and relation to allosteric regulation. *Curr Opin Struct Biol.* 2007;17(6):633-640.
23. Schueler-Furman O, Wodak SJ. Computational approaches to investigating allostery. *Curr Opin Struct Biol.* 2016;41:159-171.
24. Wodak SJ, Paci E, Dokholyan NV, et al. Allostery in its many disguises: from theory to applications. *Structure.* 2019;27(4):566-578.
25. Zheng W, Brooks BR, Thirumalai D. Allosteric transitions in biological nanomachines are described by robust normal modes of elastic networks. *Curr Protein Pept Sci.* 2009;10(2):128-132.
26. Zhuravlev PI, Papoian GA. Protein functional landscapes, dynamics, allostery: a tortuous path towards a universal theoretical framework. *Q Rev Biophys.* 2010;43(3):295-332.
27. Qiu X, Janson CA, Culp JS, et al. Crystal structure of varicella-zoster virus protease. *Proc Natl Acad Sci U S A.* 1997;94(7):2874-2879.
28. Tobi D. Large-scale analysis of the dynamics of enzymes. *Proteins.* 2013;81:1910-1918.
29. Li H, Chang YY, Yang LW, Bahar I. iGNM 2.0: the Gaussian network model database for biomolecular structural dynamics. *Nucleic Acids Res.* 2016;44(D1):D415-D422.
30. Chennubhotla C, Rader AJ, Yang LW, Bahar I. Elastic network models for understanding biomolecular machinery: from enzymes to supra-molecular assemblies. *Phys Biol.* 2005;2(4):S173-S180.
31. Eyal E, Lum G, Bahar I. The anisotropic network model web server at 2015 (ANM 2.0). *Bioinformatics.* 2015;31(9):1487-1489.
32. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol.* 1970;48(3):443-453.
33. Tobi D. Dynamics alignment: comparison of protein dynamics in the SCOP database. *Proteins.* 2012;80(4):1167-1176.
34. Tobi D. Dynamical differences of hemoglobin and the ionotropic glutamate receptor in different states revealed by a new dynamics alignment method. *Proteins.* 2017;85(8):1507-1517.
35. The R Project for Statistical Computing. <https://www.r-project.org/>
36. The PyMOL Molecular Graphics System, Version 1.8 [computer program]. Schrödinger, LLC.
37. Bennett MJ, Choe S, Eisenberg D. Domain swapping: entangling alliances between proteins. *Proc Natl Acad Sci U S A.* 1994;91(8):3127-3131.
38. Bahar I, Lezon TR, Yang LW, Eyal E. Global dynamics of proteins: bridging between structure and function. *Annu Rev Biophys.* 2010;39:23-42.
39. Liang PH, Brun KA, Feild JA, et al. Site-directed mutagenesis probing the catalytic role of arginines 165 and 166 of human cytomegalovirus protease. *Biochemistry.* 1998;37(17):5923-5929.
40. Batra R, Khayat R, Tong L. Molecular mechanism for dimerization to regulate the catalytic activity of human cytomegalovirus protease. *Nat Struct Biol.* 2001;8(9):810-817.
41. Qiu X, Culp JS, DiLella AG, et al. Unique fold and active site in cytomegalovirus protease. *Nature.* 1996;383(6597):275-279.
42. Tong L, Qian C, Massariol MJ, Bonneau PR, Cordingley MG, Lagace L. A new serine-protease fold revealed by the crystal structure of human cytomegalovirus protease. *Nature.* 1996;383(6597):272-275.
43. Marnett AB, Nomura AM, Shimba N, Ortiz de Montellano PR, Craik CS. Communication between the active sites and dimer interface of a herpesvirus protease revealed by a transition-state inhibitor. *Proc Natl Acad Sci U S A.* 2004;101(18):6870-6875.
44. Kaynak BT, Bahar I, Doruker P. Essential site scanning analysis: a new approach for detecting sites that modulate the dispersion of protein global motions. *Comput Struct Biotechnol J.* 2020;18:1577-1586.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Bern D, Tobi D. The effect of dimerization and ligand binding on the dynamics of Kaposi's sarcoma-associated herpesvirus protease. *Proteins.* 2022; 90(6):1267-1277. doi:10.1002/prot.26307