ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

Intrinsic Dynamics of the ClpXP Proteolytic Machine Using Elastic Network Models

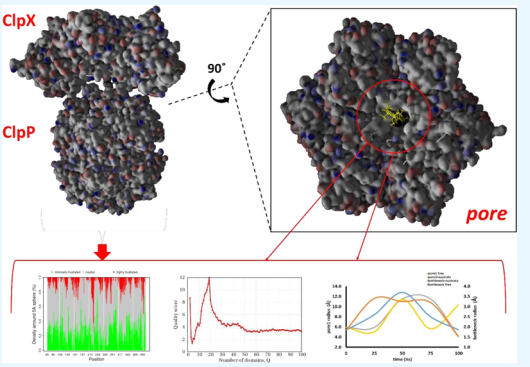
Lenin González-Paz, Carla Lossada, Maria Laura Hurtado-León, Francelys V. Fernández-Materán, José Luis Paz, Shayan Parvizi, Rafael Eduardo Cardenas Castillo, Freddy Romero,* and Ysaias J. Alvarado*

Cite This: *ACS Omega* 2023, 8, 7302−7318

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: ClpXP complex is an ATP-dependent mitochondrial matrix protease that binds, unfolds, translocates, and subsequently degrades specific protein substrates. Its mechanisms of operation are still being debated, and several have been proposed, including the sequential translocation of two residues (SC/2R), six residues (SC/6R), and even long-pass probabilistic models. Therefore, it has been suggested to employ biophysical−computational approaches that can determine the kinetics and thermodynamics of the translocation. In this sense, and based on the apparent inconsistency between structural and functional studies, we propose to apply biophysical approaches based on elastic network models (ENM) to study the intrinsic dynamics of the theoretically most probable hydrolysis mechanism. The proposed models ENM suggest that the ClpP region is decisive for the stabilization of the ClpXP complex, contributing to the flexibility of the residues adjacent to the pore, favoring the increase in pore size and, therefore, with the energy of interaction of its residues with a larger portion of the substrate. It is predicted that the complex may undergo a stable configurational change once assembled and that the deformability of the system once assembled is oriented, to increase the rigidity of the domains of each region (ClpP and ClpX) and to gain flexibility of the pore. Our predictions could suggest under the conditions of this study the mechanism of the interaction of the system, of which the substrate passes through the unfolding of the pore in parallel with a folding of the bottleneck. The variations in the distance calculated by molecular dynamics could allow the passage of a substrate with a size equivalent to ∼3 residues. The theoretical behavior of the pore and the stability and energy of binding to the substrate based on ENM models suggest that in this system, there are thermodynamic, structural, and configurational conditions that allow a possible translocation mechanism that is not strictly sequential.



INTRODUCTION

Advances in computational tools and statistical methods have made it possible to model complex protein systems in diverse cells, including proteins from the human exome, some of which are associated with disease. Studies have been done on the dynamic properties responsible for the ability of a protein to adapt to new functions in response to environmental changes, especially studies on flexibility and dynamic coupling to quantify and study the deformability of protein systems.[1] Membrane protein simulations have gained popularity in the last decade, and advances have made it possible to simulate their behavior in support of experimental findings, to elucidate protein mechanisms, and to validate protein crystal structures.[2] In fact, it has been described that proteins have a greater number of alternative minimum-energy conformations than their crystal structure. These structures, along with their deformability propensities, contain an enormous amount of information about protein function from a structural perspective.[3]

Protein homeostasis integrates the control of processes such as protein degradation. External stimuli and diseases are associated with the modulation of this type of process. Although many transcription factors and other regulatory proteins are known to be characterized by rapid proteasomal turnover and degradation, understanding the molecular dynamics that control associated biological processes remains a very active topic.[4] The dynamics of protein systems relevant to cellular homeostasis have been explored[5] including that of complex systems important for mitochondrial function and autophagy such as AAA+ motors, like the ClpXP.[6−8] However,

Received: July 9, 2022
Accepted: October 25, 2022
Published: February 14, 2023

dynamic efforts have focused particularly on studying cell cycle localization and mitochondrial fusion mechanisms.[5,6]

ClpXP are proteolytic systems that use the energy obtained through the hydrolysis of ATP to carry out various activities in cells. It has been suggested that the mechanism of action is based on a sequential model of ATP hydrolysis, in which a forceful blow is carried out that displaces two substrate residues in each translocation step (SC/2R). However, it has been pointed out that this type of proteolytic complex does not work strictly sequentially, and it has been described that it can carry out longer translocation steps than that of SC/2R. Indeed, alternative kinetic pathways such as SC/6R, and long-pass probabilistic models for ClpX translocation have been proposed. Therefore, it has been suggested to apply biophysical approaches to help determine the kinetics and thermodynamic stability of translocations.[9]

In this sense and based on the apparent inconsistency between the reported structural and functional studies, we propose to apply biophysical approaches based on elastic network models (ENM) to study the intrinsic dynamics of the theoretically most probable hydrolysis mechanism from experimental data.[10] In addition, although classical molecular dynamics simulations can describe conformational changes, they often require prohibitively high computational cost.[3] ENM are well established for modeling the collective motion of proteins, allowing for the comparison of protein motion to gain insight into structure−function relationships. Comparing dynamics has been shown to be a viable way for gaining greater understanding of the mechanisms employed by proteins for their function.[11]

The low-frequency conformations or modes of ENM-based approaches correspond to the lowest-energy dynamic conformations of classical molecular dynamics (MD) simulations, both in terms of the directions and the relative amplitudes of the motions of the atoms obtained in the conformations. The computation of conformations by ENM requires few minutes; therefore, this approach has allowed us to apply such directional information in applications such as driving MD simulations, predicting receptor flexibility in docking algorithms, performing flexibility adjustments of diverse molecular structures and generating efficiently and in less time pathways of conformational change.[12] That is why, studying the relationship between the function and dynamics of proteolytic systems or machines such as ClpXP, using the ENM approach, is relevant to contribute to the understanding of the structure−function relationship and dynamics of this protein complex. Therefore, this study will address (1) the behavior of the ClpXP proteolytic machine, (2) the determining regions for the associated dynamics, and (3) the possible theoretical mechanism to which this type of system best fits under the ENM model approach.

MATERIALS AND METHODS

Selection of Structures in Databases. In this regard, the crystal structure of a Clp class-1 type transmembrane protease was considered, specifically the ClpX region (PDB: 6PP8) and the ClpX-ClpP complex bound to substrate and ATP-γ-S (PDB: 6POS).[10] These structures were chosen because the dynamics of ClpXP protein systems have been reported to be relevant to processes such as cellular homeostasis,[5] mitochondrial function, and autophagy.[6−8] However, dynamic efforts have focused particularly on studying the localization of these systems in the cell cycle rather than at the structural level.[5,6] The structures used were obtained in PDB format from the Protein Data Bank (https://www.rcsb.org/). As usual, water molecules and co-crystallization elements were removed, and subsequently, different files in PDB format of all of the structures used were created using the Molegro Molecular Virtual (MMV)-v.7.0 package and the Chimera software.[13,14]

Study of Frustration and Structural and Configurational Deformation of ClpXP Complex. Proteins are biologically optimized to fold and be stable to fulfill their function. Therefore, it is not surprising to find some degree of energetic conflict within their local structure. These conflicts allow proteins to explore different conformations within their native set and thus allow "biological function" to emerge. In this sense, to obtain information on configurational patterns, dynamics, and transitions of complex systems such as ClpXP, frustration and its role in the functional dynamics of said complex were examined.[15] For this, the Frustratometer software (http://frustratómetro.qb.fcen.uba.ar/)[15−17] was used, a model based on the energy landscape theory that allows quantifying the degree of local protein frustration. The energy landscape theory of protein folding is a statistical description of the potential surface area of a protein. This model suggests that the most realistic model of a protein is a heteropolymer that is minimally frustrated toward the native state. This principle states that the energy of the protein decreases as the protein assumes conformations progressively similar to the native state. Conformational changes and their effect on topology become a key factor governing protein folding reactions. Frustratometer is based on the Associative Memory, Water Mediated, Structure, and Energy Model (AWSEM). AWSEM provides a nonadditive, transferable force field based on a coarse-grained approach (similar to elastic network models (ENM)) to predict the structure conformation of protein systems. AWSEM also contains terms like hydrogen bonding, and computational terms for local structure bias that take into account the effects of bodies that are modulated from the local sequence. Specifically, a "local frustration index" allows to quantify the contribution of a residue or a pair of residues to the total or local energy of a given structure, especially when they are displaced from their native location. The local frustration index makes it possible to determine the "highly frustrated" regions at the energy level, as well as the fraction or density of native contacts in each class of frustration (high, minimal, or neutral) around a sphere of 5 Å of each Cα (density around a 5 Å sphere) using a Debye−Hückel potential. In this case, the method includes long-range electrostatic interactions, which are modeled with the Debye−Hückel potential to include the dielectric effect of the solvent and charge−charge interactions by mobile ions in the solvent

$$V_{DH} = K_{elect} \sum_{i<j} \frac{q_i q_j}{\epsilon_r r_{ij}} e^{-r_{ij}/l_D} \quad (1)$$

where q_i and q_j are the charges of residues i and j that are separated by a distance r_{ij}, respectively, ε_r is the dielectric constant of the medium (1 in vacuum), and l_D is the Debye−Hückel screening length that takes into account the temperature, the ionic strength of the medium and its dielectric constant. These are set to typical physiological values of T = 25 °C, ε_r = 80 for water I = 0.1 M, and l_D = 10 Å. $K_{elect} = \kappa\varepsilon_r$, where κ is the electrostatic constant and represents the electrostatic strength of the system. For more details of the method, we recommend refs 15−17.

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

The changes or the variation of energy reflects the energetic contributions of the conformations of the polypeptide chain. This index is suitable for tertiary structures and allows studying the variation of energy from randomizing parameters such as the distances $r_{i,j}$ and the densities ρ_i of the interacting amino acids.

$$F_{ij}^{c} = \frac{E_{i,j}^{N} - \langle E_{i',j'}^{U} \rangle}{\sqrt{1/N \sum_{k=1}^{N} (E_{i',j'}^{U} - \langle E_{i',j'}^{U} \rangle)^2}} \quad (2)$$

when $E_{i,j}^{N} - \langle E_{i',j'}^{U} \rangle$ the native energy would not be discriminated from the typical energy of a random interaction in the molten globule and $F_{ij}^{c} \approx 0$. The pairwise frustration index, F_{ij}^{c}, is a Z-score of the native pair energy ($E_{i,j}^{N}$) compared with a distribution of N decoy energies ($E_{i,j}^{N}$, for $k = 1, \cdots, N$) for the pair of residues i and j. This scheme effectively evaluates the native pair with respect to a set of structural decoys that might be encountered in the folding process. This calculated frustration index is called "configurational frustration". For more details of this method, we recommend refs 15−17.

Likewise, the biomolecular structures of the ClpXP complex were simulated as a system of interacting amino acid networks to obtain information on complex phenomena such as folding, unfolding, stability, as well as the role of residues at a structural and functional level. For this, webPSN (http://webpsn.hpc.unimo.it/wpsn.php) was used, which is a high-performance server for the study of biological systems. webPSN uses a mixed protein structure network (PSN) and an ENM (PSN-ENM), to predict the network and conformation obtained from the interaction between protein residues (structural communication).[19,20]

Protein structural network (PSN) analysis and the search for the shortest intra- and intermolecular communication pathways (PSN path) have been described as novel tools aimed at studying protein structures based on graph theory. PSN has been introduced to represent molecular systems as networks of amino acids, their interactions, and the functional implications of such networks in relation to intra- and intermolecular communication. The PSN concept also provides information on the properties of protein structures, for the study of folding, the contribution of residues to the free energy of binding (in protein-protein interaction), as well as the prediction of important residues. For more details on the implementation of the PSN analysis and the algorithms described in this work, we recommend the work of the team of Seeber et al.[19−23] Specifically, PSN is modeled from the atomic coordinates of the supplied protein system residues, by representing the network nodes of interacting amino acids. This model assumes that a pair of nodes are connected if the value of the interaction percentage (between them) is greater than or equal to a given interaction strength cutoff

$$I_{ij} = \frac{n_{ij}}{\sqrt{N_i N_j}} 100 \quad (3)$$

where I_{ij} is the interaction percentage of nodes i and j, n_{ij} is the number of side-chain atom pairs within a given distance cutoff (4.5 Å as a default), and N_i and N_j are, respectively, the normalization factors for residues i and j, which take into account the differences in size of the different nodes, as well as the propensity for a greater number of contacts or interactions with other nodes in the protein system considered. I_{ij} are calculated for all node pairs excluding $j = i \pm n$, where n is a given neighbor cutoff (2 as default), and each node pair with an I_{ij} value greater than or equal to a given I_{min} cutoff is connected by an edge. In this model, nodes with the highest number of interactions between pairs of (interacting) atoms will be assumed to be edge-connected, and to represent stronger interactions between residues. At a given I_{min} cutoff, those nodes that realize more than a given number of edges (4 as default) are called hubs. The percentage of interaction of a hub node is

$$I_i = \frac{n_{ij}}{N_i} \times 100 \quad (4)$$

where I_i is the hub interaction percentage of node i, n_{ij} is the number of side-chain atom pairs within a given distance cutoff, and N_i is the normalization factor of residue i. Node interconnectivity is finally used to highlight cluster-forming nodes. For details, see refs 19, 21−24.

Models based on elastic networks are a coarse-grained (CG) normal mode analysis (NMA) technique, which allows us to describe aspects such as the vibrational dynamics of proteins in a state of minimum energy. The CG-type models allow simulating the behavior of complex protein systems from a simplified representation. To simulate the structures of minimum energy in these models, the molecules are represented as "pseudo-atoms" that approximate groups of atoms (as a complete amino acid residue), instead of being represented by individual atoms as in classical dynamics. This simplified approach allows the system's degrees of freedom to be reduced, which translates into longer simulation times. With this technique, the total structure of the protein is modeled using a reduced subset of atoms (generally Cα-type atoms), with coordinates that can be obtained from structural determinations (such as X-ray crystallography, NMR, etc.) or from simulations. molecular. For further details, see refs 19, 21−24. Highly cooperative and low-frequency global/essential modes in proteins represent a key aspect of the functional dynamics of proteins and are the basis for the diffusion of methods such as the NMA to be able to infer this type of collective modes. The robustness shown by global mode-based analyses with respect to predicting details at atomic coordinates or specific interatomic interactions, as well as their insensitivity to specific (force-field-specific) energy functions and parameters has allowed the development of models or more simplified descriptions, and with low computational cost, such as CG-ENM-based models applied to protein structures. ENM is an approach that is based on the shape of the global modes of the network of contacts between residues (which is a geometric construction) defined by the general shape or native contact topology of the protein. In the last decade, tools based on the ENM-NMA approach have contributed significantly to the understanding of the collective dynamics of proteins, and their relationship with biological functions. NMA allows us to quickly describe the collective functional movements of complex protein systems, including multichain macromolecules. Each normal mode comprises both a strain vector and a frequency. These tools allow us to calculate both the direction of atomic displacement and the relative amplitude of the movement, therefore, the NMA approaches (in Cartesian coordinates) are used to model the flexibility of proteins. NMA is a technique to investigate the vibrational motion of a harmonic oscillating system in the immediate vicinity of its equilibrium. For further details, see refs 21, 22, 24. From these principles, the WebPSN tool allows

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

us to calculate the force constant of the interactions, for which it uses two versions of the ENM algorithm to calculate the cross-correlation between the movements of the Cα atoms and the pairwise interactions between the Cα atoms (linear cut-ENM and Kovacs-ENM, respectively). In this tool, the interactions between pairs of nodes are described by means of a *Hooke* harmonic potential (of a single term) and by calculating the total energy of the system through a simple Hamiltonian potential

$$E = \sum_{i \neq j} k_{ij}(d_{ij} - d_{ij}^0)^2 \tag{5}$$

where d_{ij} and d_{ij}^0 are the instantaneous and equilibrium distances between Cα-atoms i and j, respectively, and k_{ij} is a force constant, whose definition varies depending on the type of ENM used. In the equation, the second derivatives of the harmonic potential are contained in a Hessian matrix (H) of $3N \times 3N$, whose diagonalization allows obtaining a set of $3N - 6$ eigenvectors (with frequency other than zero), and with associated eigenvalues. For details, see refs 13, 21, 22, 24.

ENM approaches implement two methods to calculate the cross-correlation of atomic motion. The first method, called "linear cutoff ENM", assumes a force constant of 1 for interactions between pairs of atoms (typically Cα) within a defined cutoff distance, while adjacent atoms are assigned a constant strength of 10. The second method, called "Kovacs-ENM", allows estimating the force constant as a parameter dependent on the distance of the particles that can interact

$$k_{ij} = C\left(\frac{d_{ij}^0}{d_{ij}}\right)^6 \tag{6}$$

where C is constant (with a default value of 40 kcal mol^{-1} Å^{-2}). Finally, the cross-correlations of the movements for the trajectory filtering are obtained from the covariance matrix (eq 11). The structural perturbation method with ENM is a technique that has been useful to characterize allosteric interaction diagrams in the context of conformations of minimum-energy or lower-frequency modes according to the ENM approach. In this approach, amino acid positions that most significantly influence the dynamics of protein systems are predicted by systematically perturbing all of the "springs" connecting interacting Cα atoms and then measuring that residue's response to such perturbations based on a given mode m. The disturbance response is calculated as

$$\delta\omega_m = v_m^T \cdot \delta H \cdot v_m \tag{7}$$

where ν_m is the eigenvector of mode m, v_m^T is its transpose, and δH is the Hessian matrix of the perturbation to the energy of the elastic network

$$\delta E = \frac{1}{2}\sum_{i \neq j} \delta k_{ij}(d_{ij} - d_{ij}^0)^2 \tag{8}$$

The response $\delta\omega_m$ is proportional to the elastic energy of the springs that are connected to the ith residue when they are perturbed by an arbitrary value (typically 0.1), with the aim of being able to define the nodes that are most critical for the dynamics of a certain mode. Specifically, a node is defined as a connection point or a set of vertices within a network, or graph of the structure of a protein, commonly each protein node is represented by the Cα atom of each residue.[25,26] Moreover, involvement coefficients I between the ENM modes and the displacement vector between a given structure/frame T and a reference structure R can be computed according to the following equation[25−27]

$$I_m = \frac{\sum_{n,i=1}^{3N} v_{mn}\Delta r_i}{\sqrt{\sum_{n=1}^{3N} v_{mn}^2 \sum_{i=1}^{3N} \Delta r_i^2}} \tag{9}$$

where $\Delta r_i = r_i^T - r_i^R$ and $r_i^{T,R}$ is the ith coordinate in the two conformers and v_{mn} is the nth element of eigenvector m. By default, the computation is done for all 3N-6 modes, and only the values of I greater than an arbitrary threshold (i.e., 0.2) are output. The cumulative square overlap (CSO) between all modes and the displacement vector is computed according to the following equation

$$\mathrm{CSO} = \sqrt{\sum_{m=1}^{3N-6} I_m^2} \tag{10}$$

Finally, residue correlation C_{ij} is computed as

$$C_{ij} = \frac{\sum_{l=1}^{N} \frac{v_{il}v_{jl}}{\lambda_l}}{\left(\sum_{m=1}^{N} \frac{v_{im}v_{jm}}{\lambda_m}\right)^{1/2}\left(\sum_{n=1}^{N} \frac{v_{in}v_{jn}}{\lambda_n}\right)^{1/2}} \tag{11}$$

where C_{ij} denotes the correlation between particles i and j, M is the number of modes considered for computation (the first 10 nonzero frequency modes), and ν_{xy} and λ_y are, respectively, the xth element and the associated eigenvalue of the yth mode. For further details, see refs 18, 22, 28, 29.

Finally, the SPECTRUS (SPECTral-based Rigid Units Subdivision) (http://spectrus.sissa.it/#home)[30] server was used, this tool allows a protein system to be broken down into its potential "quasi-rigid domains", also calculating the distance fluctuations between pairs of amino acids. More specifically, MD trajectories treated with the SPECTRUS approach allow us to compare the functional dynamics of protein systems with different degrees of structural similarity. Similarly, this method uses ENM to reliably reproduce structural fluctuations.[30] Strategies based on the decomposition of quasi-rigid domains assume that, in protein systems that are genuinely rigid, the distances between two constitutive points (or amino acid residues) tend to be conserved during movement in space. Consequently, the SPECTRUS method starts from the calculation of the distance fluctuations for each pair of amino acids, a and b

$$f_{a,b} = \sqrt{\langle d_{a,b}^2 \rangle - \langle d_{a,b} \rangle^2} \tag{12}$$

where $d_{a,b}$ is the Cα atoms distance and the angled brackets denote the average over representative conformers from available crystal structures or sampled from MD trajectories. It has been described that the fluctuation in the quasi-rigid regions can be used to define a "quality score" that allows us to identify subdivisions associated with significant low-mobility (quasi-rigid) regions based on the balance of the distance fluctuations within and between domains. This balance (between distance fluctuations within and between domains) is accepted as crucial for the optimal clustering of protein rigid domains, and its validation through internal and external clustering validation allows cluster quality scores to be created.[31] In this sense, the quality score represents a general internal quantitative criterion for the decomposition of low-mobility regions and allows the identification of innate or

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

intrinsic quasi-rigid regions. Specifically, it has been shown that the quality score can convey how compact, not very mobile, and/or quasi-rigid the protein groups or regions are compared to reference cases. The higher the value of the quality score, the greater the probability of a more quasi-rigid system. For more details of the method, we recommend refs 30, 31.

Molecular Dynamics of Regions of Interest of the ClpXP Complex. For illustrative purposes and due to the computational cost, it was decided to carry out a preliminary study of the dynamic behavior only of the pore1 of the ClpXP complex (PDB: 6PP8). The modeling of the structure was established from the center of pore1 with a search box of a radius of 50 Å. This was done to cover both the surface of the pore1 and the substrate present in the crystal structure, as well as the bottleneck of the channel. We assume that the center of the crystal structure where the substrate is found is the best option to start the region to simulate, as has been suggested for the prediction of the center and size of the cavities.[32] To predict the region to be simulated, the CurPocket method of the CB-Dock server was used.[32,33] In this method, an integration between the curvature factor of the surface $\left(\frac{V_{\text{sphere}}}{V_{\text{out}}}\right)^2$ and the surface S accessible to the solvent is computed as

$$G = \iint_s \left(\frac{V_{\text{sphere}}}{V_{\text{out}}}\right)^2 \mathrm{d}s \tag{13}$$

In this model, surface area is dependent on curvature. The curvature factor $\left(\frac{V_{\text{sphere}}}{V_{\text{out}}}\right)^2$ is defined using a sphere centered at each sampling point on the solvent-accessible surface (S) and calculating the squared ratio of the volume of the sphere (V_{sphere}) to the volume exiting the solvent-accessible surface (V_{out}). Therefore, convex, flat, and concave surfaces are different in predicting hydrophobic contributions. The radius of the sphere (1.2 Å) and the radius of atoms were set by default according to the derivation of the proposed model.[32,34] The method generates a quadratic relation of the set of points to represent the solvent-accessible surface using the method developed by Shrake and Rupley.[34,35] Each point represents an integral element of the surface area. The method then places the points on a 0.1 Å grid and calculates the volumes of the spheres at each point by accumulating the volume of the grid. Finally, the model groups the square ratio products and the area of the surface element. The curvature factor implements the k-discrete oriented polytopes (k-DOPs) method to detect the superposition of spheres quickly.[34] The predicted region of the protein was modeled and prepared using the MMV-v.7.0 package and Chimera software.[14] ATP molecules are not considered in the simulation because they are outside the boundary of the proposed region. As usual, all water molecules as well as co-crystallization elements were removed and the PDB files were separated into different files using the MMV-v.7.0 package.[36]

Subsequently, simulations were carried out to sample the minimum-energy conformations and predict structural alterations of the pore1 in the presence and absence of substrate, as well as the stability of the binding of the substrate in the pore1. For a pore1−substrate complex, the proposed MD system was represented by three phases: relaxation, equilibrium (composed of two equilibrium runs), and a production phase to obtain the final trajectories of interest as suggested above.[37−43] The MD simulation applied to the crystal structures considered was carried out using an explicit water model. For them, we started with the solvation of the system in a box with a molecular size margin of 8 Å. The MD system also consisted of one copy of pore1 (with or without substrate). Amber99SB-ILDN force field was applied with TIP3P water. Ions were added proportionally to neutralize the system and simulate physiological conditions at a concentration of 0.2 M. Specifically, Na^+ ions were used as cations, Cl^- ions for anionic effect. Periodic boundary conditions were also applied and Berendsen's algorithm was used for coupling at constant temperature and pressure (300 K and 1 atm, respectively).

First, an MD cycle was performed using a 5000-step steep descent, followed by conjugate gradient minimization, also with 5000 steps applying positional constraints on the atoms of the pore1 complex (with and without substrate). Subsequently, a 100 ps simulation was performed with the positions of the atoms in the complex with a force constant of 10 kcal mol^{-1} $Å^{-2}$ to allow the water molecules to diffuse through the system and reach equilibrium. The water molecules were considered as rigid particles to perform a simulation step of 2 fs. The particle mesh Ewald method (PME) was used to calculate the electrostatic contribution of nonbonded interactions, this was done with a cutoff distance of 14 Å during a time period of 1 fs. Regarding the van der Waals interaction, the cutoff distance was 14 Å. After applying the MD cycle with higher descent (to prepare the systems), the canonical NVT environment set at 100 ps was used, in this case, the system was kept at constant temperature (thermalized at 300 K). Subsequently, and after the NVT cycle, a 100 ps cycle was followed with the NPT (isothermal−isobaric) environment to equilibrate the system at constant pressure and temperature (1 atm and 300 K, respectively). Additionally, the SHAKE algorithm (to compensate for geometric link constraints) was used for 2 fs. Finally, for the sampling of the minimum-energy structures, the NPT environment was established as the production phase, applying an MD cycle at 300 K and for 100 ns. Structures in PDB format were generated every 10 ns (up to a total of 100 ns) to be used in the following analyses. All MD simulations and additional adjustments were performed using cosgene/myPresto. Cosgene/myPresto which is available at https://www.mypresto5.jp/en/.[13,24,44,45] The thermodynamically most probable and favored position was calculated as well as the relative binding energy of the substrate at each point of minimum energy corresponding to each pore conformation after dynamics. For this, the MMV_7.0 package was used, calculating the thermodynamic mean from the MolDock, Rerank, and PLANTS functions.[46]

Additionally, to study the fluctuation of the pore1 and the bottleneck in terms of radius variations after the dynamics, MOLE 2.5 (https://mole.upol.cz/) was used, this tool allows the rapid and automated location and characterization of cavities, pockets, channels, tunnels and pores in complex structures, including macromolecular protein systems. MOLE is an optimized and complete algorithm that allows channels to be calculated using its "channels mode" (default), with which merged tunnels and pores can be predicted. On the other hand, it offers the "pore mode", designed to determine and characterize transmembrane pores. To calculate the pores, the tool defines the region of the membrane according to the information contained in the Orientations of Proteins in Membranes (OPM) database (http://opm.phar.umich.edu/) and applying the predictions of the MEMEMBED program

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

(which allows us to study structures without membrane annotation in OPM). Specifically, MOLE performs seven steps for prediction: (1) application of the Delaunay triangulation/Voronoi diagram for atomic centers; (2) construction of molecular surfaces; (3) cavity determination; (4) determination of possible starting points; (5) and end of channels; (6) channel modeling and (7) filtering of the located channels. For this, and in relation to the cavities, the "Radius of the probe" and "Interior threshold" parameters allow us to define the surfaces (modeled with the α shape) and the cavities (β shape) inside them, respectively. In addition, other parameters aimed at the description of channels such as the "Radius of origin" and "Radius of surface coverage" allow us to optimize the position of the start point to the closest end point on the surface, respectively. For further details, see refs 47−49.

The "search space calculation" module of the MMV-v.7.0 package was also used to predict the radius of a solid sphere and define the center of the pore and the bottleneck, to measure its variations.[34] Specifically, the MMV tool defines a solid sphere as the boundary where the search is centered. For this, MMV defines a spherical search space, its center, and its radius.[50,51] This solid sphere method is similar to the rolling probe method, which essentially works by rolling a virtual probe or ball of a given radius around the van der Waals surface of the macromolecule.[52−54] This approach enables grid-based cavity prediction by creating a discrete grid in the region of interest. Followed by locating a solid sphere (probe) of adjustable radius in the search region, and checking for sphere overlaps determined by the van der Waals radius of the protein atoms. The points calculated to describe the grid where the probe collides with the spheres of the protein atom are referred to as part of the inaccessible volume; all other points are called accessible. In this way, each accessible region point is checked to determine its belonging to the cavity.[55]

The widely used Molecular Mechanics/Poisson−Boltzmann (MM/PBSA) surface area-based approach[56−59] was considered as a thermodynamic integration method, to determine the relative binding energy of ligand−protein complexes formed by the substrate (ligand) and the region of the pore1 (protein) by calculating the free energy of binding (ΔG_{bind}) of the MD trajectories. The minimum-energy structure at each point of the minimized system was subjected to the MM/PBSA score implemented in AMBER16. The formulas for the estimation of the binding free energy of MM/PBSA are shown in eqs 14−19

$$\Delta G_{bind} = G_{complex} - (G_{receptor} + G_{ligand}) \quad (14)$$

$$= \Delta H - T\Delta S \quad (15)$$

$$\approx \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad (16)$$

(14)

$$\Delta E_{MM} = \Delta E_{bonded} + \Delta E_{ele} + \Delta E_{vdw} \quad (17)$$

$$\Delta G_{solv} = \Delta G_{polar} + \Delta G_{nonpolar} \quad (18)$$

$$\Delta G_{nonpolar} = \gamma \cdot \Delta SASA + \beta \quad (19)$$

where ΔG_{bind} is the total binding free energy. It is the free energy difference between the bound state ($G_{complex}$) and the free state ($G_{receptor} + G_{ligand}$) and can be also represented by the summation of the enthalpy part (ΔH) and the entropy part ($-T\Delta S$). In this work, the enthalpy changes were computed by the MM/PBSA approaches, whereas the entropy changes were neglected because of the expensive computational cost and low accuracy. The enthalpy part can be further decomposed into the molecular mechanical energy (ΔE_{MM}) and the solvation free energy (ΔG_{solv}). The ΔE_{MM} term includes the intramolecular (ΔE_{bonded}), electrostatic (ΔE_{ele}), and van der Waals (ΔE_{vdW}) energies. The ΔG_{solv} term contains both the polar (ΔG_{polar}) and nonpolar ($\Delta G_{non\ polar}$) contributions, where the polar contributions are accounted for by the Poisson−Boltzmann (PB) model and the nonpolar are assumed proportional to the solvent-accessible surface area (SASA) (see Table S1).[60] On the other hand, the algorithm based on the size-modified Poisson−Boltzmann equation (SMPBE) (https://web.uwm.edu/smpbs/) was used, as well as the APBS program (https://server.poissonboltzmann.org/) to be able to calculate the energy components and predict the electrostatic charge associated with solvation and free binding energy.[56] The ΔG_{bind} of the complexes was calculated on frames taken after a cycle of 100 ns. Molegro Molecular Viewer (MMV-7.0) was used.

RESULTS AND DISCUSSION

In Figure 1, the local environment of the ClpX system can be seen (highlighted in red) with sufficiently destabilizing

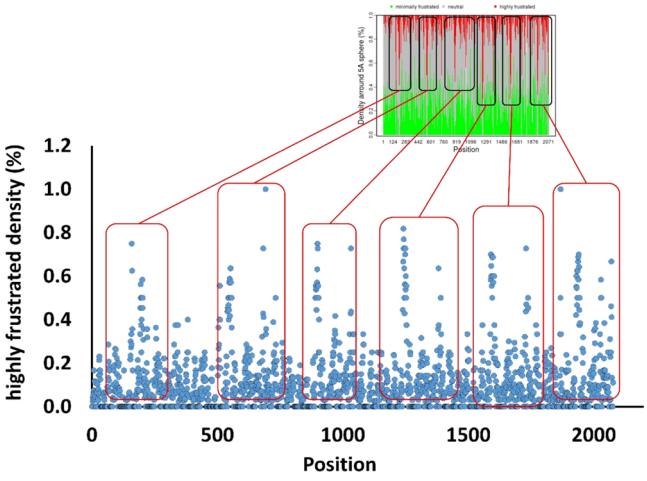


Figure 1. Energetic fingerprint of electrostatic interactions for system stabilization in the ClpX landscape. The local frustration index is shown, which allows observing the highly frustrated regions at the energy level, as well as the fraction or density in percentage expression (%) of the number of native contacts in each frustration class. The top panel shows the high (red bars), minimum (green bars), or neutral (gray bars) frustration around a simulated 5 Å sphere for each Cα (density around a 5 Å sphere) using a Debye−Hückel potential. In the bottom panel, only the high frustration density of the residuals per position for the ClpX region (blue dots) is highlighted.

configurational frustration, indicative of a highly frustrated interaction, or with an atypical energy fluctuation that compromises folding and local stability in the protein system. The values with the highest percentages of local frustration correspond to the amino acid regions modulated by electrostatic interactions predicted by the frustratometer algorithm.

Importantly, so-called long-range electrostatic interactions have been described as providing charge stabilization that channelizes the landscape for binding, and can also help guide the macromolecular docking of intrinsically disordered proteins before structure formation is complete. For other systems, long-range electrostatic interactions can also cause landscape frustration and can affect protein complex formation.[16] Currently, protocols studying the catalytic activity

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

of enzymes and proteins generally tend to omit electrostatic interactions characterized by being long-range (also known as protein electrostatic preorganization interactions) due to the difficulty associated with quantifying the optimal environment for the reactions. In this way, these observations are important because they provide information on the electrostatic preorganization of these types of protein systems, which is important to characterize the properties of their reactive state. Furthermore, it has been described that electric fields in general can influence the function of proteins.[61]

When studying the local energy contribution at the level of the subunits of the ClpX system, it was predicted that, in the characteristic destabilizing footprint, residues with a high energy fluctuation (highly frustrated) are found, including the residues suggested as keys for the formation of the structure adjacent to the pore called the IGF loop (see Figures 2 and 3).

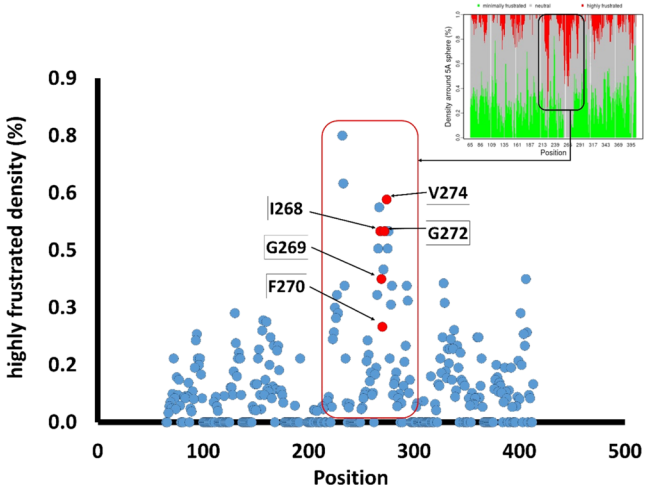


Figure 2. Energetic contribution of electrostatic interactions for the stabilization of the subunits of the ClpX system. The local frustration index is shown, which allows observing the highly frustrated residuals at the energy level (blue dots), as well as the fraction or density in percentage expression (%) of the number of native contacts in each frustration class. The top panel shows the high (red bars), minimum (green bars), or neutral (gray bars) frustration around a simulated 5 Å sphere for each Cα (density around a 5 Å sphere) using a Debye−Hückel potential. For illustrative purposes only, the bottom panel highlights the high frustration density of residues per position on chain A. Residues with high frustration values are shown, including residues suggested as keys to the formation of the pore-adjacent structure called the IGF loop (dots highlighted in red on the graph).

It is important to point out that models based on elastic networks for determining the configurational frustration of protein systems allow predicting intrinsic changes in the local energy contribution of residues that are not easily detectable with classical molecular dynamics methods. Especially since it has been described that an evolved sequence of a protein system can provide a minimally frustrated energy landscape, which results in the folding of a protein on a relatively short time scale, typically microseconds to milliseconds.[62,63]

Both the ClpX and the ClpP systems presented a sufficiently destabilizing and characteristic configurational frustration. In these two protein systems, residues with a high energy fluctuation are exposed adjacent to the pore (see Figure 4). These observations are consistent with the fact that neither the kinetic folding of proteins nor their mutational robustness has been reported to preclude the possibility that some conflicting

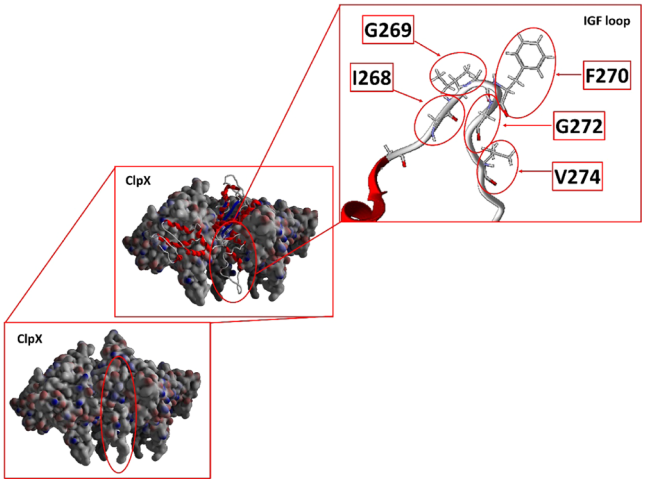


Figure 3. Three-dimensional (3D) representation of the ClpX module, highlighting the IGF loop. The region where residues with high frustration values are located within the ClpX module are shown, including residues suggested as keys for the formation of the pore-adjacent structure called IGF loop. The diagram in the bottom panel shows the ClpX module with hydrophobic surface, followed by the middle panel where the A chain in secondary structure is shown for illustrative purposes. The top panel shows the secondary structure and residues associated with the IGF loop.

signal frustration may be present locally. This kind of local frustration, if tolerable in the system, could arise naturally and could be a functionally useful adaptation. This is why the possible adaptive value of a molecule having a spatially localized frustration arises from the way in which such frustration can shape protein dynamics for specific biological functions.[64,65]

However, despite the behavior shown in Figure 4, a greater energetic frustration or instability of the regions associated with the IGF loop in the ClpX system was predicted (see Figure 3), in comparison to the energetic frustration of the total contribution of the key residues of the region ClpP determinants for loop formation (see Figure 5). On the other hand, a stabilization of the IGF loop-associated regions in the ClpX system was predicted after binding to the ClpP region. These results show that the ClpP system is decisive in contributing to the generation of a minimally frustrated system, or rather more stable, without undergoing significant changes in terms of its intrinsic energetic frustration (see Figure 5). This is important because the interactions that maintain macroscopic rigid systems have been described as extremely strong, which allows macroscopic machines to maintain their structural integrity over a wide range of conditions. Therefore, it has been pointed out that rigid macroscopic components tend to experience negligible changes in entropy.[66]

Stabilization of the complex from a highly frustrated system toward a minimally frustrated system by contribution of the ClpP region was predicted. This phenomenon is evidenced by determining the individual energy fluctuation of each of the residues associated with the IGF loop before and after the formation of the ClpXP complex (see Figure 6). This is important because it has been reported that native contacts whose energies are deep enough can influence the energetically stabilizing part of the distribution,[67] which is key because residues in the pore1 loop have been described adjust the properties of the central channel.[68] Thus, pore alterations have

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

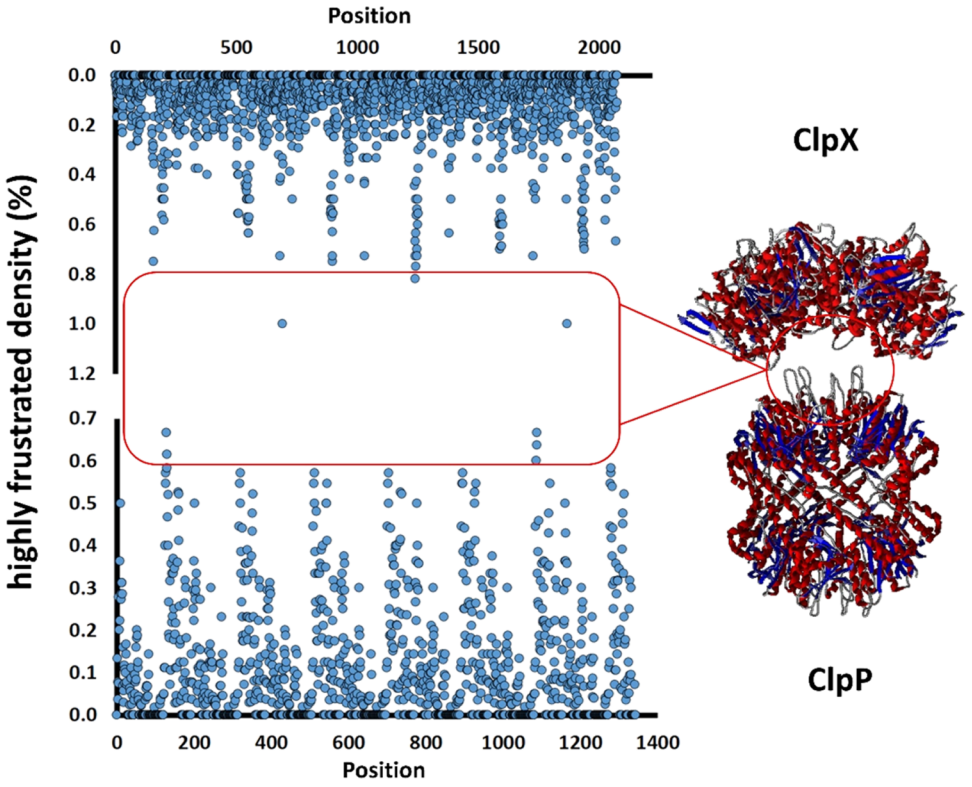


Figure 4. Energetic frustration of the residues of the ClpXP system. The local frustration index of the entire ClpXP system is shown, which allows observing the highly frustrated regions at the energy level, as well as the fraction or density in percentage expression (%) of the number of native contacts in each calculated frustration class by a Debye−Hückel potential. For illustrative purposes, the central region of the secondary structure of the ClpXP system is shown where the highly frustrated residues are located (blue dots).

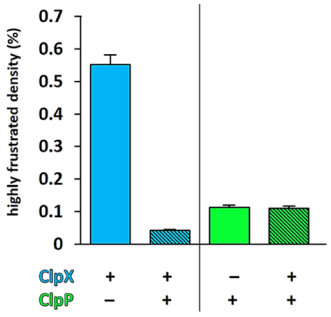


Figure 5. Intrinsic energy frustration of the ClpXP system. Specifically shown are the regions associated with the IGF loop in the ClpX system, bound or not, to the ClpP region. A bar graph of the average local high frustration of total residues associated with the IGF loop is shown. The calculations were made considering the individual and joint energy contribution of the ClpX, ClpP regions, and the ClpXP complex. The values represent the fraction or density of frustration in percentage expression (%) of the number of native contacts calculated by a Debye−Hückel potential. The bottom panel represents the contribution (+) or not (−) of the system (as the case may be) to the calculation of the frustration density.

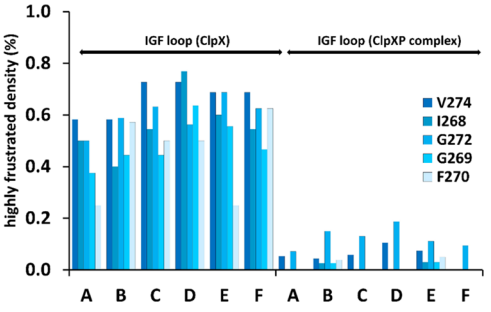


Figure 6. Individual energy fluctuation of IGF loop-associated residues before and after ClpXP complex formation. The letters A, B, C, D, E, and F on the X axis correspond to the ClpX subunits. A bar graph of the high local frustration of IGF loop-associated residues (frustration was calculated from the frustration index of each residue) before and after ClpXP complex formation is shown. The values represent the fraction or density of frustration in percentage expression (%) of the number of native contacts calculated by a Debye−Hückel potential.

the potential to influence both the chemical and mechanical properties of the engine.[68]

Although these results would allow us to infer that there may be a significant energetic contribution of the proposed amino acid residues (V274, I268, G272, G269, and F270) to the energetic and structural deformability of pore1, these predictions could be misleading. This is due to the presence of values with nonsignificant differences. Specifically, the increase in the frustration of considered amino acid residues as keys for pore1 (see Figure 7) on complex formation is not significantly high (see Figure 8). In contrast, the decrease in frustration of IGF loop amino acid residues was much greater (see Figure 6), demonstrating that pore interactions do not fully compensate for IGF loop stabilization. In this sense, it is suggested to incorporate a greater number of residues to the calculation, in addition to those proposed by the crystallographic description,[10] to increase the number of residues that contribute or not, with the fluctuation observed at a global level.

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

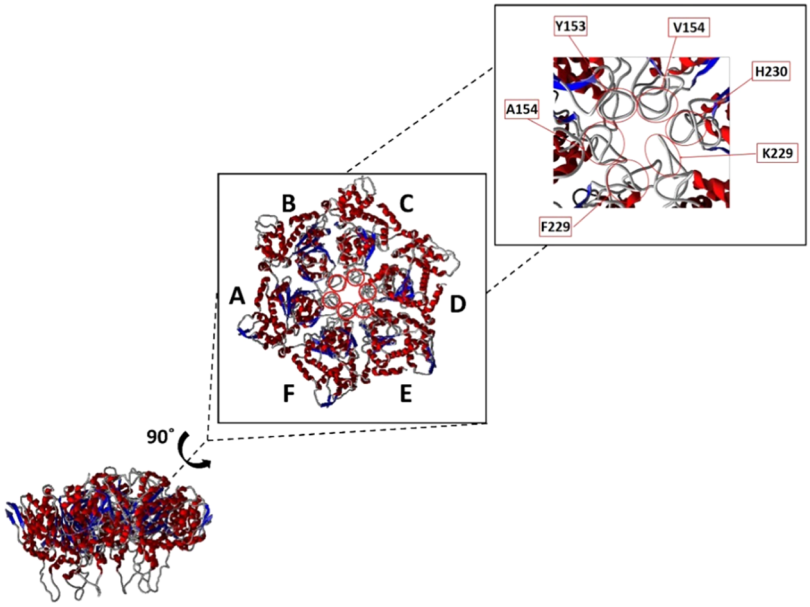


Figure 7. 3D representation of pore1 highlighting residues associated with substrate interaction. The pore1 region is shown where the residues associated with the interaction with the substrates within the ClpX module are located. The bottom panel diagram shows the secondary structure of the ClpX module seen from the side, followed by the middle panel showing the ClpX module rotated 90°, indicating the location of each of its chains. The top panel shows the secondary structure and residues associated with pore1 associated with the interaction with the substrates, and considered to study their energetic frustration.

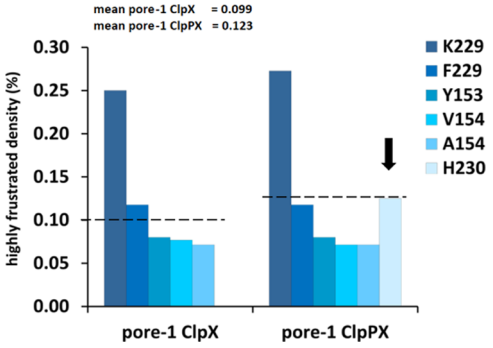


Figure 8. Energy fluctuation at the pore level before and after the formation of the ClpXP complex. A bar graph of the high local frustration of pore1-associated residues (frustration was calculated from the frustration index of each residue) before and after ClpXP complex formation is shown. The values represent the fraction or density of frustration in percentage expression (%) of the number of native contacts calculated by a Debye−Hückel potential.

On the other hand, it is worth highlighting the sensitivity of the model to slight changes in the local and individual energy contribution of the waste (see Figure 8), and in a strict sense, at the level of the pore1 (and only with the information of the relevant waste) (see Figure 7), an approximation can be seen energy behavior contrary to that observed in the adjacent regions. Specifically, key pore-associated residues exhibit minimally frustrated energy fluctuation, compared with higher energy frustration or local instability following ClpXP complex formation (see Figure 7).

These predictions relate to the existence of compensatory interactions between interacting partners in a complex protein system, which has been shown to be important in understanding the mechanisms of interaction.[15] Furthermore, preferences in the interactions involved have been described to be greatly affected by additional binding, indicating that the minimal frustration of general folding is compensated by specific binding. In other words, the conflict with the total folding of the minimum frustration is compensated by the minimum frustration of the unión.[69]

As expected, and as predicted by energy fluctuation, the isolated ClpX system presented greater structural flexibility compared to the ClpP system, which presented greater structural quasi-rigidity indicative of a more stable system. In addition, the ClpXP complex presented a quasi-rigidity lower than that predicted for the ClpP system and higher than that of the ClpX system. This allows it to be inferred that the ClpP system contributes to the stability in terms of quasi-rigidity of the ClpX region. It is also observed how the flexibility of the system fluctuates as the complex assembles, with a rigidity that decreases slightly as substrate and ATP are incorporated (see Figure 9). The energy fluctuation predicted from the configurational frustration calculation is associated with the stability and deformability of the systems in terms of the flexibility and rigidity of the protein domains. An approach based on models of elastic networks allows observing the intrinsic stability of protein systems. In this sense, predicting the flexibility of the quasi-rigid domains is crucial to know the effect of the interactions between the domains of the ClpXP complex, as well as between the elements and biomolecules associated with the ClpXP complex, as has already been predicted for other protein systems.[15,30,69]

However, as indicated for the local contribution of the residues to the frustration of the pore1 (see Figure 8), since values with nonsignificant differences are present, it is suggested to incorporate a greater number of residues into the calculation. This allows for a better approximation to the contribution of residues to the stability and stiffness of the system. We also highlight the sensitivity of the SPECTRUS model to predict slight changes in the local and individual contribution of residues to stiffness (see Figure 9).[30,31]

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

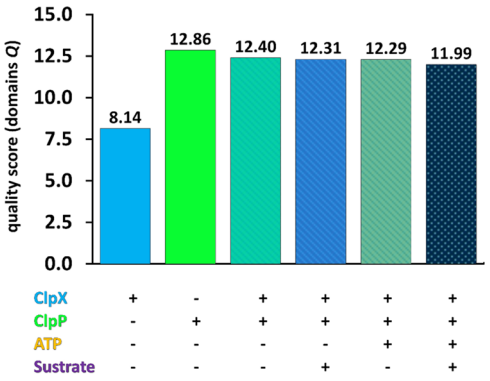


Figure 9. Flexibility of the quasi-rigid domains of the ClpXP complex. The fluctuation of quasi-rigid systems is shown in terms of the quality score offered by the SPECTRUS method for various combinations of complexes considered here. This approach allows subdivisions associated with low-mobility (quasi-rigid) regions to be identified and scored based on the balance between distance fluctuations within and between domains (between pairs of amino acids). The higher the value of the quality score, the higher the probability of a more nearly rigid system. The lower panel represents the contribution (+) or not (−) of the system (as the case may be) to the calculation of the quality score.

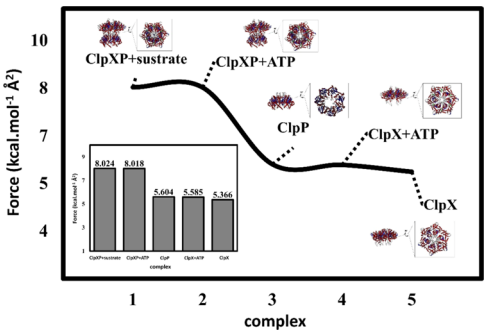


Figure 10. Interaction strength of the ClpXP complex in structural communication. For illustrative purposes only, a possible progression or trajectory of the values of the force constant is shown using a spline-type curve to smooth its oscillations. The bottom panel shows the graphical representation of the force constant values in a bar chart. This prediction was made from the ENM approaches proposed in this study, which allows calculating the cross-correlation of atomic movement and a force constant associated with the distance of the pairs of particles (Cα atoms) that interact in the protein systems studied.

In fact, contributions to conformational changes mediated by the effect of the interaction with molecules such as ATP can also be detected with this type of model based on elastic networks.[70] This is important because it has been described that exothermic chemical reactions such as that generated in the hydrolysis of ATP can be an important source of forces that drive changes and conformational transitions in proteins, including fluctuations in the rigidity of the systems.[70] In addition, it has been described that some protein systems can behave as molecular switches using the binding of nucleotides such as ATP to alternate between different active states,[71] and to drive the translocation of molecules through cooperative conformational changes.[72] It has even been shown that the presence of ATP in protein systems can generate proteins with distinct states,[73] as well as diverse topographical populations.[74] Specifically, it has been proposed in ATPase-type models that the conformational changes may be associated with the closing and opening movements caused by the binding and hydrolysis of ATP, respectively, due to the effect of a rearrangement of the network of hydrogen bonds around ATP.[75]

For illustrative purposes, we sought to determine a possible progression or trajectory of the values of the force constant, adjusting it with a spline-type curve to smooth its fluctuations or oscillations, in parallel with its graphic representation in a bar diagram (see Figure 10). Especially because the ENM approaches proposed here implement two methods to calculate the cross-correlation of atomic motion and from there allow us to predict the force constant associated with the distance of the pairs of particles (Cα atoms) that interact in the protein systems studied.[19−23,25,26,28,29] It was observed that the isolated ClpP and ClpX systems (with and without ATP) presented a lower interaction force than the ClpXP complex with ATP, followed by the complex coupled to the substrate (see Figure 10). Likewise, a greater stability of the ClpP system was predicted compared to ClpX (see Figure 10). Although these predictions based on theoretical interaction strength require experimental demonstration, they are interestingly related to experimental studies of translocation kinetics showing that after prior assembly of the ClpXP complex, subsequent addition of ATP can mediate substrate degradation,[76] furthermore, ClpP has been observed to initially form an encounter complex with ClpX, which is then stabilized by unimolecular coupling of IGF loops with ClpP.[77] This corresponds to the dynamic behavior of other protein systems that present a progression in their functioning represented by accompanying cycles. In initial instances, they can modulate the interaction with ATP molecules and then continue with the recruitment of substrates with various degrees of specificity.[78] This is important because it has been described that when residues stably intercalate into the incoming amino acid (or polypeptide) chain within the central channel, the pore loops in AAA+ systems translate the generated force into movements of the domain that are propelled by the hydrolysis of ATP within the motor to the substrate and thus are responsible for mediating both ATP coupling and hydrolysis, as well as substrate remodeling. Furthermore, the binding of ATP within the subunits at the top establishes stabilizing interactions with the neighboring subunit, and consequently, its pore1 loop rearranges to interact with the substrate.[68]

The predictions about the quasi-stiffness distribution of the system demonstrate that although ClpP is a theoretically much stiffer system than ClpX (see Figure 9), both ClpP and ClpX present similar values of the ENM force constants (see Figure 10). The similarity between the force constants of the studied systems (ClpX and ClpP) applying the ENM method, allows us to describe the stability of these systems through the coordinates of C atoms that interact through a harmonic potential, with a constant of force inversely proportional to the distance between the interacting C atoms, as has already been predicted for other protein systems.[19]

To predict pore volume fluctuations over 100 ns of simulation, an explicit water aqueous medium with physiological density was modeled and the MOLE 2.0 algorithm was used, which can calculate pores from channel detection. With this method, the so-called "pore criterion" is used, which stipulates that the end points of the pore must be farther apart than the average length of the channels that formed the pore (see Figure 11). Similarly, pore fluctuation over time was compared using the MMV, which applies a probe from the

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

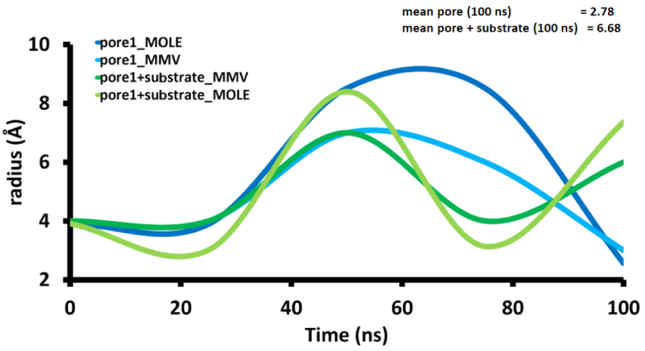


Figure 11. Fluctuation of the pore in physiological medium as a function of time. A comparative analysis between the trajectories of the radius of pore1 in the presence and absence of substrate predicted with a specialized algorithm based on the Delaunay triangulation/Voronoi diagram (MOLE) is shown, with a rolling solid sphere method on the van der Waals surface (MMV).

center of the pore. In this method, the distance from the center of the pore to the contact with the closest residue group simulated on an electrostatic surface was used as a criterion to delimit the pore (see Figure 11).

Although a mean of ≈5 Å was predicted, the pore means at 100 ns with (green) and without (blue) substrates were 6.68 and 2.78 Å, respectively. Indicative of an unfolding of the pore in the presence of the substrate (see Figure 11). This corresponds to a state of higher energy frustration in the complexed state, in which the pore has greater flexibility to facilitate interaction with the substrate (see Figure 12).

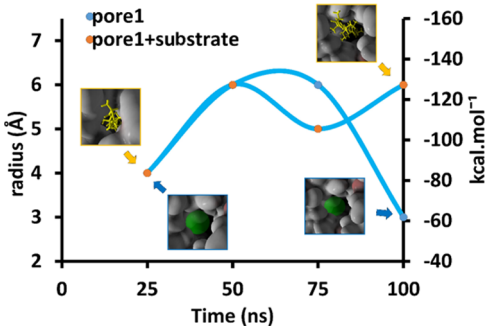


Figure 12. Fluctuation of the pore and energy of interaction with the substrate in physiological medium as a function of time. The comparative analysis between the relative energy of binding (ΔG_{bind}) of the MD trajectories of the ligand−protein complexes formed by the substrate (ligand) and the pore1 region (protein) with the Molecular Mechanics/Poisson−Boltzmann surface area approach is shown (MM/PBSA); vs mean pore radius fluctuation in the presence and absence of substrate calculated with Delaunay triangulation/Voronoi diagram and solid sphere method on van der Waals surface. The trajectories are shown as a function of each point or structure of minimum energy during the simulation period of 100 ns.

It is important to note that the measurement of the size of the cavities of interest (in this case both the pore and the bottleneck) was performed as usual using the coordinates of the atoms of the final snapshot of the protein at 100 ns MD simulation.[79] After the complete MD simulation cycle at 100 ns, the considered methods did not present significant differences in the arithmetic mean of the free pore radius (MOLE ≈ 5.5 Å and MMV ≈ 4.8 Å) (t = 0.46, GL = 8, p = 0.655, α = 0.01) and with substrate (MOLE ≈ 5.2 Å and MMV ≈ 5.0 Å) (t = 0.12, GL = 8, p = 0.902, α = 0.01). Regarding the arithmetic mean of the bottleneck after the complete MD simulation cycle at 100 ns, the considered methods also did not present significant differences between the free bottleneck (MOLE ≈ 2.60 Å and MMV ≈ 2.64 Å) (t = 0.05, GL = 8, p = 0.958, α = 0.01) and with substrate (MOLE ≈ 2.78 Å and MMV ≈ 2.50 Å) (t = −0.41, GL = 8, p = 0.687, α = 0.01). On the other hand, the values of each method for the minimum-energy structures (specifically at 100 ns MD) of the pore radius (MOLE ≈ 2.5 Å and MMV ≈ 3.0 Å) and bottleneck (MOLE ≈ 7, 3 Å and MMV ≈ 6.0 Å) were similar with a difference of ≈ 0.4 and ≈ 1.3, respectively. While the values of each method for the minimum-energy structures (specifically at 100 ns MD) of the free bottleneck radius (MOLE ≈ 1.8 Å and MMV ≈ 1.5 Å) and with substrate (MOLE ≈ 1, 4 Å and MMV ≈ 1.8 Å) were also similar between methods with a difference of ≈0.3 and ≈0.4, respectively. In the same way, the values of the pore size and the bottleneck were calculated at each point of minimum energy along the trajectory (see Supporting Information Tables S2 and S3), as has also been suggested.[80] Each method (MOLE and MMV) was able to independently reproduce the trend of the other under the conditions of this study. Both MOLE and MMV predicted the pore and bottleneck size of the reported crystallographic structure in the presence of the substrate (in terms of radius).[10] However, to validate the predictions of the sampled minimum-energy structures for the calculation of the pore1 and bottleneck radius, the protein system was examined during a longer simulation period (200 ns). It was observed that after the simulation period at 200 ns, the previously predicted trend was maintained in terms of the fluctuations of the pore1 and bottleneck radius (in the presence and absence of substrate). At the same time, the reported fluctuations in the crystallographic structure (in terms of the radius) for the pore and bottleneck in the presence of the substrate were predicted (see Supporting Information Figures S1 and S2).

This corresponds to what has been described in protein systems where, depending on the orientation and the distance between the side chains, it is likely that the stabilized loops of the pore will acquire some degree of flexibility to support the peptide space.[81] It has even been pointed out that the intrinsic flexibility of proteasomal receptors and the dynamics of their ubiquitin interactions are likely paramount to substrate processing.[82] It is important to point out that despite the fluctuating behavior of the dynamics obtained, the simulation period used has been widely recommended in several studies with similar oscillating behaviors and with trajectories that tended to be defined or discriminated at 100 ns with protease-type enzymes[83−87] including systems associated with Clp[88] and even smaller simulation periods.[89−91] In addition, it is important to note that the simulation conditions used at 100 ns over pore1 in the presence of the substrate reproduced the reported radius (6−8 Å) for the crystallographic structure considered in this study.[10] On the other hand, it is highlighted that despite having used two different methods to measure the radius of the pore1, and that these yielded results with different magnitudes, both were able to reproduce (at 100 ns) the previously reported radius of the pore1, and with a similar trend between the trajectories of both methods. This allows us to infer that both the specialized algorithm based on the Delaunay triangulation/Voronoi diagram[47−49] and the rolling probe method based on the van der Waals surface[51,55] allow us

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

to make an approximate prediction and with similar trajectories of the experimental radius reported for the pore1.[10]

However, to validate the predicted behavior, the fluctuation of the system at 200 ns was examined, as has also been suggested for this type of protein systems,[92] maintaining a behavior similar to that predicted at 100 ns of simulation for both the pore and the neck of the free bottle and with substrate. Minimum-energy trajectories at 200 ns were observed that oscillate within what was previously predicted and consistent with the respective radius of the crystal structure (see Supporting Information Figures S1 and S2). Although the time scales accessible for MDs prevent the direct observation of events on microsecond scales with reliable statistical sampling, the behavior of fluctuations that occur in short times is capable of providing reproducible information in terms of structural flexibility, as has been pointed out.[92] In fact, the reproducibility of trajectories on the proposed time scales is in excellent agreement with that identified in smFRET (single-molecule Förster resonance energy transfer) experiments for functional motions of Clp-like structures.[93] As well as to be able to know the effect of the early perturbation in the dynamics of pores, and in simulations of Clp systems that include the elimination of nucleotides and/or the substrates.[92]

Studying the relative energy of binding between the amino acid residues of the pore and the substrate, and its relationship with variations in the volume of the pore (as a function of time), similar results to those predicted were obtained in terms of binding strength. It was predicted that the larger the pore size, the higher the binding energy with the substrate (see Figure 12). It was observed that after a simulation at 100 ns in a physiological medium, the free pore (without substrate) underwent folding, while the presence of the substrate predicted an unfolding of the pore. These results correspond with the increase in pore flexibility previously predicted by the comparative analysis of pore volume over time, and with the energetic fluctuation in terms of energetic frustration (see Figure 12). It has been reported that this type of system can present more compact pores that can increase in size to favor the degradation of unfolded proteins.[94,95] An important aspect because it is known that the molecular architecture of Clp-like systems limits the activity of the enzyme to peptide degradation, with a size restriction that is reinforced by the size of the axial channel,[96−99] it is likely that different unfolded substrates translocate into different conformations, and that the size of the cavities for translocation steps may also vary with sequence.[99,100] Indeed, it has been described that the domains of Clp-like systems can unravel to allow the pores to widen to accommodate larger substrates, and then fold back to allow the pores to close again and maintain intimate contacts with smaller substrates.[99,101] This mechanism of pore expansion-contraction via a "snake jaws model" has already been suggested, in which the pore size is controlled by the size of the substrates and the conformation, and structure of the ClpX subunits.[101] Second, it has been reported that the pore loops exhibit large-amplitude fluctuations on the microsecond time frame and change their conformation upon substrate engagement and ATP hydrolysis, indicating that these motions might drive substrate translocation.[9]

To study the contribution of pore flexibility in the interaction with the substrate, pore size fluctuation with and without substrate was examined throughout the simulation. It was also compared with the region called "bottleneck" corresponding to the narrowest region of the channel, considering the presence and absence of the substrate. Although a conformational unfolding of both the pore and the bottleneck was predicted at ≈50 ns of simulation. The pore and the bottleneck in the free state (without substrate) showed a conformational change at 100 ns of the simulation with a radius of ≈4 and ≈1.5 Å, respectively, while the pore and the bottleneck in the presence of the substrate presented a radius at 100 ns of simulation of ≈8 and ≈2 Å, respectively (see Figure 13).

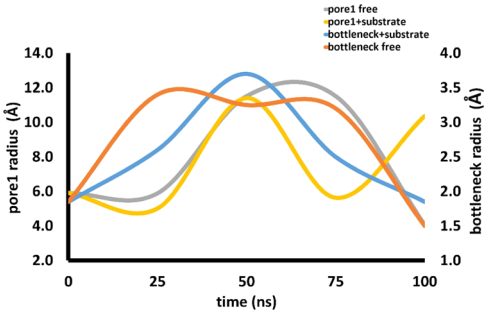


Figure 13. Pore and bottleneck radius fluctuation with and without substrate at 100 ns simulation. The comparison between the trajectories of the bottleneck radius and pore1 with and without substrate as a function of time is shown. The values correspond to the mean of the predictions made with the Delaunay triangulation/Voronoi plot and with the solid sphere method on the van der Waals surface.

These results are important because our predictions reproduced the reported pore and bottleneck radius[10] and could suggest under the conditions of this study that the mechanism of the interaction of the ClpX system with the substrate passes through unfolding of the pore in parallel with folding of the bottleneck. This could regulate the passage of the peptide through the channel, with an amplitude that could theoretically allow the passage of ≥3 residues through the pore during the simulated period. Specifically, the predicted pore distance in the presence and absence of substrate and its relationship to the binding strength of the complexes and relative binding energies show a propensity for long translocation steps. During the simulation in aqueous medium, a pore unfolding with a mean distance of up to ∼0.6 nm was predicted in the presence of the substrate. An increase in the pore from the free state to the complexed state (with substrate) of between 2 and 3 nm was predicted. The variations in the distance calculated by molecular dynamics during a period of 100 ns could theoretically allow the passage of substrate with a size equivalent to about three residues, as has been suggested.[10] These results indicate that the translocation step for this system could occupy variable lengths of the substrate, thanks to splittings that would favor longer steps that presumably represent kinetic bursts of this fundamental step as described.[10] The simulation periods applied here (up to 200 ns) (see Supporting Information Figure S1) show a cyclical trend in the fluctuations of the simulated pore1 and bottleneck system with and without substrate. This shows the relevance of applying prolonged simulation periods, especially because when comparing the ratio between the pore radius with and without substrate, as well as the ratio of the bottleneck with and without substrate, the fluctuation of these regions with tendency to fold or unfold inversely proportional between these two regions of the ClpX

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

system during the considered simulation period (see Figure 14).

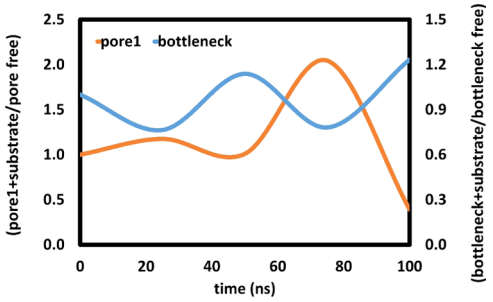


Figure 14. Ratio between pore and bottleneck fluctuation at 100 ns simulation. A comparison between trajectories of the "pore1+substrate/pore1 free" ratio versus the "bottleneck+substrate/bottleneck free" ratio as a function of time is shown.

Thus, ClpXP could operate through a probabilistic mechanism in which more than three residues were translocated in each unfolding cycle. Especially since a major problem in describing the mechanism of Clp-like systems has to do with the duration of the deployment (increased distance) and translocation (decreased distance) steps. Specifically, while an SC/2R-type model predicts the passage of two residues, the fluctuations in pore size1 predicted in this study correspond to the basic translocation step reported experimentally for ClpXP systems by optical capture, which is associated with the passage of up to six residues, with even longer oscillating steps (see Supporting Information Figure S2).[10,102]

Our simulations predict fluctuations potentially related to unfolding steps at the pore1 level (increase in distance) and translocation at the bottleneck level (decrease in distance) in a time between 100 and 200 ns (see Figures 13 and 14 and Figures S1 and S2). This is important because it has been pointed out that the mechanism described here is not strictly due to a step of ≈6 residues that is the product of three SC/2R substeps.[10] In fact, this mechanism (SC/2R) is not favorable because its time constant from the beginning to the end of each translocation step requires between 0.1 and 0.2 s (≈1×10^8 ns) in optical traps.[10,102]

In summary, the models proposed here allowed to establish a framework of deformability of the ClpXP system considered, as well as the contributions of its subunits, in the presence and absence of substrate, and ATP, on said collective movements, fluctuations of residues, and conformational and energetic changes (local and global). This opens the doors for future research aimed at studying possible conformational and energetic pathways between the different states (folded and unfolded) of simulated complexes, as has been suggested.[103] The possibility of characterizing the movements in this type of biomedically relevant protein systems has also made it possible to examine the vibratory movements of systems such as kinesin, myosin and F1-ATPase, exposing limitations in ENM models to capture conformational changes in associated globular regions to nucleotides.[104] This aspect could be circumvented if the determination of the quasi-stiffness regions is sought as shown in this study since the model showed sensitivity to ATP contributions as reported.[70−75] These approaches are important because the vital nature of tightly controlling the activity of Cp-like systems and their control mechanisms associated with conformational folding[105] has been demonstrated, and how perturbing the degrees of freedom of such machinery could alter the essential stages such as, for example, molecular coupling.[106] All of these are key aspects with therapeutic interest.[7,8]

CONCLUSIONS

The proposed ENM suggest that the ClpP region is decisive for the stabilization of the ClpXP complex, contributing to the flexibility of the residues adjacent to the pore, favoring the increase in pore size and, therefore, with the energy of interaction of its residues with a larger portion of the substrate. When analyzing the behavior of the elements of the ClpXP complex with ENM, it is suggested that the complex may undergo a stable configurational change once assembled and that the deformability of the system once assembled is oriented, on the one hand, to the increase in the rigidity of the domains of each region (ClpP and ClpX) and, on the other hand, to the flexibility gain of the pore. Subsequently, and according to the ENM approach, the interaction between the complex and the substrate could occur before the interaction with ATP, modifying the volume of the pore and generating a thermodynamically favorable bond with the substrate. It was also observed that the subsequent incorporation of ATP contributes to the dynamization of the complex in terms of an increase in its global flexibility. In terms of the kinetics associated with the increase in pore size, it was observed that the calculated trajectories required ≈ 100 ns under the conditions of this study to simulate the fluctuations presented. Our predictions suggest that the mechanism of the interaction of the ClpX system with the substrate passes through the unfolding of the pore in parallel with a folding of the bottleneck. This mechanism of pore unfolding in parallel with a folding of the bottleneck was predicted in a period of time incompatible with the times suggested at the experimental level for the hydrolysis models type SC/2R. The conclusion is that due to the theoretical behavior of the pore, along with the greater stability and binding energy to the substrate under the presented conditions, simultaneous translocation of >2 residues may be possible, and not just in a strictly sequential manner.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c04347.

Detailed parameters for the MM/PBSA calculation procedures (Table S1); pore and bottleneck radius fluctuation with and without substrate at 200 ns simulation (Figure S1); axial channel dimensions (Figure S2); values of the pore size at each point of minimum energy along the trajectory at 100 ns (Table S2); and values of the bottleneck at each point of minimum energy along the trajectory at 100 ns (Table S3) (PDF)

AUTHOR INFORMATION

Corresponding Authors

Freddy Romero − *Pulmonary, Critical Care and Sleep Medicine, Baylor College of Medicine, Houston, Texas 77030, United States*; Email: Freddy.Romerovasquez@bcm.edu

Ysaias J. Alvarado − *Centro de Biomedicina Molecular (CBM), Laboratorio de Química Biofísica Teórica y*

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

Experimental (LQBTE), Instituto Venezolano de Investigaciones Cientificas (IVIC), 4001 Maracaibo, Zulia, República Bolivariana de Venezuela; orcid.org/0000-0002-2709-409X; Email: alvaradoysaias@gmail.com

Authors

Lenin González-Paz − *Facultad Experimental de Ciencias (FEC), Departamento de Biología, Laboratorio de Genética y Biología Molecular (LGBM), Universidad del Zulia (LUZ), 4001 Maracaibo, Zulia, República Bolivariana de Venezuela; Centro de Biomedicina Molecular (CBM). Laboratorio de Biocomputación (LB), Instituto Venezolano de Investigaciones Científicas (IVIC), 4001 Maracaibo, Zulia, República Bolivariana de Venezuela*

Carla Lossada − *Centro de Biomedicina Molecular (CBM). Laboratorio de Biocomputación (LB), Instituto Venezolano de Investigaciones Científicas (IVIC), 4001 Maracaibo, Zulia, República Bolivariana de Venezuela*

Maria Laura Hurtado-León − *Facultad Experimental de Ciencias (FEC), Departamento de Biología, Laboratorio de Genética y Biología Molecular (LGBM), Universidad del Zulia (LUZ), 4001 Maracaibo, Zulia, República Bolivariana de Venezuela*

Francelys V. Fernández-Materán − *Centro de Biomedicina Molecular (CBM). Laboratorio de Biocomputación (LB), Instituto Venezolano de Investigaciones Científicas (IVIC), 4001 Maracaibo, Zulia, República Bolivariana de Venezuela*

José Luis Paz − *Departamento Académico de Química Inorgánica, Facultad de Química e Ingeniería Química, Universidad Nacional Mayor de San Marcos, 15081 Lima, Perú*

Shayan Parvizi − *Pulmonary, Critical Care and Sleep Medicine, Baylor College of Medicine, Houston, Texas 77030, United States;* orcid.org/0000-0001-5090-7303

Rafael Eduardo Cardenas Castillo − *Pulmonary, Critical Care and Sleep Medicine, Baylor College of Medicine, Houston, Texas 77030, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c04347

Author Contributions

L.G.-P. and Y.J.A. contributed to conceptualization, data curation, formal analysis, investigation, methodology, and writing—original draft. M.L.H.-L., C.L., and F.V.F.-M. performed investigation. J.L.P. carried out supervision, validation, and writing—original draft. S.P. and R.E.C.C. performed supervision and writing—review & editing. F.R. contributed to conceptualization, data curation, formal analysis, investigation, methodology, and writing—review & editing.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sector.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors express their deep gratitude to all of the reviewers of this study for their critical reading and recommendations that guaranteed the increase in the quality of the manuscript.

■ REFERENCES

(1) Campitelli, P.; Modi, T.; Kumar, S.; Ozkan, S. B. The role of conformational dynamics and allostery in modulating protein evolution. *Annu. Rev. Biophys.* **2020**, *49*, 267−288.
(2) Goossens, K.; De Winter, H. Molecular dynamics simulations of membrane proteins: An overview. *J. Chem. Inf. Model.* **2018**, *58*, 2193−2202.
(3) Smith, Z.; Ravindra, P.; Wang, Y.; Cooley, R. M.; Tiwary, P. Discovering protein conformational flexibility through artificial-intelligence-aided molecular dynamics. *J. Phys. Chem. B* **2020**, *124*, 8221−8229.
(4) Savitski, M. M.; Zinn, N.; Faelth-Savitski, M.; Poeckel, D.; Gade, S.; Becher, I.; Bantscheff, M.; et al. Multiplexed proteome dynamics profiling reveals mechanisms controlling protein homeostasis. *Cell* **2018**, *173*, 260−274.
(5) Wu, N. N.; Zhang, Y.; Ren, J. Mitophagy, mitochondrial dynamics, and homeostasis in cardiovascular aging. *Oxid. Med. Cell. Longevity* **2019**, *2019*, No. 9825061.
(6) McGrath, P. T.; Iniesta, A. A.; Ryan, K. R.; Shapiro, L.; McAdams, H. H. A dynamically localized protease complex and a polar specificity factor control a cell cycle master regulator. *Cell* **2006**, *124*, 535−547.
(7) Szczepanowska, K.; Senft, K.; Heidler, J.; Herholz, M.; Kukat, A.; Höhne, M. N.; Trifunovic, A.; et al. A salvage pathway maintains highly functional respiratory complex I. *Nat. Commun.* **2020**, *11*, No. 1643.
(8) Pryde, K. R.; Taanman, J. W.; Schapira, A. H. A LON-ClpP proteolytic axis degrades complex I to extinguish ROS production in depolarized mitochondria. *Cells Rep.* **2016**, *17*, 2522−2531.
(9) Lin, J.; Shorter, J.; Lucius, A. L. AAA+ proteins: one motor, multiple ways to work. *Biochem. Soc. Trans.* **2022**, *50*, 895−906.
(10) Fei, X.; Bell, T. A.; Jenni, S.; Stinson, B. M.; Baker, T. A.; Harrison, S. C.; Sauer, R. T. Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. *eLife* **2020**, *9*, No. e52774.
(11) Fuglebakk, E.; Tiwari, S. P.; Reuter, N. Comparing the intrinsic dynamics of multiple protein structures using elastic network models. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850*, 911−922.
(12) Ahmed, A.; Rippmann, F.; Barnickel, G.; Gohlke, H. A normal mode-based geometric simulation approach for exploring biologically relevant conformational transitions in proteins. *J. Chem. Inf. Model.* **2011**, *51*, 1604−1622.
(13) González-Paz, L.; Hurtao-León, M. L.; Lossada, C.; Fernández-Materán, F. V.; Vera-Villalobos, J.; Loroño, M.; Paz, J. L.; Jeffreys, L.; Alvarado, Y. J. Comparative study of the interaction of ivermectin with proteins of interest associated with SARS-CoV-2: A computational and biophysical approach. *Biophys. Chem.* **2021**, *278*, No. 106677.
(14) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(15) Rausch, A. O.; Freiberger, M. I.; Leonetti, C. O.; Luna, D. M.; Radusky, L. G.; Wolynes, P. G.; Ferreiro, D. U.; Gonzalo Parra, R. FrustratometeR: an R-package to compute local frustration in protein structures, point mutants and MD simulations. *Bioinformatics* **2021**, *37*, 3038−3040.
(16) Parra, R. G.; Schafer, N. P.; Radusky, L. G.; Tsai, M. Y.; Guzovsky, A. B.; Wolynes, P. G.; Ferreiro, D. U. Protein Frustratometer 2: a tool to localize energetic frustration in protein molecules, now with electrostatics. *Nucleic Acids Res.* **2016**, *44*, W356−W360.
(17) Jenik, M.; Parra, R. G.; Radusky, L. G.; Turjanski, A.; Wolynes, P. G.; Ferreiro, D. U. Protein frustratometer: a tool to localize energetic frustration in protein molecules. *Nucleic Acids Res.* **2012**, *40*, W348−W351.
(18) Chen, J.; Schafer, N. P.; Wolynes, P. G.; Clementi, C. Localizing frustration in proteins using all-atom energy functions. *J. Phys. Chem. B* **2019**, *123*, 4497−4504.

(19) Felline, A.; Seeber, M.; Fanelli, F. webPSN v2.0: a webserver to infer fingerprints of structural communication in biomacromolecules. *Nucleic Acids Res.* **2020**, *48*, W94−W103.
(20) Felline, A.; Seeber, M.; Fanelli, F. PSNtools for standalone and web-based structure network analyses of conformational ensembles. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 640−649.
(21) Seeber, M.; Cecchini, M.; Rao, F.; Settanni, G.; Caflisch, A. Wordom: a program for efficient analysis of molecular dynamics simulations. *Bioinformatics* **2007**, *23*, 2625−2627.
(22) Seeber, M.; Felline, A.; Raimondi, F.; Muff, S.; Friedman, R.; Rao, F.; Caflisch, A.; Fanelli, F. Wordom: a user-friendly program for the analysis of molecular structures, trajectories, and free energy surfaces. *J. Comput. Chem.* **2011**, *32*, 1183−1194.
(23) Seeber, M.; Felline, A.; Raimondi, F.; Mariani, S.; Fanelli, F. WebPSN: a web server for high-throughput investigation of structural communication in biomacromolecules. *Bioinformatics* **2015**, *31*, 779−781.
(24) González-Paz, L.; Hurtado-León, M. L.; Lossada, C.; Fernández-Materán, F. V.; Vera-Villalobos, J.; Loroño, M.; Paz, J. L.; Jeffreys, L.; Alvarado, Y. J. Structural deformability induced in proteins of potential interest associated with COVID-19 by binding of homologues present in ivermectin: Comparative study based in elastic networks models. *J. Mol. Liq.* **2021**, *340*, No. 117284.
(25) Doruker, P.; Atilgan, A. R.; Bahar, I. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: Application to α-amylase inhibitor. *Proteins: Struct., Funct., Bioinf.* **2000**, *40*, 512−524.
(26) Atilgan, A. R.; Durell, S. R.; Jernigan, R. L.; Demirel, M. C.; Keskin, O.; Bahar, I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* **2001**, *80*, 505−515.
(27) Marques, O.; Sanejouand, Y. H. Hinge-bending motion in citrate synthase arising from normal mode calculations. *Proteins: Struct., Funct., Bioinf.* **1995**, *23*, 557−560.
(28) Dykeman, E. C.; Sankey, O. F. Normal mode analysis and applications in biological physics. *J. Phys.: Condens. Matter* **2010**, *22*, No. 423202.
(29) Emekli, U.; Schneidman-Duhovny, D.; Wolfson, H. J.; Nussinov, R.; Haliloglu, T. HingeProt: automated prediction of hinges in protein structures. *Proteins: Struct., Funct., Bioinf.* **2008**, *70*, 1219−1227.
(30) Ponzoni, L.; Polles, G.; Carnevale, V.; Micheletti, C. SPECTRUS: A dimensionality reduction approach for identifying dynamical domains in protein complexes from limited structural datasets. *Structure* **2015**, *23*, 1516−1525.
(31) Liu, Y.; Li, Z.; Xiong, H.; Gao, X.; Wu, J.; Wu, S. Understanding and enhancement of internal clustering validation measures. *IEEE Trans. Cybern.* **2013**, *43*, 982−994.
(32) Liu, Y.; Grimm, M.; Dai, W. T.; Hou, M. C.; Xiao, Z. X.; Cao, Y. CB-Dock: a web server for cavity detection-guided protein−ligand blind docking. *Acta Pharmacol. Sin.* **2020**, *41*, 138−144.
(33) Liu, Y.; Yang, X.; Gan, J.; Chen, S.; Xiao, Z. X.; Cao, Y. CB-Dock2: improved protein−ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Res.* **2022**, *50*, W159−W164.
(34) Cao, Y.; Li, L. Improved protein−ligand binding affinity prediction by using a curvature-dependent surface-area model. *Bioinformatics* **2014**, *30*, 1674−1680.
(35) Dupuis, N. F.; Holmstrom, E.; Nesbitt, D. Molecular-crowding effects on single-molecule RNA folding/unfolding thermodynamics and kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 8464−8469.
(36) Bitencourt-Ferreira, G.; de Azevedo, W. F., Jr. Molegro Virtual Docker for Docking. *Methods Mol. Biol.* **2019**, *2053*, 149−167.
(37) Xavier Senra, M. V.; Fonseca, A. L. New tyrosinases with putative action against contaminants of emerging concern. *Proteins: Struct., Funct., Bioinf.* **2021**, *89*, 1180−1192.
(38) Mahdian, S.; Zarrabi, M.; Panahi, Y.; Dabbagh, S. Repurposing FDA-Approved Drugs to Fight COVID-19 Using In Silico Methods: Targeting SARS-CoV-2 RdRp Enzyme and Host Cell Receptors (ACE2, CD147) through Virtual Screening and Molecular Dynamic Simulations. In *Informatics in Medicine Unlocked*; Elsevier, 2021; Vol. *23*, 100541.
(39) Patel, N.; Dubins, D. N.; Pomès, R.; Chalikian, T. V. Size dependence of cavity volume: a molecular dynamics study. *Biophys. Chem.* **2012**, *161*, 46−49.
(40) Grahl, M. V.; Alcará, A. M.; Perin, A.; Moro, C. F.; Pinto, É.; Feltes, B. C.; Ghilardi, I. M.; Rodrigues, F.; Dorn, M.; da Costa, J. C.; Norberto de Souza, O.; Ligabue-Braun, R. Evaluation of Drug Repositioning by Molecular Docking of Pharmaceutical Resources Available in the Brazilian Healthcare System against SARS-CoV-2. In *Informatics in Medicine Unlocked*; Elsevier, 2021; Vol. *23*, 100539.
(41) Gautam, V.; Chong, W. L.; Chin, S. P.; Zain, S. M.; Rahman, N. A.; Vao-soongnern, V.; Lee, V. S. Loop dynamics behind the affinity of DARPins towards ERK2: Molecular dynamics simulations (MDs) and elastic network model (ENM). *J. Mol. Liq.* **2019**, *274*, 612−620.
(42) Madhavi, W. M.; Weerasinghe, S.; Momot, K. Reorientational dynamics of molecules in liquid methane: A molecular dynamics simulation study. *J. Mol. Liq.* **2021**, *324*, No. 114727.
(43) Yonezawa, Y. Electrostatic properties of water models evaluated by a long-range potential based solely on the Wolf charge-neutral condition. *Chem. Phys. Lett.* **2013**, *556*, 308−314.
(44) Kasahara, K.; Terazawa, H.; Itaya, H.; Goto, S.; Nakamura, H.; Takahashi, T.; Higo, J. myPresto/omegagene 2020: a molecular dynamics simulation engine for virtual system coupled sampling. *Biophys. Physicobiol.* **2020**, *17*, 140−146.
(45) González-Paz, L. A.; Lossada, C.; Fernández-Materán, F. V.; Paz, J. L.; Vera-Villalobos, J.; Alvarado, Y. J. Can non-steroidal anti-inflammatory drugs affect the interaction between receptor binding domain of SARS-COV-2 spike and the human ACE2 receptor? A computational biophysical study. *Front. Phys.* **2020**, *8*, No. 526.
(46) González-Paz, L.; Lossada, C.; Moncayo, L. S.; Romero, F.; Paz, J. L.; Vera-Villalobos, J.; Pérez, A.; Portillo, E.; San-Blas, E.; Alvarado, Y. J. A bioinformatics study of structural perturbation of 3CL-protease and the HR2-domain of SARS-CoV-2 induced by synergistic interaction with ivermectins. *Biointerface Res. Appl. Chem.* **2021**, *11*, 9813−9826.
(47) Pravda, L.; Sehnal, D.; Toušek, D.; Navrátilová, V.; Bazgier, V.; Berka, K.; Svobodová Vareková, R.; Koca, J.; Otyepka, M. MOLEonline: a web-based tool for analyzing channels, tunnels, and pores (2018 update). *Nucleic Acids Res.* **2018**, *46*, W368−W373.
(48) Sehnal, D.; Svobodová Vařeková, R.; Berka, K.; Pravda, L.; Navrátilová, V.; Banáš, P.; Koča, J.; et al. MOLE 2.0: advanced approach for analysis of biomacromolecular channels. *J. Cheminf.* **2013**, *5*, No. 39.
(49) Berka, K.; Hanák, O.; Sehnal, D.; Banáš, P.; Navratilova, V.; Jaiswal, D.; Otyepka, M.; et al. MOLE online 2.0: Interactive web-based analysis of biomacromolecular channels. *Nucleic Acids Res.* **2012**, *40*, W222−W227.
(50) Botelho, F. D.; dos Santos, M. C.; Gonçalves, A. D. S.; Kuca, K.; Valis, M.; LaPlante, S. R.; França, T. C. C.; de Almeida, J. S. Ligand-based virtual screening, molecular docking, molecular dynamics, and MM-PBSA calculations towards the identification of potential novel ricin inhibitors. *Toxins* **2020**, *12*, No. 746.
(51) Vianna, C. P.; de Azevedo, W. F. Identification of new potential Mycobacterium tuberculosis shikimate kinase inhibitors through molecular docking simulations. *J. Mol. Model.* **2012**, *18*, 755−764.
(52) Lee, B.; Richards, F. M. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **1971**, *55*, 379−400.
(53) Richards, F. M. Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151−176.
(54) Connolly, M. L. Analytical molecular surface calculation. *J. App. Crystallogr.* **1983**, *16*, 548.
(55) Thomsen, R.; Christensen, M. H. MolDock: a new technique for high-accuracy molecular docking. *J. Med. Chem.* **2006**, *49*, 3315−3321.
(56) Xie, Y.; Ying, J.; Xie, D. SMPBS: Web server for computing biomolecular electrostatics using finite element solvers of size

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318

modified Poisson-Boltzmann equation. *J. Comput. Chem.* **2017**, *38*, 541−552.
(57) Sen Gupta, P. S.; Biswal, S.; Panda, S. K.; Ray, A. K.; Rana, M. K. Binding mechanism and structural insights into the identified protein target of COVID-19 and importin-α with in-vitro effective drug ivermectin. *J. Biomol. Struct. Dyn.* **2022**, *40*, 2217−2226.
(58) Mosquera-Yuqui, F.; Lopez-Guerra, N.; Moncayo-Palacio, E. A. Targeting the 3CLpro and RdRp of SARS-CoV-2 with phytochemicals from medicinal plants of the Andean Region: molecular docking and molecular dynamics simulations. *J. Biomol. Struct. Dyn.* **2022**, *40*, 2010−2023.
(59) Kalhor, H.; Sadeghi, S.; Abolhasani, H.; Kalhor, R.; Rahimi, H. Repurposing of the approved small molecule drugs in order to inhibit SARS-CoV-2 S protein and human ACE2 interaction through virtual screening approaches. *J. Biomol. Struct. Dyn.* **2022**, *40*, 1299−1315.
(60) Wang, Z.; Wang, X.; Li, Y.; Lei, T.; Wang, E.; Li, D.; Kang, Y.; Zhu, F.; Hou, T. farPPI: a webserver for accurate prediction of protein-ligand binding structures for small-molecule PPI inhibitors by MM/PB(GB)SA methods. *Bioinformatics* **2019**, *35*, 1777−1779.
(61) Hennefarth, M. R.; Alexandrova, A. N. Advances in optimizing enzyme electrostatic preorganization. *Curr. Opin. Struct. Biol.* **2022**, *72*, 1−8.
(62) Bellissent-Funel, M. C.; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.; Sterpone, F.; van der Spoel, D.; Xu, Y.; Garcia, A. E. Water Determines the Structure and Dynamics of Proteins. *Chem. Rev.* **2016**, *116*, 7673−7697.
(63) Chu, W. T.; Yan, Z.; Chu, X.; Zheng, X.; Liu, Z.; Xu, L.; Zhang, K.; Wang, J. Physics of biomolecular recognition and conformational dynamics. *Rep. Prog. Phys.* **2021**, *84*, No. 126601.
(64) Ferreiro, D. U.; Hegler, J. A.; Komives, E. A.; Wolynes, P. G. Localizing frustration in native proteins and protein assemblies. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 19819−19824.
(65) Gianni, S.; Freiberger, M. I.; Jemth, P.; Ferreiro, D. U.; Wolynes, P. G.; Fuxreiter, M. Fuzziness and frustration in the energy landscape of protein folding, function, and assembly. *Acc. Chem. Res.* **2021**, *54*, 1251−1259.
(66) Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. Biomolecular dynamics: order-disorder transitions and energy landscapes. Reports on progress in physics. *Phys. Soc.* **2012**, *75*, No. 076601.
(67) Potoyan, D. A.; Zheng, W.; Komives, E. A.; Wolynes, P. G. Molecular stripping in the NF-κB/IκB/DNA genetic regulatory network. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 110−115.
(68) Puchades, C.; Sandate, C. R.; Lander, G. C. The molecular principles governing the activity and functional diversity of AAA+ proteins. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 43−58.
(69) Yan, Z.; Wang, J. Funneled energy landscape unifies principles of protein binding and evolution. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 27218−27223.
(70) Liu, J.; Sankar, K.; Wang, Y.; Jia, K.; Jernigan, R. L. Directional force originating from ATP hydrolysis drives the GroEL conformational change. *Biophys. J.* **2017**, *112*, 1561−1570.
(71) Lamers, M. H.; Georgijevic, D.; Lebbink, J. H.; Winterwerp, H. H.; Agianian, B.; De Wind, N.; Sixma, T. K. ATP increases the affinity between MutS ATPase domains: Implications for ATP hydrolysis and conformational changes. *J. Biol. Chem.* **2004**, *279*, 43879−43885.
(72) Mancini, E. J.; Kainov, D. E.; Grimes, J. M.; Tuma, R.; Bamford, D. H.; Stuart, D. I. Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell* **2004**, *118*, 743−755.
(73) Shih, W. M.; Gryczynski, Z.; Lakowicz, J. R.; Spudich, J. A. A FRET-based sensor reveals large ATP hydrolysis−induced conformational changes and three distinct states of the molecular motor myosin. *Cell* **2000**, *102*, 683−694.
(74) Chada, N.; Chattrakun, K.; Marsh, B. P.; Mao, C.; Bariya, P.; King, G. M. Single-molecule observation of nucleotide induced conformational changes in basal SecA-ATP hydrolysis. *Sci. Adv.* **2018**, *4*, No. eaat8797.
(75) Ito, Y.; Ikeguchi, M. Mechanism of the $\alpha\beta$ conformational change in F1-ATPase after ATP hydrolysis: free-energy simulations. *Biophys. J.* **2015**, *108*, 85−97.
(76) Ortega, J.; Lee, H. S.; Maurizi, M. R.; Steven, A. C. Alternating translocation of protein substrates from both ends of ClpXP protease. *EMBO J.* **2002**, *21*, 4938−4949.
(77) Amor, A. J.; Schmitz, K. R.; Baker, T. A.; Sauer, R. T. Roles of the ClpX IGF loops in ClpP association, dissociation, and protein degradation. *Protein Sci.* **2019**, *28*, 756−765.
(78) Keramisanou, D.; Vasantha Kumar, M. V.; Boose, N.; Abzalimov, R. R.; Gelis, I. Assembly mechanism of early Hsp90-Cdc37-kinase complexes. *Sci. Adv.* **2022**, *8*, No. eabm9294.
(79) Elghobashi-Meinhardt, N. Cholesterol transport in wild-type NPC1 and P691S: molecular dynamics simulations reveal changes in dynamical behavior. *Int. J. Mol. Sci.* **2020**, *21*, 2962.
(80) Berka, K.; Paloncýová, M.; Anzenbacher, P.; Otyepka, M. Behavior of human cytochromes P450 on lipid membranes. *J. Phys. Chem. B.* **2013**, *117*, 11556−11564.
(81) Rizo, A. N.; Lin, J.; Gates, S. N.; Tse, E.; Bart, S. M.; Castellano, L. M.; DiMaio, F.; Shorter, J.; Southworth, D. R. Structural basis for substrate gripping and translocation by the ClpB AAA+ disaggregase. *Nat. Commun.* **2019**, *10*, No. 2393.
(82) Chen, X.; Htet, Z. M.; López-Alfonzo, E.; Martin, A.; Walters, K. J. Proteasome interaction with ubiquitinated substrates: From mechanisms to therapies. *FEBS J.* **2021**, *288*, 5231−5251.
(83) Gogoi, B.; Chowdhury, P.; Goswami, N.; Gogoi, N.; Naiya, T.; Chetia, P.; Handique, P. J.; et al. Identification of potential plant-based inhibitor against viral proteases of SARS-CoV-2 through molecular docking, MM-PBSA binding energy calculations and molecular dynamics simulation. *Mol. Diversity* **2021**, *25*, 1963−1977.
(84) Yadav, P.; Rana, M.; Chowdhury, P. DFT and MD simulation investigation of favipiravir as an emerging antiviral option against viral protease (3CLpro) of SARS-CoV-2. *J. Mol. Struct.* **2021**, *1246*, No. 131253.
(85) Li, J.; Jiang, L.; Cao, X.; Wu, Y.; Lu, F.; Liu, F.; Liu, Y.; Liu, Y. Improving the activity and stability of Bacillus clausii alkaline protease using directed evolution and molecular dynamics simulation. *Enzyme Microb. Technol.* **2021**, *147*, No. 109787.
(86) Aljarba, N. H.; Hasnain, M. S.; Bin-Meferij, M. M.; Alkahtani, S. An In-silico investigation of potential natural polyphenols for the targeting of COVID main protease inhibitor. *J. King Saud Univ., Sci.* **2022**, *34*, No. 102214.
(87) Majumder, R.; Mandal, M. Screening of plant-based natural compounds as a potential COVID-19 main protease inhibitor: an in silico docking and molecular dynamics simulation approach. *J. Biomol. Struct. Dyn.* **2022**, *40*, 696−711.
(88) Motiwala, T.; Akumadu, B. O.; Zuma, S.; Mfeka, M. S.; Chen, W.; Achilonu, I.; Syed, K.; Khoza, T. Caseinolytic Proteins (Clp) in the Genus Klebsiella: Special Focus on ClpK. *Molecules* **2022**, *27*, 200.
(89) Li, T.; Pang, W.; Wang, J.; Zhao, Z.; zhang, X.; Cheng, L. Docking-based 3D-QSAR, molecular dynamics simulation studies and virtual screening of novel ONC201 analogues targeting Mitochondrial ClpP. *J. Mol. Struct.* **2021**, *1245*, No. 131025.
(90) Zhang, J.; Luo, B.; Sui, J.; Qiu, Z.; Huang, J.; Yang, T.; Luo, Y. IMP075 targeting ClpP for colon cancer therapy in vivo and in vitro. *Biochem. Pharmacol.* **2022**, *204*, No. 115232.
(91) Jing, S.; Wang, L.; Wang, T.; Fan, L.; Chen, L.; Xiang, H.; Shi, Y.; Wang, D. Myricetin protects mice against MRSA-related lethal pneumonia by targeting ClpP. *Biochem. Pharmacol.* **2021**, *192*, No. 114753.
(92) Damre, M.; Dayananda, A.; Varikoti, R. A.; Stan, G.; Dima, R. I. Factors underlying asymmetric pore dynamics of disaggregase and microtubule-severing AAA+ machines. *Biophys. J.* **2021**, *120*, 3437−3454.
(93) Mazal, H.; Iljina, M.; Riven, I.; Haran, G. Ultrafast Brownian−ratchet mechanism for protein translocation by a AAA+ machine. *bioRxiv*2020, 11, DOI: 10.1126/sciadv.abg4674.
(94) Li, D. H. S.; Chung, Y. S.; Gloyd, M.; Joseph, E.; Ghirlando, R.; Wright, G. D.; Cheng, Y. Q.; Maurizi, M. R.; Guarné, A.; Ortega, J.

Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: a model for the ClpX/ClpA-bound state of ClpP. *Chem. Biol.* **2010**, *17*, 959−969.
(95) Lee, B. G.; Park, E. Y.; Lee, K. E.; Jeon, H.; Sung, K. H.; Paulsen, H.; Rübsamen-Schaeff, H.; Brötz-Oesterhelt, H.; Song, H. K. Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nat. Struct. Mol. Biol.* **2010**, *17*, 471−47.
(96) Kim, D. Y.; Kim, K. K. The structural basis for the activation and peptide recognition of bacterial ClpP. *J. Mol. Biol.* **2008**, *379*, 760−771.
(97) Bewley, M. C.; Graziano, V.; Griffin, K.; Flanaga, J. M. The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. *J. Struct. Biol.* **2006**, *153*, 113.
(98) Ingvarsson, H.; Maté, M. J.; Högbom, M.; Portnoï, D.; Benaroudj, N.; Alzari, P. M.; Ortiz-Lombardia, M.; Unge, T. Insights into the inter-ring plasticity of caseinolytic proteases from the X-ray structure of Mycobacterium tuberculosis ClpP1. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2007**, *63*, 249−259.
(99) Baker, T. A.; Sauer, R. T. ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochim. Biophys. Acta, Mol. Cell Res.* **2012**, *1823*, 15−28.
(100) Barkow, S. R.; Levchenko, I.; Baker, T. A.; Sauer, R. T. Polypeptide translocation by the AAA+ ClpXP protease machine. *Chem. Biol.* **2009**, *16*, 605−612.
(101) Glynn, S. E.; Martin, A.; Nager, A. R.; Baker, T. A.; Sauer, R. T. Crystal structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. *Cell* **2009**, *139*, 744−756.
(102) Olivares, A. O.; Kotamarthi, H. C.; Stein, B. J.; Sauer, R. T.; Baker, T. A. Effect of directional pulling on mechanical protein degradation by ATP-dependent proteolytic machines. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E6306−E6313.
(103) Kimber, M. S.; Yu, A. Y. H.; Borg, M.; Leung, E.; Chan, H. S.; Houry, W. A. Structural and theoretical studies indicate that the cylindrical protease ClpP samples extended and compact conformations. *Structure* **2010**, *18*, 798−808.
(104) Hwang, W.; Karplus, M. Structural basis for power stroke vs. Brownian ratchet mechanisms of motor proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 19777−19785.
(105) Oguchi, Y.; Kummer, E.; Seyffer, F.; Berynskyy, M.; Anstett, B.; Zahn, R.; Bukau, B.; et al. A tightly regulated molecular toggle controls AAA+ disaggregase. *Nat. Struct. Mol. Biol.* **2012**, *19*, 1338−1346.
(106) Yu, J. Coordination and control inside simple biomolecular machines. *Protein Conform. Dyn.* **2014**, *805*, 353−384.

https://doi.org/10.1021/acsomega.2c04347
ACS Omega 2023, 8, 7302−7318