JOURNAL OF CHEMICAL INFORMATION AND MODELING

ARTICLE

pubs.acs.org/jcim

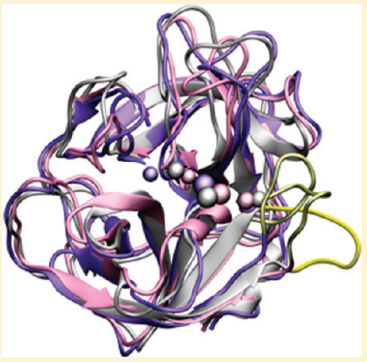
Characterizing the Dynamics and Ligand-Specific Interactions in the Human Leukocyte Elastase through Molecular Dynamics Simulations

Sìlvia G. Estácio,* Rui Moreira, and Rita C. Guedes

Research Institute for Medicines and Pharmaceutical Sciences (iMed.UL), Faculty of Pharmacy, University of Lisbon, Av. Prof. Gama Pinto, 1649-003 Lisbon, Portugal

Supporting Information

ABSTRACT: The human leukocyte elastase (HLE), a neutrophil serine protease of the chymotrypsin superfamily, is a major therapeutic target for a number of inflammatory diseases, such as chronic obstructive pulmonary disease (COPD). In this work, we present a comparative explicit water molecular dynamics (MD) study on the free and inhibitor-bound HLE. Knowledge of the flexibility and conformational changes induced by this irreversible inhibitor, whether in a prebound stage or covalently bound at the enzyme binding site, encases fundamental biological interest and is particularly relevant to ongoing structure-based drug design studies. Our results suggest that HLE operates by an induced-fit mechanism with direct intervention of a surface loop which is open toward the solvent in the free enzyme and closed while in the presence of the ligand. MM-PBSA free energy calculations furthermore elucidate the energetic contributions to the distinct conformations adopted by this loop. Additionally, a survey of the major contributions to the inhibitor binding free energies was attained. Our findings enforce the need to account for HLE flexibility, whether through the use of MD-generated ensembles of HLE conformations as targets for molecular docking or via sophisticated flexible-docking algorithms. We anticipate that inclusion of the observed HLE dynamic behavior into future drug design methodologies will have a relevant impact in the development of novel, more efficient, inhibitors.



1. INTRODUCTION

The human leukocyte elastase (HLE, EC 3.4.21.37) is a small serine protease, a member of the chymotrypsin family, which can degrade a wide variety of protein substrates, most notoriously elastin, the specific protein of elastic fibers. Serine proteases with a chymotrypsin fold possess an identical fold consisting of two asymmetric β-barrels and a C-terminal α-helix. Each barrel consists of six antiparallel β-sheets. The barrels form a cleft in which the so-called catalytic triad amino acids, Ser195, His57, and Asp102, are located.[1] This identification of the amino acids follows the chymotrypsinogen numbering of Hartley and Shotton,[2] which is derived from the alignment of the protein chain with the chymotrypsinogen sequence. The catalytic triad cleaves a peptide bond between the carbonyl group of the N-terminal amino acid and the amide group of the C-terminal amino acid. The functional enzyme contains 218 amino acids, and its S_1 binding pocket is hydrophobic in nature due to the presence of Val190, Phe192, Ala213, Val216, and Phe228. The S_2 pocket, lined by Phe215, Leu99, and the imidazole ring of His57, is also quite hydrophobic. The preference for small hydrophobic residues such as valine, leucine, norvaline, and isoleucine at P_1 is dictated by the hydrophobic nature of the S_1 pocket and the presence of large bulky residues at positions 190 and 216.

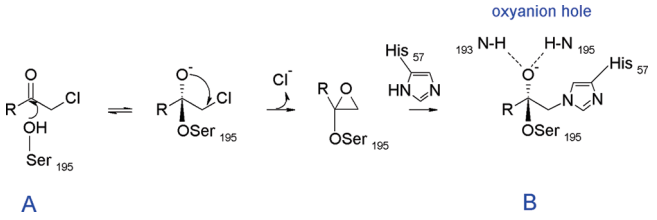
The common general catalytic mechanism of serine proteases starts with a nucleophilic attack of the hydroxyl group of Ser195 on the carbonyl carbon of the scissile (cleaved) peptide bond of the substrate. His57 enhances the nucleophilicity of Ser195 by removing the hydroxyl proton. Asp102 provides electrostatic stabilization to the protonated His57. The tetrahedral intermediate formed has its oxyanion stabilized by hydrogen bonds formed with the backbone NH groups of Ser195 and Gly193 in the so-called oxyanion hole. The protonated side chain (imidazolium ion) of His57 assists the collapse of the tetrahedral intermediate through protonation of the amine leaving group. The peptide bond is simultaneously broken and an acyl−enzyme complex is formed. The latter is hydrolyzed through attack by a water molecule assisted by general base catalysis by His57 forming a new tetrahedral intermediate. Breakdown of this intermediate allows Ser195 to recover its hydrogen from His57 and gives a final noncovalent enzyme−product complex.[3]

Though the S_1 pocket is understood to have a fundamental role in the substrate specificity of individual chymotrypsin-like serine proteases, the model by which substrate specificities are assigned exclusively to the amino acid compositions and shapes of the S_1 pockets has been demonstrated to be incomplete. Surface loops surrounding the substrate binding site exhibit diversity which has been accounted to contribute to the specificity of the activity of serine proteases.[1,4,5]

The imbalance between the amounts of HLE and its endogenous inhibitors leads to excessive elastin proteolysis and destruction of connective tissue, thus having a major role in

Received: February 15, 2011
Published: June 04, 2011

Scheme 1. Proposed Mechanism of Inhibition of Serine Proteases by Peptidyl Chloromethyl Ketones[18,19]



the development of a number of inflammatory diseases, such as chronic obstructive pulmonary disease, pulmonary emphysema, adult respiratory distress syndrome, chronic bronchitis, and rheumatoid arthritis.[6,7] The human leukocyte elastase is, therefore, an attractive therapeutic target, having prompted an intense search for suitable inhibitors, with numerous peptide-based as well as nonpeptidyl ligands being thoroughly examined.[8−10] Despite the high medicinal relevance and the many inhibitors already characterized, only a few crystal structures of HLE have been published,[11−16] and to the best of our knowledge at the time of writing this work, no structure of the native enzyme has been reported.

Peptide chloromethyl ketones such as the methoxy succinyl-Ala-Ala-Pro-Ala chloromethyl ketone have been shown to be effective in vivo in arresting the development of experimentally induced emphysema in animal models of the disease.[17] Most importantly, they have also served as a standard of comparison for newly developed inhibitors.[3] Thus, the structure of the HLE−chloromethyl ketone complex, in defining optimal peptide binding to the enzyme, is a most useful model in the development of safer inhibitors.

Most efficient inhibitors of serine proteases contain an electrophilic warhead that traps the nucleophilic serine residue with the concomitant formation of a relatively stable enzyme−inhibitor adduct in interaction with the catalytic center.[20] The chloromethyl ketone inhibitor is an irreversible, transition-state-like, inhibitor containing as a chemically reactive functionality the α-chloromethyl ketone moiety. The electron-withdrawing chloride increases the electrophilicity of the ketone carbonyl carbon, thus activating it toward nucleophilic attack by Ser195 (Scheme 1). Once the inhibitor is covalently bound to the enzyme through Ser195, it functions as an alkylating agent of the active site histidine residue, His57. The final enzyme−inhibitor complex is stabilized by two hydrogen bonds with the backbone NH groups of Ser195 and Gly193 in the oxyanion hole.

The design of new efficient HLE inhibitors should rely on an improved understanding of its binding site structural plasticity and conformational changes induced by ligand interaction and binding. This is particularly relevant given previous acknowledgment by Wei et al.[12] of a small conformational change observed for HLE upon binding to a chloromethyl ketone inhibitor. Molecular dynamics (MD) has become a sought after theoretical tool for studying the structure and dynamics of biomolecular systems within the atomic resolution provided by classical mechanics.[21] Drug design methodologies based on the use of receptor models extracted from the crystal structures of ligand−protein complexes restrict the compounds on the hit lists to conform to the size and shape of those X-ray binding sites. Molecular docking studies undertaken in the past to identify suitable HLE inhibitors might be biased, since they were usually performed against a receptor model extracted from a ligand−HLE complex crystallographic structure. The extended conformational sampling of the enzyme binding site obtained through MD allows overcoming the limitations associated with the sparse availability of crystal structures and the disclosure of features not evident for those database structures. MD simulations can ultimately provide access to the evolution in time of almost any physically definable observable within force field accuracy and finite sampling limitations. Two MD studies on HLE have, previously, been carried out by Steinbrecher et al.[22,23] to predict ligand binding free energies. However, no structural or dynamical analyses were performed.

The aim of this work is to evaluate HLE's structural and dynamical behavior and relative stability while in the presence of the well-known methoxy succinyl-Ala-Ala-Pro-Ala chloromethyl ketone irreversible inhibitor, whether in a noncovalently bound pre-inhibition state (Scheme 1A) or covalently bound at the active site (Scheme 1B). With that purpose, explicit water MD simulations were performed on three distinct systems: free HLE, HLE in interaction with the inhibitor in a noncovalent pre-catalysis complex, and HLE covalently bound to the inhibitor in agreement with the known mechanism of inhibition (Scheme 1). Identification of the flexibility and conformational changes induced by the chloromethyl ketone will ultimately contribute to the rational design of new HLE inhibitors with an improved selectivity.

2. METHODS

2.1. Molecular Dynamics Simulations. The crystal structure of the human leukocyte elastase covalently bound to a methoxy succinyl-Ala-Ala-Pro-Ala chloromethyl ketone inhibitor [Protein Data Bank (PDB) code: 1HNE[11]] was used to build the three HLE models.

The protonation states of all enzyme's ionizable residues at a pH of 7.4 were assigned using MOE[24] and the enzyme was protonated accordingly with the pdb2gmx functionality of the GROMACS 3.3 package.[25,26] His57 was protonated at the Nδ1 atom. The GROMOS96 53a6 force field[27] was used. HLE contains four disulfide bridges—Cys42−Cys58, Cys163−Cys201, Cys168−Cys182, and Cys191−Cys220—which were retained for the three enzyme models. The topology of the noncovalently bound form of the inhibitor was generated using the PRODRG web server, version 2.5 beta,[28] which is able to generate the GROMOS96 53a6 force field[27] parameters for the HETATM groups in the pdb file. The topology of the covalently bound inhibitor was also generated using the PRODRG web server having as input the inhibitor bound to two residues in the active site, His57 and Ser195, as in the HLE pdb X-ray structure. The inhibitor force field parameters thus generated were included in the GROMOS96 53a6 force field[27] parameter files as a new residue type before a topology for the entire system, HLE with the covalently bound inhibitor, was generated by GROMACS. Partial charges for both inhibitors were assigned having as references comparable groups within the building blocks of the force field. Force field parameters and partial charges for both inhibitor forms, whether covalent or noncovalently bound to the enzyme, are available in the Supporting Information (Figure S1 and Table S1).

Bond lengths were constrained with LINCS[29] and with SETTLE[30] for water. A time step of integration of 2 fs was used within the leapfrog integration scheme. The free HLE and

inhibitor-bound (covalent or noncovalently) HLE systems were solvated in, respectively, 11 580 and 11 515 explicit SPC water molecules in periodic rectangular simulation boxes. The box dimensions were chosen to provide at least a 10 Å buffer of water molecules around the solute. To neutralize and prepare the systems under a physiological ionic concentration (0.15 M), 35 or 34 sodium and 45 or 44 chloride ions were added by replacing water molecules. The fully solvated systems with rectangular periodic boundary conditions were then subjected to steepest descent minimization runs to remove clashes between atoms. Initial velocities for each system were then randomly assigned from a Maxwell−Boltzmann distribution at 300 K followed by a positional restrained dynamics with heavy harmonic restraints on the protein heavy atoms (1000 kJ mol^{-1} nm^{-2}) to allow relaxation of the solvent, ions, and/or noncovalent inhibitor around the enzyme. A short run of 400 ps in which each system was heated to 300 K and allowed to relax at that temperature through coupling with a Berendsen thermostat[31] followed. To ensure full relaxation of the system properties, a further 400 ps of isothermal−isobaric ensemble equilibration followed. Pressure control was implemented using the Berendsen barostat[31] with a reference pressure of 1 atm, 1.0 ps relaxation time, and isothermal compressibility of 4.5×10^{-5} bar^{-1}. Temperature control was set using the Berendsen thermostat at 300 K. Fully unrestrained production runs were performed with a time step of 2 fs and using the isothermal−isobaric ensemble at $T = 300$ K and $p = 1$ bar. The production runs were 120 ns long for the free and noncovalently bound HLE systems and 75 ns long in the case of the covalently bound system. The temperature and the pressure were held constant using a Nose−Hoover thermostat[32,33] with a coupling constant (relaxation time) of 0.1 ps and a semi-isotropic Parrinello−Rahman barostat[34,35] with a coupling constant (relaxation constant) of 0.5 ps. The free HLE or inhibitor-bound complexes of HLE and the solvent (including ions) were independently coupled to the heat bath. A particle mesh Ewald scheme[36,37] was used to calculate the electrostatic interactions with a 10 Å cutoff for the real space. A cutoff of 14 Å was used for the short-range van der Waals interactions (Lennard-Jones terms). Neighbor searching was carried out up to 10 Å and updated every five steps.

2.2. Essential Dynamics. The extraction of data on the local flexibility of HLE from the MD simulations was performed after essential dynamics analysis. Essential dynamics (ED) is a widely applied technique based on the principal component analysis (PCA) of conformational ensembles that allows the identification of the most relevant or correlated motions of groups of residues of a protein along a trajectory generated by molecular dynamics simulation.[38,39]

The enzyme coordinates were extracted at 10 ps intervals over the last, stable 30 ns of each MD trajectory for the bulk water free and inhibitor-bound HLE systems. Global translational and rotational movements during the trajectories were removed through superimposition and a least-squares fitting routine on the enzyme Cα atoms. The covariance matrix of Cα atoms was constructed and then diagonalized by using g_covar and g_anaeig within the GROMACS package,[25,26] yielding a set of eigenvectors and their respective eigenvalues. Each eigenvector represents one single direction of collective motion, whereas the corresponding eigenvalue represents the amplitude of motion along that vector. The local flexibility of the enzyme was reported as the root-mean-square fluctuation (RMSF) on the positions of the Cα atoms calculated from the coordinates of the system in the essential subspace. The deviations of the HLE amino acids Cα atoms from their time-averaged positions were obtained as averages over the eigenvectors representing ∼80% of the total mobility of the system (i.e., the total sum of eigenvalues for the system).

2.3. RMSD Clustering. In order to select a reduced set of representative models of the HLE binding site when in interaction with the covalent or noncovalently bound inhibitor, root-mean-square deviation (RMSD) conformational clustering was performed using the gromos method[40] implemented in GROMACS (g_cluster).[25,26] Enzyme configurations were extracted every 10 ps over the last 30 ns of each system MD trajectory. The RMSD clustering analysis, subsequent to a process of structural-fitting which includes superimposition (to remove overall translation) and a least-squares-fitting procedure (to remove overall rotations) using the entire enzyme Cα atoms, was performed on a set of residues that define the enzyme binding site area, including the catalytic triad and S_1 and S_2 pocket residues: His57, Leu99, Asp102, Val190, Phe192, Ser195, Ala213, Phe215, Val216, and Phe228. These residues were clustered using the atom-positional RMSD of all atoms (including hydrogen atoms and side chains) as the similarity criterion. A value of 1.4 Å was chosen as the optimal RMSD cutoff after evaluation of the dependence of the cluster populations against the total number of clusters in the range 1.0−1.5 Å. In the gromos clustering algorithm, the conformation with the highest number of neighbors, identified within the chosen RMSD cutoff, is chosen as the center of the first cluster. All the neighbors of this conformation are removed from the ensemble of conformations. The center of the second cluster is then determined in the same way, and the procedure is repeated until each structure has been assigned to a cluster.[40]

2.4. MM-PBSA Free Energy Calculations. Total binding free energies for both enzyme−inhibitor complexes were calculated through the molecular mechanics Poisson−Boltzmann surface area method.[41,42] For each system, a single trajectory approach[43] was used and the calculations were performed on 300 snapshot structures extracted at 100 ps intervals over the last 30 ns of the system's trajectory, where the complex appeared to gain a stable conformation.

In the MM-PBSA approach an interaction free energy is defined as

$$\Delta G_b = G_{complex} - G_{enzyme} - G_{inhibitor} \quad (1)$$

$$G = E_{MM} + G_{polar,\,solv} + G_{nonpolar,\,solv} - TS \quad (2)$$

$$E_{MM} = E_{int} + E_{elec} + E_{vdW} \quad (3)$$

The molecular mechanics contribution to the interaction free energy, E_{MM}, includes electrostatic, van der Waals, and internal energies (which cancel out in eq 1 due to the use of the single trajectory approach). E_{elec} was calculated using the APBS software package,[44] and E_{vdW} was calculated in GROMACS.[25,26] The electrostatic contribution to the solvation energy, $G_{polar,solv}$, was obtained through solution of the Poisson−Boltzmann (PB) equations, which switches on an electrostatics continuum. The APBS software package[44] was used with a grid spacing of 0.5 Å and a solution of 150 mM NaCl (consistent with the ionic concentration of the MD simulations). The interior dielectric constant was set to 1, while the dielectric constant of water was set to 78.4. The atomic charges and radii used in the PB

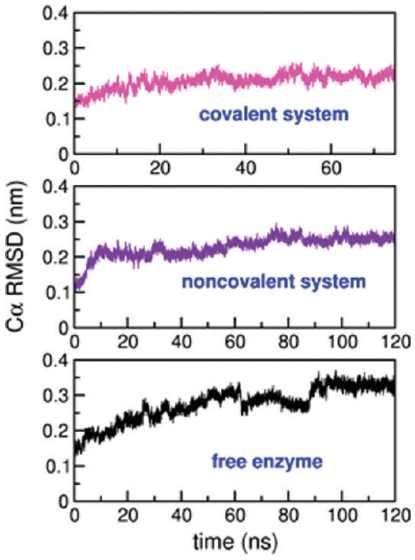


Figure 1. Time dependence of the HLE Cα atoms RMSDs from the reference X-ray structure for, from bottom to top, the free HLE, noncovalently bound HLE, and covalently bound HLE MD simulations.

calculations were derived from the GROMOS96 53a6[27] force field. The charges were directly taken from the force field through the PDB2PQR software.[45] The atomic radii were computed from the force field Lennard-Jones parameters. The nonpolar contribution to the solvation energy, which accounts for the burial of the solvent-accessible surface area (SASA) upon binding, was described according to $G_{nonpolar,solv} = \gamma \times SASA \times \beta$. The solvent-accessible surface area was calculated in GROMACS[25,26] with a 1.4 Å radius probe. γ and β were set to 2.2 kJ mol^{-1} nm^{-2} and 3.8 kJ mol^{-1}, respectively.[46] *TS* is the product of temperature (300 K) and solute entropy. Entropy changes associated with the interactions between the enzyme and the inhibitor were derived from normal-mode analysis[47] carried out on five MD snapshots extracted from the last, equilibrium, 30 ns of each complex MD trajectory. Conformations for the enzyme, the inhibitor, and the enzyme−inhibitor thus obtained were subjected to rigorous energy minimization until the maximum force acting on an atom was <10^{-6} kJ mol^{-1} nm^{-1}. Normal mode analysis was performed by calculating and diagonalizing the mass-weighted Hessian matrix, and the frequencies of the normal modes thus obtained were used to calculate vibrational entropies. These calculations were performed with GROMACS,[25,26] compiled in double precision.

Regarding the approximation used to split the covalently bound system into enzyme and inhibitor moieties, this was attained by disrupting the covalent bonds between His57 Nε2 and the inhibitor CBU and Ser195 Oγ and the inhibitor CBR atoms (vide Figure S1, Supporting Information).

3. RESULTS AND DISCUSSION

3.1. Overall Stability of the MD-Simulated HLE Systems. Long and stable trajectories were obtained for the three HLE systems.

Figure 1 shows the time dependence of the RMSDs for the enzyme Cα atoms from the X-ray reference enzyme structure over the production phase of each simulation. The Cα RMSD values for the covalently bound system simulation reached an equilibrium plateau after approximately 45 ns and from this point another 30 ns was added to the trajectory. Runs for the free enzyme and the noncovalently bound enzyme were further extended to 120 ns. In both cases, equilibrium is considered beyond 90 ns.

The RMSDs of the enzyme Cα atoms from the original X-ray structure indicate some relaxation of the enzyme structure and illustrate the comparatively more pronounced stability of the inhibitor-bound complexes. For the free HLE bulk water MD simulation, the enzyme Cα RMSD at equilibrium fluctuated about a mean of 3 Å, whereas for both inhibitor complexes the enzyme Cα RMSD fluctuated about a lower mean of 2−2.5 Å. This might reflect a general conformational lability of the enzyme with the inhibitor inducing a stabilizing effect due to its interactions within the active site.

The reasonably low Cα RMSD values are an indication that the force field and simulation protocol used in these simulations are adequate for investigating the current HLE systems.

3.2. Essential Dynamics Analysis and Inhibitor-Induced Changes of the Enzyme Flexibility. It is important to investigate whether the conformational changes induced by the inhibitor, of which the lower Cα RMSD values are an indication, are accompanied by significant changes in the local flexibility of the enzyme. In fact, most enzyme activity is based on a precise balance between structural plasticity and stability.[48,49] Knowledge of the comparative dynamics of the residues lining the binding site area of the three systems may also prove interesting to drug design investigations. The atomic mobility of each active site amino acid can be related to its binding propensity, with strong effective binding residues generally being much less flexible than weak ones. The flexible regions of HLE were, therefore, checked by computing the atomic fluctuations along the equilibrium segment of each MD trajectory.

Before proceeding and in order to simplify subsequent analyses of specific residues at the active site, the relative positions of selected S_1, S_2, and catalytic triad amino acids for a MD snapshot of the covalent complex are depicted in Figure 2. Some HLE surface loops are also highlighted. In this work two identification systems for the amino acids are in use: the chymotrypsinogen numbering,[2] which allows comparison with other chymotrypsin-fold serine proteases, and the presently designated MD numbering related with a renumbering of the HLE sequence (in the original X-ray structure the chymotrypsinogen-numbered HLE chain starts with a residue numbered 16 while the models used for MD simulations are automatically renumbered starting in 1).

Essential dynamics analysis was performed over the stable MD trajectories in order to identify possible relevant displacements of groups of residues and give emphasis to the direction and amplitude of the most dominant HLE motions. This analysis was restricted to the enzyme Cα atoms given that, in most cases, such reduction in complexity retains the relevant information about the protein dynamics.[38] In the sequence of the ED analysis carried out on each system MD trajectory, evaluation of the number of eigenvectors which represent ∼80% of the total mobility of the enzyme in each system gave a total of 33 for the free enzyme, 42 for the enzyme in the noncovalently bound system, and 40 for the covalent case. As such, there is no evidence for the existence of clean-cut collective motions of amino acids in either system. There is a widespread representation, in terms of eigenvectors, of the essential motions in the enzyme which account for a high degree of noncooperativity between HLE's

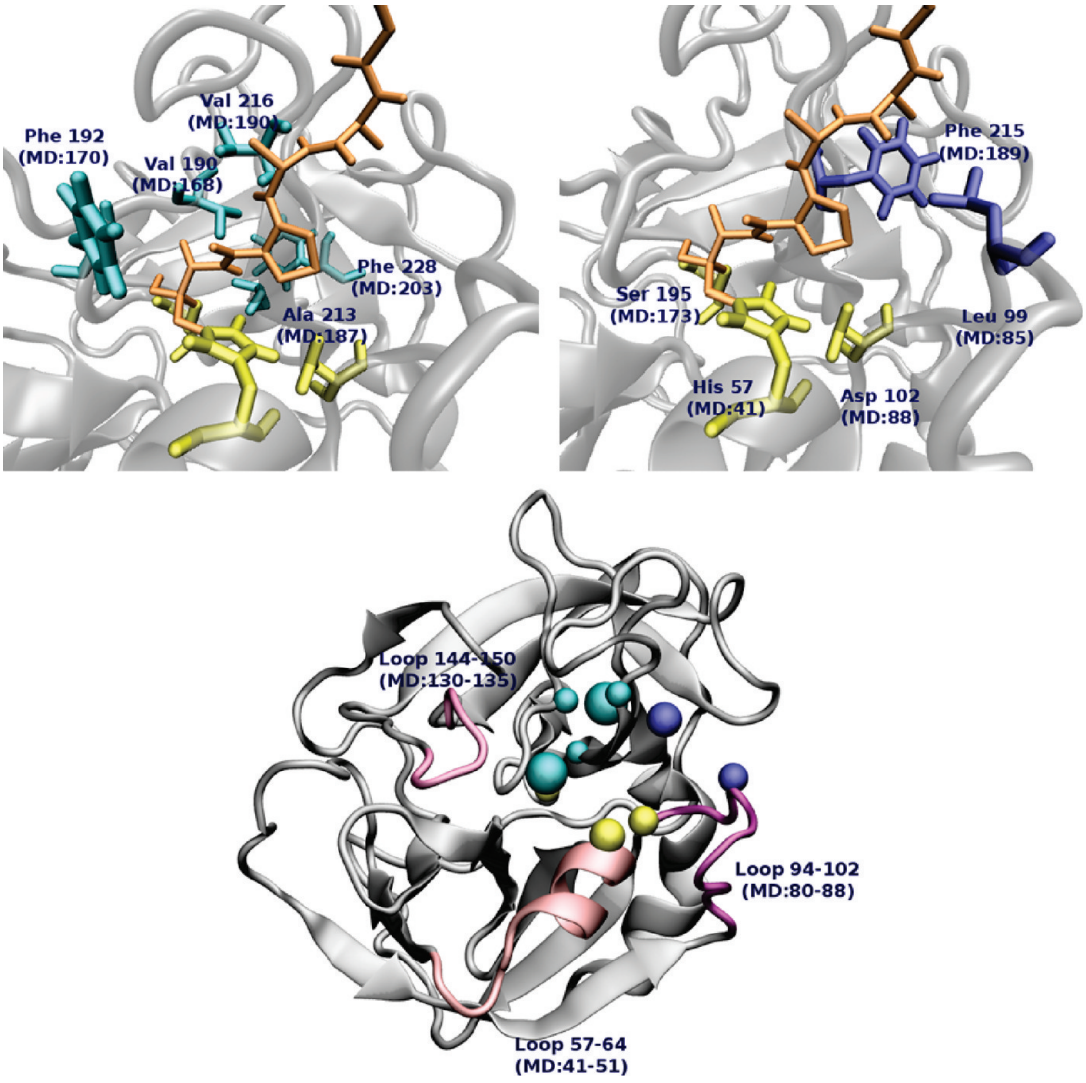


Figure 2. Residues lining the HLE binding site (50 ns MD snapshot of the covalently bound HLE system with the inhibitor represented in orange). Top left: S_1 pocket residues in light blue. Top right: S_2 pocket residues in dark blue and catalytic triad in yellow. Bottom: Some HLE surface loops highlighted in pink tones. Residues lining the binding site are represented as spheres.

amino acids motions. Moreover, it is worth mentioning that the presence of both forms of the inhibitor increases the complexity of the HLE motions given that a larger number of eigenvectors (40 and 42) is required to explain the ~80% variance in the corresponding MD trajectories.

Figure 3 shows the HLE amino acids displacements as a function of the residues number (MD numbering scheme) calculated for each Cα atom and averaged over the first eigenvectors accounting for the most significant part of the total mobility of the enzyme in each system. The small dark blue squares in Figure 3 point to active site S_1, S_2, and catalytic triad residues. Whereas some regions become more ordered due to the presence of the inhibitor, whether covalent or noncovalently bound to the enzyme, others become less. The results attest that some of the most ordered regions in both inhibitor complexes are involved in inhibitor−enzyme interactions within the binding site.

It is noticeable that among the binding site residues indicated in Figure 3 there are discrepancies in terms of their relative flexibilities, which is especially evident for the free HLE. Namely, the amino acids His57, Leu99, Val190, Phe192, and Ser195, respectively residues 41, 85, 168, 170, and 173 in Figure 3 according to the MD numbering scheme, are the most plastic binding site residues. It has previously been noticed that the binding pockets of many proteins have a dual rigid/plastic character related to the need to allow access to and accommodate the ligand.[48] The high Leu99 plasticity is related to its location on a surface loop adjacent to the HLE binding site.

Among the S_1 residues in focus (vide Figure 2, top left), the most significant differences in flexibility among the three systems occur for both Val190 (residue 168 in Figure 3 according to the MD residue identification) and Phe192 (residue 170 in Figure 3). Both amino acids equilibrium positions are particularly restricted by the presence of the inhibitor. While the Val190 position extends itself into the base of the pocket, modulating its specificity profile, Phe192 is located at the entrance of the S_1 pocket, which might explain the privileged interactions of both residues with the inhibitor in the complexes. The catalytic triad amino acids (residues 41, 88, and 173 in MD numbering in Figure 3) are, as expected, stabilized by the presence of the

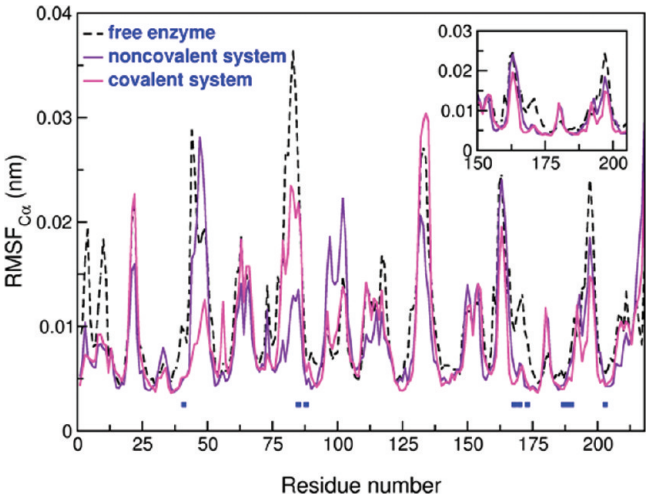


Figure 3. Cα RMSF per HLE residue. Residue displacement calculated for each Cα atom as a function of the residue number and averaged over the first 33 (free HLE), 42 (noncovalently bound HLE), and 40 (covalently bound HLE) eigenvectors of the total (the chosen eigenvectors represent ∼80% of the total mobility of HLE during each MD simulation). The small squares signal S_1, S_2 and catalytic triad amino acids (MD residue numbering).

inhibitor. These observations can be further complemented with the evolution in time of the RMSDs of the active site residues from the original X-ray enzyme structure (Figure S2, Supporting Information).

Despite the observed increased rigidity induced by the inhibitor, the lack of drastic changes in the fluctuations of the binding site residues upon binding to the inhibitor may be related with the shallow nature of the HLE binding site, which makes for few inhibitor−enzyme interactions. Averhoff et al.[50] acknowledged the absence of a clear substrate recognition motif in HLE.

Surface loop regions (vide Figure 2, bottom) have the higher deviations from their time-average positions, with the Tyr94-Asp95-Pro96-Val97-Asn98-Leu99-Leu100-Asn101-Asp102 loop (in chymotrypsinogen numbering, residues 80−88 in Figure 3) having the highest displacements of Cα atoms within the free enzyme MD simulation. The presence of the inhibitor significantly reduces the loop RMS fluctuations, with the most noticeable effect observed for the noncovalently bound inhibitor system. Another flexible surface loop, formed by residues Leu144-Gly145-Arg147-Asn148-Arg149-Gly150 (residues 130−135 in Figure 3), fluctuates more in the covalently bound inhibitor system than in the free enzyme, which suggests that those amino acids are not directly involved in inhibitor−enzyme interactions. The loop is, however, fairly fixed in the noncovalent system, possibly due to the more external position of the inhibitor and its extended chain, which may, somehow, block its fluctuations. This loop corresponds to loop D in trypsin and is included in Perona and Craik's[1] loops classification for chymotrypsin-fold serine proteases. The high plasticity of the loop His57-Cys58-Val59-Ala60-Asn61-Val62-Asn62A-Val62B-Arg63-Ala64 (41−51 in MD numbering in Figure 3) in the free and noncovalently bound systems is significantly reduced for the covalently bound enzyme. This is probably a cooperative effect in which rigidity is transmitted along that segment once the His57 position becomes fixed upon establishing a covalent bond with the inhibitor. That loop corresponds to loop B in trypsin.[1]

The inhibitor stabilizing effect is still evident for the segment 16−30 (residues 1−15 in Figure 3) which occupies a superficial position on the enzyme.

ED analyses can be further extended so as to identify similarities or dissimilarities among the enzyme dynamical motions in the three MD trajectories. With this purpose, only the first five eigenvectors encompassing 52% (free HLE), 44% (noncovalently bound HLE), and 45% (covalently bound HLE) of all fluctuations were considered. The essential motions captured by these few eigenvectors are easily identifiable through representation of the individual Cα RMSDs by residue along each eigenvector (Figure S3, Supporting Information). In the free enzyme, eigenvector 1 (which accounts for 27% of the total variance along the MD trajectory) represents, almost exclusively, motion associated with the loop 94−102 (80−88 in Figure S3). The first eigenvector of the covalently bound system accounts for motions of both loops 94−102 and 144−150 (130−135 in Figure S3), while in the noncovalently bound enzyme the same eigenvector represents, among others, the fluctuations of loop 57−64 (41−51 in Figure S3). In this particular case, motion associated with the loop 94−102 has a low frequency. It is also worth mentioning that, apart from the eigenvector 1 associated with the free enzyme, none of the other eigenvectors considered represents a perfectly localized dynamical motion within the enzyme. The similarities between the essential motions of the enzyme in the three systems analyzed can be obtained by using the definition of similarity index[51]

$$\gamma_{AB} = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{n}(V_i^A \cdot V_j^B)^2 \qquad (4)$$

which is derived from the scalar products of the essential eigenvectors accounting for a given variance in the trajectory (in this particular case, ∼50%). The similarity index between the essential motions of the free and noncovalently bound enzymes, 0.14, indicates a high degree of dissimilarity. Comparison between the free and covalently bound forms of the enzyme, on the contrary, points to a certain degree of subspace overlap with a value of 0.58, which is most certainly associated with the fact that the five essential eigenvectors span fluctuations over similar regions in the enzyme.

3.3. RMSD-Based Clustering. An RMSD-based clustering analysis focusing on 10 residues lining the binding site enables a comparison between the relative structural populations of the HLE active site in the three distinct MD trajectories. The central conformations representative of the three most dominant clusters on each of the simulations illustrate the range of structural variability in the active site region (Figure 4).

The clustering analysis reveals some relevant features of the HLE binding site dynamics, namely its noticeable flexibility, which is attested by the significant amount of clusters obtained (Table 1). Such structural variability is reduced in the presence of both inhibitors, with the total number of clusters being reduced from 61 in the free HLE MD simulation to 54 in the noncovalently bound complex and 46 in the covalently bound system MD simulations. In terms of the comparative dynamics between the inhibitor-bound systems, it should be noticed that in both cases the majority of the active site conformational changes is represented by approximately the same number of clusters, ∼20, which indicates that their meaningful active site dynamics is similar.

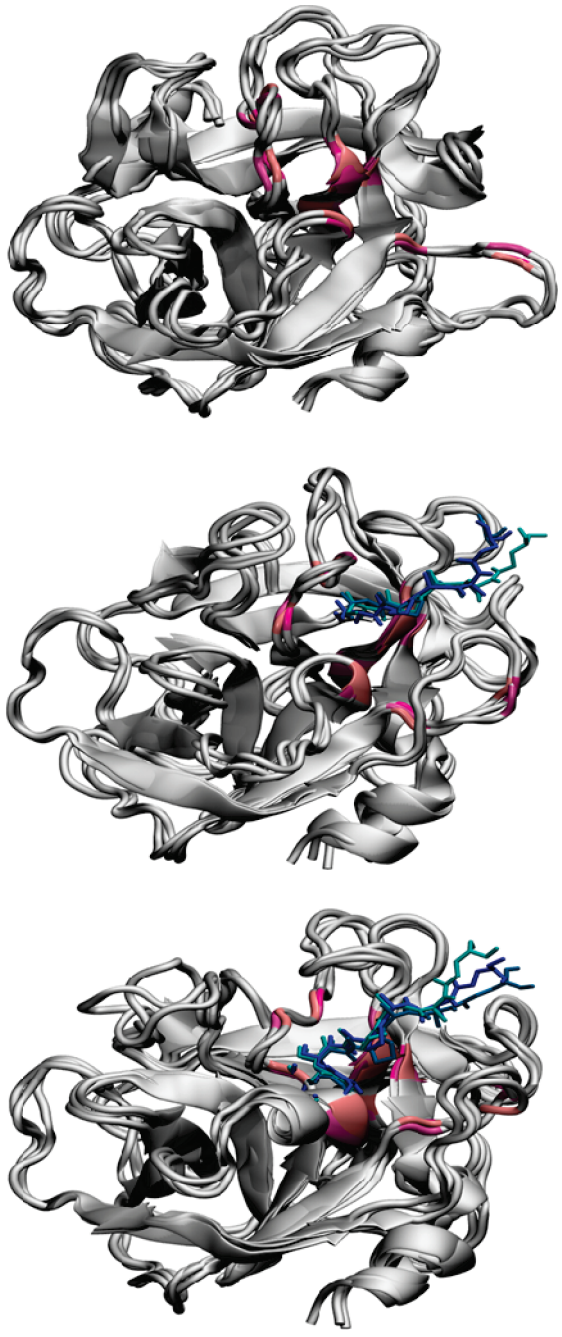


Figure 4. Representation of the central conformations representative of the three most dominant clusters for each of the simulations (from top to bottom: free HLE, noncovalently bound HLE, and covalently bound HLE). The binding site amino acids used for clustering are highlighted in pink tones and the inhibitor in blue.

In Figure 4 the MD relaxation of the free enzyme is apparent in the representative structures of the most dominant clusters of conformations. In the free enzyme there appears to exist a reorientation of some surface loops toward the solvent.

The effect of the covalently bound inhibitor translates into a more ordered active site at the catalytic triad histidine position (vide Figure 4, bottom): a α-helix protruding from His57 is formed upon binding to the inhibitor. This second order structure along the loop 57−64 (41−51 in MD numbering) is also responsible for its increased rigidity as observed for the enzyme Cα RMSFs in Figure 3. In order to further assess the stability of this HLE segment, secondary structure assignments for the individual amino acid residues 57−64 were made by using the DSSP algorithm[52] over 3 × 300 enzyme configurations extracted with 100 ps intervals from each simulation equilibrium trajectory. This analysis accounted for the presence of helices in all analyzed enzyme configurations of the covalently bound system. For the noncovalently bound HLE, that value amounts to only 15% of the equilibrium configurations, while in the free enzyme the presence of a helical structure for the sequence 57−64 is restricted to only 3% of the configurations. However, secondary structure analysis can be made more general so as to include another element, the hydrogen-bonded turn, which is a shorter version of a helix when the hydrogen-bonding pattern is too short. The sequence 57−64 adopts a turn nature for 85% of the noncovalently bound HLE configurations analyzed and for 43% of the free enzyme configurations. Analysis of the secondary structure of this proteic sequence for a set of available X-ray structures of HLE with a covalently bound inhibitor (1HNE, 2RG3, 1H1B, and 1B0F pdb entries) revealed an omnipresent helical structure that may be attributed to stabilization by the covalent ligand. In the X-ray structure of a noncovalent complex between HLE and a fragment of a 107-mer peptide protease inhibitor (2Z7F pdb entry), this particular sequence bearing His57 also adopts a helical structure despite the noncovalent nature of the interaction between HLE and the ligand. However, the fact that the inhibitor establishes six hydrogen bonds with the enzyme chain in the primary contact region may account for an overstabilization leading to an increase of secondary structure at the active site, which is less pronounced for the noncovalent system in this work.

Table 1. Total Number of Clusters and Number of Clusters Representing 90% of the Ensemble for Each System[a]

system	number of clusters	90% ensemble
free HLE	61	26
noncovalent system	54	21
covalent system	46	20

[a] A 1.4 Å RMSD cutoff was used.

3.4. Role of Surface Loops in the Enzyme Catalytic Mechanism. In order to gain insight into the conformational changes that the enzyme undergoes upon binding to the inhibitor, immediately before or after formation of the covalent adduct, the representative structures of the most populated cluster of conformations in each system MD trajectory were superimposed (Figure 5).

Differences in conformation for residues 94−102 belonging to a loop homologous to the surface segment loop C in trypsin[1,50] are noticeable between the free and the inhibitor-bound HLE systems. According to Perona and Craik,[1] this loop is invariably positioned to directly contact the substrate on the N-terminal side of the scissile bond in chymotrypsin-fold serine proteases. The importance of this surface loop has previously been acknowledged by Averhoff et al.,[50] who identified asparagine at position 98 as being essential for HLE specificity, contributing to a correct positioning of Leu99, which is part of the S_2 pocket. The loop is shown in evidence in yellow in Figure 5, with the darker shade of yellow highlighting its positioning in the free enzyme. Calculation of the average Cα RMSDs of the residues in this loop for the

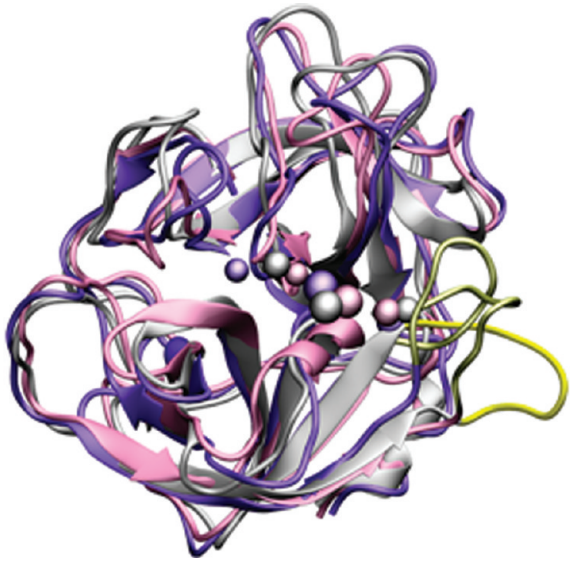


Figure 5. Superposition of the central representative enzyme structures of the most dominant cluster in each MD simulation (gray, free HLE; violet, noncovalently bound HLE; pink, covalently bound HLE). In yellow, focus is given to the surface loop 94−102 (in chymotrypsin position numbers), with the darker shade of yellow highlighting that loop in the free HLE MD simulation. Spheres indicate the positions of the catalytic triad residues in the three systems.

stable MD trajectories having as a reference those residues in the original crystal structure (after removing overall translation and rotations through structural-fitting using the enzyme Cα atoms) supports this observation: 0.98 ± 0.08 nm for the free HLE, 0.43 ± 0.02 nm for the noncovalently bound system, and 0.25 ± 0.06 nm for the covalent complex. It appears that in the free HLE MD trajectory, the loop, which is directed toward the inhibitor on both inhibitor-bound complexes, relaxes and protrudes toward the solvent. Given these RMSD values it may be assumed that the conformational differences in the aforementioned loop between the free and the inhibitor-bound enzymes hint for the existence of an induced-fit mechanism.[53] In an induced-fit mechanism, the catalytically active conformation of the enzyme is achieved only upon binding to the substrate. In such a framework, rearrangement of the loop 94−102 might be associated with a transition to a catalytically productive conformation of the binding site region.

Wei et al.[52,54] reported a small induced-fit conformational change of HLE upon binding to a valine chloromethyl ketone inhibitor. Contrarily to the present work, however, the authors did not compare conformations for both the free and bound HLE but, otherwise, reported an X-ray analysis on the conformational changes induced in HLE by both peptidyl chloromethyl ketone and turkey ovomucoid inhibitors. Observation of only a small, induced-fit, structural divergence among inhibitors was not wholly unexpected, since the chloromethyl ketone peptide chain bounds HLE in a conformation similar to the corresponding P1−P4 residues of the turkey ovomucoid inhibitor.

Some other authors have acknowledged the existence of a variety of serine proteases that act by an induced-fit mechanism, including the human cytomegalovirus protease,[55−57] the complement protease factor D,[58] and the prolyl endopeptidase.[59]

In order to investigate this assumption further, it is essential to verify whether the loop 94−102 contraction favored in the presence of the inhibitor (substrate) is accompanied by the

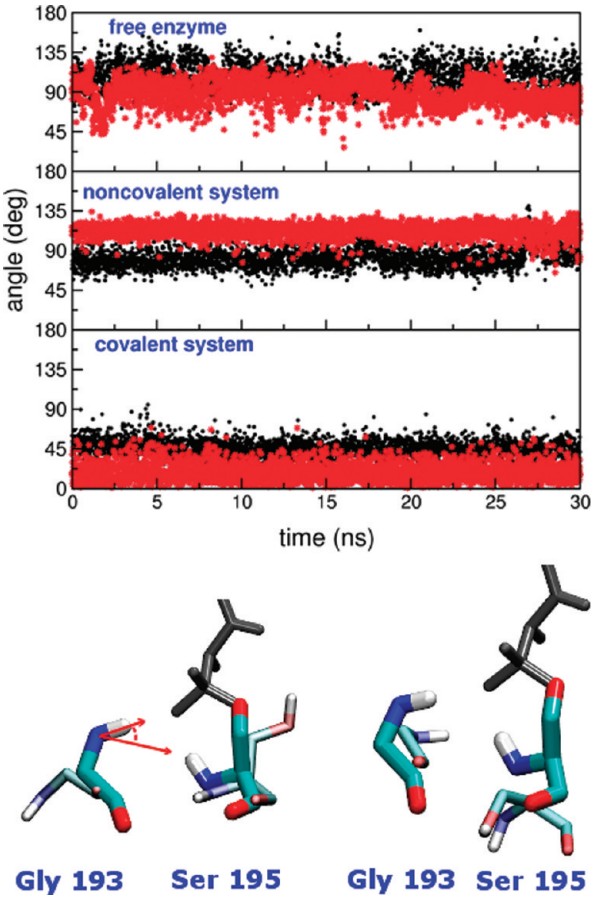


Figure 6. Top: Representation of $\angle[H_{NH}(Gly193)-N_{NH}(Gly193)\cdots H_{NH}(Ser195)]$ (in black) and $\angle[H_{NH}(Ser195)-N_{NH}(Ser195)\cdots H_{NH}(Gly193)]$ (in red) time evolution for the stable MD trajectory segments. Bottom: Comparison between MD snapshots for the oxyanion holes of the covalently bound enzyme (thicker bonds, inhibitor attached) and the free enzyme (lighter bonds, left) or the noncovalently bound system (lighter bonds, right). Schematization of one of the calculated angles is additionally given.

formation of a catalytically productive/competent oxyanion hole otherwise absent in the free enzyme or if in fact the free enzyme is already catalytically active.[60,61] With that purpose, two angles characterizing the oxyanion hole were defined, $\angle[H_{NH}(Gly193)-N_{NH}(Gly193)\cdots H_{NH}(Ser195)]$ and $\angle[H_{NH}(Ser195)-N_{NH}(Ser195)\cdots H_{NH}(Gly193)]$ (both with vertices in the amides N atoms).

Evolution in time of those angles over the last 30 ns of each MD trajectory is shown in Figure 6. If the oxyanion hole is in its active conformation, both angles should assume values fairly below 90°, indicating an inward common orientation of both protons. For the covalent system both angles assume average values of 46 ± 9° and 19 ± 10°, respectively, while for the free enzyme MD run they assume values around 82 ± 11° and 112 ± 8°. In the noncovalent system, the average angles have values of 108 ± 16° and 89 ± 14°, respectively. From the graph and the MD snapshots immediately below, which represent all three systems at the oxyanion hole site (representative conformations of the most populated clusters), it is evident that both in the free and noncovalently bound systems the oxyanion hole is not in its productive or catalytically competent form. Though the inhibitor presence in the noncovalently bound inhibitor system contributes to

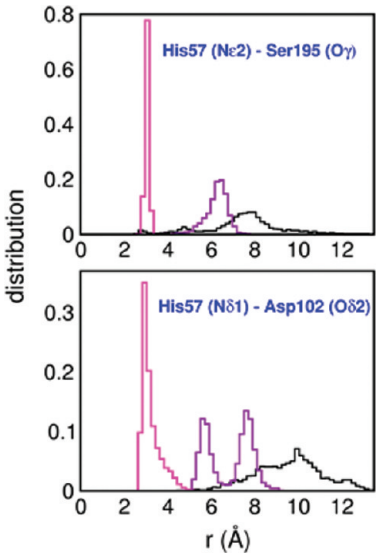


Figure 7. Distribution of some important distances between active site amino acids obtained for the equilibrium MD trajectories (black, free HLE; violet, noncovalently bound HLE; magenta, covalently bound HLE).

a similar contraction of the loop 94−102, absence of the transition-like tetrahedral enzyme−inhibitor adduct does not favor the formation of a oxyanion hole. Calculation of the heavy-atom hydrogen-bond distances between the oxyanion hole residues, Gly193 and Ser195, and the inhibitor alanine oxygen at P_1 for both inhibitor-bound systems confirms the presence of the catalytically productive oxyanion hole in the covalent system and its absence in the noncovalent counterpart: $d[N_{NH}(Gly193)\cdots O(AlaP_1)]_{coval} = 2.8 \pm 0.2$ Å, $d[N_{NH}(Gly193)\cdots O(AlaP_1)]_{noncoval} = 3.2 \pm 1.1$ Å, $d[N_{NH}(Ser195)\cdots O(AlaP_1)]_{coval} = 3.0 \pm 0.2$ Å, and $d[N_{NH}(Ser195)\cdots O(AlaP_1)]_{noncoval} = 5.8 \pm 0.6$ Å. In the crystal structure (1HNE pdb entry) those H-bond heavy-atom distances amount to 3.1 Å.[11]

In Figure 7, comparison between the distributions of two important binding site hydrogen bonds in terms of their heavy-atom distances obtained over the equilibrium MD trajectories for the three systems gives an additional perspective into the structural changes induced by the presence of the inhibitor. Hydrogen bonds between His57 Nδ1−H and Asp102 Oδ2 and between Ser195 Oγ−H and His57 Nε2 are generally observed among chymotrypsin-like serine proteases−inhibitor complexes.[62] In the HLE crystal structure considered as a basis for this work, the later hydrogen bond is absent due to the covalent bonds between those atoms and the inhibitor.

On the top panel of Figure 7, the distance between the His57 Nε2 and the Ser195 Oγ atoms is, as expected, much lower for the covalent system due to involvement of those atoms in covalent bonds with the irreversible inhibitor. In the original crystal structure that distance amounts to 3.0 Å, and the MD distribution peaks around the same value. The free enzyme presents the most noticeable separation among those atoms, with a maximum distribution around 8 Å. In the noncovalently bound system, that heavy-atom distance lies in an intermediate value of ∼6 Å. The noncovalently bound inhibitor thus confines the His57 Nε2 and the Oγ Ser195 atoms in closer proximity than in the free enzyme. Contraction of the loop 94−102 for the noncovalently bound enzyme, which resembles the same loop placement for the covalent system, may contribute to a better approximation between those atoms.

Table 2. MM-PBSA Binding Free Energies in kcal mol^{-1}[a]

free energy contribution	noncovalent	covalent
ΔE_{elec}	−36.2 (9.1)	−48.2 (9.0)
ΔE_{vdW}	−33.9 (3.0)	−28.8 (3.2)
$\Delta G_{solv,polar}$	50.3 (8.4)	25.4 (9.2)
$\Delta G_{solv,nonpolar}$	−3.6 (1.6)	−4.0 (0.2)
$-T\Delta S$	−5.0 (6.5)	18.1 (7.0)
$\Delta G_{binding}$	−28.4 (9.1)	−37.5 (8.5)

[a] Average contributions to the binding free energies in both inhibitor-bound complexes calculated over the last 30 ns of each system MD trajectory corresponding to equilibrium. Errors, given in parentheses, correspond to mean standard deviations. ΔE_{elec} and ΔE_{vdW} are the molecular mechanics electrostatic and van der Waals contributions, $\Delta G_{solv,polar}$ is the polar component of the solvation free energy, whereas $\Delta G_{nonpolar,solv}$ is the nonpolar component. $T\Delta S$ is the entropic contribution. $\Delta G_{binding}$ is the total binding energy. Errors in $\Delta G_{binding}$ were obtained by error propagation of the error in "$\Delta G_{binding}$ without entropic contribution" (calculated as the average of the individual $\Delta E_{elec} + \Delta E_{vdW} + \Delta G_{polar,solv} + \Delta G_{nonpolar,solv}$ contributions for each MD snapshot) and the error in $-T\Delta S$.

On the bottom panel of Figure 7, the distance between the His57 Nδ1 atom and the Asp102 Oδ2 atom in the covalently bound enzyme, with a maximum around 3 Å, points to the existence of a mechanism-related hydrogen bond, His57 Nδ1−H···Oδ2 Asp102, that is absent for the other two systems. It is an additional evidence for the absence of a productive alignment of the active site amino acids in the free and noncovalently bound enzymes. Asp102 is located over the loop 94−102, so its catalytically favorable binding site position is also a consequence of the loop conformation. In the crystal structure that hydrogen bond heavy atoms distance amounts to 2.7 Å.[11] Distribution of this distance for the noncovalently bound enzyme is bimodal, which reflects the flip of Asp102 carboxylate group along the 30 ns of equilibrium trajectory. In the free enzyme, once more, the distance among those atoms reaches the highest value.

3.5. Inhibitor−Enzyme Binding Free Energies. In order to investigate and quantify the individual contributions, driving forces, and the affinity for the binding process, the end-point free energies of interaction of both inhibitor systems were calculated through the MM-PBSA methodology. The simple thermodynamic cycle and single-trajectory postprocessing allows the efficient computation of the various contributions. Results are registered in Table 2.

Values in the table are model-dependent and reflect the inherent approximations implemented in the MM-PBSA formalism. The numbers represent only qualitative trends. Despite having been used as an approximation in some previous works,[63,64] it is reasonable to expect that the splitting procedure between inhibitor/enzyme configurations used in the covalent system MM-PBSA analysis might bring some additional errors into the calculations. However, the similarities identified so far between the inhibitor−enzyme interactions at the active site in both systems attest to the reasonability of a comparative analysis of the free energy results.

The estimated interaction free energies of the noncovalent and covalent complexes are ∼−28.4 and ∼−37.5 kcal mol^{-1}, respectively. These values should, however, be interpreted with caution, given the fact that the enzyme−inhibitor splitting procedure used for the covalent complex (vide Methods, section 2.4) does not properly take into account the chemical modification which is a main stabilizing factor in that system.

Table 3. MM-PBSA Free Energies, in kcal mol^{-1}, of HLE and HLE's Loop 94−102 for Both the Free and Noncovalently Bound Forms of the Enzyme[b]

	HLE		loop 94−102	
free energy contribution	free	noncovalent	free	noncovalent
E_{int}	1450.9 (31.5)	1438.5 (26.7)	72.5 (6.3)	69.2 (6.2)
E_{elec}	−31218.4 (76.8)	−31297.0 (58.6)	−1832.5 (41.5)	−1931.2 (20.3)
E_{vdW}	−1436.8 (22.5)	−1420.2 (19.6)	−68.1 (8.5)	−66.3 (5.9)
$G_{solv,polar}$	−3037.7 (74.2)	−2922.5 (56.6)	132.7 (40.6)	225.5 (21.9)
$G_{solv,nonpolar}$	33.3 (0.9)	33.1 (0.8)	0.8 (0.2)	0.4 (0.2)
$-TS$	−2333.7 (5.7)	−2331.4 (4.8)	−94.9 (6.3)	−96.4 (6.0)
G_{total}	−35026.9 (39.3)	−34995.0 (36.3)	−1789.5 (12.5)	−1798.8 (11.5)

[b] Average contributions calculated over the last 30 ns of each system MD trajectory corresponding to equilibrium (300 configurations by system). Entropy calculations were restricted to five MD snapshots for each system. Errors, in parentheses, correspond to mean standard deviations. Errors in G_{total} were obtained by error propagation.

According to the calculations, each complex formation is an enthalpically driven process with weakly favorable (noncovalent system) or unfavorable (covalent system) entropic contributions. The favorable formation of both complexes is driven by the electrostatic and vdW terms of the molecular mechanics free energy and the nonpolar component of the solvation free energy. The total solvation free energy, $G_{polar,solv} + G_{nonpolar,solv}$, is unfavorable for both complexes; thus, if considering only that contribution, it appears the HLE−inhibitor complexes would rather not be formed. The molecular mechanics energy component of the binding free energy compensates for this effect, favoring the complexes over the unbound molecules. The penalty paid by the positive solvation electrostatic term is completely covered by the favorable electrostatic interactions within the resulting enzyme−inhibitor covalent complex, which is not the case for the noncovalently bound complex. In both inhibitor−bound complexes, the entropic term does not favor (covalent system) or favors only weakly (noncovalent system) the interaction between the enzyme and the inhibitor. This is due to the fact that in both bound states the enzyme and the inhibitor are restrained to relatively more fixed positions. In fact, HLE is characterized by higher entropy in its free unbound state, which is attested by, for example, the RMSD clustering results registered in Table 1. The entropic penalty paid upon binding to the inhibitor is, understandably, high for the covalent system.

Rationalizing, upon formation of the HLE−inhibitor complexes, the enzyme approaches the inhibitor in an unfavorable desolvation process, replacing solute−solvent interactions with interactions within the complex. The unfavorable change in the electrostatics of solvation is fully compensated by the favorable electrostatic charge−charge interactions within the covalently bound complex, but not in the noncovalent complex. Once the enzyme and the inhibitor are in contact, vdW interactions stabilize the complexes structures.

As a means to assess the major energetic contributions to the observed constricted or relaxed conformations of the loop 94−102, an approximation was used to extract the loop free energy from the enzyme's corresponding value. The enzyme was formally split into two parts, and MM-PBSA free energy values for HLE and "HLE with no loop 94−102" were obtained. The free energy of the loop is approximated by $G_{loop} = G_{HLE} - G_{HLE\ no\ loop}$. This approximation[63,64] was used for both the free and the noncovalently bound enzyme systems, so as to gain insight into the energetic contributions to the closed or open loop conformations, respectively. In the covalent complex the additional inhibitor−enzyme interactions hinder comparisons with the values obtained for the nonbound form of the enzyme. It should be mentioned that the split of the enzyme free energy value into "loop" and "enzyme with no loop" contributions is an approximation that adds further uncertainty to the energies calculated, and therefore, the output has to be interpreted with care. The outcome of this analysis is registered in Table 3. Here, the contribution of the internal energy E_{int} (eq 3), which includes contributions from bending and torsion terms (there is no contribution from bond stretching due to the use of bond constraints in the simulations), does not vanish and has to be taken into account.

Both loop conformations are close in energy and, therefore, close in stability. This suggests that both arrangements of the loops are equally probable.

Internal energy terms contribute similarly to both open and closed loop conformations energetics. This is also the case for the molecular mechanics energy vdW terms.

Each loop conformation is strongly stabilized by the molecular mechanics electrostatic contribution to the free energy, with the condensed loop in the noncovalently bound inhibitor system having the major electrostatic stabilization. Though solvation terms contributing to the loop free energy assume unfavorable values in both cases, this effect is less pronounced for the open conformation of the loop, which is, therefore, favored in a solvation context.

Regarding the enzyme as a whole, it is evident that polar solvation terms have a stabilizing effect on the enzyme. The most favorable solvation contribution to the free energy value is, again, observed for the free enzyme. That effect might be partially ascribed to the loop arrangement. The apolar solvation term is similar in both cases, for the whole enzyme or only the loop.

Entropic contributions cannot be described as being determinant for the loop positioning given the equivalent values obtained for both loop conformations and the corresponding enzymes when treated as a whole.

In brief, the open conformation of the loop 94−102, observed for the free enzyme MD trajectory, is favored by the polar solvation contribution to the loop free energy. On the contrary, the closed conformation of that loop in the noncovalently bound HLE is promoted by the molecular mechanics electrostatic contribution to the free energy.

4. CONCLUSIONS

It has long been acknowledged that a protein function is indissociable from its structure and dynamic nature. Characterization of an enzyme's mode of action as a basis for rational drug design should, therefore, rely on a thorough description of the conformational changes and flexibility modifications that might occur upon ligand binding.

In this study, MD simulations have been used to investigate the flexibility and relative stabilities of the inhibitor-bound and inhibitor-free human leukocyte elastase with a special focus given to the binding pocket in a bid to identify the residues responsible for the complex stability.

Analysis of specific binding pocket amino acids failed to provide evidence of specific residues, which by undergoing a high stabilization upon inhibitor binding might indicate an important role in HLE substrate specificity. This is possibly related to the superficial nature of the S_1 pocket in HLE, which hinders identification of specific residues detrimental for binding.

A major finding that emerges from our study is the suggestion that HLE operates by an induced-fit mechanism. This induced-fit mechanism involves a surface loop adjacent to the binding pocket, loop 94−102, which contracts toward the inhibitor, bringing active site residues into a catalytically productive alignment. It is widely accepted that flexible loops have fundamental roles in a variety of enzymes, allowing them to acquire their functional forms. The role of these conformational changes in the loop 94−102 for HLE specificity is unclear and should be the object of further investigation. Presumably, the conformational changes in an induced-fit promote the selection of the right substrates and the exclusion of incorrect ones.[65]

Given the aforementioned flexibility effects observed upon ligand binding, HLE may prove itself a challenging target for structure-based drug design. The conformational sampling provided by atomic-level simulations constitutes thus an essential tool for those studies enabling one to tackle problems related to the use of static ensemble-averaged crystallographic enzyme models. Scarceness of available X-ray or NMR structures further justifies the use of MD simulations to obtain suitable HLE models for subsequent drug design studies. In the future, the use of MD-generated ensembles of HLE conformations as targets for molecular docking is highly recommended. A more sophisticated alternative could be the use of fully flexible molecular docking algorithms.

■ ASSOCIATED CONTENT

Supporting Information. Figure listing the force field united atom types used for the inhibitor. Table listing the force field parameters and atomic partial charges used for the inhibitor, whether covalent or noncovalently bound to the enzyme. Figure representing the evolution in time of RMSDs of S_1 and S_2/catalytic triad residues from the reference crystal structure. Figure representing the Cα RMSF per HLE residue along the first five eigenvectors identified in the ED analyses. This information is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: sestacio@ff.ul.pt; telephone: +351 217 946 400; fax: +351 217 946 470.

■ ACKNOWLEDGMENT

This work was funded by Fundação para a Ciência e Tecnologia (FCT, Portugal) through Project PTDC/QUI/64056/2006. S.G.E. gratefully acknowledges financial support from FCT with a postdoctoral fellowship (SFRH/BPD/46313/2008).

■ REFERENCES

(1) Perona, J. J.; Craik, C. S. Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* **1997**, *272*, 29987–29990.

(2) Hartley, B. S.; Shotton, D. M. In *The Enzymes*; Academic Press, 1971; Vol. III, pp 323−373.

(3) Babine, R. E.; Bender, S. L. Molecular recognition of protein−ligand complexes: Applications to drug design. *Chem. Rev.* **1997**, *97*, 1359–1472.

(4) Takuya, O.; Hakoshima, T.; Itakura, M.; Yamamori, S.; Takahashi, M.; Hashimoto, Y.; Shosaka, S.; Kato, K. Role of loop structures of neuropsin in the activity of serine protease and regulated secretion. *J. Biol. Chem.* **2002**, *277*, 14724–14730.

(5) Hedstrom, L.; Szilagyi, L.; Rutter, L. J. Converting trypsin to chymotrypsin: The role of surface loops. *Science* **1992**, *255*, 1249–1253.

(6) Korkmaz, B.; Moreau, T.; Gauthier, F. Neutrophil elastase, proteinase 3 and cathepsin G: Physicochemical properties, activity and physiopathological functions. *Biochimie* **2008**, *90*, 227–242.

(7) Pham, C. T. N. Protease-activated receptor-2 (PAR-2) is a weak enhancer of mucin secretion by human bronchial epithelial cells in vitro. *Int. J. Biochem. Cell Biol.* **2008**, *40*, 1317–1333.

(8) Mulchande, J.; Oliveira, R.; Carrasco, M.; Gouveia, L.; Guedes, R. C.; Iley, J.; Moreira, R. 4-Oxo-β-lactams (azetidine-2,4-diones) are potent and selective inhibitors of human leukocyte elastase. *J. Med. Chem.* **2010**, *53*, 241–253.

(9) Mulchande, J.; Guedes, R. C.; Tsang, W.-Y.; Page, M. I.; Moreira, R.; Iley, J. Azetidine-2,4-diones (4-oxo-β-lactams) as scaffolds for designing elastase inhibitors. *J. Med. Chem.* **2008**, *51*, 1783–1790.

(10) Guarnieri, F.; Spencer, J. L.; Lucey, E. C.; Nugent, M. A.; Stone, F. J. A human surfactant peptide−elastase inhibitor construct as a treatment for emphysema. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 10661–10666.

(11) Navia, M. A.; McKeever, B. M.; Springer, J. P.; Lin, T. Y.; Williams, H. R.; Fluder, E. M.; Dorn, C. P.; Hoogsteen, K. Structure of human neutrophil elastase in complex with a peptide chloromethyl ketone inhibitor at 1.84 Å resolution. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 7–11.

(12) Wei, A. Z.; Mayr, I.; Bode, W. The refined 2.3 Å crystal structure of human leukocyte elastase in a complex with a valine chloromethyl ketone inhibitor. *FEBS Lett.* **1988**, *234*, 367–373.

(13) Macdonald, S. J. F.; Dowle, M. D.; Harrison, L. A.; Clarke, G. D. E.; Inglis, G. G. A.; Johnson, M. R.; Shah, C.; Smith, R. A.; Amour, A.; Fleetwood, G.; Humphreys, D. C.; Molloy, C. R.; Dixon, M.; Godward, R. E.; Wonacott, A. J.; Singh, O. M. P.; Hodgson, S. T.; Hardy, G. W. Discovery of further pyrrolidine trans-lactams as inhibitors of human neutrophil elastase (HNE) with potential as development candidates and the crystal structure of HNE complexed with an inhibitor (GW475151). *J. Med. Chem.* **2002**, *45*, 3878–3890.

(14) Koizumi, M.; Fujino, A.; Fukushima, K.; Kamimura, T.; Takimoto-Kamimura, M. Complex of human neutrophil elastase with 1/2SLPI. *J. Synchrotron Radiat.* **2008**, *15*, 308–311.

(15) Cregge, R. J.; Durham, S. L.; Farr, R. A.; Gallion, S. L.; Hare, C. M.; Hoffman, R. V.; Janusz, M. J.; Kim, H.-O.; Koehl, J. R.; Mehdi, S.; Metz, W. A.; Peet, N. P.; Pelton, J. T.; Schreuder, H. A.; Sunder, S.; Tardif, C. Inhibition of human neutrophil elastase. 4. Design, synthesis, X-ray crystallographic analysis, and structure−activity relationships for a series of P2-modified, orally active peptidyl pentafluoroethyl ketones. *J. Med. Chem.* **1998**, *41*, 2461–2480.

(16) Huang, W.; Yamamoto, Y.; Li, Y.; Dou, D.; Alliston, K. R.; Hanzlik, R. P.; Williams, T. D.; Groutas, W. C. X-ray snapshot of the

mechanism of inactivation of human neutrophil elastase by 1,2,5-thiadiazolidin-3-one 1,1-dioxide derivatives. *J. Med. Chem.* **2008**, *51*, 2003–2008.

(17) Powers, J. C. In *Modification of Proteins; Advances in Chemistry*; American Chemical Society: Washington, DC, 2011; Vol. 198, pp 347−367.

(18) Imperiali, B.; Abeles, R. H. Inhibition of serine proteases by peptidyl fluoromethyl ketones. *Biochemistry* **1986**, *25*, 3760–3767.

(19) Powers, J. C.; Asgian, J. L.; Ekici, O. D.; James, K. E. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem. Rev.* **2002**, *102*, 4639–4750.

(20) Ohbayahi, C. Current synthetic inhibitors of human neutrophil elastase in 2005. *Expert Opin. Ther. Pat.* **2005**, *15*, 759–771.

(21) Adcock, S. A.; McCammon, J. A. Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chem. Rev.* **2006**, *106*, 1589–1615.

(22) Steinbrecher, T.; Case, D. A.; Labahn, A. A multistep approach to structure-based drug design: Studying ligand binding at the human neutrophil elastase. *J. Med. Chem.* **2006**, *49*, 1837–1844.

(23) Steinbrecher, T.; Hrenn, A.; Dormann, K. L.; Merfort, I.; Labahn, A. Bornyl (3,4,5-trihydroxy)-cinnamate—An optimized human neutrophil elastase inhibitor designed by free energy calculations. *Bioorg. Med. Chem.* **2008**, *16*, 2385–2390.

(24) *Molecular Operating Environment (MOE)*, v2009.10; Chemical Computing Group: Montreal, Quebec, Canada, 2009.

(25) Lindahl, E.; Hess, B.; van der Spoel, D. J. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306–317.

(26) van der Spoel, D. J.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(27) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(28) Aalten, A. W.; van Schuettelkopf, D. M. F. PRODRG—A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D* **2004**, *60*, 1355–1363.

(29) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

(30) Miyamoto, S.; Kollman, P. A. SETTLE: An analytical version of the shake and rattle. *J. Comput. Chem.* **1992**, *13*, 952–962.

(31) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(32) Nose, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255–268.

(33) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev.* **1985**, *31*, 1695–1697.

(34) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(35) Nose, S.; Klein, M. L. Structural transformations in solid nitrogen at high pressure. *Phys. Rev. Lett.* **1983**, *50*, 1207–1210.

(36) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald-an N. log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(37) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. A smooth particle mesh ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(38) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential dynamics of proteins. *Proteins: Struct. Funct. Genet.* **1993**, *17*, 412–425.

(39) Stepanova, M. Dynamics of essential collective motions in proteins: Theory. *Phys. Rev. E* **2007**, *76*, 051918.

(40) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide folding: When simulation meets experiment. *Angew. Chem., Int. Ed.* **1999**, *38*, 236–240.

(41) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897.

(42) Swanson, J. M. J.; Henchman, R. H.; McCammon, J. A. Revisiting free energy calculations: A theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophys. J.* **2004**, *86*, 67–74.

(43) Amaro, R. E.; Cheng, X.; Ivanov, I.; Xu, D.; McCammon, J. A. Characterizing loop dynamics and ligand recognition in human and avian-type influenza neuraminidases via generalized Born molecular dynamics and end-point free energy calculations. *J. Am. Chem. Soc.* **2009**, *131*, 4702–4709.

(44) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10037–10041.

(45) Dolinsky, T.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G.; Baker, N. PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **2004**, *35*, W522–W525.

(46) Wang, W.; Kollman, P. A. Computational study of protein specificity: The molecular basis of HIV-1 protease drug resistance. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 14937–14942.

(47) Brooks, B.; Karplus, M. Harmonic dynamics of proteins: Normal modes and fluctuations in bovine pancreatic trypsin inhinitor. *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80*, 6571–6575.

(48) Carlson, H. Protein flexibility is an important component of structure-based drug discovery. *Curr. Pharm. Des.* **2002**, *8*, 1571–1578.

(49) Shu, Q.; Frieden, C. Relation of enzyme activity to local/global stability of murine adenosine deaminase: 19F NMR studies. *J. Mol. Biol.* **2005**, *345*, 599–610.

(50) Averhoff, P.; Kolbe, M.; Zychlinsky, A.; Weinrauch, Y. Single residue determines the specificity of neutrophil elastase for *Sighella* virulence factors. *J. Mol. Biol.* **2008**, *377*, 1053–1066.

(51) Noy, A.; Pérez, A.; Lankas, F.; Javier Luque, F.; Orozco, M. Relative flexibility of DNA and RNA: A molecular dynamics study. *J. Mol. Biol.* **2004**, *343*, 627–638.

(52) Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637.

(53) Koshland, D. E., Jr. Correlation of structure and function in enzyme action. *Science* **1963**, *142*, 1533–1541.

(54) Bode, W.; Meyer, E., Jr.; Powers, J. C. Human leukocyte and porcine pancreatic elastase: X-ray crystal structures, mechanism, substrate specificity, and mechanism-based inhibitors. *Biochemistry* **1989**, *28*, 1951–1963.

(55) de Oliveira, C. A. F.; Guimarães, C. R. W.; Barreiro, G.; de Alencastro, R. B. Investigation of the induced-fit mechanism and catalytic activity of the human cytomegalovirus protease homodimer via molecular dynamics simulations. *Proteins: Struct. Funct. Bioinf.* **2003**, *52*, 484–491.

(56) LaPlante, S. R.; Cameron, D. R.; Aubry, N.; Bonneau, P. R.; Déziel, R.; Grand-Maître, C.; Ogilvie, W. W.; Kawai, S. H. The conformation of a peptidyl methyl ketone inhibitor bound to the human cytomegalovirus protease. *Angew. Chem., Int. Ed.* **1998**, *37*, 2729–2732.

(57) LaPlante, S. R.; Bonneau, P. R.; Aubry, N.; Cameron, D. R.; Déziel, R.; Grand-Maître, C.; Plouffe, C.; Tong, L.; Kawai, S. H. Characterization of the human cytomegalovirus protease as an induced-fit serine protease and the implications to the design of mechanism-based inhibitors. *J. Am. Chem. Soc.* **1999**, *121*, 2974–2986.

(58) Taylor, F. R.; Bixler, S. A.; Budman, J. I.; Wen, D.; Karpusas, M.; Ryan, S. T.; Jaworski, G. J.; Safari-Fard, A.; Pollard, S.; Whitty, A. Induced fit activation mechanism of the exceptionally specific serine protease, complement factor D. *Biochemistry* **1999**, *38*, 2849–2859.

(59) Li, M.; Chen, C.; Davies, D. R.; Chiu, T. K. Induced-fit mechanism for prolyl endopeptidases. *J. Biol. Chem.* **2010**, *285*, 21487–21495.

(60) Aleshin, A. E.; Shiryaev, S. A.; Strongin, A. Y.; Liddington, R. C. Structural evidence for regulation and specificity of flaviviral proteases and evolution of the *Flaviviridae* fold. *Protein Sci.* **2007**, *16*, 795–806.

(61) Ekonomiuk, D.; Caflisch, A. Structural evidence for regulation and specificity of flaviviral proteases and evolution of the *Flaviviridae* fold. *Protein Sci.* **2009**, *18*, 1003–1011.

(62) Hedstrom, L. Serine protease mechanism and specificity. *Chem. Rev.* **2002**, *102*, 4501–4524.

(63) Haider, S.; Parkinson, G.; Neidle, S. Molecular dynamics and principal component analysis of human telomeric quadruplex multimers. *Biophys. J.* **2008**, *95*, 296–311.

(64) Hazel, P.; Parkinson, G.; Neidle, S. Predictive modeling of topological and loop variations in dimeric DNA quadruplex structures. *Nucleic Acids Res.* **2006**, *34*, 2117–2127.

(65) Johnson, K. A. Role of induced fit in enzyme specificity: A molecular forward/reverse switch. *J. Biol. Chem.* **2008**, *283*, 26297–26301.