

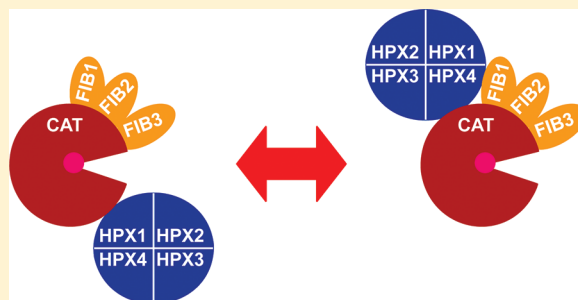
Alternative Interdomain Configurations of the Full-Length MMP-2 Enzyme Explored by Molecular Dynamics Simulations

Natalia Díaz* and Dimas Suárez

Departamento de Química Física y Analítica, Julián Clavería 8, Universidad de Oviedo, Oviedo (Asturias), 33006 Spain

S Supporting Information

ABSTRACT: Conformational freedom between the different domains of the matrix metalloproteinase family of enzymes has been repeatedly invoked to explain the mechanism of hydrolysis of some of their most complex macromolecular substrates. This proposed interdomain motion has been experimentally confirmed to occur in solution for matrix metalloproteinases MMP-1, MMP-9, and MMP-12. In this work, we computationally assess the likely conformational freedom in aqueous solution of the full-length form of the MMP-2 enzyme in the absence of its pro-peptide domain. To this end, we perform molecular dynamics (MD) simulations and approximate free energy analyses in four different arrangements of the protein domains that correspond to (a) the compact conformation observed in the X-ray structure; (b) an initially elongated structure in which the hemopexin (HPX) domain is separated from the catalytic (CAT) and fibronectin domains; and (c–d) two alternative conformations suggested by protein–protein docking calculations. Overall, our results indicate that the interdomain flexibility is very likely a general property of the MMP-2 enzyme in solution.



INTRODUCTION

Multidomain Structure of the Matrix Metalloproteinases (MMPs). MMPs are a family of zinc and calcium-dependent endopeptidases involved in a wide variety of biological functions, both in physiological and pathological conditions.^{1–5} Their catalytic (CAT) domain (~170 residues) hydrolyzes all of the macromolecular components located in the pericellular environment. It displays a characteristic fold including a five-stranded β -sheet, three α -helices, and several bridging loops. But most of the MMPs are multidomain enzymes, and besides the CAT domain, they usually present a pro-peptide and an hemopexin-like domain (HPX).^{1,6–8} The propeptide, which is a small N-terminal domain (~80 residues) that blocks the access to the active site in the latent pro-MMPs, is proteolytically removed from the CAT domain upon enzyme activation. The C-terminal HPX domain (~210 residues) forms a β -propeller structure with four blades (HPX1–HPX4) arranged around a central axis, each composed of one α -helix and a four-stranded antiparallel β -sheet.^{8,9} HPX is considered to be important for the activation of some pro-MMPs, the inhibition of the active forms, their dimerization, and their attachment to the cell surface.¹⁰ It also mediates the binding and cleavage of substrates like chemokines or the triple helix of collagen by providing secondary binding sites named exosites.^{6,10–12}

HPX is joined to CAT by a proline-rich linker (LNK) also termed the hinge region. For many MMPs this linker is relatively short (~16–20 residues), whereas for others, it constitutes a highly glycosylated domain (~64 residues). LNK is of critical importance for the degradation of complex substrates like collagen, which seems to require the concerted action of

the CAT and HPX domains.^{12,13} Enzymes like MMP-2 and MMP-9 have an additional and characteristic domain consisting of three tandem copies of a 58 amino acid residue fibronectin type II-like module (FIB1–FIB3), which are inserted into the sequence of CAT.¹⁴ Each FIB module comprises two short double-stranded antiparallel β -sheets, which are connected through a short α -helix and arranged approximately perpendicular to each other, defining thus a hydrophobic pocket that is accessible from the outside. The FIB structures are stabilized by two disulfide bonds, and they are linked to each other and to the CAT domain by flexible polypeptide linkers. Different analyses have confirmed that the FIB modules account for the affinity of MMP-2 and MMP-9 for gelatin, type I and IV collagen, elastin, and laminin.^{15–18} Finally, those MMPs bearing a transmembrane α -helix and a cytoplasmic domain give rise to the MT-MMP subfamily of enzymes.^{6,7}

Evidences for Interdomain Flexibility in the MMPs. Conformational freedom between the different domains of the MMPs, especially between the CAT and the HPX domains, has been suggested as a general property for this family of enzymes.^{19,20} Both the narrowness of the active site cleft, which requires that peptide substrates adopt an extended conformation for efficient catalytic activity, and the broad range of structurally diverse proteins that can be processed by the MMPs have led researchers to assume some degree of reciprocal interdomain reorientation during their catalytic action. This is remarkably

Received: November 17, 2011

Revised: February 5, 2012

Published: February 11, 2012

the case for the hydrolysis of collagen given that the MMPs must unwind the triple helix prior to the cleavage of its individual α -chains.²¹ Moreover interdomain flexibility might also facilitate molecular ratcheting along matrix molecules for progressive degradation and/or movement.^{22,23} Thus it has been shown that activated MMP-1, which requires the presence of HPX for collagen hydrolysis, moves progressively on the collagen fibril following a “bias diffusion mechanism” dependent on collagen proteolysis.²⁴ In addition, spectroscopic analyses suggest the ability of MMP-2 to diffuse laterally on the gelatin surface, what seems greatly facilitated by the C-terminal HPX domain.²⁵

The slight differences in the reciprocal orientation of the CAT and HPX domains observed in the MMP-1 crystal structures, depending whether or not the pro-domain is present, have been taken as a first evidence that the two domains could experience some relative mobility.^{19,26} More recently, small-angle X-ray scattering, atomic force microscopy imaging, and nuclear magnetic resonance (NMR) experiments confirmed the ability of the MMP-1, MMP-9, and MMP-12 enzymes to reorient their CAT and HXP domains in the absence of solid-state packing forces.^{20,27,28} Atomic force microscopy images have also revealed significant changes in the MMP-9 interdomain conformation upon binding to a collagen molecule.²³ Therefore it seems likely that the compact arrangement observed in the X-ray structures of the full-length MMP enzymes could be in equilibrium with other more open structures in solution. Thus, as already pointed out by some authors, the CAT and HPX domains of the MMPs could be packed as separate entities in the crystal structures with a connecting flexible linker peptide or “hinge”.^{6,12}

Molecular modeling studies have also suggested that alternative conformations to that observed in the crystal structures of the full-length MMPs could be accessible in solution. Thus structural studies performed for MMP-1 and MMP-2 have confirmed that a noticeable motion of their HPX domain relative to CAT is required for these enzymes to form an inhibitory complex with the so-called tissue inhibitors of metalloproteinases (TIMPs).²⁹ On the other hand, a molecular dynamics (MD) simulation of a single MMP-2 enzyme in aqueous solution has shown that extended configurations could be structurally and energetically feasible following a conformational change of the LNK region with a concomitant rearrangement of the HPX domain with respect to the CAT and FIB ones.³⁰ Subsequently, protein–protein docking calculations have been used to explore alternative conformations accessible to MMP-2.³¹ In the most stable docking solutions, the HPX domain is placed at the rear part of the CAT domain.

Goals of This Study. In this work, we pursue to further analyze the structure and stability of various possible configurations of the full-length MMP-2 enzyme in aqueous solution by means of MD simulations and approximate free energy calculations. We selected four different configurations of the active MMP-2 enzyme that mainly differ in the relative position of HPX with respect to the CAT-FIB domains and, consequently, in the folding of the linker region. The structural and dynamical analyses of the four MD simulations will allow us to characterize the likely evolution of the initial configurations and the stability of the main interdomain contacts. In addition, by computing the approximate free energies of the configurations along the different trajectories by means of the molecular-mechanics Poisson–Boltzmann (MM-PB) approach complemented with configurational entropy estimations using the

Schlitter's equation, we will be able to assess their relative stability in aqueous solution.

■ EXPERIMENTAL METHODS

Setup of the Models. In Figure S1 of the Supporting Information we display the initial geometries of the four different models built for the full-length MMP-2 enzyme in its active form:

- The X model corresponds to the compact arrangement observed in the 1CK7 X-ray structure, where the first blade of HPX contacts the CAT domain. We removed the propetide domain (Pro₃₁-Asn₁₀₉) and placed the resulting N-terminal coil (Tyr₁₁₀-Lys₁₁₆) in the hydrophobic cavity defined by the Ω -loop and the α 3-helix.^{14,32} The missing residues in the linker region that connects the CAT and the HPX domains were built with the MolIDE1.0 program.³³
- The second model was obtained from a snapshot selected from the final portion of our previous 100 ns MD simulation of the full-length enzyme (model E in Figure S1, Supporting Information). In this previous simulation, we observed a transition from the X-ray compact structure to a more extended form characterized by the absence of interdomain interactions between the HPX and the CAT and FIB domains. However, the simulation was halted once that the transient extended configuration was observed because the original solvent box, which was dimensioned for the compact form of the MMP-2 enzyme, could not avoid artifacts due to interactions between mirror images of the protein molecule and therefore, it was not adequate to properly sample the extended configuration. In the present work, the LNK region has an initial end-to-end distance of 46.4 Å (Asp₄₅₀-Ca \cdots Ile₄₆₈-Ca, MMP-2 numbering),³⁰ and a much larger solvent box is used to find out whether or not the extended conformation of the MMP-2 is dynamically stable.
- The third and fourth selected models of the full-length MMP-2, D-1 and D-2, correspond to the most likely docking solutions from our previous protein–protein docking calculations using the PyDock methodology.^{34–37} Initial coordinates for these docking calculations were obtained from a snapshot extracted from the previous MD simulation of the full-length MMP-2 enzyme. We also note that the linker region comprising the Asp₄₅₀-Ile₄₆₈ residues, which were not taken into account during the pyDock calculations, was then rebuilt through a series of model-building operations and molecular mechanics calculations in which the most stable docking solutions were combined with MD snapshots obtained from an independent simulation of the isolated linker peptide. In the final docked structures, the fourth blade of the HPX domain is interacting with the rear of the CAT and FIB domains simultaneously (model D-1) or with the rear of the CAT domain only (model D-2), with a partially folded linker (Asp₄₅₀-Ca \cdots Ile₄₆₈-Ca distances of 26.5 and 15.7 Å, respectively).

All of the MMPs hold a zinc ion bound by the conserved motif HEXXHXXGXXH³⁸ in the deep active site cleft. A second zinc ion and a number of calcium ions also play a role in stabilizing the global structure of CAT and fine-tuning the access to the primary binding sites (S₃-S₃') located along the active site

area.^{39,40} The four models of the full length MMP-2 enzyme considered in this work present two zinc ions and two calcium ions bound to the CAT domain of the enzyme, whereas a third calcium ion is placed at the entrance of the central channel observed in the hemopexin domain.⁴⁰ Zn₁, which is the catalytic zinc ion, is modeled in a distorted trigonal–bipyramid geometry coordinated by the three histidines from the conserved motif and two water molecules.⁴¹ Zn₂, which is the “structural” zinc ion, is coordinated by three histidines and one aspartic acid in a tetrahedral environment in all of the crystal structures.⁸

For the X model, we kept all of the water molecules observed in the experimental structure in direct interaction with the protein atoms. For the E model, all of the water molecules placed in a 6.0 Å shell around the protein atoms were conserved from the MD snapshot (~10200 water molecules). Concerning the D-1 and D-2 configurations, where all of the explicit water molecules were removed prior to the protein–protein docking calculations, we employed the DOWSER program to place selected water molecules.⁴² Finally, each “pre-solvated” model was centered in an octahedral solvent box of TIP3P water molecules that extended 15 Å from the protein surface. Nine Na⁺ counterions were placed at the edges of each solvent box to neutralize the systems. This resulted in systems containing 120195, 257892, 112053, and 100599 atoms for the X, E, D-1, and D-2 models, respectively.

The all-atom AMBER parm03 force field was used to represent the protein atoms.⁴³ For the zinc atoms, we employed a bonded model based on the previously developed parametrization within the parm94 force field.⁴⁰ Following the parm03 prescriptions, we reparametrized the atomic charges of the Zn ions and their ligands employing the RESP protocol and the B3LYP/cc-pVTZ electrostatic potential computed using the IEFPCM continuum solvent model ($\epsilon = 4$).⁴³ For the calcium ions, we employed the nonbonded parametrization proposed by Aqvist.⁴⁴ However, we have observed in previous calculations that the coordination spheres around the different calcium ions are not stable with the AMBER parm03 force field because the three Ca²⁺ tend to unbound the enzyme after 4–5 ns of MD simulation.³⁰ Thus, we imposed the following harmonic constraints on the distances between the calcium ions and selected oxygen atoms in their ligands: Ca₁...O@Asp₁₈₅, Ca₁...O@Gly₁₈₆, Ca₁...O@Asp₁₈₈, Ca₁...O@Leu₁₉₀, Ca₂...O@Asp₁₆₈, Ca₂...O@Gly₂₀₀, Ca₂...O@Gly₂₀₂, Ca₃...O@Asp₄₇₆, Ca₃...O@Asp₅₂₁, Ca₃...O@Asp₅₆₉, Ca₃...O@Asp₆₁₈. For these restraints, we defined a reference distance of 2.8 Å and a small force constant of 5.4 kcal/(mol Å²).

MD Simulations. Energy minimizations and MD simulations were performed with the NAMD program.⁴⁵ Solvent molecules and Na⁺ counterions were initially relaxed by means of energy minimizations and 100 ps of MD. Then, the full systems were minimized to remove bad contacts in the initial geometry and heated gradually to 300 K during 120 ps of MD. The SHAKE algorithm was used to constrain all of the R-H bonds, and periodic boundary conditions were applied to simulate a continuous system. A nonbonded cutoff of 10.0 Å was used whereas the particle-mesh Ewald (PME) method with a grid spacing of ~1 Å was employed to include the contributions of long-range interactions. Langevin dynamics was employed to control the temperature (300 K) using a damping factor of 2 ps⁻¹, whereas pressure control (1 atm) employed the Berendsen bath coupling.

A 200 ns trajectory was computed for each model with a time step of 2 fs. Coordinates were saved every 2 ps, and snapshots

extracted every 4 ps were considered for analysis using the ptraj program included in the Amber10 package and some other specific software developed locally. To analyze the relative position between the different domains along the four simulations, we computed the Euler angles that characterize the relative orientation of two rigid coordinate systems placed at the center of mass of the considered domains. For the CAT domain, the x -axis is parallel to the axis running along the α 2-helix, and the third reference point that defines the xy plane is located at the center of mass of the backbone residues of the helix α 3. For each FIB module, its x -axis connects the center of mass of the two characteristic disulfide linkages, whereas the third reference point is placed at the small β -sheet. Finally, the x -axis to describe the movement of HPX runs along the central channel of the β -propeller structure with the third reference point situated at the disulfide linkage that connects the first and the last residue of this domain.

Energetic Analyses. To compare the relative stability of the four full-length MMP-2 models, we estimated various free energy components by performing MM-PB calculations^{46,47} over 4000 snapshots extracted from each MD simulation every 50 ps. The snapshots were postprocessed through the removal of all solvent molecules and counterions. The MM-PB energy, which should be interpreted as a physically based scoring function, was computed according to the following equation:

$$G_{\text{MM-PB}} = E_{\text{MM}} + \Delta G_{\text{solv}}^{\text{PB}} + \Delta H_{\text{solute-solvent}}^{\text{vdW}} + \gamma A \quad (1)$$

where E_{MM} is the molecular mechanics energy, $\Delta G_{\text{solv}}^{\text{PB}}$ is the electrostatic solvation energy obtained from Poisson–Boltzmann calculations,⁴⁸ and $\Delta H_{\text{solute-solvent}}^{\text{vdW}}$ is the solute–solvent van der Waals energy contributing to the nonpolar part of solvation energy.⁴⁷ The cavitation free energy contribution to the nonpolar solvation energy is determined by a molecular surface area dependent term (γA).

The SANDER program included in Amber10 was used to compute (no cutoff) the molecular mechanics energy terms (E_{MM}), while the electrostatic contributions to the solvation free energy ($G_{\text{PB-elec}}$) were determined using the PBSA program also available in Amber10. In the PB calculations, atomic charges and radii were taken from the parm03 representation. The linearized PB equation was solved on a cubic lattice by using an iterative finite-difference method. The cubic lattice had a grid spacing of 0.33 Å, and the points at the boundary of the grid were set to the sum of Debye–Hückel potentials. The van der Waals interaction energies between solute and solvent atoms were determined for a water shell of 12 Å thickness around the solute with no cutoff using SANDER. To estimate the cavitation energy, the surface tension proportionality constant γ was set to 69 cal mol⁻¹ Å⁻², and the molecular surface area was determined using the MOLSURF program included in Amber10, applying Bondi radii for the solute atoms and a water probe radius of 1.4 Å.

The configurational entropy of the protein molecules was estimated by the following approximate formula introduced by Schlitter⁴⁹ based on a quantum-mechanical one-dimensional harmonic approximation:

$$S_{\text{config}} = \frac{1}{2} R \ln \left(\det \left[1 + \frac{k_B T \epsilon^2}{\hbar^2} \sigma' \right] \right) \quad (2)$$

where σ' is the mass-weighted covariance matrix calculated directly in terms of Cartesian coordinates of the solute atoms.^{50,51}

This formula turns out to be computationally very efficient, allowing thus the estimation of the entropy of large macromolecules.

To remove the overall translation of the center of mass and the overall rotation of the protein, all of the MD snapshots employed in the entropy calculations were superposed on top of each other using a least-squares fit. Due to sampling issues, the S_{config} values depend on the length of the MD trajectory that is analyzed, although they usually tend to a limiting value as the simulation time increases. Thus, we fitted the various S_{config} plots by using the following empirical relationship that has been proposed in a previous work:⁵²

$$S_{\text{config}}(t) = S_{\text{config}}^{\infty} - \frac{A}{t^{2/3}}$$

where t is the simulation time, and the limiting entropy value ($S_{\text{config}}^{\infty}$) and A are both fitting parameters that are optimized by a nonlinear regression analysis.

RESULTS

Overall Structure of the Enzyme. Changes in the internal geometry and the relative position of the various domains that build up the active full-length MMP-2 enzyme were first analyzed by computing the mean values of the root-mean-squared deviations (rmsd) of the coordinates of the heavy atoms with respect to the 1CK7 crystal structure. The results are collected in Figure 1 and in Table S1 in the Supporting Information, whereas the time evolution of the rmsd for the backbone atoms of CAT, FIB, and HPX is displayed in Figure S2. According to the rmsd values obtained for the CAT domain, 1.1–1.5 Å for the backbone heavy atoms, the internal structure of this domain remains close to that observed in the 1CK7 structure in all of the simulations, the X and D-1 configurations presenting the lowest results (1.1 ± 0.1 Å). In the same way, the internal structure of the HPX domain is quite similar to the crystallographic one (backbone rmsd values of 1.6–2.0 Å), the largest deviations being observed for HPX4 along the E and D-2 trajectories (2.4 ± 0.3 Å backbone atoms). In contrast, the large rmsd values computed for the whole FIB domain warn us that its internal geometry changes during the simulations, especially in the case of the E and D-2 configurations (5.2 ± 0.7 and 5.1 ± 0.3 Å, respectively). Segregation of these rmsd values into three different contributions, one for each FIB module, points out that the FIB intramodule rearrangements are generally moderate (rmsd = 1.5–2.0 Å), but they are accompanied by a considerable degree of intermodule displacement. Most remarkably, the largest changes both in internal structure and relative positioning correspond to the third FIB module (FIB3) during the E and D-2 configurations (e.g., rmsd FIB3 values are 3.6 ± 0.4 and 3.3 ± 0.2 Å, respectively). These relatively large rmsd values for FIB3 are most likely due to the lack of FIB...HPX interdomain interactions in these two enzyme configurations as well as to the absence of the propeptide domain in the simulated full-length MMP2 enzyme (FIB3 interacts with the propeptide in the reference 1CK7 ray structure of the pro-MMP-2 enzyme).

Concerning protein flexibility as measured by the rmsd's with respect to the average structure obtained for each trajectory, the RMSF values in Table S2 of the Supporting Information further confirm that the architecture of the CAT and HPX domains remains quite stable in all the simulations (0.6–0.8 and 0.9–1.1 Å for the backbone atoms, respectively). The FIB domain shows a

significant flexibility for the E and D-1 configurations when compared with the X and D-2 ones (2.0–2.2 Å vs 1.2–1.3 Å). This larger flexibility of the whole FIB domain can be traced back to changes in the internal structure of the modules, mainly in FIB3 for the E trajectory, and to some degree of reorientation among the FIB modules.

To further characterize the time evolution of the interdomain arrangements during the MD simulations, we monitored the distances between the center of mass of the various domains (Figure S3 of the Supporting Information). In the case of the E simulation, the HPX domain is initially separated from CAT and FIB (e.g., $R_{\text{CM}}(\text{CAT}\cdots\text{HPX}) = \sim 75$ Å) given that the linker region is fully stretched, but such an extended conformation turned out to be stable only during the first 25 ns of simulation time and, subsequently, the HPX and CAT domains approached each other ($R_{\text{CM}}(\text{CAT}\cdots\text{HPX}) = \sim 40$ Å) and formed direct protein–protein contacts during the last 50 ns. In contrast, the CAT...FIB distance along the E simulation remains quite stable, the resulting average value (27.9 ± 0.4 Å) being quite similar to those computed for the other three trajectories. On the other hand, the D-1 and D-2 simulations show moderate fluctuations in their interdomain distances (e.g., $R_{\text{CM}}(\text{CAT}\cdots\text{HPX})$ has mean values of 33.2 ± 1.8 and 32.0 ± 0.6 Å for D-1 and D-2, respectively), which indicates that the overall interdomain configurations in these simulations are well-maintained. Even more stable values of the interdomain separation were observed in the case of the compact X arrangement all along the simulation time (e.g., $R_{\text{CM}}(\text{CAT}\cdots\text{HPX}) = 39.6 \pm 0.4$ Å). Interestingly, the average interdomain distances for D-1 and D-2 suggest that these models are slightly more compact than the X one. This is also confirmed by computing the radius of gyration for the full enzyme: 27.7 ± 0.2 Å for X, 26.5 ± 0.3 Å for D-1, 26.9 ± 0.3 Å for D-2, and 36.0 ± 4.0 Å for E.

To characterize the relative orientation of the protein domains regardless of their separation, we computed the Euler angles defined by a reference coordinate system placed at the CAT domain and a series of analogous rigid coordinate systems placed at the center of mass of the HPX domain and at each of the three FIB modules. The time evolution of the resulting ϕ , θ , and ψ angles is shown in the form of polar plots collected in Figures S4–S7 of the Supporting Information for the various pairs of coordinate systems (CAT...HPX, CAT...FIB1, etc.). Thus, we were able to characterize an important rotational motion of the HPX domain with respect to CAT that occurs along the E simulation concomitant with the reduction in the HPX...CAT distance (see Figure S4). Both the significant displacement and reorientation of HPX explain the large RMSF obtained for the whole protein in the E simulation. In contrast, no substantial rearrangement in the initial HPX orientation with respect to the CAT domain is observed in the X, D-1, and D-2 trajectories. This means that the characteristic orientation of the HPX domain predicted by our previous Pydock calculations, which is quite different to that of the X model, turned out to be essentially stable during the MD simulations as the corresponding Euler angles slightly fluctuate along the simulations (see Figure S4). Concerning the positioning of the three fibronectin modules, Figure S5 shows that FIB1 adopts a slightly different orientation in the D-1 trajectory as compared with the other three simulations. More pronounced rotational effects occur at the FIB2 module, which is notably reoriented in the D-1 trajectory and experiences rigid rotational oscillations of large amplitude ($\pm 20^\circ$; see Figure S6). Similarly, the

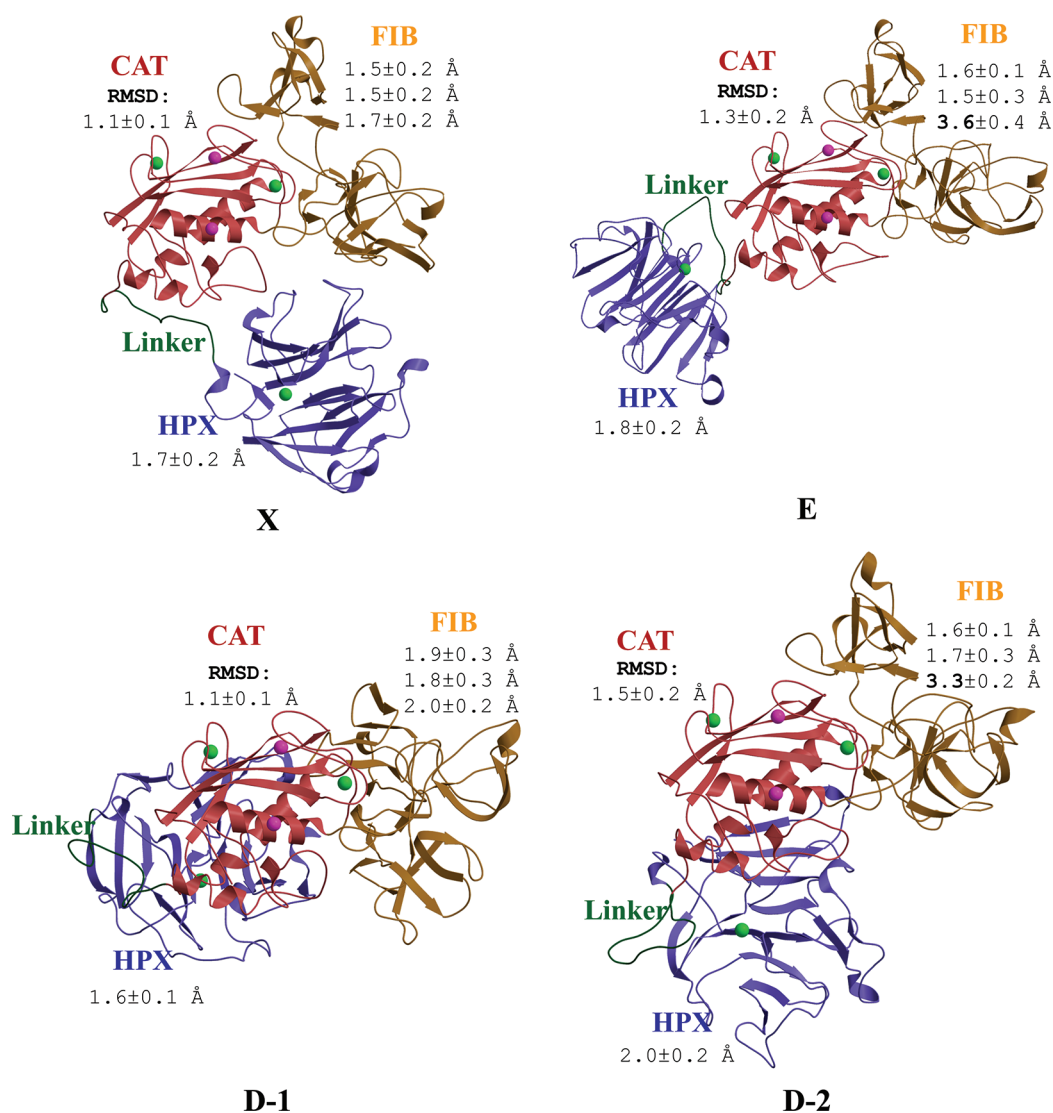


Figure 1. Secondary structure for the different domains of snapshots obtained along the four simulations performed for the full-length MMP-2 enzyme. Zn²⁺ (in purple) and Ca²⁺ (in green) ions are shown as spheres. Average rmsd values and standard deviations (in Å) are computed for the backbone atoms of the catalytic (CAT), the three modules of the fibronectin domain (FIB), and the hemopexin domain (HPX) along each 200 ns trajectory.

rotational motions of the FIB3 module fluctuate significantly along the E and D-1 trajectories.

Interdomain Residue Contacts. Clearly, the relative positioning of the CAT, FIB, and HPX domains in the four configurations of the full-length MMP-2 enzyme examined in this work can be traced back to different interdomain residue contacts. Similarly, the flexibility of the interdomain motions along the 200 ns simulations is directly linked to the stability of those protein–protein contacts. Hence, in this section, we characterize the main interdomain contacts observed in the four simulations in aqueous solution and compare them with similar contacts that are present in the crystallographical 1CK7 structure.

Let us first be reminded that the experimental 1CK7 configuration corresponding to the pro-MMP-2 enzyme in the solid state is characterized by several polar and hydrophobic interdomain contacts that stabilize a quite compact arrangement (radius of gyration = 27.6 Å) of the different domains. For example, the loop connecting β 1 and α 1 in the CAT domain contacts the first module of the FIB domain

(Thr₁₃₉–O γ ...Tyr₂₃₂–O η 2.4 Å, Glu₁₃₈–O ϵ ...Arg₂₂₂–N η 3.0 Å, Leu₁₃₅...Val₂₂₀ 5.5 Å), whereas residues in α 1 interact with the second FIB module (Pro₁₃₃–O...Thr₃₃₅–N 3.0 Å, Tyr₁₃₁...Tyr₃₄₁ 5.1 Å). In addition, α 3 in CAT contacts the linker region between CAT and HPX (Gln₄₃₅–N ϵ ...Leu₄₆₁–O 2.6 Å, Gln₄₃₅–N...Pro₄₆₃–O 2.7 Å). Some residues in the characteristic Ω -loop of CAT are close to the first blade of the HPX domain (Asn₄₃₀–O...Arg₄₉₁–N η 2.9 Å, Pro₄₁₇...Pro₅₀₆ 5.1 Å), and a cluster of hydrophobic residues from the final segment of the linker and the first blade of HPX (Pro₄₆₆...Phe₄₉₂ 5.4 Å, Pro₄₆₆...Trp₄₉₄ 5.8 Å, Ile₄₆₈...Trp₄₉₄ 4.6 Å, Ile₄₆₈...Pro₅₀₃ 4.6 Å) contributes to maintain the relative orientation between them. Finally a salt bridge has also been observed between the FIB and the HPX domains (Glu₂₄₃–O ϵ ...Arg₅₅₀–N η 2.9 Å) that connects the first FIB module with the second HPX blade.

The stability in aqueous solution of the interdomain contacts observed in the 1CK7 structure can be directly assessed by the X simulation. In fact several of the crystallographical contacts remained perfectly stable (see Table S3 in the Supporting Information). Thus, the Pro₁₃₃–O...Thr₃₃₅–N contact between

CAT and FIB2, the $\text{Gln}_{435}\cdots\text{Pro}_{463}\cdots\text{O}$ interaction between CAT and the linker, and the hydrophobic contacts between the final part of the linker and HPX1 ($\text{Pro}_{466}\cdots\text{Phe}_{492}$ and $\text{Ile}_{468}\cdots\text{Trp}_{494}$), present all occupation percentages of 90–100% all along the X simulation. However, other interactions, like the $\text{Glu}_{243}\cdots\text{Oe}\cdots\text{Arg}_{550}\cdots\text{N}\zeta$ salt-bridge, are progressively weakened or even lost along the trajectory. Thus, it can be concluded that the interdomain arrangement characteristic of the 1CK7 structure is quite well-maintained along the X simulation thanks to a subset of the crystallographically observed protein–protein contacts that were dynamically stable in this particular MD simulation.

As mentioned above, the starting points of the D-1 and D-2 simulations were derived from rigid docking calculations that probed the ability of the HPX domain to give favorable contacts with the molecular surface of the CAT-FIB ensemble. In the docking solution corresponding to the D-1 configuration, HPX is located at the rear of the CAT domain (the front part of CAT being located at the Zn1 active site) with its fourth blade interacting with CAT and FIB2 residues mainly via hydrophobic contacts. In this starting structure, the partially folded linker chain interacts with CAT residues through polar contacts and with HPX via hydrophobic contacts. As the secondary structure and the residue side chains were basically frozen in the previous docking solutions, it is clear that the 200 ns MD simulation should favor the sampling and relaxation of the protein–protein contacts. Hence, Table S4 of the Supporting Information reveals that the nature and identity of the interdomain direct contacts evolves during the MD simulation, providing thus a refined description of the protein–protein interactions characteristic of the D-1 configuration. For instance, the linker region suffers a conformational transition during the first 50 ns that resulted in the formation of a very stable hydrophobic cluster between the linker (Val_{464} and Ile_{468}) residues and the first blade of HPX (Ile_{473} , Phe_{475} , Phe_{492} , Trp_{494} , Pro_{506}). On the other hand, the MD simulation also reveals that the HPX domain, which is slightly rotated in the second half of the trajectory, prefers to form some polar contacts with the FIB2 module (e.g., the $\text{FIB2}\cdots\text{HPX4}$ contact $\text{Tyr}_{314}\cdots\text{O}\eta\cdots\text{Asn}_{611}\cdots\text{N}$) and that the CAT and FIB domains can give new polar interactions like $\text{Lys}_{187}\cdots\text{N}\cdots\text{Asp}_{370}\cdots\text{O}$ and $\text{Asp}_{210}\cdots\text{O}\cdots\text{Lys}_{372}\cdots\text{N}\zeta$ (see Table S4).

In the starting D-2 geometry, the fourth blade of HPX interacts only with the rear of the CAT domain, basically through polar contacts (e.g., $\text{Gln}_{123}\cdots\text{Ne}\cdots\text{Ser}_{652}\cdots\text{O}\gamma$, $\text{Asp}_{153}\cdots\text{O}\delta\cdots\text{Lys}_{633}\cdots\text{N}\zeta$, etc.). In the same structure, the linker residues interact mainly with HPX (Ile_{473} , Arg_{491} , and Lys_{489}). With respect to this initial picture, the MD relaxation of the D-2 configuration resulted again in the rearrangement and/or formation of new polar and hydrophobic contacts among the protein domains that are collected in Table S5 of the Supporting Information. Thus, the linker region reorients itself to give many polar and hydrophobic contacts with either CAT or HPX residues all along the simulation (e.g., $\text{Glu}_{467}\cdots\text{O}\cdots\text{Lys}_{489}\cdots\text{N}\zeta$, $\text{Gln}_{435}\cdots\text{N}\cdots\text{Pro}_{463}\cdots\text{O}$, $\text{Ile}_{468}\cdots\text{Trp}_{494}$, $\text{Leu}_{135}\cdots\text{Val}_{220}$, etc.), further stabilizing the relative positioning of the CAT and HPX domains. In addition, the $\text{CAT}\cdots\text{HPX4}$ interaction is reinforced thanks to the formation of a new $\text{Asp}_{142}\cdots\text{O}\delta\cdots\text{Ser}_{647}\cdots\text{O}\gamma$ contact (see Table S5). On the other hand, the $\text{CAT}\cdots\text{FIB1}$ contact is maintained by the $\text{Asp}_{136}\cdots\text{O}\delta\cdots\text{Arg}_{222}\cdots\text{N}\eta$ salt-bridge and the $\text{Asp}_{136}\cdots\text{O}\delta\cdots\text{Tyr}_{232}\cdots\text{O}\eta$ H-bond, whereas Arg_{315} controls the orientation of FIB2 with

respect to CAT thanks to the stability of the $\text{Tyr}_{131}\cdots\text{O}\cdots\text{Arg}_{315}\cdots\text{N}\eta$ and $\text{Asp}_{209}\cdots\text{O}\delta\cdots\text{Arg}_{315}\cdots\text{N}\eta$ contacts all along the trajectory.

At the beginning of the E simulation, the HPX domain is entirely surrounded by water molecules as its center of mass is placed at 75.8 and 87.7 Å away from the center of mass of the CAT and FIB domains, respectively. However, as already commented, the HPX and CAT-FIB domains approach each other as the linker region connecting them folds back during the second half of the simulation. Thus, we observed the formation of a direct $\text{Glu}_{443}\cdots\text{Oe}\cdots\text{Lys}_{633}\cdots\text{N}\zeta$ salt-bridge between $\alpha 3$ in CAT and HPX4 during the last 50 ns of the trajectory (32% of occupancy and 5.8 Å). On the other hand, some of the initial contacts between the CAT domain and residues from FIB1 ($\text{Leu}_{135}\cdots\text{Leu}_{215}$, $\text{Leu}_{135}\cdots\text{Val}_{220}$, $\text{Asp}_{136}\cdots\text{O}\delta\cdots\text{Arg}_{222}\cdots\text{N}\eta$), FIB2 ($\text{Tyr}_{131}\cdots\text{N}\cdots\text{Gly}_{313}\cdots\text{O}$, $\text{Tyr}_{131}\cdots\text{Tyr}_{314}$, $\text{Pro}_{133}\cdots\text{O}\cdots\text{Thr}_{335}\cdots\text{N}$, $\text{Pro}_{137}\cdots\text{Tyr}_{314}$, ...) and FIB3 ($\text{Asp}_{210}\cdots\text{O}\delta\cdots\text{Lys}_{372}\cdots\text{N}\zeta$, $\text{Leu}_{212}\cdots\text{Pro}_{391}$, $\text{Thr}_{214}\cdots\text{O}\gamma\cdots\text{Gln}_{393}\cdots\text{O}$) are preserved all along the E trajectory (see Table S6 of the Supporting Information).

Energetic Analyses. To investigate the relative stability of the X, D-1, D-2, and E configurations of the full-length MMP-2 enzyme, we evaluated their approximate free energy (G_{MMPB}) in aqueous solution as described in Methods. The evolution of the G_{MMPB} values, which combine the MM potential energy of the protein with the Poisson–Boltzmann (PB) solvation energy of the protein and the nonpolar contributions, is displayed in Figure 2 for each simulation.

Although the G_{MMPB} values represent *approximate* free energies that lack any solute entropic contribution (see below), they clearly suggest that the full-length MMP-2 macromolecule relaxes from the conformation in the crystal structure with higher energy to a more stable conformation in aqueous solution. Thus, we see in Figure 2 that the MM-PB energy tended to significantly decrease during the first ~40 ns of the X simulation, which was started from the 1CK7 solid state structure. After this initial stage, the G_{MMPB} plot of the X trajectory passes through upward and downward phases during the rest of the trajectory. Most likely, the initial decrease in G_{MMPB} is due to the relaxation of the protein structural motifs that are most solvent accessible as well as that of the linker chain connecting the CAT and HPX domains, whose initial conformation was generated by model building as commented above in the Methods section. On the other hand, we also see in Figure 2 that the G_{MMPB} energy of the E trajectory fluctuates all along the 200 ns trajectory around a very stable accumulated average. This behavior was not entirely unexpected, partly because the starting point of the E configuration (including the linker region and the closest shell of water molecules around the protein) was taken from a MD snapshot of a previous trajectory. However, the observed constancy of G_{MMPB} regardless of the large interdomain rearrangements occurring during the E trajectory suggests that the elongation/shortening of the $\text{HPX}\cdots\text{CAT-FIB}$ separation does not imply large changes in the combined potential and solvation energy of the MMP-2 enzyme.

Concerning the docking-based simulations, our energetic analyses indicate that the D-1 configuration is significantly stabilized during the first quarter of the 200 ns trajectory whereas the D-2 plot equilibrates faster. In the two simulations, however, it can be observed that the accumulated mean value slowly decreases and the energy trend fluctuates during the second half of the trajectory, which is more accentuated in the case of the D-2 simulation. Even though the internal geometry of the protein domains (excepting the linker region) in the

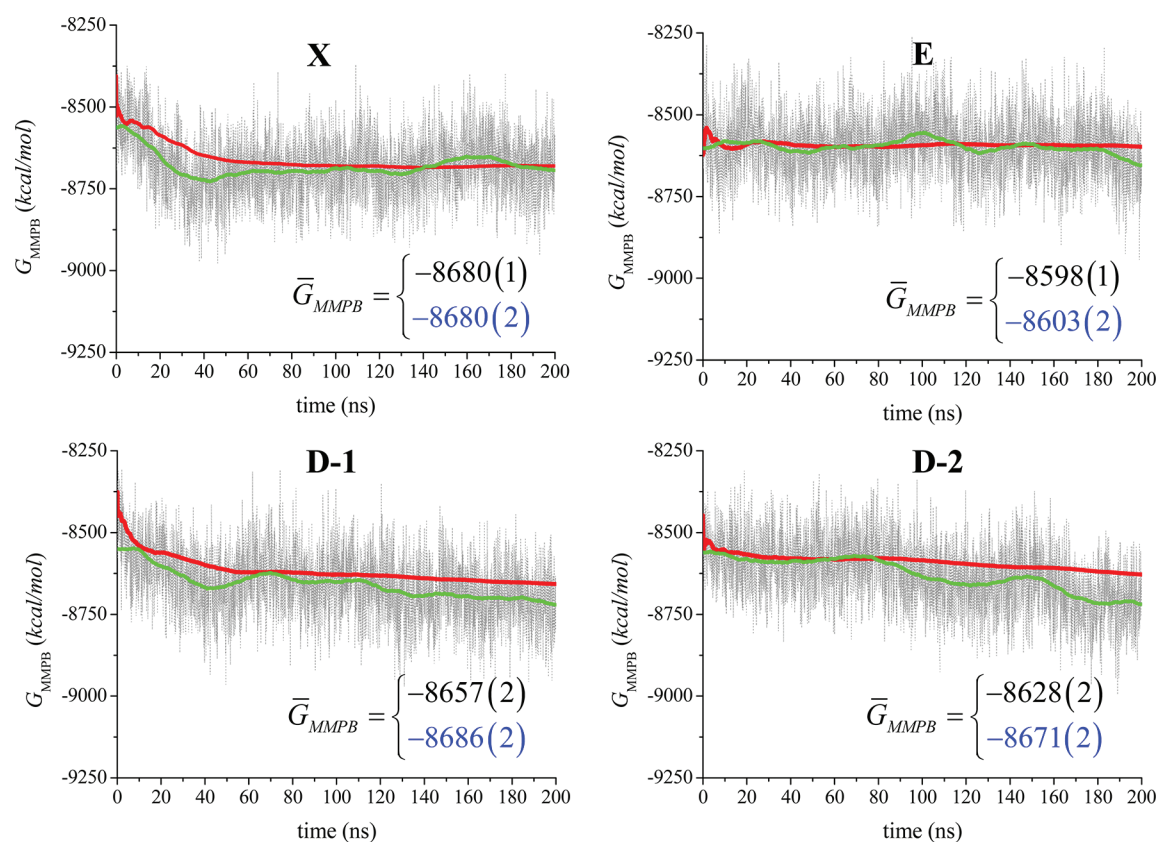


Figure 2. Solute energy plus solvation free energy (G_{MMPB} in kcal/mol) plots for the various MMP-2 configurations. The red and green curves represent the evolution of the accumulated average and the adjacent average, respectively. The mean value and standard error (in parentheses) for the whole trajectory and for the last 100 ns of simulation time (in blue) are also indicated.

starting point of the docking configurations were also taken from a previously equilibrated MD simulation, the observed trends in their G_{MMPB} plots can be explained in terms of various dynamical changes that take place during the simulations as the relaxation of the linker region, the rupture and formation of residue contacts among the protein domains, the slight re-orientation and rigid rotation motions of the protein domains, and so forth.

To energetically score the X, E, D-1, and D-2 configurations using the G_{MMPB} data from their respective MD simulations, we have to take into account that the average G_{MMPB} values may be sensitive to long-time-scale dynamic effects. Therefore, we computed the mean values of G_{MMPB} for the whole 200 ns trajectories and for their second halves (100–200 ns time interval). In any case, the resulting \bar{G}_{MMPB} values, which are collected in Figure 2, indicate that the E configuration, which evolves from an initially extended conformation toward a more compact one, is much less stable than the other models by ~ 70 – 90 kcal/mol. Assigning the most stable configuration in terms of the relative \bar{G}_{MMPB} energies depends on the time interval used for averaging. Thus, the X model has the lowest \bar{G}_{MMPB} value in the 0–200 ns interval, followed by the D-1 and D-2 models that have higher \bar{G}_{MMPB} values by 23 and 52 kcal/mol, respectively. But if we restrict the averaging to the second half of the trajectories, when the G_{MMPB} energies exhibit a more stable behavior, then the D-1 configuration turns out to be the most stable one by only 6 and 15 kcal/mol with respect to the X and D-2 trajectories.

Although the MD-averaged G_{MMPB} values constitute a physically based scoring function, it is clear that we cannot

predict reliable free energy differences among the X, D-1, and D-2 configurations due to the limited sampling of the inter-domain motions. Furthermore, our G_{MMPB} calculations ignore the role of the solute configurational entropy that needs also to be taken into account. To this end, we also estimated the configurational entropy contribution (S_{config}) associated to the protein atoms by means of the Schlitter approximation. Thus, Figure 3 shows the S_{config} plots versus simulation time either for the whole MMP-2 enzyme or its various protein domains that resulted from the accumulated entropy values ongoing from the first MD frame to the last one using a 5 ns offset. To improve the statistical convergence, the calculations were carried out over the 200 ns trajectories. Unfortunately, the S_{config} curves still exhibit a nonzero slope at the end of the simulation (see Figure 3), and therefore, we resorted to the estimation of the $S_{\text{config}}^{\infty}$ limiting values using the extrapolation formula that is described in the Methods section. The resulting $S_{\text{config}}^{\infty}$ values allow us to discriminate between the E trajectory, which has an estimated configurational entropy of 34.5 kcal/(K mol), and the other three models, which have $S_{\text{config}}^{\infty}$ values that are below ~ 2.0 kcal/(K mol). In addition, the D-1 configuration ($S_{\text{config}}^{\infty} = 32.7$ kcal/(K mol)) seems to be stabilized by configurational entropy with respect to the X and D-2 configurations (32.2 and 32.0 kcal/(K mol)). When we compare the entropy plots and the $S_{\text{config}}^{\infty}$ data segregated into protein domain contributions, the largest entropy differences (± 0.4 – 0.5 kcal/(K mol)) across the various MMP-2 configurations arise in the entropy of their HPX and FIB domains. In principle, the entropy due to the interdomain motions can also be estimated by subtraction of the domain contributions from the total $S_{\text{config}}^{\infty}$ value, which

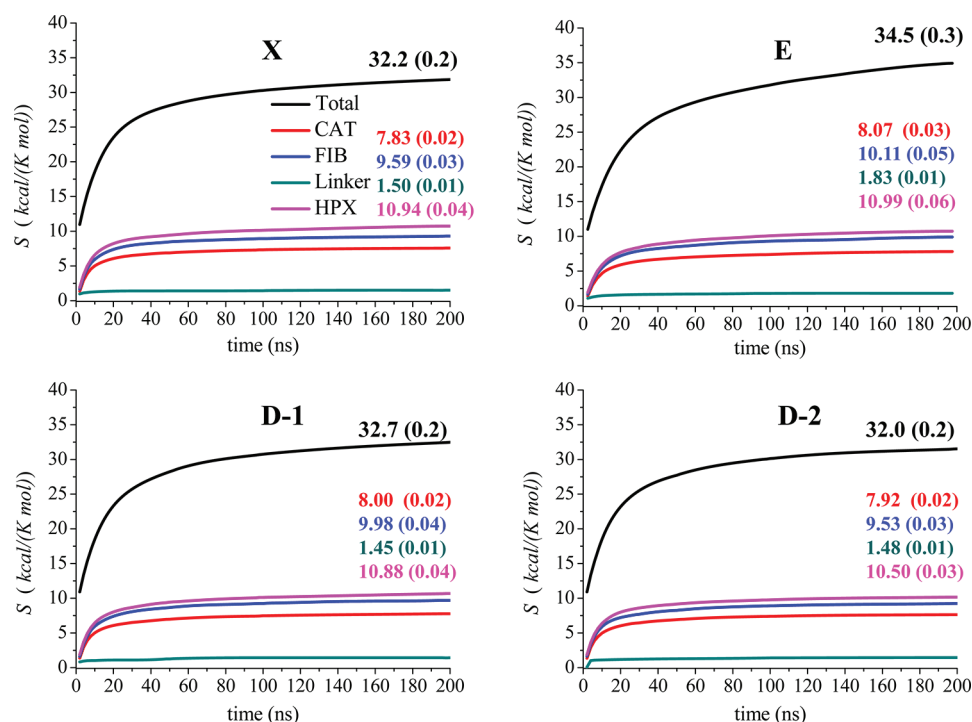


Figure 3. Configurational entropy plots (in kcal/(K mol)) obtained with the Schlitter's formula for the full-length form and the separate protein domains of the MMP-2 enzyme in the various configurations. The values of the $S^{\infty}_{\text{config}}$ fitting parameters are also indicated for each plot (standard errors in parentheses).

results in entropy values of 2.3, 3.5, 2.4, and 2.6 kcal/(K mol) for the X, E, D-1, and D-2 configurations, respectively. Taken all together, our entropy estimations clearly suggest that differences in the inter and intradomain mobility can have a notable entropic impact in the relative stability of the MMP-2 models. Particularly, the flexibility of the HPX domain with respect to the rest of the enzyme that is characteristic of the E configuration, apparently results in a large entropy increase. In terms of free energies at 300 K, these entropy differences observed would amount to tens and even hundreds of kcal/mol. However, it must be emphasized that the Schlitter approximation and the closely related quasi-harmonic method provide an *upper* limit to the true entropy given that supralinear correlations among the system variables are ignored and that the multim minima potential energy surface is exceedingly smoothed by defining only one minimum and ignoring any anharmonicity (large uncertainties can also arise due to the arbitrariness in the choice of reference atoms for the preliminary structure superposition).⁵³ Therefore, it is very likely that the entropy plots in Figure 3 would largely overestimate the actual entropy differences among the MMP-2 configurations, which prevents us from computing reliable free energy terms.

DISCUSSION

The first remarkable result of our simulations is that the interdomain arrangement characteristic of the 1CK7 structure is essentially preserved along the 200 ns X trajectory thanks to a small number of highly persistent CAT...LNK (e.g., Gln₄₃₅-N...Pro₄₆₃-O) and CAT...FIB (e.g., Pro₁₃₃-O...Thr₃₃₅-N) interdomain contacts. Nonetheless, the inherent flexibility and modular constitution of the FIB domain result in appreciable changes in the relative positioning and orientation of the three FIB modules along the X simulation, most likely in response to the absence of the N-terminal propeptide domain. A very

similar behavior concerning the FIB dynamics has been observed in our previous 100 ns MD simulation started at the 1CK7 X-ray configuration using similar settings.³⁰ However the previous simulation resulted also in the loss of various interdomain HPX...CAT and HPX...FIB contacts, with a concomitant conformational change of the linker region connecting the HPX and CAT domains that resulted in a fully elongated form of the MMP-2 enzyme. This transition has not been observed in the present X configuration, which exhibits only moderate structural and energetic changes after the initial 40 ns of simulation time.

The different outcome of the two independent MD simulations started at the 1CK7 structure seems in consonance with similar variations between individual MD simulations that have been detected in other systems.^{54,55} In fact, it can be reasonably expected that many extensive unbiased simulations would be required to sample the conformational space more efficiently and extract relevant information about the potential interdomain rearrangements of the MMP-2 enzyme. As this approach is still very costly in terms of computer time, other techniques can be used like biased MD protocols⁵⁶ or protein-protein docking methodologies³⁷ specifically aimed at exploring the conformational landscape available to multidomain proteins. In the case of the MMP-2 enzyme, we have previously applied the Pydock methodology for rigid protein-protein docking calculations,³¹ showing that the HPX domain of the MMP-2 enzyme could interact through a hydrophobic region located in the fourth blade with the rear part of the CAT domain or with the FIB1-FIB2 modules and that those alternative conformations are compatible with the molecular structure of the linker region. The D-1 and D-2 simulations reported in this study allow us to further examine the stability in aqueous solution of the most favored protein-protein docked structures. As expected, the extensive MD sampling

with explicit solvent has resulted in the relaxation of the protein–protein contacts and in the slight reorientation of the protein domains with reference to the initial docked structures (see Tables S4–S5 and Figures S3–S6 of the Supporting Information). In this way, we note that the D-1 and D-2 simulations have substantially refined the molecular details of the protein–protein docking solutions, but they have also validated them given that the MD trajectories do not depart too much from the starting docked structures.

Certainly, the calculation of the MM-PB energies and the Schlitter's entropies for the MMP-2 models allow us to make only semiquantitative assessments of their relative stability owing to the approximations involved in the energy and entropy calculations and the limited sampling. Nevertheless, we believe that these energetic analyses strongly suggest that the “docked-based” D-1 and D-2 configurations of the active MMP-2 could have a similar stability in aqueous solution to that of the X model representing the compact arrangement observed in the X-ray 1CK7 structure. Moreover, the D-1 configuration arises as the most likely configuration among the three compact models, attending to the combined G_{MMPB} and S_{config} calculations. It may also be interesting to note that the CAT domain is more accessible in the D-1 configuration than in D-2 or X as characterized by average solvent accessible surface area measurements (see Table S7 of the Supporting Information), thus suggesting that the placement of the HPX domain at the rear of the CAT domain and in simultaneous interaction with FIB could be favorable for catalysis. However, by taking into account the statistical uncertainty of the calculations, neither the X or D-2 models can be ruled out as they have a similar stability. Therefore, on the basis of the present calculations, it could well be that the MMP-2 enzyme would populate various interdomain arrangements.

A common feature of the X, D-1, and D-2 configurations is that the relatively long linker chain of residues is folded to accommodate a compact arrangement of the HPX and CAT-FIB domains. Could a fully extended conformation of the MMP-2 linker chain be stable in aqueous solution? This question is partially answered by the E simulation in which the initial structure has an elongated HPX...CAT-FIB disposition as the linker is effectively extended. This model was solvated in a very large solvent box to discard any potential artifact from periodic boundary conditions. During the first 25 ns of the E simulation, the elongated conformation was preserved, and subsequently, the linker chain folds back along 100 ns of simulation. As this dynamical evolution is not accompanied by any significant change in the combined potential and solvation energy of the MMP-2 protein, it seems likely that the E trajectory simply wanders through the most probable dynamical pathway of the linker region, which leads to its refolding and the formation of weak contacts between the HPX and CAT domains at the end of the trajectory. Interestingly the G_{MMPB} calculations, which account for solvent effects and protein–protein interactions, indicate that the final interdomain arrangement observed in the E trajectory turns out to be much less stable than those of the X, D-1, and D-2 configurations. This suggests that, by resuming the E trajectory, this model could eventually evolve toward a more stable interdomain arrangement. However, consideration of configurational entropy by means of the Schlitter's method warn us that entropy can largely stabilize the relative wandering of the HPX domain coupled with the unfolding/refolding of the linker chain. Taking into account that the interconversion between different compact

forms like those sampled along the X and D-1 trajectories would require the transition through more extended arrangements, it seems that configurational entropy could compensate the energetic penalty associated to the loss of protein–protein contacts. Unfortunately, since the Schlitter's method is known to systematically overestimate entropy effects and these calculations also depend on the length of the simulations, a quantitative assessment of the actual free energy difference between the elongated MMP-2 conformation and the more compact ones is still not feasible.

Another interesting result from our simulations is that the FIB modules adopt varying conformations in all of the simulations, as confirmed by the observed shifts in the Euler angles defining the relative orientation between the FIB modules and the CAT domain. The flexibility shown by the FIB domain in the different simulations could be relevant for the interaction of the MMP-2 enzyme with other proteins during protein–protein recognition processes. In fact the FIB domain of MMP-2 has been previously characterized as the collagen binding domain, and it seems crucial for the collagenolytic activity performed by this enzyme.^{18,57}

CONCLUSIONS

On the basis of our computational results, we conclude that it is very likely that, in addition to the compact interdomain disposition observed in the available crystal structures, alternative configurations as those generated by protein–protein docking calculations (D-1 and D-2) would also be structurally and energetically accessible in aqueous solution to the active form of the isolated MMP-2 enzyme. We also propose that other extended and flexible forms, which would be stabilized by entropic effects, could be involved in the interconversion among the compact structures. As mentioned in the Introduction, a similar picture has been obtained by experimental measurements performed for the MMP-12 enzyme, showing that about half of the conformations observed in solution are as compact as the solid state structure, but with different relative orientations of the CAT and HPX domains, while the rest of the conformations are more extended and some even highly extended.²⁸ Although the existence of these alternative configurations in which the HPX domain reorients around the CAT-FIB domains has still to be confirmed experimentally for the MMP-2 enzyme, the models presented in this work could also be useful for modeling protein assemblies involved in the regulation and/or the collagenolytic activity of the MMP-2.

ASSOCIATED CONTENT

Supporting Information

Figures and tables as explained here: Figure S1 shows a ribbon model of the starting structure for each trajectory. Table S1 and Figure S2 report the average and the time evolution rmsd values. Table S2 collect the average RMSF values. Figure S3–S7 display the time evolution of the interdomain distances and the polar plots of the Euler angles. Tables S3–S6 summarize the main interdomain contacts. Table S7 contain the average SASA values for the different domains. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Tel.: +34-985103468. Fax: +34-985103125. E-mail: diazfnatalia@uniovi.es.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by the CTQ2007-63266 (MEC, Spain) grant. The authors thankfully acknowledge the computer resources, technical expertise, and assistance provided by the Barcelona Supercomputing Center—Centro Nacional de Supercomputación.

■ REFERENCES

- (1) Nagase, H.; Woessner, J. F. J. *J. Biol. Chem.* **1999**, *274*, 21491–21494.
- (2) Steffensen, B.; Häkkinen, L.; Larjava, H. *Crit. Rev. Oral Biol. Med.* **2001**, *12*, 373–398.
- (3) López-Otin, C.; Matrisian, L. M. *Nat. Rev. Cancer* **2007**, *7*, 800–808.
- (4) Rodríguez, D.; Morrison, C. J.; Overall, C. M. *Biochim. Biophys. Acta* **2010**, *1803*, 39–54.
- (5) Butler, G. S.; Overall, C. M. *Biochemistry* **2009**, *48*, 10830–10845.
- (6) Overall, C. M. *Mol. Biotechnol.* **2002**, *22*, 51–86.
- (7) Bode, W.; Maskos, K. *Biol. Chem.* **2003**, *384*, 863–872.
- (8) Maskos, K. *Biochimie* **2005**, *87*, 249–263.
- (9) Li, J.; Brick, P.; O'Hare, M. C.; Skarzynski, T.; Lloyd, L. F.; Curry, V. A.; Clark, I. M.; Bigg, H. F.; Hazleman, B. L.; Cawston, T. E.; Blow, D. M. *Structure* **1995**, *3*, 541–549.
- (10) Piccard, H.; Van den Steen, P. E.; Opdenakker, G. *J. Leukocyte Biol.* **2007**, *81*, 870–892.
- (11) Bode, W. *Structure* **1995**, *3*, 527–530.
- (12) Murphy, G.; Knäuper, V. *Matrix Biol.* **1997**, *15*, 511–518.
- (13) Knäuper, V.; Docherty, A. J.; Smith, B.; Tschesche, H.; Murphy, G. *FEBS Lett.* **1997**, *405*, 60–64.
- (14) Morgunova, E.; Tuuttila, A.; Bergmann, U.; Isupov, M.; Lindqvist, Y.; Schneider, G.; Tryggvason, K. *Science* **1999**, *284*, 1667–1670.
- (15) Steffensen, B.; Wallon, U. M.; Overall, C. M. *J. Biol. Chem.* **1995**, *270*, 11555–11566.
- (16) Gehrmann, M. L.; Douglas, J. T.; Bányai, L.; Hedvig, T.; Patthy, L.; Llinás, M. *J. Biol. Chem.* **2004**, *279*, 46921–46929.
- (17) Xu, X.; Wang, Y.; Lauer-Fields, J. L.; Fields, G. B.; Steffensen, B. *Matrix Biol.* **2004**, *23*, 171–181.
- (18) Xu, X.; Mikhailova, M.; Llangovan, U.; Chen, Z.; Yu, A.; Pal, S.; Hinck, A. P.; Bjorn Steffensen, B. *Biochemistry* **2009**, *48*, 5822–5831.
- (19) Jozic, D.; Bourenkov, G.; Lim, N.-H.; Visse, R.; Nagase, H.; Bode, W.; Maskos, K. *J. Biol. Chem.* **2005**, *280*, 9578–9585.
- (20) Bertini, I.; Fragai, M.; Luchinat, C.; Melikyan, M.; Mylonas, E.; Sarti, N.; Svergun, D. I. *J. Biol. Chem.* **2009**, *284*, 12821–12828.
- (21) Chung, L.; Dinakaran, D.; Yoshida, N.; Lauer-Fields, J. L.; Fields, G. B.; Visse, R.; Nagase, H. *EMBO J.* **2004**, *23*, 3020–3030.
- (22) Overall, C. M.; Butler, G. S. *Structure* **2007**, *15*, 1159–1161.
- (23) Rosenblum, G.; Van den Steen, P. E.; Cohen, S. R.; Bitler, A.; Opdenakker, G.; Sagi, I. *PLoS ONE* **2010**, *5*, e11043.
- (24) Saffarian, S.; Collier, I. E.; Marmer, B. L.; Elson, E. L.; Goldberg, G. *Science* **2004**, *206*, 108–111.
- (25) Collier, I. E.; Saffarian, S.; Marmer, B. L.; Elson, E. L.; Goldberg, G. *Biophys. J.* **2001**, *81*, 2370–2377.
- (26) Iyer, S.; Visse, R.; Nagase, H.; Acharya, K. R. *J. Mol. Biol.* **2006**, *362*, 78–88.
- (27) Rosenblum, G.; Van den Steen, P. E.; Cohen, S. R.; Grossman, J. G.; Frenkel, J.; Sertchook, R.; Slack, N.; Strange, R. W.; Opdenakker, G.; Sagi, I. *Structure* **2007**, *15*, 1227–1236.
- (28) Bertini, I.; Calderone, V.; Fragai, M.; Jaiswal, R.; Luchinat, C.; Melikyan, M.; Mylonas, E.; Svergun, D. I. *J. Am. Chem. Soc.* **2008**, *130*, 7011–7021.
- (29) Remacle, A. G.; Shiryayev, S. A.; Radichev, I. A.; Rozanov, D. V.; Stec, B.; Strongin, A. Y. *J. Biol. Chem.* **2011**, *286*, 21002–21012.
- (30) Díaz, N.; Suárez, D.; Valdés, H. *J. Am. Chem. Soc.* **2008**, *130*, 14070–14071.
- (31) Valdés, H.; Díaz, N.; Suárez, D.; Fernández-Recio, J. *J. Chem. Theory Comput.* **2010**, *6*, 2204–2213.
- (32) Díaz, N.; Suárez, D. *Proteins: Struct., Funct., Bioinf.* **2008**, *72*, 50–61.
- (33) Canutescu, A. A.; Dunbrack, R. L. Jr. *Bioinformatics* **2005**, *21*, 2914–2916.
- (34) Fernández-Recio, J.; Totrov, M.; Skorodumov, C.; Abagyan, R. *Proteins* **2005**, *58*, 134–143.
- (35) Fernández-Recio, J.; Totrov, M.; Abagyan, R. *J. Mol. Biol.* **2004**, *335*, 843–865.
- (36) Cheng, T. M.; Blundell, T. L.; Fernández-Recio, J. *Proteins* **2007**, *68*, 503–515.
- (37) Cheng, T. M.; Blundell, T. L.; Fernández-Recio, J. *BMC Bioinf.* **2008**, *9*, 441–453.
- (38) Tallant, C.; Marrero, A.; Gomis-Rüth, F. X. *Biochim. Biophys. Acta* **2010**, *1803*, 20–28.
- (39) Lowry, C. L.; McGeehan, G.; LeVine, H. *Proteins: Struct., Funct., Genet.* **1992**, *12*, 42–48.
- (40) Díaz, N.; Suárez, D. *Biochemistry* **2007**, *46*, 8943–8952.
- (41) Díaz, N.; Suárez, D. *J. Phys. Chem. B* **2006**, *110*, 24222–24230.
- (42) Zhang, L.; Hermans, J. *Proteins: Struct., Funct., Genet.* **1996**, *24*, 433–438.
- (43) Duan, Y.; Wu, C.; Zhang, W.; Chowdhury, S.; Yang, R.; Caldwell, J.; Cieplak, P.; Wang, J.; Lee, M.; Luo, R.; Xiong, G.; Lee, T.; Kollman, P. J. *Comput. Chem.* **2003**, *24*, 1999–2012.
- (44) Åqvist, J. *J. Phys. Chem. B* **1990**, *94*, 8021–8024.
- (45) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (46) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. *Acc. Chem. Res.* **2000**, *33*, 889–897.
- (47) Gohlke, H.; Case, D. A. *J. Comput. Chem.* **2003**, *25*, 238–250.
- (48) Sharp, K.; Honig, B. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *19*, 301–332.
- (49) Schlitter, J. *J. Chem. Phys. Lett.* **1993**, *215*, 617–621.
- (50) Schäfer, H.; Mark, A. E.; van Gunsteren, W. F. *J. Chem. Phys.* **2000**, *113*, 7809–7817.
- (51) Schäfer, H.; Daura, X.; Mark, A. E.; van Gunsteren, W. F. *Proteins: Struct., Funct., Genet.* **2001**, *43*, 45–56.
- (52) Harris, S. A.; Gavathiotis, E.; Searle, M. S.; Orozco, M.; Loughton, C. A. *J. Am. Chem. Soc.* **2001**, *123*, 12658–12663.
- (53) Suárez, E.; Díaz, N.; Suárez, D. *J. Chem. Theory Comput.* **2011**, *7*, 2638–2653.
- (54) Skjaerven, L.; Grant, B.; Muga, A.; Teigen, K.; McCammon, J. A.; Reuter, N.; Martinez, A. *PLoS ONE* **2011**, *7*, e1002004.
- (55) Caves, L. S. D.; Evanseck, J. D.; Karplus, M. *Protein Sci.* **1998**, *7*, 649–666.
- (56) Abrams, C. F.; Vanden-Eijnden, E. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 4961–4966.
- (57) Gioia, M.; Monaco, S.; Fasciglione, G. F.; Coletti, A.; Modesti, A.; Marini, S.; Coletta, M. *J. Mol. Biol.* **2007**, *368*, 1101–1113.