JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

The Large Scale Conformational Change of the Human DPP III–Substrate Prefers the "Closed" Form

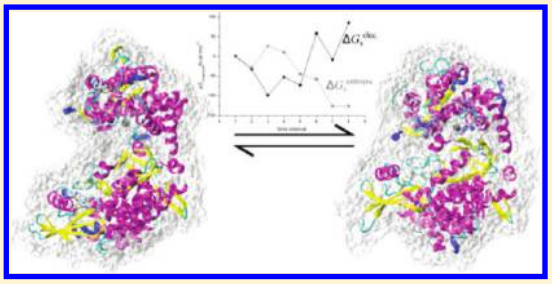
Antonija Tomić, Miguel González, and Sanja Tomić*

Division of Physical Chemistry, Ruđer Bošković Institute, Bijenička 54, 10 000 Zagreb, Croatia

Departament de Química Física and IQTC, Universitat de Barcelona, C/Martí i Franquès, 1, 08028, Barcelona, Spain

Supporting Information

ABSTRACT: Human dipeptidyl peptidase III (DPP III) is a two domain metallo-peptidase from the M49 family. The wide interdomain cleft and broad substrate specificity suggest that this enzyme could experience significant conformational change. Long (>100 ns) molecular dynamics (MD) simulations of DPP III revealed large range conformational changes of the protein, suggesting the *pre-existing equilibrium* model for a substrate binding. The binding free energy calculations revealed tighter binding of the preferred synthetic substrate Arg–Arg–2–naphtylamide to the "closed" than to the "open" DPP III conformation. Our assumption that Asp372 plays a crucial role in the large scale interdomain closure was proved by the MD simulations of the Asp372Ala variant. During the same simulation time, the variant remained more "open" than the wild type protein. Apparently, Ala was not as efficient as Asp in establishing the interdomain interactions. According to the MM–PBSA calculations, the electrostatic component of the free energy of solvation turned out to be higher for the "closed" protein than for its less compact form. However, the gain in entropy due to water released from the interdomain cleft nicely balanced this negative effect.



INTRODUCTION

The dipeptidyl peptidase III is a zinc–exopeptidase that hydrolyzes dipeptide from the N terminus of its substrates, which consist of three or more amino acid residues. Its broad specificity toward peptides of varying lengths and compositions still has not been properly understood. According to the MEROPS database, it belongs to the M49 peptidase family, which is distributed over four kingdoms of organisms: eubacteria, protists, fungi, and animals.[1] Besides its role in intracellular protein catabolism, there is evidence for its participation in pain modulation and defense against oxidative stress.[2−4]

Crystal structures of the yeast C130S mutant and human DPP III (PDB codes 3CSK and 3FVY, respectively) have been determined. Although their sequence identity is only about 35%, their 3D structures, consisting of two domains forming a cleft, are very similar, with RMSD = 2.17 Å (calculated for 676 amino acids residues).[5] Despite the striking similarity of the two proteins' 3D structures, their active site positions (close to the interdomain cleft), and the Zn^{2+} coordination geometries, the catalytic activity of the yeast DPP III is, in comparison to its human counterpart, reduced. The aim of our study was to determine possible reasons for the broad substrate specificity and catalytic efficiency of the human DPP III.

The presence of the wide interdomain cleft, together with large enzyme promiscuity, suggests that the human DPP III could experience significant dynamics fluctuations and internal motions.

There are numerous examples showing the direct correlation between the internal motion of an enzyme and its activity.[6−12] For example, the catalytic activities of many enzymes are closely associated with the motion of a loop that opens and closes the active site and eventually positions key residues in contact with the bound ligand. Not so long ago, the concept of conformational substates was not only preserved for an enzyme–substrate complex, yet there are numerous examples indicating that conformational exchange exists in the free enzyme as well.[13] The so-called *pre-existing equilibrium* model assumes interconversion of a protein within an ensemble of conformations.[14] The ligand selectively binds to one of them, shifting the pre-existing equilibrium toward the binding conformation. This concept differs from the *induced fit* model, where substrate binding induces changes in the protein structure to bring the amino acid residues from the binding pocket in an orientation which enables proper binding.[15]

In order to elucidate reasons for its broad substrate specificity and efficient catalytic performance, we studied the behavior of human DPP III during its simulation in explicit water and determined the relative stability of the distinct conformations that the protein has adopted during the simulation. In order to determine the role of Asp372 in the large scale motion of its domains, the Asp372Ala variant was also studied by molecular dynamics simulations. Further on, we calculated the binding free energies for the preferred synthetic substrate Arg–Arg–2–

Received: March 14, 2012
Published: June 1, 2012

dx.doi.org/10.1021/ci300141k | *J. Chem. Inf. Model.* 2012, 52, 1583–1594

naphtylamide (RRNA) binding to both the "closed" and "open" DPP III conformations.

Just very recently, after submission of our publication, the X-ray structure (PDB code 3T6B) of human DPP III bound to the opioid peptide tynorphyn appeared in the Protein Data Bank. The experimentally determined structure of the complex revealed a fully "closed" conformation of DPP III, differing from the experimental structure of the free protein (PDB code 3FVY). In this paper, we compare the available experimental results with the results of our calculations (which showed that in solution DPP III exists in "open" and "closed" conformations) and propose an explanation for protein behavior in solution.

COMPUTATIONAL METHODS

System Preparation. The recently determined crystal structure of the ligand-free human dipeptidyl peptidase III (DPP III) was extracted from the Brookhaven Protein Databank (ID: 3FVY). The crystal structure of the protein consists of a polypeptide and a zinc atom in the active site. The amino acid residues missing in the crystal structure, Pro224–Asp227, were added, and the shape of this region was modeled using the program Modeller9v2.[16] Arg and Lys residues were positively charged (+1), and all Glu and Asp residues were negatively charged (−1), as expected under physiological conditions. The initial protonation of the protein was determined by the WHATIF-server (at pH 7.0; http://swift.cmbi.ru.nl/servers/html/index.html), and protonation of His residues was checked according to their ability to form hydrogen bonds with adjacent amino acid residues. The ff03 force field of Duan et al.[17] was used throughout the molecular modeling study performed with the AMBER10[18,19] and GROMACS 4.5.3[20−24] packages. The nonbonding Zn^{2+} parameters, as described in our previous work[25] and modified according to a PDB survey, were used.[26]

Simulations Performed within the AMBER10 Suite. The ligand-free protein structure as well as its complex with the synthetic substrate Arg–Arg–2–napthyamide (RRNA) were placed in the truncated octahedron filled with TIP3P water molecules. In order to neutralize the system, the sodium ions were placed in the vicinity of the negatively charged amino acid residues at the protein surface. The resulting system, consisting of ∼98 000 atoms (∼29 000 molecules of water), was simulated using periodic boundary conditions. The electrostatic interactions were calculated using the particle-mesh Ewald method.[27,28]

Prior to molecular dynamics (MD) simulations, the protein geometry was optimized in three cycles with different constraints. In the first cycle (1500 steps), only water molecules were relaxed, while the protein and zinc atoms were constrained using the harmonic potential with a force constant of 32 kcal/(mol Å^2). In the second (2500 steps) and the third cycle (1500 steps), the same force was applied to the zinc atom while the protein backbone was constrained with 32 and 10 kcal/(molÅ^2), respectively. During the first period of equilibration (100 ps of gentle heating from 0 to 300 K), the *NVT* ensemble was used, while all of the following simulations (water density adjustment and productive simulations) were performed at constant temperature and pressure (300 K and 1 atm, the *NPT* ensemble). The temperature was held constant using Langevin dynamics with a collision frequency of 1 ps^{-1}. Bonds involving hydrogen atoms were constrained using the SHAKE algorithm.

Two different DPP III–RRNA complexes were considered; in one of them, the substrate was bound to the "open" (initial) conformation of the enzyme, and in the other one it was bound to the "closed" (obtained after 72 ns of MD simulations) enzyme form. The procedure of substrate docking was described in our previous publication (see Supporting Information, Technical Details).[29] Both complexes, were simulated for 30 ns with a time step of 1 fs, and structures were sampled every 1 ps.

After 50 and 71 ns of MD simulation, in order to enable a more efficient conformational space search, the temperature of the ligand-free DPP III was elevated to 400 K, followed by cooling back to 300 K. Both the heating and cooling intervals were 500-ps-long, and the time step was 1 fs. In between, the system was simulated at room temperature (*NPT*, 300 K and 1 atm, time step 2 fs). Finally, 29 ns of the MD simulation at 300 K were carried out using the *NPT* ensemble and a time step of 2 fs. Altogether, the MD simulation resulted with the 101-ns-long trajectory, with structures sampled every 1 ps.

Simulations Performed within GROMACS 4.5.3 Suite. The ligand-free DPP III and its D372A mutant were placed in the dodecahedron box filled with TIP3P water molecules, and Na^+ ions were added in order to neutralize the systems. The resulting systems, consisting of ∼98 000 atoms (∼29 000 molecules of water), were simulated using periodic boundary conditions. The electrostatic interactions were calculated using the particle-mesh Ewald method.[27,28]

Prior to MD simulations, the protein geometry was optimized using a steepest descent algorithm. The MD simulations were performed at constant temperature and pressure (300 K and 1 bar, the *NPT* ensemble) using the velocity-Verlet algorithm[30] and a time step of 1 fs. The temperature was held constant using Nosé–Hoover temperature coupling[31,32] with a coupling time constant of 1 ps. For isotropic pressure control, the Martyna–Tuckerman–Tobias–Klein (MTTK)[33] procedure, with a coupling constant of 1 ps, was used. After 50 ns of *NPT* simulations, with protein structures and energies sampled every 1 ps, the system with the D372A mutant was heated to 400 K (500 ps), followed by 500 ps of simulations, during which the temperature of the system decreased to 300 K. The final 20 ns of the MD simulations were carried out using the *NPT* ensemble and a time step of 1 fs. Altogether, MD simulations resulted with the 50- and 71-ns-long trajectories for the wild type and D372A mutant, respectively.

The Relative Free Energy Calculations. *Ligand–Free DPP III.* Several methods were applied in order to evaluate the stability of the conformations that the protein had adopted during the simulation.

In order to determine the change in Gibbs potential energy while the protein changes from the "open" to the "closed" conformation, we performed Molecular Mechanics Poisson–Boltzmann Surface Area (MM–PBSA) calculations[34] on eight 2-ns-long trajectories representing clusters of DPP III structures with similar radii of gyration. Each cluster consisted of 100 structures extracted every 20 ps from the MD trajectory.

The enthalpic part of the Gibbs free energy, approximated with the sum of nonbonding conformational energy and the solvation free energy, was calculated using the MM–PBSA approach as implemented in the AMBER10 suit of programs with a concentration of the singly charged counterions of 0.05 M. The polar component of solvation enthalpy was calculated using the Poisson–Boltzmann method, and the nonpolar

component was determined by $\Delta H_{nonpol} = \gamma SASA + \beta$, where the solvent accessible surface area (SASA) was calculated with the MolSurf program.[35] The surface tension γ and the offset β were set to the standard values of 5.42×10^{-3} kcal/(mol Å^2) and 0.92 kcal/mol, respectively.[36] The calculations were accomplished for the enzyme immersed into the solvent utilizing three different solute dielectric constants, 1.0, 2.0, and 4.0.

Entropy was estimated considering both the solvent and the enzyme contribution. Change in the enzyme conformational entropy was calculated using the normal-mode analysis as implemented in AMBER11 (module nmode). Prior to the gas phase entropy calculation, structures were minimized using the NAB (Nucleic Acid Builder) implementation of the nmode program that enables calculations of the electrostatic solvation energies by the Generalized Born (GB) method. Since the normal-mode analysis is computationally expensive, we only considered the residues lining the enzyme interdomain cleft, as a part of the enzyme that experienced the greatest environmental change upon the enzyme conformational modification. This part is characterized with the largest number of new contacts established during the protein closure, as well as with the largest solvent accessible surface area (SASA) changes. The SASA was determined by the programs Naccess V2.1.1.1[37] and PSAIA,[38] using a probe with the same radius as water (1.4 Å).

Finally, change of the entropic part of the solvent free energy was estimated from the number of water molecules in the protein first and second solvation shells.

Complexes with RRNA. The enthalpic part of the binding free energy was calculated using the MM−PBSA approach as implemented in the AMBER10 suite of programs. The calculations were performed for both complexes, with the enzyme "open" and "closed" conformations (see above). For each, three 5-ns-long sections of their trajectories were considered. The conditions described in our previous publication[29] were used.

Principal Component Analysis. Principal component analysis (PCA), a method that extracts the dominant, concentrated atomic motions during the simulation, was used in order to reduce the dimensionality of the data obtained from molecular dynamics simulation of the ligand-free DPP III and its complexes with the substrate. The covariance matrix $3N \times 3N$ (N being the number of C, Cα, N, and O atoms) generated from the trajectory using the module ptraj from the AMBER10 suite of programs was diagonalized, and the eigenvectors and eigenvalues were determined. All structures were aligned to the initial one to eliminate translations and rotations of the geometric center. In order to obtain the contribution of atomic vibrations to a particular mode, the MD trajectory was projected on the particular eigenvector and time-averaged.

Transport of Water Molecules and Substrate through the Protein. To explore the pathways for the water molecule entrance/exit from the enzyme active site, two methods were applied.

CAVER 2.0 v 0.003 software,[39,40] as a plugin for PyMol, was used for the identification and visualization of tunnels in the protein structure obtained after 101 ns of MD simulation starting from the zinc ion to the protein exterior. However, a protein is not a rigid body, and the presence of water/substrate molecules further increases its flexibility. In order to account for the protein's adaptability to the trapped water molecules, we performed random acceleration molecular dynamic (RAMD)[41−43] simulations for the DPP III system obtained after 101 ns of MD simulations. RAMD simulations were run for 250 ps, or until the distance between the enzyme center of mass and that of the water molecule became larger than 30 Å. The acceleration of the force applied to the water center of mass was 0.6 kcal/(g Å), and the time step was 1 fs. The direction of the force was kept constant for 40 time steps. If during this period a water molecule did not move more than 0.01 Å, a new direction was chosen randomly; otherwise, the same force was applied for the next 40-time-step period.

To explore the possible pathways for the substrate, RRNA, trafficking through the protein, we performed RAMD calculations using a random force acceleration of 0.1 kcal/(g Å). The WT_{72}−RRNA structure obtained after 15 ns of MD simulations was used as a starting structure for these RAMD simulations.

RESULTS AND DISCUSSION

MD Simulations of the Ligand-Free Protein. During the 101 ns of MD simulations performed using the AMBER10 programs suite, the enzyme tertiary structure changed considerably (see Figure 1). Apparently, as a result of interdomain movement, the protein globularity significantly increased (see Figure 2). Similar changes were noticed during 50 ns of simulations by GROMACS (see Figure S1, Supporting Information). A possible explanation why the ligand-free DPP III is exclusively in the "open" form in the crystal structure arises from the way the molecules are packed (see Supporting Information, Figure S2).

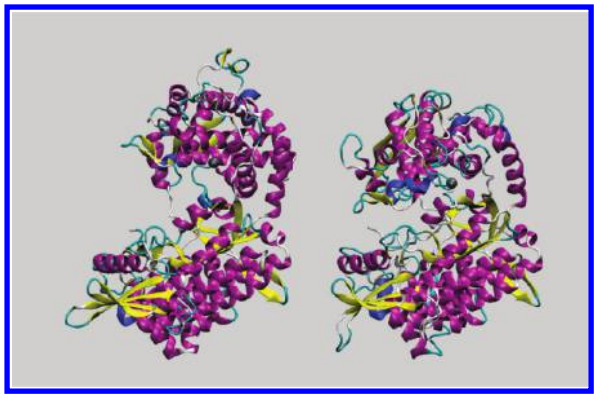


Figure 1. Initial (X-ray) structure of the human DPP III (left) and the conformation obtained after 101 ns of the MD simulation in water (right). Orientation of the lower domain was held fixed. The zinc cation is shown as a gray VdW sphere.

As a consequence, the active site became less accessible, and the number of water molecules occupying the 10 Å sphere around the central zinc ion was reduced to more than half of its initial value (Figure 3 and Figure S3, Supporting Information). However, the "upper" and "lower" domain structures (colored red and blue in Figure 4, respectively) have not changed significantly while approaching each other. Also, the overall secondary structure of the protein has not changed significantly (see Table S1, Supporting Information). RMSD values of these two domains plotted versus time indicate their stabilization at the very beginning of the simulation. The region comprising 34 amino acid residues from the "upper" domain and mostly made of turns (colored black and yellow in Figure 4 and Figure S2 in Supporting Information, respectively) experienced the largest changes during the simulation.

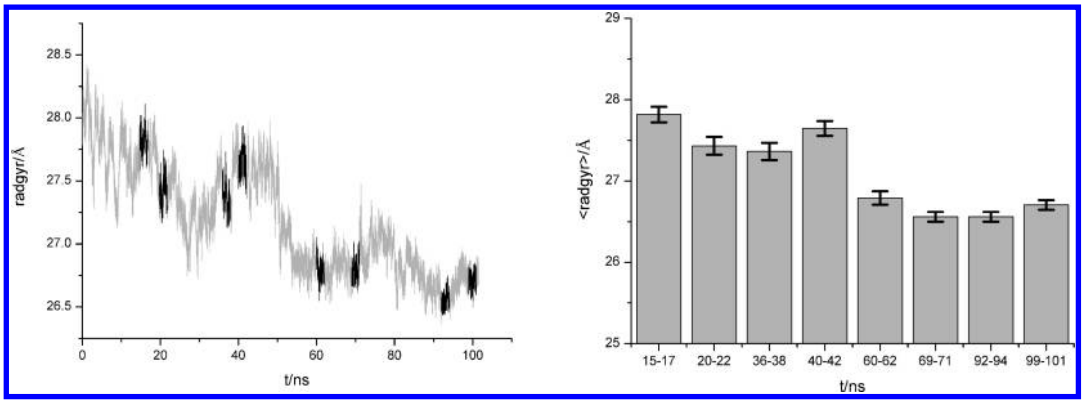


Figure 2. Eight clusters of the protein structures (colored black), selected from the 101-ns-long MD trajectory, for the MM−PBSA calculations (left). The selection was based on the radius of gyration values. Their average values and standard deviations for each of the selected clusters are shown on the plot to the right.

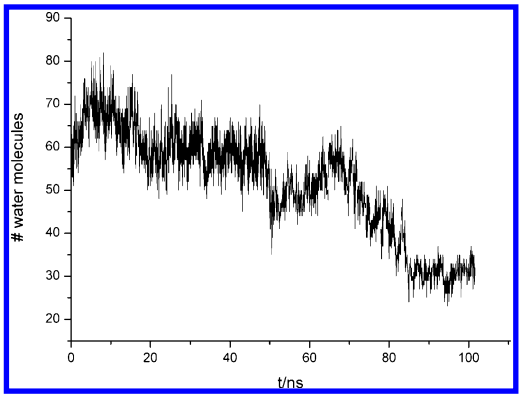


Figure 3. Number of the water molecules occupying the 10 Å sphere around the zinc ion during 101 ns of the MD simulation.

The "lower" to "upper" domain movement is significantly pronounced at the regions close to the cleft edge (regions colored red in Figure 5, bottom). Especially interesting is the translocation of Tyr318 (part of loop1 in Figure 5, bottom), the functional role of which had been experimentally proven.[44] Namely, the Tyr318 to Phe substitution resulted in a significant (2 orders of magnitude) decrease of the DPP III catalytic efficiency of the Arg−Arg−2−naphtylamide hydrolysis. During the MD simulation, the distance between its Cα atom and the zinc ion decreased from 26 to 16 Å (see Figure S4, Supporting Information). This movement was accompanied by its hydrogen bonding to Arg565, followed by formation of the CH−π interactions with Phe556, both from the protein "upper" domain. These interactions probably helped in consolidation of the active site and in stabilization of the transition state. As a result of the interdomain motion, several additional hydrogen bonds were established connecting the residues from the "upper" and "lower" domains: Ser504−Asp396, Ser500−Gln400, Gln566−Glu316, and Arg565−Pro321 (written as "upper"−"lower" domain H-bonded amino acid residue pair, respectively).

The B-factor analysis, shown in Figure 5 (top), revealed that residues located at the more distant part of the interdomain cleft are more flexible than those lining the cleft (Figure 5, bottom), where only three short regions, namely, Phe368 to Asp372, Ala414 to Thr419, and Glu667 to Leu671, exhibited larger shifts. It should be noticed that the amino acid residues from the Ala414−Thr419 region are missing in the recently determined structure of DPP III−tynorphyn (PDB code:

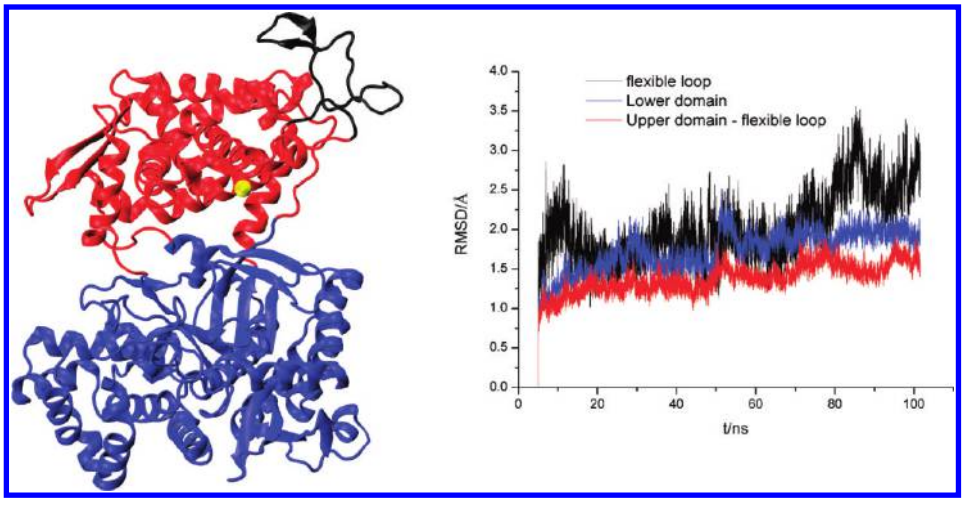


Figure 4. Parts of the DPP III (shown at the initial DPP III structure), left, and their RMSD profiles (colored accordingly), right. The "lower" domain (residues 4 to 335, 374 to 416, and 670 to 726) is colored blue, the majority of the "upper" domain (residues 336 to 373, 417 to 458, and 493 to 669) red, and the flexible loop belonging to the "upper" domain (residues 459 to 492) black. The DPP III structure obtained after 5 ns of MD simulations was used as a reference structure for RMSD analysis. The zinc cation is shown as a yellow sphere.

dx.doi.org/10.1021/ci300141k | J. Chem. Inf. Model. 2012, 52, 1583−1594

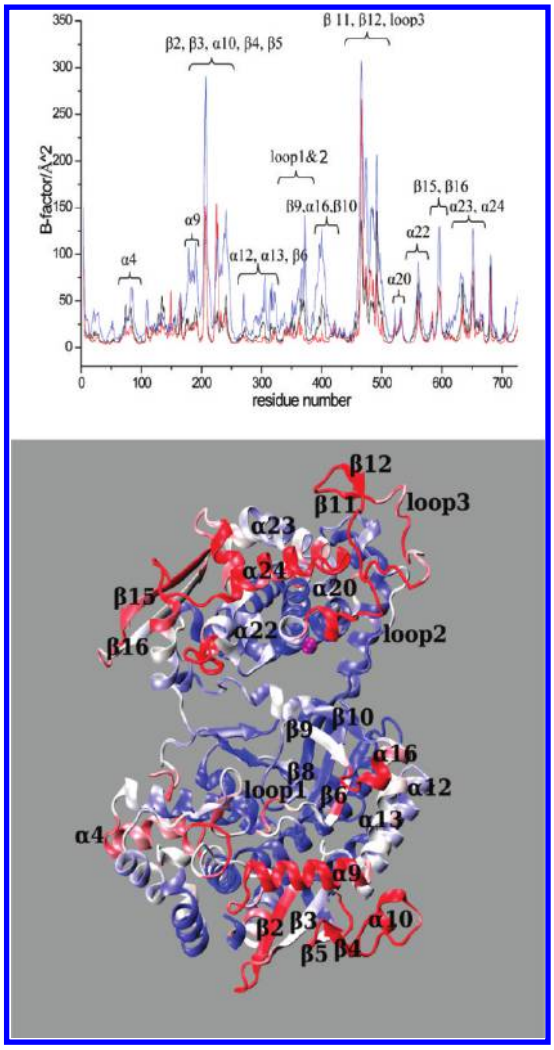


Figure 5. Top: Residue based B-factors, averaged over the DPP III backbone atoms, calculated for three segments of the MD trajectory; from the 1st to 3rd ns, from the 15th to 17th ns, and from the 92nd to 94th ns, colored blue, black, and red, respectively. Bottom: DPP III structure colored according to the B-factor profile calculated for the entire 101-ns-long MD trajectory. Values more than one standard deviation above and below the average are colored red and blue, respectively. Values in between are colored white. Except loop tags, elements of the secondary structure are derived by DSSP (as implemented at the WHATIF-server, http://swift.cmbi.ru.nl/servers/html/index.html). The zinc cation is shown as a purple sphere.

3T6Y). In order to elucidate how the long-range conformational changes of the enzyme influenced the local flexibility, B-factors were calculated for three, separated, 2-ns-long segments of the 101-ns-long MD trajectory. Apparently, during the protein transformation from the "open" to "closed" conformation, flexibility of the most flexible protein regions notably decreased.

A comparison between the calculated and the experimental (PDB code 3FVY) B-factors revealed a very good agreement (Figure S5 in Supporting Information). The largest discrepancy occurs in the most flexible regions, which correspond to the nonstructured (loops) parts of the protein. These regions are probably more rigid in the crystal than in solution due to the crystal packing effect. Indeed, in the crystal environment, some of these regions, like for example the 470−500 loop, are placed in close contact with the symmetry related molecule (Figure S2, Supporting Information).

Normal Mode Analysis. Principal component analysis of the 101-ns-long MD trajectory revealed that 81% of the protein motion could be explained by only two normal modes (see Figure 6). The three largest normal modes describe 85% of the protein motion.

The most prominent motion, described by the top three eigenvectors, corresponds to the protein domain's movement (Figure 7). The first eigenvector, accounting for 70% of the overall protein motion, describes the enzyme closure (see Figure 7a). Narrowing of the enzyme active site, either by parallel shift of the protein domains in opposite directions or by simultaneous approaching and moving apart of their opposite ends of interdomain cleft, is described by the next two eigenvectors (Figures 7b and c, respectively). The main interactions governing these large domain motions are the hydrogen bond formed between Asp372 and Lys405 accompanied by the hydrogen bond and CH−π interactions between Tyr318 and Arg565 and Phe556, respectively (Figure S4, Supporting Information). It is interesting to note that Tyr318 and Lys405 are absolutely conserved in the M49 family.

In order to further rationalize the results of our MD simulations, we performed an additional normal-mode analysis using the Delarue group server for normal-mode analysis (http://lorentz.dynstr.pasteur.fr/). We compared the first three modes determined from the MD results (i.e., those modes with the lowest frequency and the largest amplitude of oscillation) with those obtained from the server and found that they are equivalent (Figure S6 in the Supporting Information).

DPP III Conformational Change Aided by Asp372. Closer insight into the protein structures sampled during the MD simulations revealed that the hydrogen bonds formed between Asn406 and Asp372 and Glu451, at the very beginning of the MD simulation (immediately after the equilibration), aided the enzyme conformational change. These hydrogen bonds enabled the movement of Asp372 toward the "lower" domain and its hydrogen bonding to Lys405 (see Figure S4, Supporting Information) and frequently with Lys304, inducing formation of the hydrogen bonds between Lys405 and Thr374 and between Glu508 and Asn391, altogether resulting in enzyme closure (see Figure 8 and Figure S7 in Supporting Information) as well as in active site shrinkage. As mentioned above, several other hydrogen bonds as well as the CH−π interactions, between the amino acid residues more distant from the active site, e.g., the interaction between Tyr318 and Arg565 and Phe556, have been established simultaneously. As a consequence, α13 (part of which is absolutely conserved Trp300) and α16 helices from the protein "lower" domain approached the "upper" one, resulting in moving the "lower" domain five-stranded β structure (β6, β7, β8, β9, and β10 in Figure 5, bottom), part of which is Lys405, deep into the interdomain cleft.

Protein Accessibility to Solvent. The effect of the protein conformational change on the solvent accessibility was evaluated by measuring the change of the residue based solvent accessible surface area (SASA). The difference between the SASA calculated for the initial DPP III structure and the one obtained after 101 ns of MD simulations (Figure S8, Supporting Information) revealed that the binding site and the interdomain cleft accessibility to the solvent decreased while the accessibility of the protein "back", convex side, increased during the simulation.

dx.doi.org/10.1021/ci300141k | J. Chem. Inf. Model. 2012, 52, 1583−1594

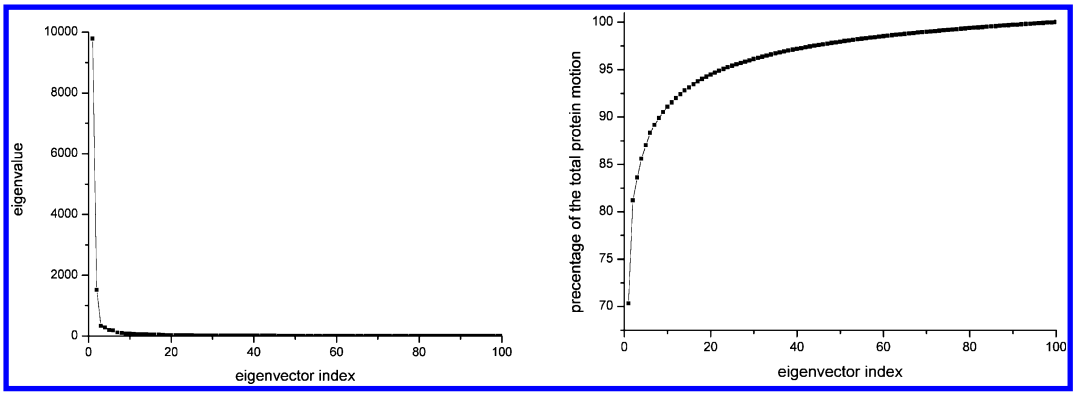


Figure 6. Eigenvalue individual values (left) and percentage of the protein motion (right) as a function of the number of eigenvectors considered. Eigenvalues are derived by the principal component analyses of the 101-ns-long MD trajectory of the ligand-free DPP III. The concerted motion specified by the first two eigenvectors accounts for ∼81% of the overall motion.

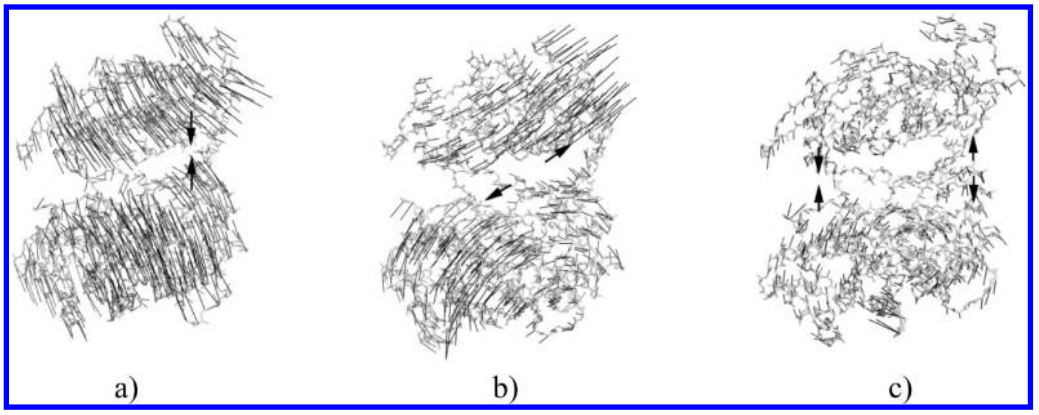


Figure 7. Three of the most important principal components (with the eigenvalues decreasing from a to c) derived from the principal component analysis of the 101-ns-long trajectory of the ligand-free DPP III. The black lines attached to each Cα atom indicate the eigenvector direction and magnitude of its eigenvalue. The black, thick arrows indicate the direction of the concentrated motion.

In order to elucidate the solvent accessibility of the active site in the "closed"" enzyme conformation, possible access routes connecting the zinc cation and external solvent environment were determined. Using the software CAVER, four main water tunnels were found (see Figure 9). The entrance of the shortest (∼21 Å long), blue tunnel, defined by residues Ile301, Lys304, Leu369, Thr370, and Asp372, is about 3.4 Å in diameter, while the entrance of the longest (∼43 Å long), red tunnel (defined by Phe109, Thr112, Tyr318, Glu555, Leu663, and Lys666), is only ∼1.1 Å in diameter and has a bottleneck shape. However, since this part of the enzyme is quite flexible (according to the B-factor analyses), the red tunnel can easily stretch to enable the entrance of a substrate as well.

Three of these four tunnels, blue, red, and yellow, were confirmed by RAMD simulations. To exit the enzyme active site, the water molecule mostly used the blue tunnel (72%). The red and yellow tunnels were used evenly.

During the simulation, Zn^{2+} was, besides with His450, His455, Glu451, and Glu508, coordinated by two water molecules. Altogether, five water molecules exchanged in the Zn^{2+} coordination sphere during the 101 ns of the MD simulation. However, never were both water molecules coordinating Zn^{2+} replaced at the same time (as can be seen in Figure 10). With the water molecules in the bulk, they interchanged through the yellow and blue channels (with even frequencies). This process was much faster at the beginning of the simulation than during the last 50 ns. A reason could be a narrowing of the entrance to the enzyme active site.

Near the zinc ion, the yellow tunnel (with the widest entrance; ∼5 Å in diameter) has a gorge lined by Asn391, His568, and Glu508, i.e., residues constituting the enzyme binding site.

System Free Energy Differences. In order to elucidate the free energy differences between the solvated "open" and "closed" protein conformations, we performed MM−PBSA calculations for the selected clusters (Figure 2).

While the intramolecular electrostatic and van der Waals energies, as well as the nonpolar component of the free energy of solvation, did not change significantly during the simulation, the electrostatic component of the solvation free energy slightly increased during the MD simulations (Table S2, Supporting Information), suggesting that the "open" conformation would be preferred in water. However, the increase of the water entropy compensated for this effect (Figure 11). Namely, the number of water molecules occupying the protein first and second solvation spheres reduced by more than 100 (Table S2, Supporting Information, last column). The difference is mostly the result of water extrusion from the interdomain cleft during the protein closure (Figure 3). Water molecules far from the protein surface (water molecules that do not belong to the first two protein solvation spheres) have larger entropy than structured water molecules close to the protein surface. Energy values for a water molecule release, quoted in the literature,[45,46] range from 1.2 kcal mol^{-1} to 2.3 kcal mol^{-1}. This implies that the energy obtained by release of the water molecules from the protein first and second solvation shells, as the enzyme

dx.doi.org/10.1021/ci300141k | J. Chem. Inf. Model. 2012, 52, 1583−1594

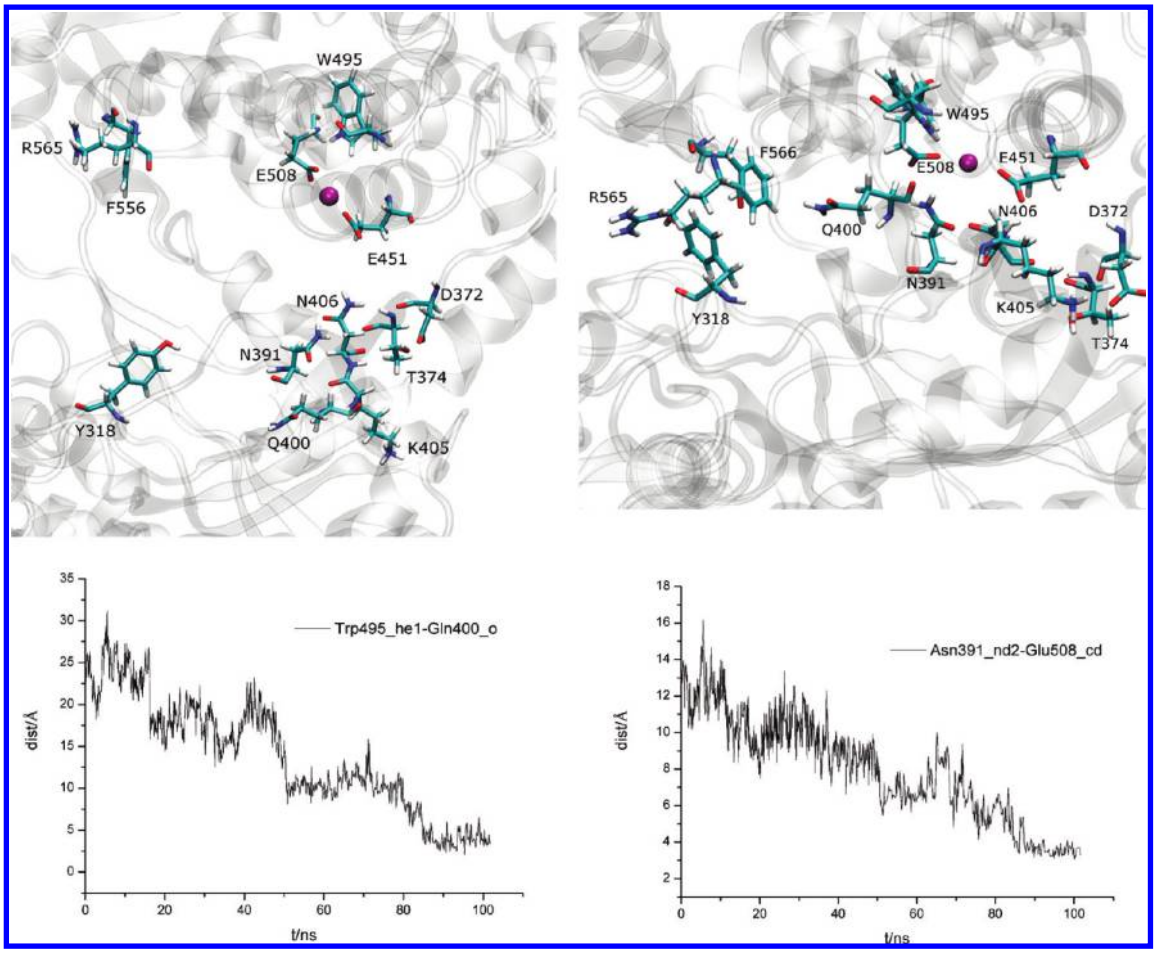


Figure 8. The DPP III binding site before and after 101 ns of MD simulations, top left and right, respectively. Closure of the enzyme binding site can be traced by monitoring distances between the amino acid residues on the cleft edge, Trp495−Gln400 and Asn391−Glu508, bottom left and right, respectively.

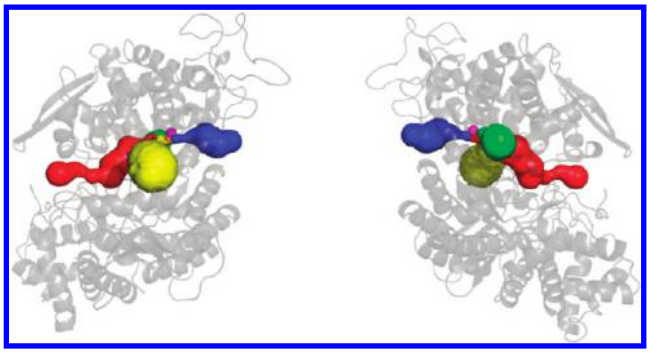


Figure 9. Four main tunnels, colored blue, red, green, and yellow, calculated by the program CAVER for the structure obtained after 101 ns of MD simulation (two different orientations are shown). The zinc cation is shown as a purple sphere.

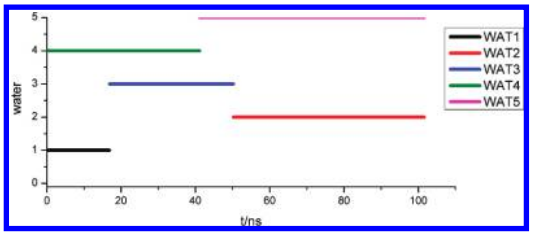


Figure 10. Resistant time (time that a particular water molecule spent within the metal coordination sphere) for the water molecules coordinating Zn^{2+} during 101 ns of MD simulation.

transforms from the "open" to the "closed" conformation, ranges from 151 to 290 kcal mol^{-1}. Taking together the enthalpic and entropic parts (the last one originated from the water molecules release) of the system free energy, it seems natural to consider that in solution the protein can exist in both the "open" and "closed" conformations.

The conformational entropy contribution was evaluated by the normal-mode analyses (Table 1). We considered only the part of the enzyme which exhibits the biggest change during the simulation, i.e., the interdomain cleft, and according to the results, there is no significant change of vibration, rotation, and translational degrees of freedom as the protein goes from the "open" to "closed" conformation.

From the simulated time (101 ns), we cannot make conclusions regarding the probabilities of the open and closed states in the equilibrium, but we may expect that, probably, both states will be relevant, as the experimental evidence supports. This hypothesis is also supported by the results of the independent normal-mode analysis performed using NOMAD (the Normal Mode server http://lorentz.immstr.pasteur.fr). Further on, the fact that the protein kept its "open" form for almost 50 ns of the 101 ns of MD simulation performed by AMBER suggests that the protein is present in both "open" and "closed" forms in solution.

dx.doi.org/10.1021/ci300141k | J. Chem. Inf. Model. 2012, 52, 1583−1594

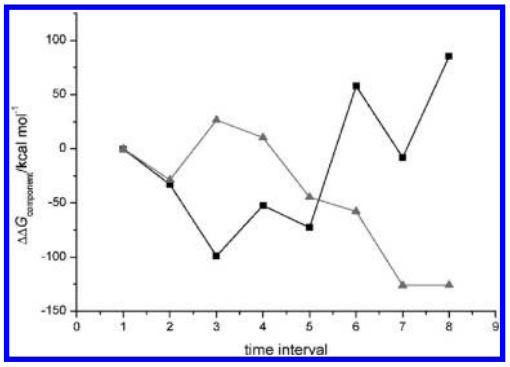


Figure 11. The electrostatic (black line and squares) and entropic (gray line and triangles) solvation free energy components calculated at eight 2-ns-long intervals sampled during the 101-ns-long MD trajectory (see Figure 2). Differences with respect to the values calculated for the first interval (15−17 ns) are displayed. Values were obtained using the solute dielectric constant ε_r = 2 (Table S2, Supporting Information) and a water molecule release energy of 1.2 kcal mol^{-1}.[44]

Bezerra et al.[47] found that binding of an opioid peptide is entropy-driven and that its binding is accompanied by a large domain motion similar to that noticed during our MD simulations. It seems reasonable to consider that peptide binding boosts protein closure and shifts the population of protein conformations in the direction of the "closed" form. Both the experiment[47] and calculations indicated that the protein closure is an entropically driven process. Isothermal titration calorimetry revealed that the penta-peptide binding, as well as the binding of several other peptides, like endomorphin-1, Leu-enkephalin, and IVYPW, are strongly endothermic processes, entropically driven. In the interdomain cleft of the experimentally determined structure of the unbound DPP III, there are approximately 60 water molecules, while in the structure of the complex most of them have been displaced. During the MD simulations of the ligand-free enzyme, about 40 water molecules were released from the enzyme binding site, due to the protein closure. Furthermore, the binding free energy calculations revealed that the closure itself is an entropically driven process.

MD Simulations of the D372A Variant. In order to examine the role of Asp372 in the large scale conformational change in detail, it was mutated to Ala, and the mutated enzyme was simulated for 71 ns. Indeed, the protein closure was much less pronounced in the Asp372Ala mutant than in the wild type enzyme (see Figure 12, left, and compare with Figure 2), i.e., the radius of gyration obtained after 71 ns of MD simulations of DPP III and its Asp372Ala variant are 26.98 Å and 27.68 Å, respectively (mean values calculated from the 29th to 30th ns are 27.06 Å and 27.71 Å, respectively, see Table S3, Supporting Information). Our assumption that the lack of hydrogen bonds between Asp372 and Asn406 and Lys405 will decrease the extent of the enzyme closure was confirmed. The reason for the partial enzyme closure happening during the first 2.5 ns of the MD simulations was the formation of the hydrogen bond between Asn406 and Glu451 (not shown). Although preserved during the whole simulation time, this hydrogen bond was insufficient for the more pronounced enzyme closure detected during simulations of the wild type enzyme (for comparison see the Figure S7 and S1), and the number of the water molecules occupying the 10 Å sphere around the zinc ion (Figure 12, right) remained constant during 71 ns of MD simulation.

MD Simulations for the Enzyme−Substrate Complex, DPP III−RRNA. In order to elucidate the relative stability of the substrate−enzyme complexes, the binding free energies were calculated for the substrate bound in both "open" (WT−RRNA) and "closed" (WT_{72}−RRNA) conformations of the protein. For this purpose, the compact protein conformation obtained after 72 ns of the MD simulations of the ligand-free enzyme was selected, and the substrate was docked into the active site. As already noticed in our previous study (of the substrate binding to the "open" protein conformation), the substrate binding hindered further protein closure; i.e., its presence disrupted the formation of the hydrogen bonds between Asp372 and the amino acid residues from the catalytic domain. Indeed, the radius of gyration even slightly increased upon substrate binding (see Figure S9, Supporting Information, for the case of substrate binding to the "closed", WT_{72}, conformation).

Already during the minimization, the Asp372 side chain reoriented, and the RRNA established two hydrogen bonds with it, through its N terminus and by the first arginine side chain (Figure 13). These interactions, present during the entire MD simulation of WT_{72}−RRNA, were also noticed in the complex with "open" DPP III.[29] As a consequence, similarly as in the complex with the "open" protein, the short flexible loop (amino acids 370 to 374) at the bottom end of the cleft moved toward the substrate binding site (into the cleft).

During the simulations, the substrate coordinated Zn^{2+} with the carbonyl oxygen of the second residue (arginine) from the N terminus. As in the free protein, the zinc ion was coordinated by two His and two Glu residues as well. While during the entire simulations Glu451 coordinated to the zinc ion monodentately, always with the same carboxyl oxygen, Glu508, in the complex with the "closed" enzyme, sometimes coordinated the zinc ion bidentately (see Figure S10, Supporting Information).

Although rather flexible, with its side chain fluctuating between two conformations, the second arginine from the RRNA amino-terminus established several strong interactions with the protein during the simulation. In one of these conformations, its side chain was hydrogen bonded to Ser504 and Glu508 (from the protein "upper" domain), and in the other one it interacted with Asn391, Arg399, and Phe404 (from the protein "lower" domain). Change of the binding pocket

Table 1. The Entropy Contribution to the System Free Energy Calculated Using the Normal Mode Analyses Implemented in the MM−PBSA Module of AMBER11 Program Suit[a]

	TS_{TRA}/kcal mol^{-1}		TS_{ROT}/kcal mol^{-1}		TS_{VIB}/kcal mol^{-1}		TS_{TOT}/kcal mol^{-1}	
t/ns	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD
1−3	17.11	0.00	17.93	0.02	3351.86	13.13	3386.90	13.13
99−101	17.11	0.00	17.85	0.00	3342.44	14.44	3377.40	14.45

[a]During minimization dielectric constant of the solvent was 80.

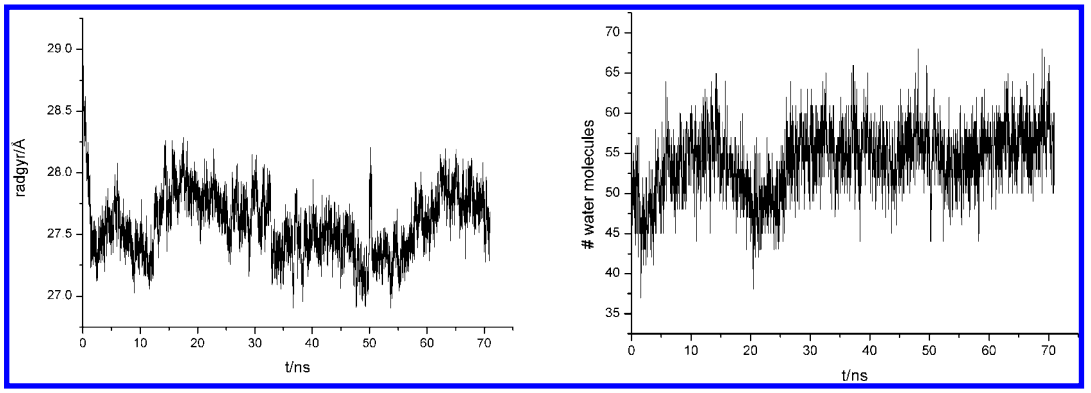


Figure 12. Radius of gyration of the D372A mutant during 71 ns of MD simulations (left) and the number of water molecules during the simulation occupying the 10 Å sphere around the zinc ion (right).

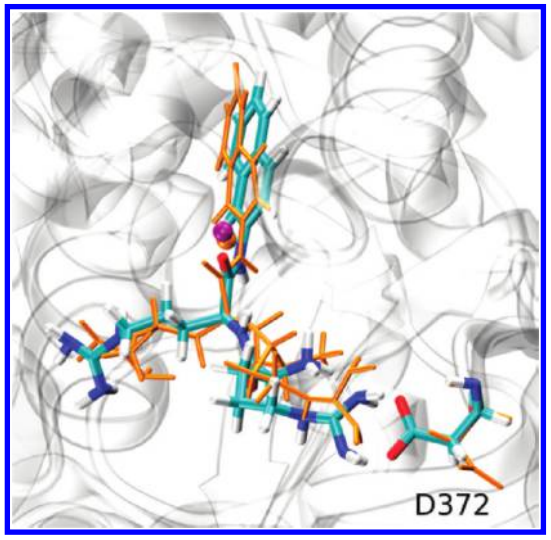


Figure 13. Overlay of the substrate binding sites in the initial (orange) and optimized (colored by atom type) WT_{72}−RRNA complex. Optimization of the initial WT_{72}−RRNA structure resulted in Asp372 reorientation. The central zinc ions are shown as spheres.

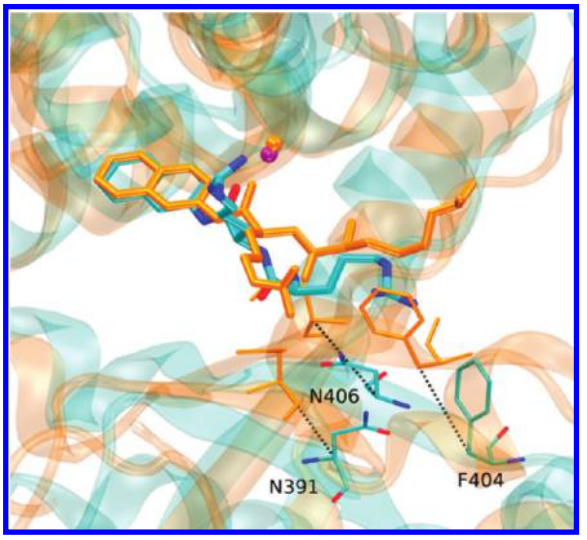


Figure 14. Overlay of two RRNA orientations (thick sticks) in the DPP III binding site, with amino acid residues Asn391, Phe404, and Asn406 (thin sticks) displayed. The RRNA and the amino acid residues colored by the atom type are extracted after 30 ns of MD simulations of the WT−RRNA complex and those, colored orange, after 30 ns of MD simulations of the WT_{72}−RRNA complex. The zinc ions are shown as spheres. Shifts of the amino acid residues from the "lower" domain upon the protein closure are represented as black dotted lines.

shape, induced by the protein closure, is indicated as a shift in positions of Asn391, Phe404, and Asn406 (~5 Å) toward the substrate, see Figure 14.

Although the RRNA molecule in both complexes, with "open" and "closed" DPP III, interacts with the same amino acid residues, interactions are stronger and more persistent in the latter. Apparently, the DPP III−RRNA complex seems to be more stable (see for example Figure S11, Supporting Information) when the protein is in its more compact form; see Table 2 for the results of the MM−PBSA calculations.

It should be noted that the absolute free energy values are meaningless and that we were interested primarily in the binding free energy difference. To be more precise, the most relevant here is the trend of the binding free energy changes.

Conserved Trp300 and His568. Tomić et al.[29] noticed previously that in the "open" WT−RRNA complex His568 fluctuated between two conformations. However, our present study has shown that during the entire simulations of the WT_{72}−RRNA complex His568 was by the hydrogen bond with Glu512 anchored in one position (see Figure 15).

The importance of the fully conserved Trp300 for the human DPP III activity, and its role in maintaining functional integrity of the enzyme's S2 subsite, had been already revealed by the

Table 2. The Binding Free Energies Calculated Using the MM−PBSA Approach (The Solute Dielectric Constant Was 2 and That of the Solvent 80)[a]

	15−20 ns		20−25 ns		25−30 ns	
system	⟨PBSA⟩/ kcal mol^{-1}	SD/ kcal mol^{-1}	⟨PBSA⟩/ kcal mol^{-1}	SD/ kcal mol^{-1}	⟨PBSA⟩/ kcal mol^{-1}	SD/ kcal mol^{-1}
WT-RRNA	−37.2	6.5	−43.74	5.32	−43.95	5.60
WT_{72}-RRNA	−56.04	5.31	−55.54	5.07	−58.60	4.52

[a]The Zn charge used in the calculations was +1.5*e*.

site-directed mutagenesis approach.[48] During our present study, we have noticed that during MD simulations of both free and ligated DPP III, Trp300 has made numerous van der Waals and CH−π interactions with surrounding nonpolar residues, and the strong hydrogen bond with Glu309. In the DPP III−RRNA complex, it also interacted with the Lys405 side chain and in

dx.doi.org/10.1021/ci300141k | J. Chem. Inf. Model. 2012, 52, 1583−1594

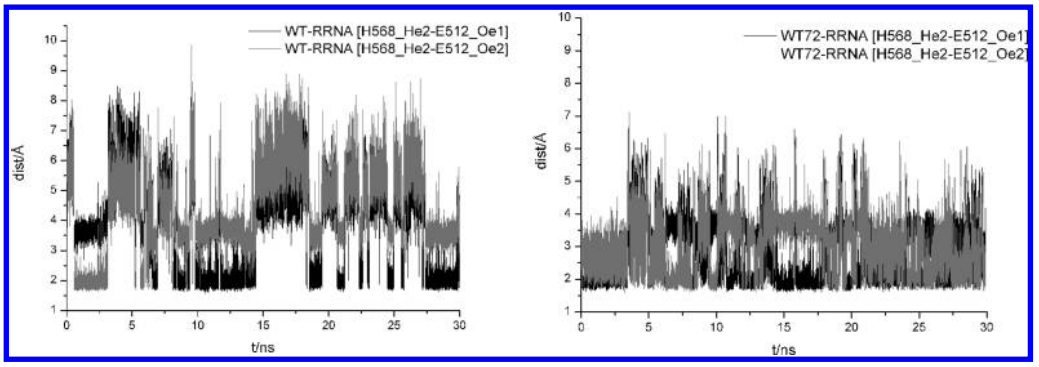


Figure 15. Distances between the hydrogen at the ε-nitrogen atom of His568 and two carboxyl oxygens of Glu512 during 30 ns of MD simulation of both the "open" and "closed" DPP III in complex with RRNA, left and right, respectively.

this way influenced the positions of Phe404 and Asn406 which interacted with the substrate and stabilized its position in the active site (see Figure S12, Supporting Information). It should be noticed that Lys405 is an absolutely and Phe404 a highly conserved amino acid residue within M49 family.

Normal Mode Analysis. Principal component analysis of the 30-ns-long MD trajectories obtained simulating DPP III−RRNA complexes, WT−RRNA and WT_{72}−RRNA, revealed that ∼60% of the overall protein motions could be explained by the top four and five eigenvectors, respectively. The top two eigenvectors for the WT−RRNA complex (see Figure S13, Supporting Information) mostly correspond to the top two eigenvectors determined for the ligand-free enzyme, but with an inversed order of importance and a much lower amplitude; i.e., the eigenvector shown in the structure in the middle of Figure 7 corresponds to the first eigenvector determined for the WT−RRNA complex, but with a 10 times smaller magnitude. However, just about 45% of the overall protein motion could be explained by these two modes. In the case of the WT_{72}−RRNA complex, none of the five eigenvectors describing ∼60% of the overall protein motion matches any of the two main eigenvectors determined for the ligand-free enzyme. Rather, they are a combination of domains motion and local movements, describing the opening and moderate reorganization of the enzyme in order to adapt the ligand.

Substrate Trafficking Trough the Protein. Results obtained using the random accelerated MD (RAMD) simulations for the possible paths for the substrate, RRNA, entry and exit in the human DPP III active site agree with those obtained using the program CAVER. Although all four tunnels shown in Figure 10 were used by the RRNA molecule to leave the DPP III active site, the blue and the yellow one were used 90% of the time, and with equal probability.

■ CONCLUSION

The present study on human DPP III, a member of the metallo−peptidase M49 family, revealed large scale conformational transitions of this two domain protein which could be described as "protein closure". A comparison of the MD simulation trajectories obtained for the ligand-free protein, wild type and the Asp372Ala mutant, clearly showed that Asp372 is playing a main role in this large scale interdomain motion. Although boosted with the cascade of strong interactions, like hydrogen bond formation between Glu451 and Asn406 at the very beginning of the MD simulation, the movement is finalized by hydrogen bonding of Asp372 and Lys405. Mutation of Asp372 to Ala enabled the formation of this hydrogen bond, and the protein closure was much less pronounced.

The obtained components of the Gibbs free energy indicated that the observed conformational change of DPP III is a result of the enthalpy−entropy compensation. The increase of the system enthalpy due to protein closure is compensated by the entropy increase due to reorganization of the water molecules, i.e., their release in a bulk.

These results are reinforced by very recent experimental findings (Bezzera et al.[47]) showing that the substrate binding is an entropic process, which is governed by the release of water molecules from the interdomain cleft, and that the enthalpic term of the binding free energy is unfavorable. Thus, according to the calculations and the experimental results, we hypothesize that in water solution the enzyme can exist in both the "open" and "closed" forms and that the peptide binding (preferable in the more compact conformation) induces only the end stage of the binding site lock, supporting the combination of the pre-existing equilibrium and the induced-fit model. This could explain the broad specificity of human DPP III for peptides of varying lengths and compositions. So, when exploring the binding of different ligands into the enzyme active site, one must be careful when choosing which of the DPP III conformations is the active one.

Experimental and computational studies indicate that the enzyme secondary structure, as well as the each domain conformation, are conserved upon the protein closure.

Calculations of the free energy difference for binding of the preferred synthetic substrate Arg−Arg−2−naphtylamide to different protein conformations revealed that it binds tighter to the "closed" than to the "open" protein. In the "closed" form, the 100% evolutionarily conserved His568 is, by the hydrogen bond with Glu512, anchored to one conformation, while during simulations of the "open" protein complex it fluctuated between two conformations. Since His568 had been indicated as important for the substrate hydrolysis, its position might be important for the catalytic efficiency of DPP III. Binding of the ligand either to the "open" or to the "closed" DPP III conformation has stopped further protein closure and additionally rigidified the amino acid residues interacting with the substrate as well as the helix connecting the "upper" and "lower" domains and the part of the "lower" domain close to the active site. This study also revealed roles of several other, in the M49 family, either absolutely, Trp300, Tyr318, Arg399 and Lys405, or highly conserved residues, Asn391, Phe404, and Asn406, all belonging to the "lower", noncatalytic domain. By their hydrogen bonding to Asp372, Lys405 and Asn406 played

a crucial role in protein closure. Together with Arg399 and Phe404, they also interacted with the substrate and stabilized its position in the enzyme active site. Tyr318 and Asn391 interacted with amino acid residues from the catalytic domain and consolidated the compact enzyme form. During the simulation, Trp300 moved more than 10 Å with respect to its initial position; its side chain established interaction with Lys405 and thus took an indirect role in substrate stabilization.

■ ASSOCIATED CONTENT

Supporting Information

Table S1 shows the enzyme secondary structure elements. Table S2 contains the components of the protein solvation free energy. Table S3 summarizes the average values of the free and complexed enzyme radius of gyration. Figures S1 and S9 depict the protein radius of gyration profiles during the simulations of the ligand-free DPP III, and its complexes with the substrate, respectively. Figure S2 shows the crystal packing of the unbound human DPP III. Figure S3 shows the position of water molecules occupying the 10 Å sphere around the zinc ion in the enzyme. Figures S4 and S7 show the change in certain distances and angles accompanying the enzyme closure. Figure S5 displays enzyme experimental and calculated B-factors. Figure S6 depicts the three lowest frequency normal modes obtained by the NOMAD. Figure S8 displays the change in protein solvent accessible surface area. Figure S10 shows the change in distances between the zinc ion and the glutamates coordinating it in the DPP III complex with RRNA. Figure S11 shows the RMSD profiles of the substrate arginine residues in the same case as before. Figure S12 depicts the position of Tpr300. Figure S13 summarizes the results of the normal-mode analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +385 1 457 1251. Fax: +385 1 468 0245. E-mail: sanja.tomic@irb.hr.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work has been supported by the Ministry of Science, Education, and Sport of the Republic of Croatia (project number 098-1191344-2860) and the Croatian National Grid Infrastructure (CRO NGI). The authors also thank Gordan Horvat for his help with GROMACS simulations and Dr. Marija Abramić for valuable suggestions. A.T. acknowledges the HPC−Europa2 Transnational Access program, and all authors acknowledge support by the Barcelona Supercomputing Center and Red Española de Supercomputación (QCM-2011-1-0015, BCV-2011-2-0009). M.G. acknowledges the projects CTQ2011-27857-C02-01 (Spanish Ministry of Science and Innovation) and 2009SGR 17 (Autonomous Goverment of Catalonia).

■ ABBREVIATIONS

DPP III, dipeptidyl−peptidase III; MD, molecular dynamics; RRNA, Arg−Arg−2−naphtylamide; SASA, solvent accessible surface area; PDB, Protein Data Bank; MM−PBSA, Molecular Mechanics Poisson−Boltzmann Surface Area; RMSD, root-mean-square deviations; radgyr, radius of gyration; PCA, principal component analysis; RAMD, random acceleration molecular dynamics.

■ REFERENCES

(1) Abramić, M.; Špoljarić, J.; Šimaga, Š. Prokaryotic homologs help to define consensus sequences in peptidase family M49. *Period. Biol.* **2004**, *106*, 161−168.

(2) Chiba, T.; Li, Y. H.; Yamane, T.; Ogikubo, O.; Fukuoka, M.; Arai, R.; Takahashi, S.; Ohtsuka, T.; Ohkubo, I.; Matsui, N. Inhibition of recombinant dipeptidyl peptidase III by synthetic hemorphin-like peptides. *Peptides* **2003**, *24*, 773−778.

(3) Chen, J. M.; Barrett, A. J. In *Handbook of Proteolytic Enzymes*; Barrett, A. J., Rawlings, N. D., Woessner, J. F., Eds.; Elsevier, Academic Press: London, 2004; Vol. *1*, Dipeptidyl-peptidase III, pp 809−812.

(4) Liu, Y.; Kern, J. T.; Walker, J. R.; Johnson, J. A.; Schultz, P. G.; Luesch, H. A genomic screen for activators of the antioxidant response element. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 5205−5210.

(5) Comparison performed using blast2seq and jFATCAT_rigid, methods, respectively implemented at RCSB PDB server (http://www.rcsb.org/pdb/workbench/workbench.do).

(6) Keskin, Z. Binding induced conformational changes of proteins correlate with their intrinsic fluctuations: a case study of antibodies. *BMC Struct. Biol.* **2007**, *7*, 31−41.

(7) Jencks, W. P. In *Catalysis in Chemistry and Enzymology*; Courier Dover Publications: New York, 1987.

(8) Hammes, G. Multiple conformational changes in enzyme catalysis. *Biochemistry* **2002**, *41*, 8221−8228.

(9) Benkovic, S. J.; Hammes-Schiffer, S. A perspective on enzyme catalysis. *Science* **2003**, *301*, 1196−1202.

(10) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. How enzymes work: analysis by modern rate theory and computer simulations. *Science* **2004**, *303*, 186−195.

(11) Eisenmesser, E. Z.; Akke, M.; Bosco, D. A.; Kern, D. Enzyme dynamics during catalysis. *Science* **2002**, *295*, 1520−1523.

(12) Austin, R. H.; Beeson, K. W.; Eisenstein, L.; Frauenfelder, H.; Gunsalus, I. C. Dynamics of ligand binding to myoglobin. *Biochemistry* **1975**, *14*, 5355−5373.

(13) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D. Intrinsic dynamics of an enzyme underlies catalysis. *Nature* **2005**, *438*, 117−121.

(14) James, L. C.; Roversi, P.; Tawfik, D. S. Antibody Multispecificity Mediated by Conformational Diversity. *Science* **2003**, *299*, 1362−1367.

(15) Koshland, D. E. Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 98−104.

(16) Sali, A.; Blundell, T. L. Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.

(17) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.

(18) Case, D. A.; Cheatham, T. E.; Darden, T. A.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.

(19) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, CA, 2008.

(20) Bekker, H.; Berendsen, H. J. C.; Dijkstra, E. J.; Achterop, S.; van Drunen, R.; van der Spoel, D.; Sijbers, A.; Keegstra, H.; Reitsma, B.; Renardus, M. K. R. In *Physics Computing 92*; de Groot, R. A., Nadrchal, J., Eds.; World Scientific: Singapore, 1993; Gromacs: A parallel computer for molecular dynamics simulations, pp 257−261.

(21) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(22) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **2001**, *7*, 306−317.
(23) van der Spoel, D.; 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.
(24) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(25) Bertoša, B.; Kojić-Prodić, B.; Wade, R. C.; Tomić, S. Mechanism of auxin interaction with auxin binding protein (ABP1): a molecular dynamics simulation study. *Biophys. J.* **2008**, *94*, 27−37.
(26) Dokmanić, I.; Šikić, M.; Tomić, S. Correlation between the metal ion type, coordination number and the amino acid residues involved in the coordination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2008**, *64*, 257−263.
(27) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald−an Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(28) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577−8592.
(29) Tomić, A.; Abramić, M.; Špoljarić, J.; Agić, D.; Smith, D. M.; Tomić, S. Human Dipeptidyl Peptidase III: Insight into ligand Binding form a Combined Experimental and Computational Approach. *J. Mol. Recognit.* **2011**, *24*, 804−814.
(30) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A computer-simulation method for the calculation of equilibrium-constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637−649.
(31) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984**, *52*, 255−268.
(32) Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(33) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* **1996**, *87*, 1117−1157.
(34) Swanson, J. M.; 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.
(35) Connolly, M. L. Analytical molecular surface calculation. *J. Appl. Crystallogr.* **1983**, *16*, 548−558.
(36) Almlöf, M.; Carlsson, J.; Åqvist, J. Improving the Accuracy of the Linear Interaction Energy Method for Solvation Free Energies. *J. Chem. Theory Comput.* **2007**, *3* (6), 2162−2175.
(37) Lee, B.; Richards, F. M. The interpretation of protein structures: estimation of static accesibility. *J. Mol. Biol.* **1971**, *55*, 379−400.
(38) Mihel, J.; Šikić, M.; Tomić, S.; Jeren, B.; Vlahoviček, K. PSAIA - Protein Structure and Interaction Analyzer. *BMC Struct. Biol.* **2008**, *8*, 21−31.
(39) Petřek, M.; Otyepka, M.; Banáš, P.; Košinová, P.; Koča, J.; Damborský, J. CAVER: A New Tool to Explore Routes from Protein Clefts, Pockets and Cavities. *BMC Bioinf.* **2006**, *7*, 316−324.
(40) Medek, P.; Beneš, P.; Kozlíková, B.; Chovancová, E.; Pavelka, A.; Szabó, T.; Zamborský, M.; Andres, F.; Klvaňa, M.; Brezovský, J.; Sochor, J.; Damborský, J. *CAVER; CAVER 2.0*; Masaryk University: Czech Republic, 2008.
(41) Schleinkofer, K.; Sudarko, P. J.; Winn, R. C.; Lüdemann, S. K.; Wade, R. C. Do mammalian cytochrome P450a show multiple ligand access pathways and ligand channeling? *EMBO Rep.* **2005**, *6*, 584−589.
(42) Lüdemann, S. K.; Lounnas, V.; Wade, R. C. How do substrates enter and products exit the buried active site of cytochrome P450cam? 2. Steered molecular dynamics and adiabatic mapping of substrate pathways. *J. Mol. Biol.* **2000**, *303*, 813−830.
(43) Winn, P. J.; Lüdemann, S. K.; Gauges, R.; Lounnas, V.; Wade, R. C. Comparison of the dynamics of substrate access channels in the three cytochrome P450s reveals different opening mechanisms and a novel functional role for buried arginine. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5361−5366.
(44) Salopek-Sondi, B.; Vukelić, B.; Špoljarić, J.; Šimaga, Š.; Vujaklija, D.; Makarević, J.; Jajčanin, N.; Abramić, M. Functional tyrosine residue in the active center of human dipeptidyl peptidase III. *Biol. Chem.* **2008**, *389* (2), 163−167.
(45) Sheu, S. Y.; Yang, D. Y. Determination of Protein Surface Hydration Shell Free Energy of Water Motion: Theoretical Study and Molecular Dynamics Simulation. *J. Phys. Chem. B* **2010**, *114*, 16558−16566.
(46) Pal, S. K.; Peon, J.; Bagchi, B.; Zewail, A. H. Biological water: Femtosecond dynamics of macromolecular hydration. *J. Phys. Chem. B* **2002**, *106*, 12376−12395.
(47) Bazzera, G. A.; Dobrovetsky, E.; Viertimayr, R.; Dong, A.; Binter, A.; Abramić, A.; Macheroux, P.; Dhe-Paganon, S.; Gruber, K. Entropy-driven binding of opioid peptides induces a large domain motion in human dipeptidyl peptidase III. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 6525−6530.
(48) Špoljarić, J.; Salopek-Sondi, B.; Vukelić, B.; Agić, D.; Šimaga, Š.; Jajčanin-Jozić, N.; Abramić, M. Absolutely conserved tryptophan in M49 family of peptidases contributes to catalysis and binding of competitive inhibitors. *Bioorg. Chem.* **2009**, *37*, 70−76.