Molecular Dynamics Study of Triosephosphate Isomerase from *Trypanosoma cruzi* in Water/Decane Mixtures

Norma Díaz-Vergara and Ángel Piñeiro*

Departamento de Fisicoquímica, Facultad de Química, Universidad Nacional Autónoma de México, UNAM, Ciudad Universitaria 04510, México D.F., Mexico

Received: October 22, 2007; In Final Form: December 10, 2007

A comprehensive study of the triosephosphate isomerase from the parasite *Trypanosoma cruzi* (TcTIM) in water, in decane, and in three water/decane mixtures was performed using molecular dynamics (MD) simulations in a time scale of 40 ns. The structure and dynamics of the enzyme, as well as the solvent molecules' distribution and mobility, were analyzed in detail. In the presence of decane, the amplitudes of the most important internal motions of the enzyme backbone were observed to depend on the solvent concentration: the higher the water concentration, the greater the amplitudes. Contrary to this trend, the amplitudes of the TcTIM motions in pure water were similar to those of the simulation with the lowest water concentration. The enzyme was observed to be almost motionless in pure decane due to a sharp increase of the number of intramolecular hydrogen bonds. This caused a contraction of the enzyme structure accompanied by a loss of secondary structure and of a decrease of the hydrophilic solvent accessible surface. A similar behavior, although to a lesser extent, was observed in the simulation at the lowest water concentration. Our results suggest that the presence of decane molecules located at specific sites of the enzyme might accelerate its internal movements, although a minimum number of water molecules is needed for the protein to keep its structure and dynamics. Altogether, this work provides new insight into protein and water behavior in organic solvents as well as into the dynamics of TcTIM itself.

1. Introduction

An interesting relationship connecting rigidity, thermal stability, and catalytic activity of enzymes has been proposed by several authors in the past: the higher the flexibility and activity of the enzyme, the lower its unfolding temperature is.[1] Therefore, it appears that, throughout evolution, proteins have sought a balance between flexibility and thermal stability, which allows them to achieve their biological function and, at the same time, to preserve a well-defined average structure. Water plays an essential role in this context.[2,3] It is well-known that water molecules located in the vicinity of a solute show differences in their organization and hence in their thermodynamic, dynamic, and electrostatic properties, when compared to bulk waters.[4−6] The water behavior at a molecular level can be strongly affected by the presence of cosolvents, and this is critical for the dynamics of macromolecules.

In the middle 1980s, Zaks and Klibanov found that the catalytic activity of several enzymes, when they are suspended in hydrophobic solvents at low water concentrations, is comparable or even higher than in aqueous solution.[7−11] On the basis of their studies, they concluded that the more hydrophobic the main solvent, the less water is needed to maintain protein activity.[10] Moreover, the same authors claimed that water acts as a lubricant, forming hydrogen bonds with accessible groups of enzymes and providing them with the flexibility necessary for catalysis.[10] Such explanations are generally accepted, although no direct evidence has been provided. The activity of enzymes under those conditions, or when they are trapped in reverse micelles,[12,13] became more than a curiosity when several biotechnological applications were proposed, taking advantage of this alternative.[14−18] To date, no experimental technique has been able to provide a detailed structural−dynamics description of this phenomenon. To survey this issue, a number of MD simulation studies of proteins in water/organic solvent mixtures have been performed.[19−29] However, the short time scale employed (7 ns for the longest trajectories) did not allow the observation of clear protein structural−dynamic differences between the several conditions considered. In 2003, Soares et al.[19] reported the first systematic study of two model proteins (ubiquitin and cutinase) in hexane, under varying hydration conditions. Although they found correlations between root-mean-square deviations (RMSD), root-mean-square fluctuations (RMSF), and the percentage of water in the system, their analysis was performed only over the last 2 ns of 4 ns long trajectories. No significant protein backbone motions are expected at that time scale, so the observed fluctuations should be mainly due to the dynamics of the amino acid side chains. A high degree of dispersion, which the authors justified by the probable unstability of the initial structures, was observed in these simulations. Yang et al.[27] performed MD simulations of surfactant-solubilized subtilisin BPN′ in octane, tetrahydrofuran, and acetonitrile at two hydration conditions and at a time scale of ∼4 ns. The enzyme did not show significant structural or dynamic solvent-dependent differences. Then, the authors focused their analysis on the partition of water molecules between the surface of the protein and the organic phase and on the hydration of the active site. Very recently, Micaelo and Soares[28] published a comprehensive study of the hydration mechanism of cutinase in hexane, diisopropyl ether, 3-pentanone, ethanol, and acetonitrile with increasing hydration levels

* Corresponding author. E-mail: fangel@usc.es. Current address: Department of Applied Physics, Faculty of Physics, University of Santiago de Compostela, E-15782 Santiago de Compostela, Spain.

10.1021/jp7102275 CCC: $40.75
Published on Web 02/27/2008

TABLE 1: List of MD Simulations Including Parameters and Results of Analysis

trajectory	no. of water molecules	No. of decane molecules	% water	simulation time (ns)	initial structure ID	total intramolecular H-bonds	radii of gyration (nm)	inter-subunit distances (nm)	hydrophobic SAS (nm^2)	hydrophilic SAS (nm^2)
RW_a	24825	0	100	20	1TCD					
RW_b	24825	0	100	40	$RW_a(20)$	379 ± 9	2.46 ± 0.01	3.58 ± 0.03	117 ± 2	93 ± 2
W8D	3579	3402	11.80	40	$RW_a(20)$	380 ± 11	2.50 ± 0.01	3.66 ± 0.03	128 ± 2	90 ± 2
W6D	2336	3208	8.45	40	$RW_a(20)$	388 ± 10	2.49 ± 0.01	3.63 ± 0.02	129 ± 2	90 ± 1
W4D	1185	2825	5.04	40	$RW_a(20)$	392 ± 9	2.392 ± 0.008	3.35 ± 0.02	127 ± 2	87 ± 1
RD	0	2587	0	40	1TCD	563 ± 11	2.345 ± 0.006	2.28 ± 0.01	122 ± 2	58.1 ± 0.9

at a time scale of 7 ns. Slight structural changes were observed in the enzyme, but no analysis of its dynamics was performed. This latter work was mainly focused on the water behavior. These representative studies revealed the lack of MD simulations at longer time scales that could provide new insight into the structure and dynamics of enzymes at low water concentrations as well as into the role of the solvent for protein function.

In the present work, 40 ns long MD simulations of triosephosphate isomerase from the parasite *Trypanosoma cruzi* (TcTIM) in several water/decane mixtures were performed. As a reference, two simulations of the same enzyme in pure decane and pure water, respectively, at the same time scale were carried out. TcTIM is a homodimeric enzyme with 251 residues per subunit, which folds in the so-called α/β-barrel motif with eight central β-strands surrounded by eight α-helices.[30] This enzyme was chosen because it was experimentally proven[31] that TIMs from other species show a high catalytic activity at low water concentrations. Besides, it is a well-liked target for drugs against American tripanosomiasis, better known as Chagas' disease.[30−37] A detailed analysis of the protein structure and dynamics, as well as of the solvent molecules' behavior, was performed under different conditions. The main objective of this work was to look for concentration-dependent differences in the protein and solvent behavior that could provide new insight into the understanding of the catalytic activity of enzymes in organic media. Additionally, the dynamic study of TcTIM in pure water by itself, one of the reference simulations in this work, could be helpful to better understand this enzyme and then to design or improve inhibitors. No MD studies of enzymes in organic solvents at low water concentrations, and either of TIM in aqueous media, at the time scale of tens of nanoseconds have been performed previously.

2. Materials and Methods

2.1. Setup of the Simulation Boxes. Six simulation systems were built up with one TcTIM molecule and different amounts of water as well as decane, forming a nonpolar phase. Decane was chosen due to its high hydrophobicity, which ensures that practically all water molecules remain close to the protein in the simulations that involve a mixture of both solvents. The systems, as well as the corresponding trajectories, were named RW_a, RW_b, W8D, W6D, W4D, and RD. RW_a consisted of 24 825 pre-equilibrated water molecules solvating the enzyme structure as taken from the Protein Data Bank (PDB ID 1TCD),[30] with no decane present. The protonation states of ionizable residues corresponded to biological pH ($\sim$7.4). The size of the simulation unit cell was chosen so that the distance between every protein atom and box walls was larger than 2 nm. This system was first simulated for 20 ns, the resulting final configuration being denoted by $RW_a(20)$. RW_a was then extended for 40 ns more, producing the trajectory named RW_b. To build up the system W8D, waters at a distance larger than 8 Å from any of the protein atoms were removed from $RW_a(20)$, and the resulting arrangement was enclosed into a dodecahedron simulation box with the walls at a minimum of 2 nm from any of the water molecules. The free space of such a box was then filled with pre-equilibrated decane molecules. The same procedure as for W8D was repeated using distances of 6 and 4 Å to give the systems W6D and W4D, respectively. This procedure assures that both the protein and the water molecules in its immediate vicinity are in equilibrium to each other in the initial conformations of these three systems. As in previous studies,[19,27] pH memory was assumed for the protein when transferred from pure water to the water/decane solutions. The system RD is analogous to RW_a but employing decane instead of water as the unique solvent. As in the case of RW_b, W8D, W6D, W4D, and RD were simulated for 40 ns. Table 1 specifies the number and relative concentration of water and decane molecules, as well as the time scale and the initial protein structure employed in each simulation.

2.2. MD Simulation Parameters. All MD simulations were performed using the GROMACS package[38−41] version 3.2.1 with the GROMOS96 (43a2) force field[42,43] for the protein and decane and the simple point charge (SPC) model for water.[44] Periodic boundary conditions with rhombic dodecahedron boxes as the basic unit cell in the NTP ensemble were used. The pressure was maintained by weak coupling[45] to a reference value of 1 bar, with a coupling time of 0.5 ps. The isothermal compressibility utilized in each simulation was the average of the water and decane values, weighted by the concentration of each solvent component. Water, decane, and protein were coupled separately to a temperature bath[45] at 298 K, with a coupling constant of 0.1 ps. Nonbonded interactions were evaluated using a twin range cutoff of 0.9 and 1.4 nm; interactions within the shorter and longer cutoffs were updated every step and every five steps, respectively. Beyond the 1.4 nm cutoff, a reaction field correction with a dielectric constant ϵ of 78 was employed for all the simulations, including those where the main solvent was decane. Since the dielectric constant applies only at large distances, no artificial effects were expected in the protein structure/dynamics nor in the water behavior. Before every MD simulation, an energy minimization using the steepest descent method was performed to avoid the most unfavorable interactions. Random initial velocities were assigned to the systems from a Maxwell−Boltzmann distribution at 298 K. The equations of motion were integrated using the leapfrog method with a time step of 2 fs. The bond lengths and angles in water were constrained using the SETTLE algorithm,[46] while the LINCS algorithm[47] was used to constrain bond lengths within the protein and decane molecules. During the MD simulations, coordinates, velocities, and energies were stored every 10 ps for further analysis.

2.3. Analysis of MD Trajectories. The viewers RASMOL[48] 2.7, VMD[49] 1.8.2, and PyMOL[50] 0.99 were employed to roughly inspect the behavior of both the enzyme and the solvent. This first screening was used to design a more detailed analysis. RMSD, taking always as a reference the X-ray structure of the protein, were calculated for all the trajectories. To study the

mobility of every residue, RMSF values were calculated over the last 20 ns of the trajectories RW_b, W8D, W6D, W4D, and RD, taking also the X-ray structure as a reference. The principal components (PC) of the protein motion were determined from the diagonalization of the variance−covariance matrix, providing the amplitude and the direction of each concerted motion (i.e., the eigenvalues and the corresponding eigenvectors). This analysis was carried out for the last 20 ns of all the trajectories except for RW_a, employing only atoms of the protein backbone. The trajectories were projected on the eigenvectors with the largest eigenvalues, and images of the two extreme enzyme−backbone conformations for each of these motions were generated. Additionally, the same analysis was also performed over the 40 ns of the whole RW_b trajectory. Cluster analysis was carried out over the last 20 ns of all the trajectories using the GROMOS method.[51] The RMSD cutoff employed to define each cluster was 0.15 nm for trajectories RW_b, W8D, W6D, and W4D and 0.10 nm for RD. A detailed screening of the protein secondary structure was also performed for all the simulations, by using the definitions of the Dictionary of Protein Secondary Structure (DPSS).[52] RMSD, RMSF, and secondary structure as a function of time were analyzed by means of programs contained in the GROMACS package[38−41] version 3.2.1, while the analysis of essential protein motions as well as the cluster determinations were performed by using version 3.3.1 of the same package. A number of simple programs were developed to study the distribution and dynamics behavior of the solvent molecules around the enzyme. For several frames of every trajectory, the number of oxygens belonging to water molecules as well as decane carbons at less than 4 Å from every TcTIM atom were calculated. Additionally, for the last 20 ns of the trajectories that correspond to systems involving water (RW_b, W8D, W6D, and W4D), the distances between every oxygen atom of water molecule and every protein atom were monitored. Such a computationally expensive analysis allows the calculation of the water molecules' residence time in the vicinity of every TcTIM residue. The number of intramolecular hydrogen bonds of the protein were also calculated for all the trajectories using a donor−acceptor atom cutoff distance of 3.5 Å and a donor−hydrogen−acceptor angle of less than 30°. The radii of gyration (RoG), as well as hydrophilic and hydrophobic contributions to the solvent accessible surface (SAS) of the enzyme, were determined throughout all the trajectories. The obtained values were averaged over the last 20 ns. Tools from the GROMACS package were employed to determine the H-bonds, RoG, and SAS.

3. Results and Discussion

3.1. Visual Analysis of Trajectories. In the time scale and under the simulation conditions employed in this work, the two TcTIM subunits remained together, and their tertiary structures were apparently conserved. However, the enzyme conformation arising from the trajectory RD after 40 ns of MD simulations seems to be smaller than the rest of the conformations, and it missed several α-helices (Figure 1). The size of the final structure obtained from W4D also appears to be slightly smaller than those obtained from RW_b, W8D, and W6D. Notably, most of the water molecules in W8D, W6D, and W4D formed a cluster in the cavity next to the interface of the two protein subunits. The fact that the cluster formation was observed in the three different trajectories suggests that it is not accidental. Several reasons could explain this grouping; the most evident arises from the combination of the high hydrophobicity of decane, the main solvent in those simulations, together with

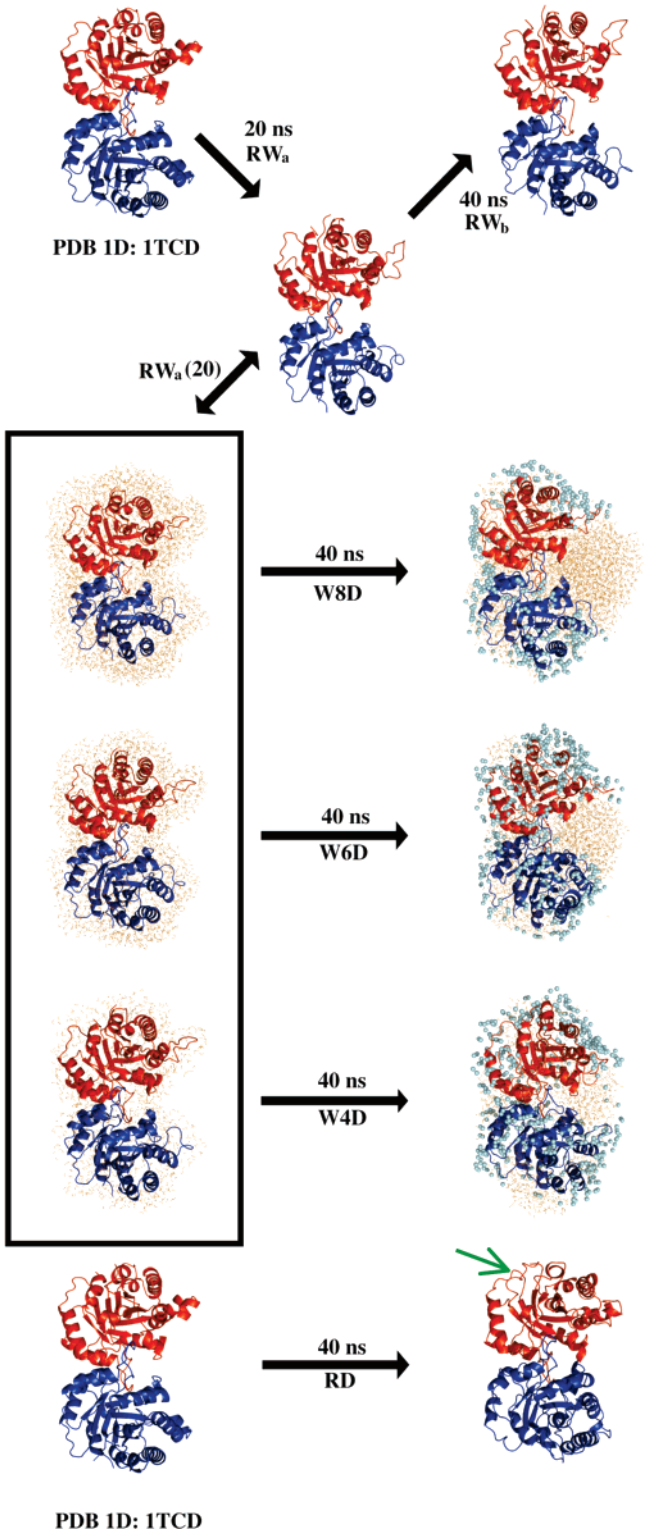

Figure 1. Initial and final conformations of the TcTIM for all the trajectories. The time scales and the name of the systems are indicated. For W8D, W6D, and W4D, all waters (orange sticks) and decane-carbon atoms (green balls) at a distance of less than 4 Å are displayed. Subunits A and B (as referred to in the text) are drawn in blue and red, respectively. The green arrow in the last conformation of RD points to the region of the enzyme that lost its secondary structure.

the geometry of the protein that provides a nucleation point around which a water cluster can grow. The enzyme internal motions could also contribute to the water aggregation. The region of the protein that is in contact with the water cluster can be considered to be relatively hydrophobic but in no way

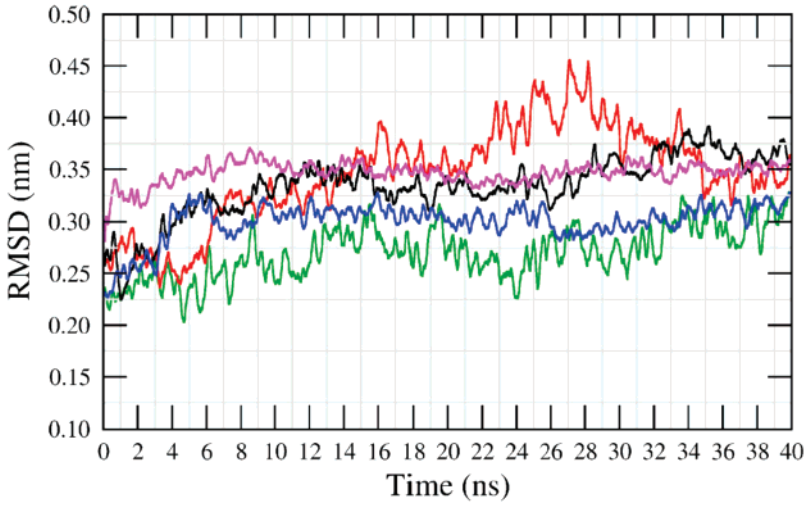

Figure 2. RMSD of the TcTIM backbone as a function of time for the trajectories: RW_b (green), W8D (red), W6D (black), W4D (blue), and RD (magenta). The crystal structure of the enzyme was employed as a reference in all cases.

more than decane. So, it seems that the major reason for the water cluster formation, precisely in the protein cavity, is the minimization of the exposure of water molecules to decane. In addition to the cluster, a number of water molecules seem to exist, individually or in small groups, fixed to specific sites of the protein.

3.2. Protein Structure and Dynamics. As described in the Materials and Methods, the simulations RW_b, W8D, W6D, and W4D started from the same initial conformation: $RW_a(20)$ (Figure 1 and Table 1). However, RW_b was a direct extension of RW_a, while in the case of W8D, W6D, and W4D, the solvent was partially substituted by decane, giving a starting configuration that is not fully equilibrated. After a few nanoseconds, the RMSD of all the trajectories exhibited fluctuations with no clear trends (Figure 2). For instance, the RMSD of W8D increased between ~22 and ~30 ns and then decreased to recover, at the end of the trajectory, a value comparable to that at 10 ns. Notably, the RMSD corresponding to the simulation with decane as the unique solvent (RD) displayed a behavior very similar to that of the other four systems, although it appeared to present less fluctuations. RMSD values are important but not specific enough because the structure deviation of the protein at every time step is represented by just one number. RMSF values for every TcTIM residue were calculated from the last 20 ns of the trajectories RW_b, W8D, W6D, W4D, and RD (Figure 3a). Accumulated RMSF values ($RMSF_{acc}$) (i.e., the sum of amino acid RMSF values for each subunit and trajectory) were also calculated (Figure 3b). It can be observed that, in general, loops 5 and 6 (between residues 129−139 and 169−179, respectively) presented the highest fluctuations. RMSF values of loop 5 are important for all subunits and trajectories, while loop 6 of subunits B in W8D and A in W6D exhibited a much higher mobility than the same loop in the other subunit (Figure 3a). The activity of this enzyme has been correlated with a movement of such a loop in just one subunit, with a period of 200−400 μs.[53,54] The movements identified in our trajectories are probably just a flavor of the complete route covered by loop 6 because only local movements of relatively high frequency could be observed in our simulations. However, the higher fluctuations in W8D and W6D when compared to those in pure water, with similar RMSD values, suggest that the presence of decane might speed up the enzyme dynamics. In general, RMSF residue profiles were similar from one simulation or subunit to another. Loops 5 and 6 were the amino acid segments with maximum fluctuations in the two subunits for all the simulations, while the β-strand segments were the regions of lowest fluctuations. Although the differences in the $RMSF_{acc}$ values were not extremely large, it is noteworthy that the maximum and minimum values corresponded to the systems W8D and RD, respectively. Except for the trajectory W6D, the second subunit always presented $RMSF_{acc}$ values slightly higher than the first one. It is evident from data shown in Figure 3a that the higher mobility of the subunit A in W6D was due to the motion of its loop 6 in that simulation.

The components with the highest amplitude of the TcTIM motion during the last 20 ns of all the trajectories were identified as explained in the Materials and Methods. These amplitudes were sorted in decreasing order, and the corresponding values were found to decay in an exponential way in all the systems (Figure 4a). Comparison between different systems revealed that, in the presence of decane, the higher the water concentration, the higher the mobility of the enzyme. Remarkably, TcTIM in pure water showed amplitudes comparable to those in W4D, the simulation with the lowest water concentration. The enzyme was clearly more rigid in the pure hydrophobic solvent. The accumulated RMSF values, when compared between the different systems, show a trend similar to that of the amplitudes of the principal motions of the enzyme. However, the PC analysis is finer than simple RMSF calculations, and the analysis was performed considering only the backbone atoms of the enzyme, whose movements are expected to be more directly related to protein function. It has been proposed in the past that the amino acid segments with high RMSF values coincide with the protein regions that first lose their structure in the unfolding processes.[55] Thus, the rigidity of the protein observed in pure decane could be related to the higher thermal stability that enzymes usually present in organic solvents.[1] In agreement with RMSF values, PC analysis showed that the most important motions of the enzyme in all the simulations involve loops 5 and 6. Probably due to the length of the trajectory, no great structural differences could be appreciated between the extremes of the projections of the trajectories on the most important eigenvectors. Attempting to obtain further insight into the protein motion, and taking advantage of the fact that the trajectory RW_b was a continuation of RW_a, PC analysis was performed for this whole trajectory (40 ns long). Again, the PC of the enzyme dynamics involved loops 5 and 6. Movie S1 shows the motion with the highest eigenvalue obtained from this simulation.

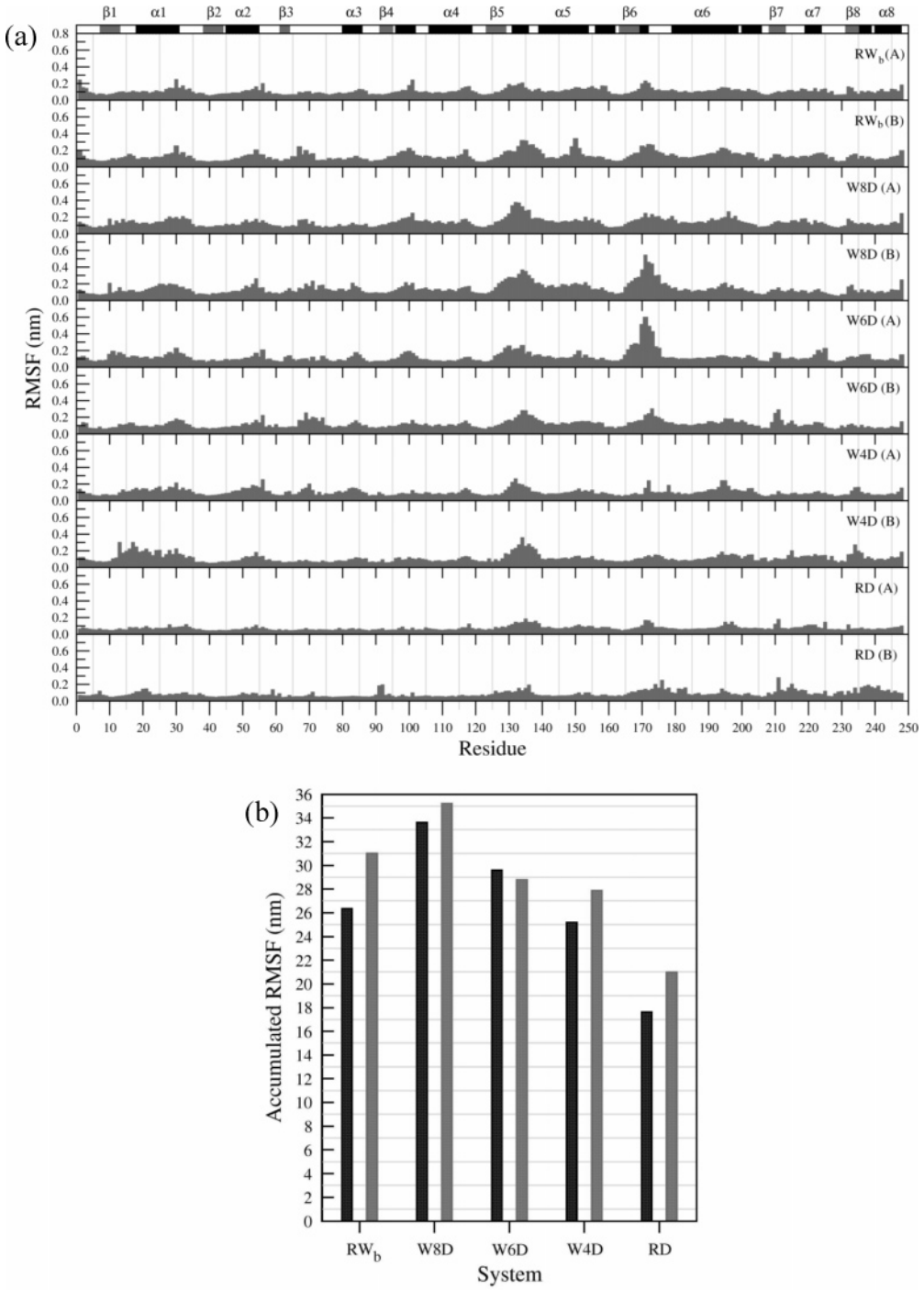

Figure 3. (a) RMSF of every TcTIM residue, calculated over the last 20 ns of each trajectory. The name of the system and subunit (A or B) are indicated at the right upper corner of every line. Segments of residues that form α-helices or *β*-strands are labeled at the top of the plot. (b) Accumulated RMSF values, calculated by summing the values of each line in panel a for subunits A (black) and B (gray).

Besides that of the loops, a relative movement of the enzyme subunits can be observed. Such a movement could be described as a shrinking−stretching of the enzyme toward the cavity that hosts the water cluster in W8D, W6D, and W4D, and it could contribute to the aggregation of water molecules in that region in the simulations where the two solvents are present. The extreme conformations corresponding to the motion observed in the movie are shown in Figure 4b, where the open and closed conformations of loop 6 can be clearly appreciated (Movie S1 shows precisely the transition pathway). Similar conformations of loop 6 have been previously reported for other TIMs, although the period of the transition has been estimated to be of the order of hundreds of microseconds.[56] The observation of this movement in our simulations could be favored by the initial enzyme structure or/and by the simulation conditions. To study the dynamics of this movement, the distance between the center of mass of the three amino acids of loop 6 that present the highest displacement, and one amino acid of loop 7 that is relatively static and located in the direction of the movement, were calculated (Figure 4c). It can be seen that, in general, the distance in subunit A is higher than the distance in subunit B. The two distances are almost constant for the trajectories W4D and RD, while they are much more variable for the other three trajectories. In particular, for RW_b, loop 6 of the second subunit seems to start from an open conformation, and it changes to a closed conformation at ∼15 ns. The net displacement of the chosen

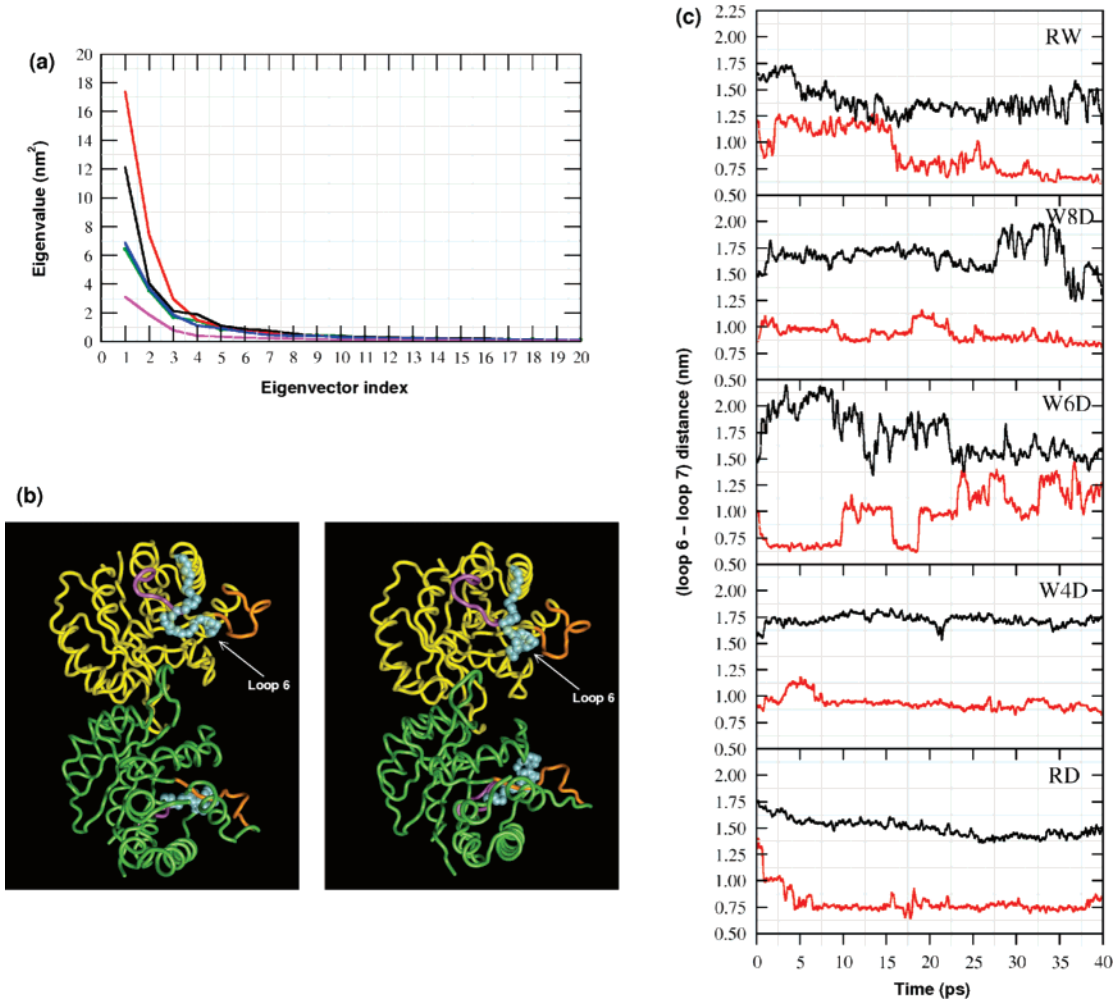

Figure 4. (a) Highest eigenvalues obtained from principal component analysis of the simulations RW_b, W8D, W6D, W4D, and RD in green, red, black, blue, and magenta, respectively. (b) The two extreme conformations of the principal TcTIM motion obtained from the trajectory RW_b. Only the backbones of subunits A (yellow) and B (green) are displayed. The residues of loops 5−7 are in orange, cyan spheres, and magenta, respectively. (c) Distance between the center of mass of the amino acids Gly170-Thr171-Gly172 (loop 6) and Val212 (loop 7) of the subunits A (black) and B (red). The name of the system is indicated at the right upper corner of every plot.

group of amino acids is ∼0.6 nm in the direction of the measurement. At least two transitions (closed−open) were observed for the same subunit in the trajectory W6D.

As described in the Materials and Methods, an analysis of structural clusters over the last 20 ns of all the trajectories was performed. The representative structures of every cluster for each trajectory were superimposed on each other to show the differences between the most visited conformations along the simulation pathways. The catalytic residues (Lys 14, His 96, and Glu 168) showed a higher mobility in the trajectories RWb, W8D, and W6D than in RD and W4D. Particularly, Lys 14 and His 96 were found to be particularly flexible in W8D and W6D, where they exposed a slight rotation around the aliphatic chain in Lys and of the carboxilic head in Glu. In spite of these features, the average structure of the site appeared to be conserved in all cases.

A detailed analysis of the protein secondary structure was performed for all the trajectories as a function of time. The number of residues that form part of the β-strands was nearly constant (∼38 for each subunit) for all the simulations. Even when the global structure of the protein after 40 ns of MD simulations was preserved, a lack of the number of residues that constitute α-helices took place in the trajectory RD (Figure 1). In this simulation, such a number fluctuated between 65 and 80 for each enzyme subunit, while for the systems containing water, it took values between 90 and 100. In particular, the shortest helices of the α/β-barrel structure (α3 and α7) in one of the TcTIM subunits, as well as α8 in the other subunit, unfolded in the trajectory RD (unfold in this context does not necessarily mean a transition to a random coil but a lack of stable H-bonds leading to a stretching or distortion of the helix. As seen in Figure 1 for the final conformation of the trajectory RD, certain order is still present in the unfolded segments, although the structure cannot be defined as a helix on the basis of the DPSS.[52]). α3 was observed to be unstable in the same subunit along the remaining four trajectories, including RW. α5 also unfolded during the trajectory W4D, again for the same subunit as for the system RD. The lack of well-defined α-helices for these segments did not appear to be critical for the stability of the protein because of their small size (5−8 residues). However, during the simulations of RD, helix α1 also unfolded in one of the protein subunits. Such a helix consists of 18 residues at the beginning of the simulations, which represent a significant percentage of the overall protein secondary structure (Figure 1).

3.3. Solvent Molecule Distribution and Dynamics. It is evident from Figure 1 that the distribution of solvent molecules around the enzyme structure was not homogeneous in the final conformation of the trajectories W8D, W6D, and W4D. However, the number of water molecules in the vicinity of every

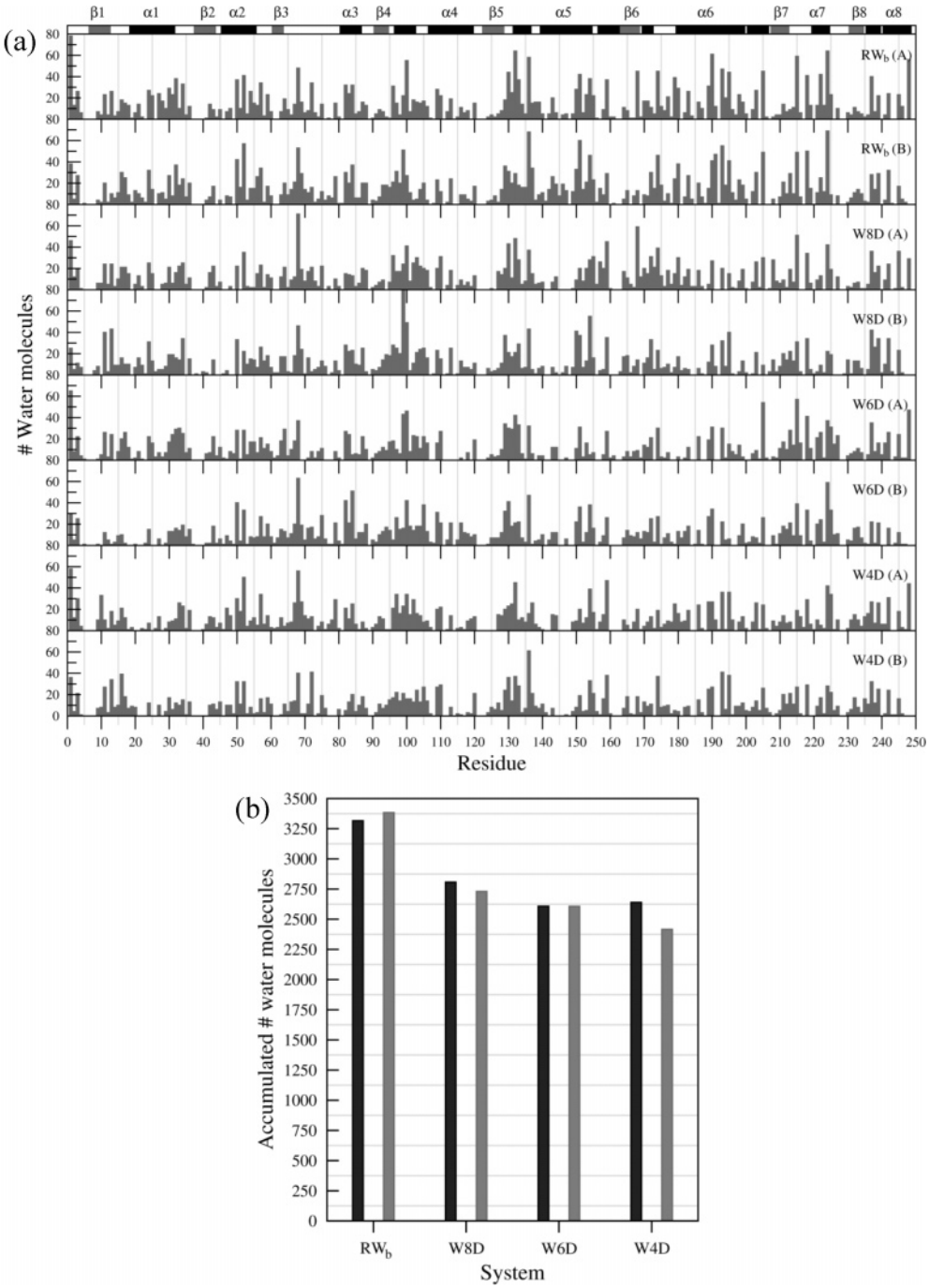

Figure 5. (a) Number of water molecules that are at a distance of less than 4 Å in the last conformation of the trajectories RW$_b$, W8D, W6D, and W4D. The name of the system and subunit (A or B) are indicated at the right upper corner of every line. Segments of residues that form α-helices or β-strands are labeled at the top of the plot. (b) Accumulated number of water molecules calculated by summing values of each line in panel a for subunits A (black) and B (gray).

protein residue was almost the same as that in RW$_b$, even though the difference in the total number of waters is important, ranging from 24 825 to 1185 (Figure 5a). The profiles for both subunits in the same system were also very similar. The accumulated number of water molecules was calculated (Figure 5b). A similar analysis was performed for the number of carbon atoms of decane molecules at less than 4 Å of every protein residue (Figure 6a). The decane atom distribution was less similar among systems and protein subunits than the water distribution. Nevertheless, several protein regions with a consistent affinity for the hydrophobic solvent could be identified such as segments of helix 6, loop 6, and loop αβ2 (Figure 6a). This result could be used to propose potential binding sites for hydrophobic ligands, as previously suggested by Mattos and Ringe.[57,58] The accumulated values of the number of decane atoms were also calculated (Figure 6b). By comparing Figures 5 and 6, it is clear that the water molecule distribution around the enzyme was more specific than the distribution of the nonpolar solvent. It is worth mentioning that, although the results presented in Figures 5 and 6 were obtained from a single conformation of each system, the same analysis was performed for a number of frames through the last 20 ns of all the trajectories with no significant variations.

The residence time values of water molecules were filtered to determine the number of waters that remained at a distance of less than 4 Å of every residue for more than 2 ns. This

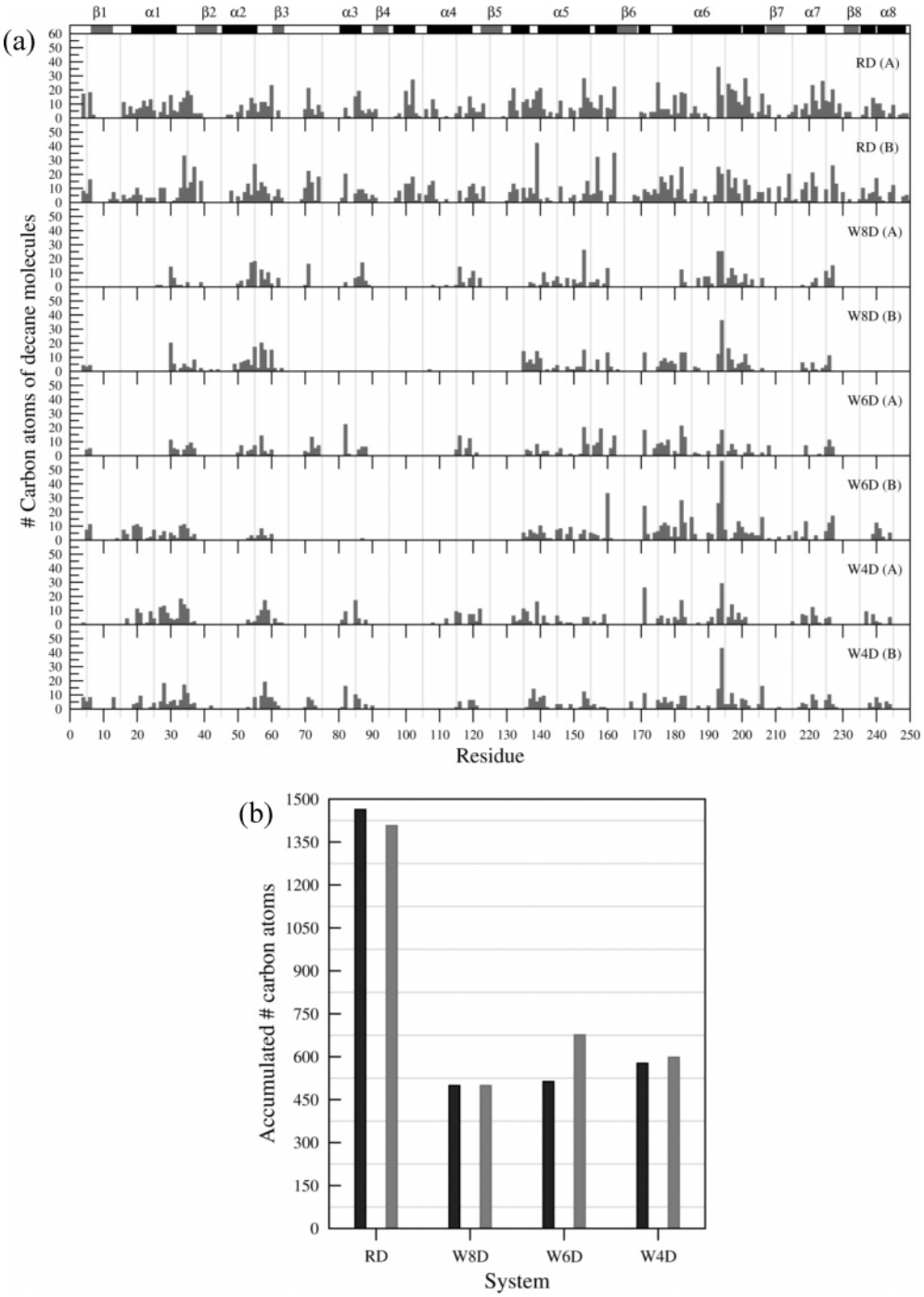

Figure 6. (a) Number of carbon atoms of decane molecules that are at a distance of less than 4 Å in the last conformation of the trajectories W8D, W6D, W4D, and RD. The name of the system and subunit (A or B) are indicated at the right upper corner of every line. Segments of residues that form α-helices or *β*-strands are labeled at the top of the plot. (b) Accumulated number of carbon atoms of decane molecules calculated by summing values of each line in panel a for subunits A (black) and B (gray).

calculation was performed for the systems RW_b, W8D, W6D, and W4D over the last 20 ns of every trajectory (Figure 7a). It could be appreciated that there were less water molecules with residence times higher than 2 ns with pure water than for the simulations with water/decane mixtures. This is reasonable due to the low partition coefficient of water in decane and also because the systems W8D, W6D, and W4D do not have many water molecules available to interact with the residues that indeed need to be hydrated. The corresponding accumulated numbers of water molecules were also calculated (see Figure 7b). The increment of the number of water molecules with long residence times when the water concentration decreases is clear. An analogous analysis was performed with a longer cutoff for the residence time, the results being similar although with a lower number of water molecules in the vicinity of each residue (data not shown).

Aproximately 15 water molecules with different residence times were found to be located in the dimer interface of the enzyme in the trajectories RW_b, W8D, W6D, and W4D, and a variable number of water molecules seeped into the active site. The presence of water molecules in contact with the catalytic residues was associated with an open conformation of loop 6. This supports the theory that such a loop modulates the entrance of the substrate to the catalytic residues. In the simulation of

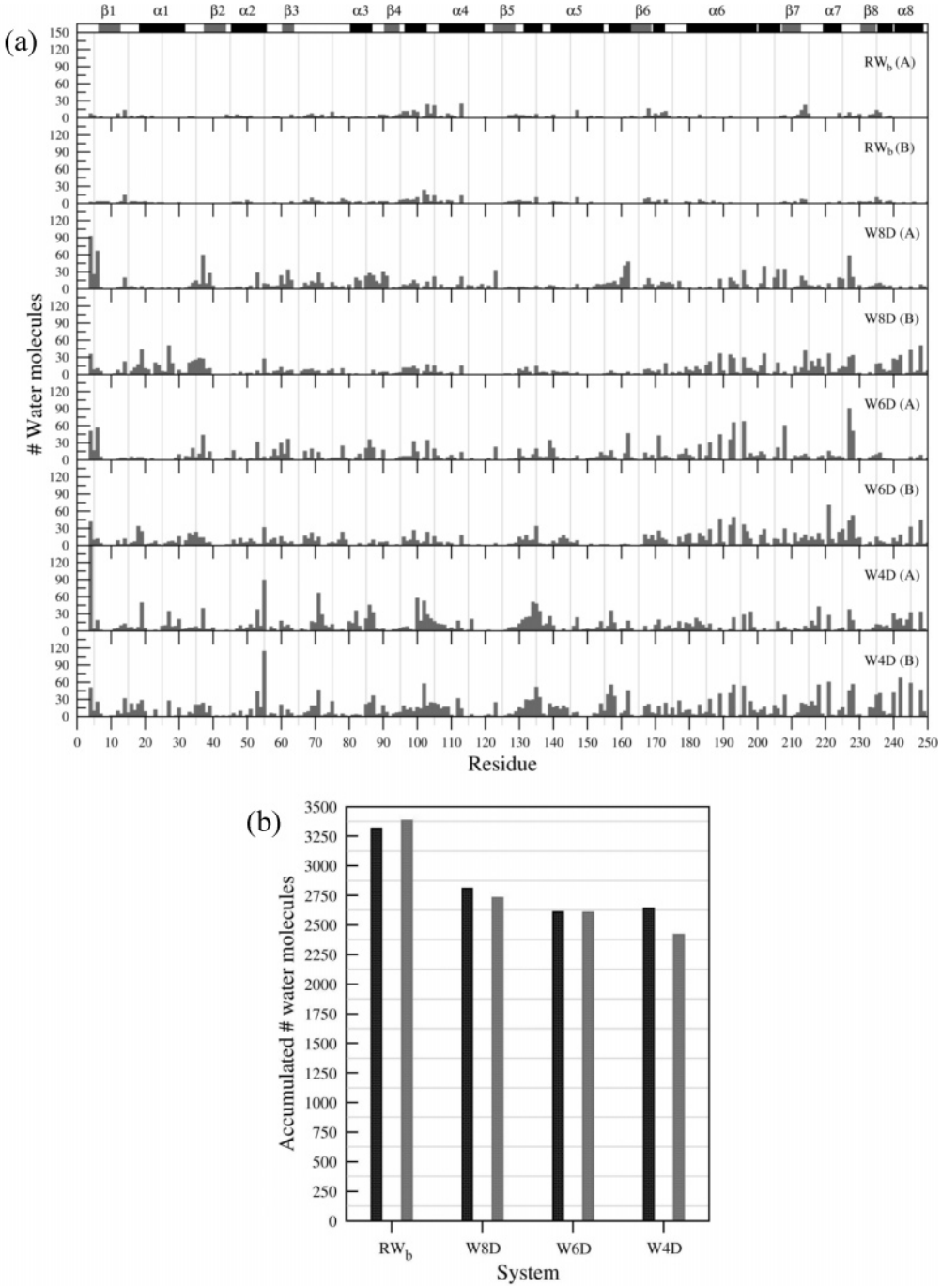

Figure 7. (a) Number of water molecules that are at a distance of less than 4 Å from every TcTIM residue, which remain for at least 2 ns during the last 20 ns of the trajectories RW$_b$, W8D, W6D, and W4D. The name of the system and subunit (A or B) are indicated at the right upper corner of every line. Segments of residues that form α-helices or β-strands are labeled at the top of the plot. (b) Accumulated number of water molecules calculated by summing values of each line in panel a for subunits A (black) and B (gray).

RD, no decane molecules were found to be close to the interface nor to the active site.

3.4. Intramolecular Hydrogen Bonds. It has been proposed that many enzymes maintain their mobility and activity at low water concentrations because, even under those conditions, they can form hydrogen bonds with the water molecules of their environment.[10] In the absence of water, the polar groups of the protein switch to interact with each other, forming intramolecular H-bonds. The numbers of these H-bonds for trajectories RW$_b$, W8D, W6D, W4D, and RD are shown in Figure 8a. The plot is very clear, showing that all the simulations involving water exhibited almost the same number of intramolecular H-bonds (see average values in Table 1). However, for the simulation with decane as a unique solvent, there were no water molecules available, and many more intramolecular H-bonds were formed. The same calculation was performed considering only the amino acids that form part of β-strands or α-helices, obtaining similar results (Figure 8b). Also in the case of RD, the increment in the number of H-bonds resulted in a contraction of the TcTIM structure, which was quantified by measuring the RoG. The average values of these radii over the last 20 ns of every trajectory are reported in Table 1. The hydrophobic contributions to the SAS did not vary significantly from one system to another, but the hydrophilic SAS was much lower in RD than in any of the other systems. Thus, the contraction of the protein structure in RD involved the formation of H-bonds between the polar

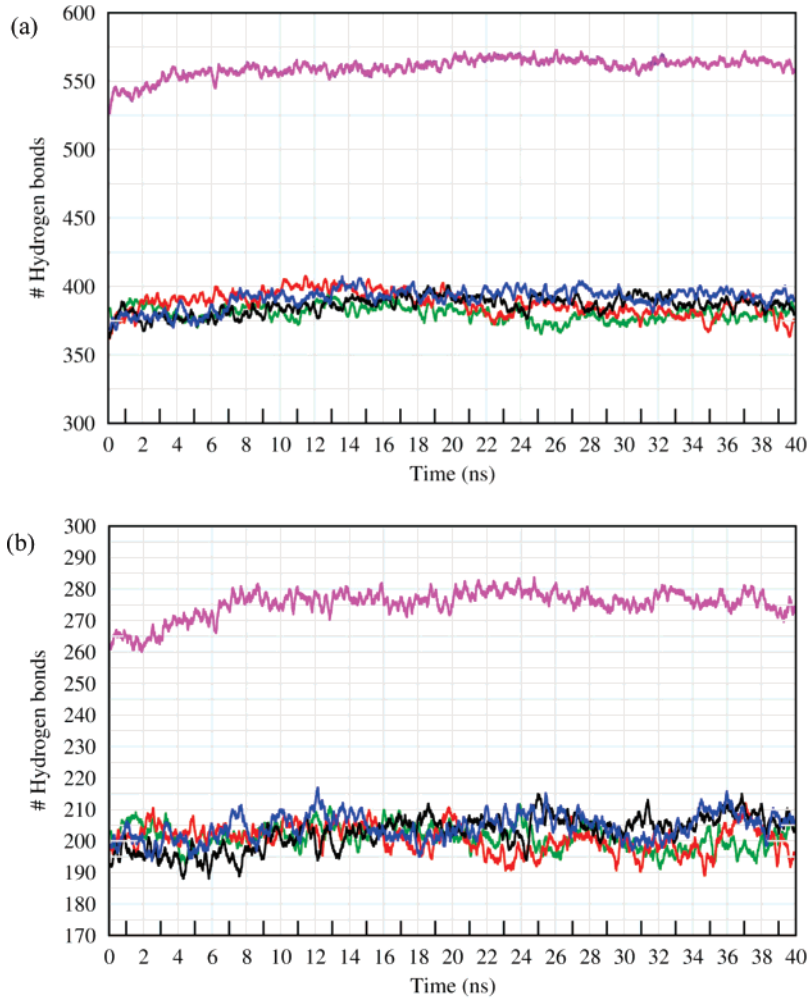

Figure 8. Number of intramolecular hydrogen bonds of (a) all TcTIM residues and (b) the TcTIM residues that form part of α-helices and *β*-strands, as a function of time, for the trajectories RW_b (green), W8D (red), W6D (black), W4D (blue), and RD (magenta).

hydrophilic groups. Simultaneously, those groups internalized in the protein to escape from the hydrophobic solvent. Unexpectedly, the enzyme also reduced its size throughout the simulation of W4D. Such a contraction cannot be explained in the same terms as for RD because the reduction of hydrophilic SAS in this system was almost negligible (see Table 1), and the numbers of H-bonds, as well as the distribution of water molecules in the vicinity of the residues, were similar to the values found for RW_b, W8D, and W6D. The distance between the centers of mass of both subunits was calculated for all the systems, their values being consistently lower for W4D and RD than for RW_b, W8D, and W6D (Table 1). Thus, the reduction of RoG in W4D was due to the two subunits approaching each other, maintaining the hydrophilic and hydrophobic SAS values. This approach also contributed to the decrease of RoG in RD, and it could be related to the previously mentioned shrinking−stretching of the enzyme structure toward the cavity. Such a motion could not be reversible due to the pressure of the decane molecules and to the low number of waters present in this simulation.

4. Conclusion

The effects of the water on the decane concentration ratio in the structure and dynamics of TcTIM, as well as the behavior of the solvent molecules in the vicinity of the enzyme, were studied in detail through MD simulations at a time scale of 40 ns. In contrast to recent MD studies performed at much shorter time scales,[27] the enzyme was found to be sensitive to the water concentration. The analysis of the protein fluctuations and of the PC of the backbone motion revealed that, in the presence of organic solvents, the higher the water concentration, the higher the mobility of the enzyme. Nevertheless, TcTIM in pure water showed amplitudes comparable to those in W4D, the simulation with the lowest water concentration, while it was clearly more rigid in pure decane. These results suggest that the enzyme dynamics speeds up at relatively low water concentrations, probably due to the presence of organic solvent molecules attached to specific sites of the protein surface, although there is a threshold in the number of water molecules needed for the protein to keep its structure and dynamics. Our analysis also provides new insight into the dynamics of TcTIM that could be helpful in better understanding this enzyme.

Acknowledgment. We thank Prof. Armando Gómez-Puyou, Universidad Nacional Autónoma de México (UNAM), Dr. Pilar Brocos (USC), Prof. Miguel Costas (UNAM), and Dr. Hao Fan (UCSF) for their critical reading of the manuscript and for very useful suggestions. This work was supported by Grants 41328-Q and J49811-Q from CONACyT and Grant IN105107 from PAPIIT-UNAM, México. We are grateful to the Dirección General de Servicios de Cómputo Académico (DGSCA) of the UNAM for computer time employed in this work and for their excellent services. N.D.-V. is grateful for financial support from CONACyT.

Supporting Information Available: Movie S1, showing the projection of the trajectory RW_b on the eigenvector with the highest eigenvalue, obtained by PC analysis. This information is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Tuena de Gomez-Puyou, M.; Gomez-Puyou, A. *Crit. Rev. Biochem. Mol. Biol.* **1998**, *33*, 53.

(2) Finney, J. L.; Poole, P. L. *Commun. Mol. Cell. Biophys.* **1984**, *2*, 129.

(3) Levy, Y.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3325.

(4) Pal, S. K.; Zewail, A. H. *Chem. Rev.* **2004**, *104*, 2099.

(5) Rupley, J. A.; Careri, G. *Adv. Protein Chem.* **1991**, *41*, 37.

(6) Rupley, J. A.; Gratton, E.; Careri, G. *Trends Biochem. Sci.* **1983**, *8*, 18.

(7) Zaks, A.; Klibanov, A. M. *Science (Washington, DC, U.S.)* **1984**, *224*, 1249.

(8) Zaks, A.; Klibanov, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 3192.

(9) Zaks, A.; Klibanov, A. M. *J. Am. Chem. Soc.* **1986**, *108*, 2767.

(10) Zaks, A.; Klibanov, A. M. *J. Biol. Chem.* **1988**, *263*, 8017.

(11) Zaks, A.; Klibanov, A. M. *J. Biol. Chem.* **1988**, *263*, 3194.

(12) Gonnelli, M.; Strambini, G. B. *J. Phys. Chem.* **1988**, *92*, 2854.

(13) Strambini, G. B.; Gonnelli, M. *J. Phys. Chem.* **1988**, *92*, 2850.

(14) Bell, G.; Halling, P. J.; Moore, B. D.; Partridge, J.; Rees, D. G. *Trends Biotechnol.* **1995**, *13*, 468.

(15) Dordick, J. S. *Appl. Biochem. Biotechnol.* **1988**, *19*, 103.

(16) Dordick, J. S. *Enzyme Microb. Technol.* **1989**, *11*, 194.

(17) Klibanov, A. M. *CHEMTECH* **1986**, *16*, 354.

(18) Klibanov, A. M. *Trends Biotechnol.* **1997**, *15*, 97.

(19) Soares, C. M.; Teixeira, V. H.; Baptista, A. M. *Biophys. J.* **2003**, *84*, 1628.

(20) Micaelo, N. M.; Teixeira, V. H.; Baptista, A. M.; Soares, C. M. *Biophys. J.* **2005**, *89*, 999.

(21) Colombo, G.; Toba, S.; Merz, K. M. *J. Am. Chem. Soc.* **1999**, *121*, 3486.

(22) Norin, M.; Haeffner, F.; Hult, K.; Edholm, O. *Biophys. J.* **1994**, *67*, 548.

(23) Peters, G. H.; van Aalten, D. M. F.; Edholm, O.; Toxvaerd, S.; Bywater, R. *Biophys. J.* **1996**, *71*, 2245.

(24) Toba, S.; Hartsough, D. S.; Merz, K. M. *J. Am. Chem. Soc.* **1996**, *118*, 6490.

(25) Zheng, Y. J.; Ornstein, R. L. *J. Am. Chem. Soc.* **1996**, *118*, 4175.

(26) Zheng, Y. J.; Ornstein, R. L. *Protein Eng.* **1996**, *9*, 485.

(27) Yang, L.; Dordick, J. S.; Garde, S. *Biophys. J.* **2004**, *87*, 812.

(28) Micaelo, N. M.; Soares, C. M. *FEBS Lett.* **2007**, *274*, 2424.

(29) Zheng, Y. J.; Ornstein, R. L. *Biopolymers* **1996**, *38*, 791.

(30) Maldonado, E.; Soriano-Garcia, M.; Moreno, A.; Cabrera, N.; Garza-Ramos, G.; de Gomez-Puyou, M. T.; Gomez-Puyou, A.; Perez-Montfort, R. *J. Mol. Biol.* **1998**, *283*, 193.

(31) Garza-Ramos, G.; Tuena de Gomez-Puyou, M.; Gomez-Puyou, A.; Gracy, R. W. *Eur. J. Biochem.* **1992**, *208*, 389.

(32) Gag, X. G.; Maldonado, E.; Perez-Montfort, R.; Garza-Ramos, G.; Tuena de Gomez-Puyou, M.; Gomez-Puyou, A.; Rodriguez-Romero, A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10062.

(33) Garza-Ramos, G.; Cabrera, N.; Saavedra-Lira, E.; Tuena de Gomez-Puyou, M.; Ostoa-Saloma, P.; Perez-Montfort, R.; Gomez-Puyou, A. *Eur. J. Biochem.* **1998**, *253*, 684.

(34) Perez-Montfort, R.; Garza-Ramos, G.; Alcantara, G. H.; Reyes-Vivas, H.; Gao, X. G.; Maldonado, E.; Tuena de Gomez-Puyou, M.; Gomez-Puyou, A. *Biochemistry* **1999**, *38*, 4114.

(35) Reyes-Vivas, H.; Hernandez-Alcantara, G.; Lopez-Velazquez, G.; Cabrera, N.; Perez-Montfort, R.; Tuena de Gomez-Puyou, M.; Gomez-Puyou, A. *Biochemistry* **2001**, *40*, 3134.

(36) Tellez-Valencia, A.; Avila-Rios, S.; Perez-Montfort, R.; Rodriguez-Romero, A.; Tuena de Gomez-Puyou, M.; Lopez-Calahorra, F.; Gomez-Puyou, A. *Biochem. Biophys. Res. Commun.* **2002**, *295*, 958.

(37) Tellez-Valencia, A.; Olivares-Illana, V.; Hernandez-Santoyo, A.; Perez-Montfort, R.; Costas, M.; Rodriguez-Romero, A.; Lopez-Calahorra, F.; Tuena de Gomez-Puyou, M.; Gomez-Puyou, A. *J. Mol. Biol.* **2004**, *341*, 1355.

(38) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43.

(39) http://www.gromacs.org.

(40) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306.

(41) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701.

(42) Schuler, L. D.; van Gunsteren, W. F. *Mol. Simul.* **2000**, *25*, 301.

(43) Daura, X.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **1998**, *19*, 535.

(44) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. *Interaction Models for Water in Relation to Protein Hydration*; Reidel D. Publishing Company: 1981.

(45) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(46) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952.

(47) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. *J. Comput. Chem.* **1997**, *18*, 1463.

(48) Sayle, R. A.; Milnerwhite, E. J. *Trends Biochem. Sci.* **1995**, *20*, 374.

(49) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.

(50) http://pymol.sourceforge.net/.

(51) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *Angew. Chem., Int. Ed.* **1999**, *38*, 236.

(52) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.

(53) Rozovsky, S.; Jogl, G.; Tong, L.; McDermott, A. E. *J. Mol. Biol.* **2001**, *310*, 271.

(54) Rozovsky, S.; McDermott, A. E. *J. Mol. Biol.* **2001**, *310*, 259.

(55) Daggett, V. *Chem. Rev.* **2006**, *106*, 1898.

(56) Wierenga, R. K.; Noble, M. E.; Davenport, R. C. *J. Mol. Biol.* **1992**, *224*, 1115.

(57) Mattos, C.; Ringe, D. *Nat. Biotechnol.* **1996**, *14*, 595.

(58) Ringe, D. *Curr. Opin. Struct. Biol.* **1995**, *5*, 825.