

# Protein Dynamics in Organic Media at Varying Water Activity Studied by Molecular Dynamics Simulation

Rasmus Wedberg,<sup>†,⊥</sup> Jens Abildskov,<sup>\*,†</sup> and Günther H. Peters<sup>\*,‡,§</sup>

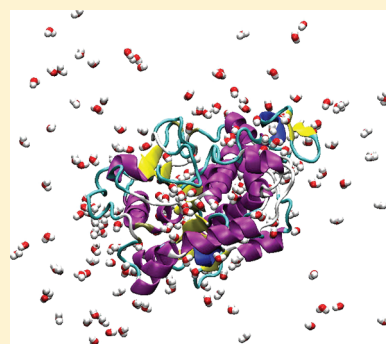
<sup>†</sup>Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søltofts Plads, Building 229, DTU, 2800 Kongens Lyngby, Denmark

<sup>‡</sup>Department of Chemistry, Technical University of Denmark, Kemitorvet, Building 207, DTU, 2800 Kongens Lyngby, Denmark

<sup>§</sup>MEMPHYS—Center for Biomembrane Physics

## S Supporting Information

**ABSTRACT:** In nonaqueous enzymology, control of enzyme hydration is commonly approached by fixing the thermodynamic water activity of the medium. In this work, we present a strategy for evaluating the water activity in molecular dynamics simulations of proteins in water/organic solvent mixtures. The method relies on determining the water content of the bulk phase and uses a combination of Kirkwood–Buff theory and free energy calculations to determine corresponding activity coefficients. We apply the method in a molecular dynamics study of *Candida antarctica* lipase B in pure water and the organic solvents methanol, *tert*-butyl alcohol, methyl *tert*-butyl ether, and hexane, each mixture at five different water activities. It is shown that similar water activity yields similar enzyme hydration in the different solvents. However, both solvent and water activity are shown to have profound effects on enzyme structure and flexibility.



## 1. INTRODUCTION

Organic solvents are widely appreciated as reaction media for enzyme catalysis<sup>1</sup> and, for example, frequently employed for increasing solubilities of substrate and products, shifting of reaction equilibria toward products, and eliminating side reactions with water. The solvent has profound effects on several reaction parameters. Enzymatic activity is typically several orders of magnitude lower in organic media, as compared to in water. Activity depends on the nature of organic solvents<sup>2</sup> and has been shown to correlate with solvent hydrophobicity measures, e.g.,  $\log P$ .<sup>3–5</sup> The medium also affects enzyme specificity.<sup>6,7</sup> This has important implications, since specificity can be a key property in biocatalysis. Enzyme stability is as well medium-dependent, and in some organic solvents, dramatically higher enzyme stability has been observed at high temperature, as compared to in water.<sup>8</sup> In addition to the effects of organic solvent, enzyme hydration has significant impact on catalytic properties.<sup>9</sup> In particular, enzymes are typically inactivated upon complete dehydration, which has been ascribed to the fact that a layer of water molecules at the enzyme surface is required for catalytic activity to be retained in organic media.

The mechanisms underlying solvent and hydration effects on enzyme activity are not yet fully understood, although several hypotheses have been discussed in the literature. These can be grouped in three major effects. First, the medium affects protein flexibility which has been probed using electron spin resonance<sup>10,11</sup> and time-resolved fluorescence anisotropy.<sup>12</sup> In those studies, flexibility was correlated with activity, and it was proposed that the water layer at the protein surface would act

as a lubricant promoting flexibility and thus activity. Second, water and organic solvent molecules bind to the active site acting as inhibitors.<sup>13–18</sup> Third, organic media differ in their ability to stabilize the strongly polar transition state.<sup>19,20</sup> In addition, medium effects on enzyme specificity have to some extent been rationalized in terms of substrate solvation thermodynamics.<sup>21–24</sup>

Interactions between protein, organic solvent, and hydration layer are difficult to probe experimentally, and molecular dynamics (MD) simulations have emerged as a useful complement to experimental studies for understanding molecular phenomena.<sup>25</sup> Several computational studies have investigated the effects of nonaqueous solvent and hydration level on protein properties such as structure and dynamics. Early studies showed that protein flexibility was lower in organic solvents than in water, but revealed no significant differences in backbone structure, possibly due to the limited simulation lengths.<sup>26–33</sup>

More recent studies have recognized the role of hydration and studied its effect on protein structure and flexibility. Yang and co-workers carried out simulations of surfactant-solubilized subtilisin BPN' at different hydration levels and found no significant effect.<sup>34</sup> On the contrary, Soares and co-workers performed simulations of *Fusarium solani* cutinase and found that the mean-square deviation with respect to the crystal structure had a “U”-shaped dependence on the hydration

**Received:** November 16, 2011

**Revised:** January 24, 2012

**Published:** February 6, 2012



Table 1. Simulations of CALB Listing Identifier (ID), Number of Water Molecules, Hydration Level, and Water Activity

hexane				MTBE			
ID	no. water <sup>a</sup>	hydration <sup>b,c</sup>	$a_w^{c,d}$	ID	no. water <sup>a</sup>	hydration <sup>b,c</sup>	$a_w^{c,d}$
H43	43	42.92 <sub>2</sub>	0.02 <sub>1</sub>	E50	50	48.7 <sub>5</sub>	0.07 <sub>3</sub>
H87	87	85.3 <sub>4</sub>	0.04 <sub>1</sub>	E100	100	94.6 <sub>5</sub>	0.13 <sub>2</sub>
H171	171	159 <sub>1</sub>	0.4 <sub>2</sub>	E171	171	152 <sub>2</sub>	0.39 <sub>5</sub>
H335	335	259.4 <sub>5</sub>	1.0 <sub>2</sub>	E335	335	249 <sub>3</sub>	1.1 <sub>1</sub>
H500	500	330 <sub>3</sub>	<i>e</i>	E500	500	310 <sub>7</sub>	<i>e</i>
methanol				tert-butyl alcohol			
ID	no. water <sup>a</sup>	hydration <sup>b,c</sup>	$a_w^{c,d}$	ID	no. water <sup>a</sup>	hydration <sup>b,c</sup>	$a_w^{c,d}$
M210	210	33 <sub>2</sub>	0.038 <sub>1</sub>	T65	65	46 <sub>2</sub>	0.024 <sub>4</sub>
M940	940	84 <sub>1</sub>	0.177 <sub>4</sub>	T130	130	76 <sub>3</sub>	0.068 <sub>6</sub>
M2500	2500	172 <sub>1</sub>	0.411 <sub>7</sub>	T210	210	98 <sub>4</sub>	0.14 <sub>1</sub>
M4200	4200	255 <sub>2</sub>	0.583 <sub>8</sub>	T700	700	187 <sub>2</sub>	0.457 <sub>7</sub>
M6135	6135	326 <sub>3</sub>	0.717 <sub>6</sub>	T2970	2970	335 <sub>2</sub>	0.874 <sub>6</sub>

<sup>a</sup>Total number of water molecules in simulation box. <sup>b</sup>Average number of water molecules within 3.5 Å of CALB. <sup>c</sup>Subscripts indicate uncertainty in the last digit, estimated from three replica simulations of each CALB/organic solvent/water system. <sup>d</sup>Determined from the simulations as explained in the text. <sup>e</sup>Cluster formation of water molecules in nonpolar solvent.

level<sup>35–37</sup> and that cutinase flexibility increased with increasing hydration. Similar observations were made by Díaz-Vergara and Piñeiro<sup>38</sup> who carried out simulations of *Trypanosoma cruzi* triosephosphate isomerase in decane/water mixtures. Cruz and co-workers also observed effects of different hydration levels in simulations of subtilisin Carlsberg.<sup>39</sup>

Trodler and Pleiss<sup>40</sup> carried out simulations of *Candida antarctica* lipase B (CALB) in water and organic solvents to correlate structure and flexibility with solvent properties. The organic solvent simulations included the crystal water molecules at the protein surface. An apparent correlation between protein flexibility and solvent log *P* was found. The authors suggested that in hydrophobic solvents, more immobile water molecules would be present at the protein surface, which in turn would restrict protein flexibility. The simulations included the same number of water molecules in the different organic solvent simulations. This resulted however in different enzyme hydration as polar solvents strip water from the enzyme surface as observed in several studies.<sup>34,37,39,40</sup>

Several authors have argued that hydration effects should be studied in terms of thermodynamic water activities,<sup>41–46</sup> since the size of the hydration layer supposedly depends on the water activity and only to a lesser extent on the organic solvent.<sup>42</sup> Valivety and co-workers<sup>44</sup> measured the activity of *Mucor miehei* lipase as a function of water activity in nonpolar organic solvents. The authors found a bell-shaped dependence with maximum enzymatic activity obtained at a water activity of 0.55, being independent of the solvent. Enzymatic activity was also rather insensitive to the solvent, as long as the water activity was kept at a fixed value. A follow-up study by Bell and co-workers<sup>46</sup> did however show that the choice of nonpolar vs polar solvent had a significant impact on the catalytic activity at fixed water activity due to interactions between enzyme and solvent molecules.

It seems far preferable to compare solvents at constant water activity. Since this has normally not been done, mechanistic interpretation of literature data for low-water systems becomes difficult. To our knowledge, there are no simulation studies of enzymes in organic media in which the medium water activity is considered explicitly. Branco and co-workers<sup>47</sup> simulated CALB in gaseous water/argon mixtures and evaluated water activity by

assuming ideality. This approach is not applicable in organic solvents whose aqueous mixtures are nonideal.

The current study describes MD simulations of CALB in water and four organic solvents, namely, methanol, *tert*-butyl alcohol, methyl *tert*-butyl ether (MTBE), and hexane. The solvents, except for methanol, are commonly used in nonaqueous biocatalytic systems with CALB.<sup>48</sup> Tertiary alcohols are sometimes quoted as the preferred solvents since CALB has been found to be especially stable in these solvents.<sup>49–52</sup> Calculations of substrate solvation thermodynamics suggest that high conversion should be possible in MTBE for some reactions. Methanol is not used as a solvent for CALB as it has been observed to inactivate the enzyme.<sup>50,53</sup> By also considering methanol, the solvents studied cover a wide range of water/solvent nonideality. Experiments have shown that catalytic activity of CALB in organic media depends on the water activity. Usually, the activity is maximal at very low water activity,<sup>18,54–58</sup> though a bell-shaped dependence is seen for certain reactions.<sup>59,60</sup> However, the effects of solvents and changes in water activities are not fully understood on the molecular scale. Therefore, we present an MD study to gain further insight into the interplay among protein structure/dynamics, protein hydration, and solvent properties. First, a method is presented that allows determination of the bulk water activity in MD simulations of proteins in nonaqueous media. With the methodology in place, the effect of different solvents and water activity on the structure and dynamics of CALB is investigated.

## 2. METHODS

**2.1. Molecular Dynamics Simulations.** MD simulations were carried out of CALB in hexane, MTBE, methanol, *tert*-butyl alcohol, and pure water. For each of the four organic solvents, simulations were carried out at five different hydration levels corresponding to water activities in the range from 0 to 1 depending on the organic solvent (Table 1). Analyses focus on structure and dynamics of both CALB and solvent species at the protein surface.

The crystal structure coordinates of CALB were obtained from the protein data bank.<sup>61</sup> As several structures are available, the best resolved one, PDB entry 1TCA,<sup>62</sup> which has a resolution of 1.55 Å, was used. The entry includes coordinates

for CALB and 286 crystal water molecules. For simulations including fewer water molecules, those with the lowest B-factors were retained. For simulations including 286 or more water molecules, all crystal waters were retained, and additional water molecules were introduced using VMD in combination with the plug-in SOLVATE.<sup>63</sup> For the organic solvent simulations, the CALB–water complex was placed in a cubic simulation cell containing organic solvent. The organic solvent molecule coordinates were taken from the last frame of an MD simulation of pure hexane, MTBE, methanol, or *tert*-butyl alcohol of at least 500 ps. Solvent molecules closer than 2.5 Å to the CALB/crystal water complex were removed. For the simulations of CALB in pure water, the entire simulation box was built using SOLVATE. The numbers of water molecules included in each simulation are listed in Table 1. The simulations of CALB carried out in pure water included 9500 water molecules. In all simulations, it was ensured that enough solvent molecules were included in the simulation cell, such that the protein did not interact with its periodic images.

For the water-miscible solvents methanol and *tert*-butyl alcohol, the number of water molecules in each system was selected in order to achieve water activities spanning the range from 0 to 1. The number of water molecules required to attain a given activity is not known a priori, since the molecules at equilibrium are distributed over the bulk phase and the protein surface, as discussed in section 3.1. The selection was therefore guided by preliminarily performed CALB simulations and activity calculations using UNIFAC.<sup>64</sup> As will be shown in section 3.1, the prepared systems span the activity range quite well for all four organic solvents.

Like other lipases, CALB has a catalytic center with a trypsin-like triad (Ser...His...acidic residue).<sup>65,66</sup> The active site histidine (His224) was defined as neutral with the proton placed on N<sub>δ1</sub>, allowing for the essential hydrogen bond with Asp187.

Asp134 was also defined as neutral, in accordance with the suggestion of Uppenberg and co-workers.<sup>62</sup> The pK<sub>a</sub> of Asp134 was estimated as 7.25 using PROPKA 2.0,<sup>67,68</sup> which furthermore supports the treatment of Asp134 as neutral. Moreover, structural inspection revealed that the residues Asp134, Thr40, Ser153, and Gln157 lie in a plane and their orientation is stabilized by a hydrogen network: Thr40(O<sub>γ1</sub>)–Gln157(N<sub>ε2</sub>) (3.1 Å), Thr40(O<sub>γ1</sub>)–Gln157(O<sub>ε1</sub>) (3.6 Å), Asp134(O<sub>δ2</sub>)–Gln157(O<sub>ε1</sub>) (2.6 Å), Ser153(O)–Gln157(N<sub>ε2</sub>) (3.1 Å). The close distance between Asp134(O<sub>δ2</sub>) and Gln157(O<sub>ε1</sub>) suggests that Asp134 is protonated. The hydrogen network appears essential for the structure of the oxyanion hole, in which the side chain of Thr40 participates. We also investigated if rotation of Gln157 would support a charged Asp134. However, reorientation of Gln157 will result in unfavorable interactions between Ser153(O) and Thr40(O<sub>γ1</sub>) making this orientation unlikely.

All remaining Arg, Asp, Glu, and Lys residues were charged according to their protonation state at pH 7 resulting in a total protein charge of zero. The ionic strength was zero in all simulations. The CHARMM27 force field<sup>69,70</sup> was used to model the protein, as well as hexane, methanol, and *tert*-butyl alcohol molecules, while the CHARMM35 ether force field<sup>71</sup> was used for MTBE. For water, the TIP3P model with flexible bonds<sup>69</sup> was employed. For *tert*-butyl alcohol and MTBE, missing parameters for the tertiary C atoms were taken from similar atom types and are listed in Tables S1–S5 (Supporting Information).

The simulations were carried out in the NPT ensemble (atom number, pressure, and temperature constant) using the MD simulation program NAMD.<sup>72</sup> The velocity Verlet algorithm with a 1 fs step size was employed to integrate the equations of motion. Lennard-Jones forces were evaluated using a 12 Å cutoff in combination with a switching function starting at 10 Å that ensures that the potential smoothly approaches zero at 12 Å. Nonbonded forces were evaluated using a pair list with an outer radius of 14 Å. Periodic boundary conditions were employed in the *x*, *y*, and *z* directions, and electrostatic forces were evaluated using the particle mesh Ewald summation method with a grid spacing smaller than 1 Å. Temperature and pressure were maintained at 323.15 K (50 °C) and 1 atm, respectively, using the Langevin thermostat with a damping constant of 5 ps<sup>−1</sup> and the Langevin piston with a period of 200 fs and a decay constant of 500 fs. Coordinates were saved every 500 fs.

Prior to simulation, a 500 step conjugate gradient minimization of the configurational energy was carried out. During minimization, the C<sub>α</sub> atoms were constrained to their crystal structure positions. The C<sub>α</sub> atoms were constrained also during the first nanosecond of simulation, which was followed by another nanosecond run with the C<sub>α</sub> atoms restrained by a harmonic potential with a force constant of 1 kcal/mol/Å<sup>2</sup>. This procedure allowed for relaxation of the solvent molecules in contact with the protein surface. This was followed by unconstrained simulation of 18 ns, of which the final 10 ns were used for analysis.

Three replica simulations starting from different initial velocities were carried out of each system in order to estimate statistical uncertainties in the results. For the pure water system, five such replica simulations were carried out.

**2.2. Activity Coefficients for Binary Water/Organic Solvent Mixtures.** The water activity, *a<sub>w</sub>*, is given by *a<sub>w</sub>* = *γ<sub>w</sub>*(*x<sub>w</sub>*)/*γ<sub>w</sub>*<sup>∞</sup>, where *γ<sub>w</sub>*(*x<sub>w</sub>*) denotes the water activity coefficient in the binary water/organic solvent mixture at composition *x<sub>w</sub>*. For MTBE and hexane, *γ<sub>w</sub>*(*x<sub>w</sub>*) was approximated by *γ<sub>w</sub>*<sup>∞</sup>, which is the activity coefficient of water at infinite dilution. This is justified since all values of bulk *x<sub>w</sub>* in hexane and MTBE in this study were less than 0.02.

Values for *γ<sub>w</sub>*<sup>∞</sup> in the four organic solvents were determined by an alchemical free energy perturbation approach,<sup>73</sup> where a single water molecule was built in a simulation cell containing pure methanol (1000), *tert*-butyl alcohol (600), MTBE (500), hexane (500), or water (999), where the number of molecules in the simulation cell is given in parentheses. The systems were equilibrated for at least 500 ps before the free energy calculations were initiated. The simulations were carried out with the interaction parameter *λ* set to 0, 0.05, 0.1, ..., 0.95, 0.975, 0.99, 0.999, and 1. Lennard-Jones interactions were gradually decoupled as *λ* changed from 0 to 1, employing soft core scaling<sup>74</sup> with a shift parameter of 5 Å<sup>2</sup>. Electrostatic interactions were gradually decoupled as *λ* changed from 0 to 0.5. At each *λ* value, the system was equilibrated for 10 ps, and the production periods were 400 ps when *λ* < 0.5 and 200 ps when *λ* ≥ 0.5. For each system, two simulations which were carried out with *λ* either increasing or decreasing were consistent within the statistical uncertainties. The simulations yielded the Gibbs energy Δ*G*<sub>sol</sub> for introducing a water molecule in the different solvents. The activity can be expressed



via the difference in water chemical potential in mixture,  $\mu_w$ , and in pure water,  $\mu_w^o$ , as

$$k_B T \ln x_w \gamma_w = \mu_w - \mu_w^o = \Delta G_{\text{solv}} - \Delta G_{\text{water}} + k_B T \ln \frac{\rho_w}{\rho_w^o} \quad (1)$$

where  $k_B$ ,  $T$ ,  $\rho_w$ , and  $\rho_w^o$  respectively denote Boltzmann constant, temperature, and water density in a mixture and in pure water. In the second equality, Widom's expression for the chemical potential<sup>75</sup> is employed. In the limit as  $x_w \rightarrow 0$ , eq 1 becomes

$$\ln \gamma_w^\infty = \frac{\Delta G_{\text{solv}} - \Delta G_{\text{water}}}{k_B T} + \ln \frac{\rho_{\text{solv}}^o}{\rho_w^o} \quad (2)$$

Calculated values of  $\ln \gamma_w^\infty$  are listed in Table 2.

**Table 2.** Density,  $\ln \gamma_w^\infty$ , and Viscosity for the Studied Solvents

solvent	density (mol/dm <sup>3</sup> )		$\ln \gamma_w^\infty$		viscosity (10 <sup>-3</sup> Pa·s)
	MD <sup>a</sup>	exp. <sup>b</sup> [88]	MD <sup>a</sup>	UNIFAC <sup>c</sup> [64]	
hexane	6.92	7.61	8.82 <sub>6</sub>	6.7	0.29
methanol	23.3	24.6	0.48 <sub>9</sub>	0.5	0.54
MTBE	8.04	8.35	4.82 <sub>8</sub>	2.5	0.33
<i>tert</i> -butyl alcohol	10.8	10.6	1.5 <sub>2</sub>	1.3	4.57
water	56.5	55.3	—	—	0.91

<sup>a</sup>Simulation value obtained in present work, subscripts indicating uncertainty in last digit. <sup>b</sup>Experimental values measured at 298 K taken from the DIPPR database.<sup>88</sup> <sup>c</sup>UNIFAC<sup>64</sup> predictions at 323 K.

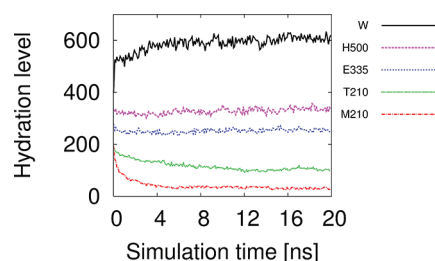
For methanol and *tert*-butyl alcohol, values of  $\ln \gamma_w$  for nondilute mixtures were obtained by employing an excess Gibbs energy ( $G^E$ ) model for the corresponding binary mixture. The  $G^E$  models were obtained using a fluctuation solution theory<sup>76,77</sup> approach to analyze MD simulations of the water/methanol and water/*tert*-butyl alcohol mixtures as described previously.<sup>78–84</sup> In this approach, the binary mixtures were simulated at a few compositions, and the pair radial distribution functions,  $g_{ij}(r)$ , where  $ij$  denotes the molecular pair, were calculated. The total correlation function integrals,  $H_{ij}$ , which are defined by

$$H_{ij} = \rho \int_0^\infty r^2 (g_{ij}(r) - 1) dr \quad (3)$$

were obtained by integrating the radial distribution functions. We have previously developed a computational method for extending radial distribution functions from simulations to infinite spatial separation and demonstrated that the accuracy of the obtained integrals improves significantly when compared to other methods.<sup>82–84</sup> From the evaluated total correlation function integrals, composition derivatives of the activity coefficients,  $(\partial \ln \gamma_1 / \partial x_1)_{N_2, T, P}$  were calculated. The derivatives were smoothed using the modified Margules model for  $G^E$ , using an objective function described previously<sup>78</sup> and including the values of  $\gamma_w^\infty$  calculated by free energy perturbation. The regressed curves and parameters are provided in the Supporting Information and respectively shown in Figure S1 and listed in Table S6.

### 3. RESULTS

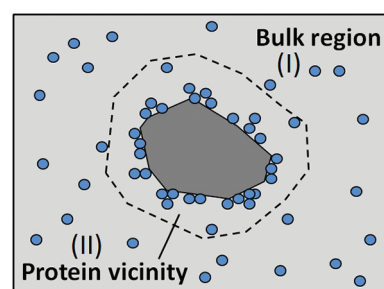
**3.1. Bulk Water Concentration and Activity.** The number of water molecules whose O atom is within 3.5 Å of any non-hydrogen atom of CALB defines here the hydration level. The distance corresponds approximately to the first minimum of radial distribution functions of water molecules around solvent-exposed protein residues.<sup>85</sup> In methanol and *tert*-butyl alcohol simulations, the hydration level decreased initially and stabilized at an approximately constant level (Figure 1), as the molecules of the water shell around CALB



**Figure 1.** Hydration level vs simulation time for simulations of the systems W (water; solid), H500 (hexane; long-dashed), E335 (MTBE; short-dashed), T210 (*tert*-butyl alcohol; dotted), and M210 (methanol; dash-dotted). The curves correspond to individual simulations.

mixed with the bulk solvent. A constant level was reached after approximately 4 and 10 ns for simulations carried out in methanol and *tert*-butyl alcohol, respectively. In hexane and MTBE, essentially all water molecules remained near the protein surface (Figure 1). Only a few water molecules were seen to escape from the hydration layer around the protein and diffuse through the bulk medium. For the pure water simulations, the hydration level increased slightly in the beginning but reached a constant level after 4 ns (Figure 1), which is consistent with an initial increase in solvent-accessible surface area (discussed in Section 3.4).

In order to quantify the distribution of water in the systems, the simulation cell was divided into a “bulk” region and a “protein vicinity” region, which will be referred to as region I and II, respectively (Figure 2). A water or organic solvent



**Figure 2.** Illustration of how the simulation cell is divided into a bulk region (I) and a protein vicinity region (II).

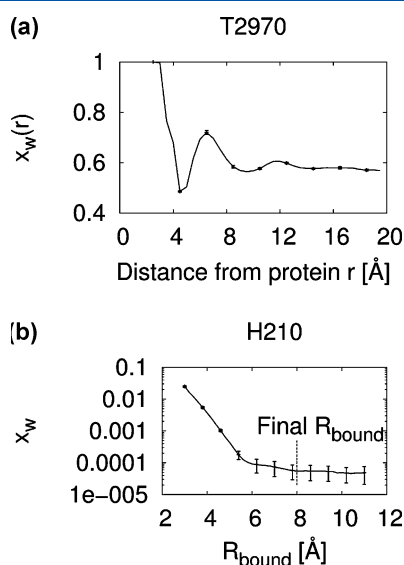
molecule was defined to be in region I when its distance to the protein surface was greater than a selected boundary distance  $R_{\text{bound}}$ . Likewise, the molecule was by definition in region II when its distance to the protein surface was smaller than  $R_{\text{bound}}$ . For a water molecule, the distance to the protein surface was defined as the distance from the O atom to the closest non-hydrogen atom of the protein. For methanol, the distance was

measured from the C atom, and for MTBE and *tert*-butyl alcohol, the tertiary C atom was used. For hexane, one of the two C atoms in the middle of the chain was used consistently.

In order to find appropriate values of  $R_{\text{bound}}$ , the average fraction of water molecules was evaluated as a function of the distance  $r$  to the protein surface according to

$$x_w(r) = \frac{\langle N_w(r) \rangle}{\langle N_w(r) \rangle + \langle N_s(r) \rangle} \quad (4)$$

where  $N_w(r)$  and  $N_s(r)$  denote respectively the number of water and organic solvent molecules with a distance to the protein surface (as defined above) between  $r - 0.25$  Å and  $r + 0.25$  Å, and where  $\langle \dots \rangle$  denotes average over time and the three replica simulations of each system. The calculation was carried out for representative systems including those with the highest and lowest water content for each organic solvent. The functions  $x_w(r)$  varied when  $r$  was small but approached fairly constant values as  $r$  increased, as shown for a representative system in Figure 3a. For simplicity, a single value of  $R_{\text{bound}} = 10$



**Figure 3.** (a) Function  $x_w(r)$  evaluated using a shell thickness of 0.5 Å for the system T2970. (b) Calculated bulk water fraction as a function of the parameter  $R_{\text{bound}}$  for the system H210. The finally selected value of  $R_{\text{bound}}$  is indicated. The curves in (a) and (b) are averages over three replica simulations, and standard error estimates are shown for selected values of  $r$  or  $R_{\text{bound}}$ .

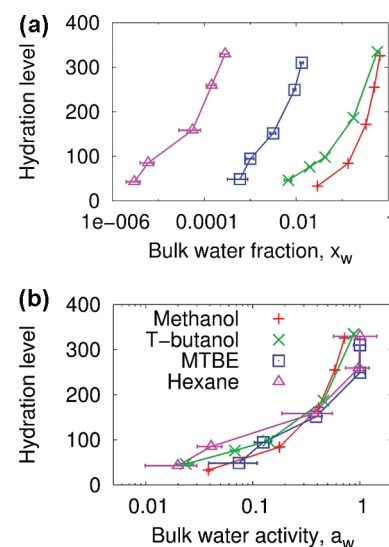
Å was used for all simulations carried out in methanol, *tert*-butyl alcohol, and MTBE. In some systems,  $x_w(r)$  had minor variations also for  $r > 10$  Å. These were however unimportant for the calculation of the bulk water fraction,  $x_w$ ; attempting larger values of  $R_{\text{bound}}$  changed  $x_w$  by at most a few percent, which was within the statistical uncertainties.

For hexane,  $R_{\text{bound}}$  was selected carefully for each system in order to improve statistical precision of the  $x_w$  estimates. This was accomplished by evaluating  $x_w$  as a function of  $R_{\text{bound}}$  along with standard error estimates based on the three replica simulations of each system. From this function, a region was identified where the variation of  $x_w$  was small compared with the corresponding standard errors, as exemplified in Figure 3b. The smallest value of  $R_{\text{bound}}$  of this region was used for the final evaluation of the bulk water fraction. This procedure resulted in

values in the range 4.8–10 Å, where the value of  $R_{\text{bound}}$  increased with the system water content.

The water activities were determined using the activity coefficients obtained as described in section 2.2, and activities are listed in Table 1. For H500 and E500, the calculated values of  $a_w$  significantly exceed unity, which appears anomalous. However, in those systems, clusters of water molecules were observed in the bulk region, which suggests that the medium was oversaturated with water and thus perhaps not a one-phase solution at equilibrium. Corresponding calculated values of  $a_w$  are thus not reported in Table 1. In the following analysis, we assume that the true  $a_w$  of H500 and E500 was very close to unity.

**3.2. Water Adsorption at the CALB Surface.** Water adsorption isotherms for CALB in the different solvents are shown in Figure 4a–b. The hydration level quantifies the



**Figure 4.** Hydration level of CALB vs (a) bulk water fraction and (b) bulk water activity. Results are shown for the methanol (+), *tert*-butyl alcohol (x), MTBE (□), and hexane (Δ). Uncertainties are shown in the  $x$ -values, while those in the  $y$ -values are too small to be displayed. Data points are joined by lines to guide the eye.

amount of adsorbed water. With the bulk water fraction,  $x_w$ , used to quantify the bulk water content (Figure 4a), the solvents give rise to four different isotherms. At similar hydration levels,  $x_w$  correlated with solvent polarity, which is consistent with previous simulation results.<sup>34,37,40</sup> If the bulk water content instead is shown in terms of the bulk water activity,  $a_w$  (Figure 4b and Table 1), the isotherm is less sensitive to the nature of the organic solvent.

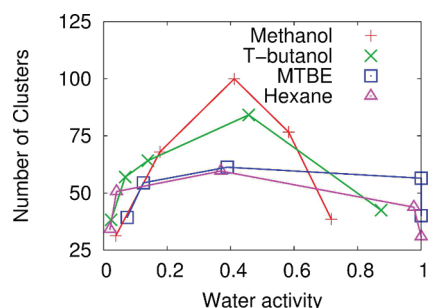
Qualitatively the isotherms are similar to experimental ones;<sup>42,45,86</sup> a steep increase at low  $a_w$  followed by a slow increase at medium  $a_w$ , and again a steep increase at high  $a_w$  (Figure 4b). This three-step behavior is most pronounced in hexane and MTBE while the isotherm in methanol seems to be convex over the entire range. The isotherm in *tert*-butyl alcohol is in between the ones in hexane and methanol.

It is striking that at low  $a_w$ , the data points of Figure 4b are scattered with no apparent trend, while at high  $a_w$ , the isotherms of MTBE and hexane lie clearly to the right of those of methanol and *tert*-butyl alcohol. Competition between water and organic solvent molecules for adsorbing to the surface of CALB may account for such differences. At low  $a_w$ , water

adsorption occurs mainly at hydrophilic sites,<sup>35,37,47</sup> and polar solvent molecules may replace water at such sites and result in lower hydration levels. At high  $a_w$ , competition for adsorption to the hydrophobic surface may be more relevant as most hydrophilic sites are occupied. Nonpolar organic molecules may be more competitive for binding in this region. The results may also be distorted by the apparently large uncertainties in activity estimates, in particular for hexane and MTBE (Table 1). Differences between adsorption isotherms in different solvents at high  $a_w$  have been observed in experiments performed on bovine serum albumin, lactoglobulin, chymotrypsinogen, and alcohol oxidase.<sup>42</sup> The authors observed higher adsorption levels in nonpolar solvents.

The differences between isotherms in the different solvents discussed above are relatively small and often explained by uncertainty in the activity estimates. Hence, fixing  $a_w$  yields hydration levels approximately independent of solvent. This is consistent with water adsorption isotherms observed in experiments<sup>42</sup> and validates the common assumption that the hydration level can be controlled in experiments by controlling the water activity.<sup>45</sup> Thus, macroscopic hydration levels are similar when water activity varies. However, there are subtle differences in what could be termed microscopic hydration from solvent to solvent, that is, in the distributions of adsorbed water over the surface and in the residence times of water molecules.

To further quantify the role of water molecules, a water cluster analysis was performed. Clusters were identified by joining water molecules with an O–O distance less than 3.5 Å. Clusters originating from water molecules in the first solvation shell of CALB were counted. These clusters contain at least one water molecule in contact with the protein but may as well contain water molecules outside the first solvation shell. The average number of such clusters is shown in Figure 5. A bell-



**Figure 5.** Average number of water clusters originating from the surface of CALB vs water activity. Results are shown for methanol (+), *tert*-butyl alcohol (x), MTBE (□), and hexane (Δ). Data points are connected by lines to guide the eye, and error bars are suppressed as they are smaller than the symbols.

shaped dependence on the water activity was obtained, with a maximum attained at around 0.4–0.5. The bell shape suggests that at low  $a_w$ , individual water molecules or small water clusters bind to specific sites on the protein surface. As  $a_w$  increases, the number of clusters at the surface increases until the clusters start to percolate at an  $a_w$  of 0.4–0.5. This is similar to adsorption isotherms for CALB in the gas phase extracted from simulations and determined in experiments.<sup>47</sup>

While all four organic solvents yielded bell-shaped curves attaining the maximum at roughly the same hydration level, the maximum number of clusters was different. This number was

~100 for methanol, ~80 for *tert*-butyl alcohol, and ~60 for hexane and MTBE. The decrease in cluster size at fixed water activity with polarity reflects that the surface water molecules were organized in fewer and consequently larger clusters in nonpolar solvents (Figure 5). Hence, the adsorption was more cooperative in nonpolar solvents. Figure 5 shows also that the number of water clusters drops faster with  $a_w$  in methanol and *tert*-butyl alcohol than in MTBE and hexane, once the maximum has been passed. This may be attributed to the fact that in the polar solvents more water molecules are present beyond the first solvation shell at similar  $a_w$ , which results in a higher degree of percolation.

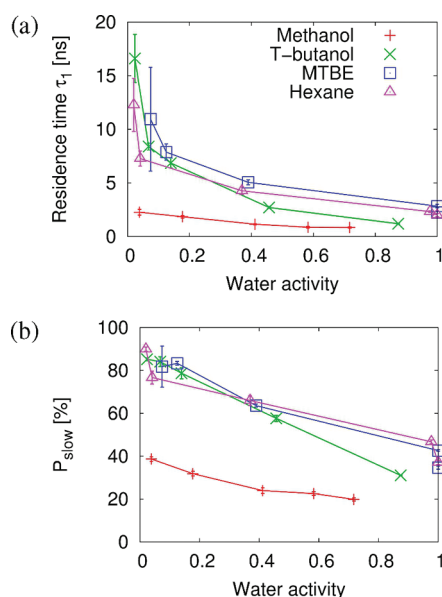
**3.3. Solvent Molecule Dynamics.** It has been observed that the dynamic properties of the hydration layer are of importance for protein dynamics. For instance, Trodler and Pleiss<sup>40</sup> found that protein flexibility anticorrelated with the number of immobile water molecules at the surface. In the present study, solvent molecule dynamics is assessed by computing their mean residence times at the CALB surface. The definition of residence time used follows that of Makarov et al.<sup>87</sup> and Schröder et al.<sup>85</sup> in which one considers the binary function  $\chi_{i,\alpha}(t)$  where  $i$  and  $\alpha$  denote a particular solvent molecule and protein residue, respectively, and where  $t$  denotes the time. The function is unity when the solvent molecule is within a cutoff distance from the residue, and zero otherwise. The analyses were carried out for water and organic solvent molecules. A 3.5 Å cutoff was used for water molecules measured from the water oxygen atom to the closest non-hydrogen atom of a particular residue. For organic solvent molecules, the cutoff distance was measured from the “central” C atom (see section 3.1). A 6.5 Å cutoff was applied, since radial distribution functions for *tert*-butyl alcohol, MTBE, and hexane around protein residues had their first minima approximately at this distance (data not shown). The same cutoff was for simplicity employed for methanol although the corresponding radial distribution functions had their first minima at approximately 5 Å. The autocorrelation function  $\rho_\alpha(t)$  for the sequence  $\chi_{i,\alpha}(t)$  corresponding to the residue  $\alpha$  was evaluated for  $t \leq 2.5$  ns for all residues  $\alpha$  being at least 25% exposed to the solvent, as indicated by the average solvent-accessible surface area (see section 3.4 for details). The overall autocorrelation function  $\rho(t)$  for adsorbed solvent molecules was calculated as the average  $\rho_\alpha(t)$  over all considered residues, and the biexponential model of Makarov et al.<sup>87</sup> was fitted to the function  $\rho(t)$ . The model, which reproduced the sampled  $\rho(t)$  well, is defined by

$$\rho_{\text{bi-exp}}(t) = a_1 e^{-k_1 t} + a_2 e^{-k_2 t} \quad (5)$$

where  $a_1$ ,  $a_2$ ,  $k_1$ , and  $k_2$  are adjustable parameters.  $\tau_1 \equiv k_1^{-1}$  and  $\tau_2 \equiv k_2^{-1}$  are the residence times of respectively slowly and rapidly exchanged solvent molecules. The relative populations of slowly and rapidly exchanged water molecules,  $P_{\text{slow}}$  and  $P_{\text{rapid}}$ , can be estimated from the relative magnitudes of  $a_1$  and  $a_2$ .<sup>87</sup>

In the pure water simulations, the values  $\tau_1 = 0.6 \pm 0.2$  ns,  $\tau_2 = 40 \pm 5$  ps, and  $P_{\text{slow}} = 14.1 \pm 0.8\%$  were obtained. In organic solvent,  $\tau_1$  decreased with increasing water activity and approached the pure water value at high  $a_w$  (Figure 6a).  $\tau_2$  followed the same trend (data not shown). The shortest residence times ( $\tau_1$  and  $\tau_2$ ) were obtained in methanol. Residence times obtained in *tert*-butyl alcohol, MTBE, and hexane did not differ significantly from one another. Since



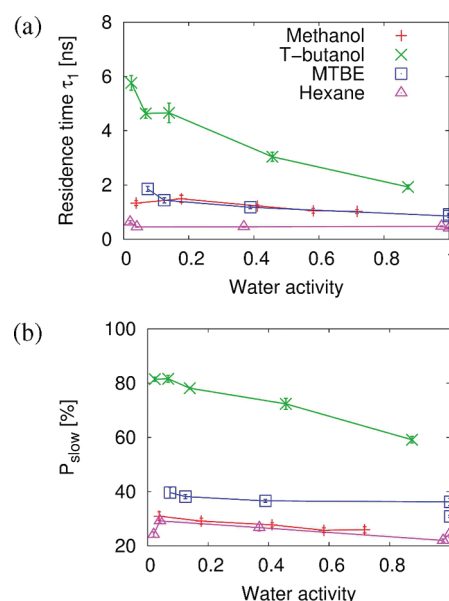


**Figure 6.** (a) Residence times ( $\tau_1$ ) and (b) population of slowly exchanged water molecules vs water activity in methanol (+), *tert*-butyl alcohol ( $\times$ ), MTBE ( $\square$ ), and hexane ( $\triangle$ ). Standard error estimates were based on three replica simulations which were started from different initial velocities, and data points are joined by lines to guide the eye.

water molecules interact more favorably with polar solvents like methanol, the corresponding short residence times may be expected. In simulations of cutinase by Micaelo and Soares,<sup>37</sup> residence times correlated negatively with hydration level, which is consistent with the present results. In their study, residence times correlated as well negatively with solvent polarity. This is only partially the case here since rather similar results were obtained in *tert*-butyl alcohol, MTBE, and hexane.

The fraction  $P_{\text{slow}}$  also decreased with increasing water activity (Figure 6b). The lowest values of  $P_{\text{slow}}$  were obtained in methanol. Values obtained in *tert*-butyl alcohol, MTBE, and hexane were higher and not significantly different from one another.  $P_{\text{slow}}$  approached 40% in methanol at low  $a_w$ , while it approached a value of 75–90% in the other solvents. Higher percentages of loosely bound water are expected in methanol, since methanol/water mixtures contain more water molecules than the nonpolar solvents when the bulk  $a_w$  are similar, as shown in section 3.2. The results demonstrate that the dynamical properties of the surface water depend on the water activity as well as the organic solvent.

In Figure 7a and b, the residence times of organic solvent molecules are compared at different water activities. For methanol, MTBE, and hexane molecules, the residence times were nearly independent of  $a_w$ .  $\tau_1$  was around 1–2 ns for methanol and MTBE, and  $\tau_2$  was 1 order of magnitude lower than  $\tau_1$  (results for  $\tau_2$  not shown). The fraction of slowly exchanged molecules  $P_{\text{slow}}$  was around 25 and 38% for methanol and MTBE molecules, respectively, and approximately independent of  $a_w$  (Figure 7b). The residence times of hexane molecules were consistently shorter and also rather insensitive to  $a_w$ . These short residence times were probably due to the fact that hexane is lacking a polar group and therefore cannot associate to the polar portion of the protein surface. For *tert*-butyl alcohol molecules, both  $\tau_1$  and  $\tau_2$  decreased with increasing  $a_w$  (Figure 7a). At low  $a_w$ ,  $\tau_1$  and  $\tau_2$  were 5.7 and 0.36 ns, respectively, which is significantly



**Figure 7.** (a) Residence times ( $\tau_1$ ) and (b) population of slowly exchanged organic molecules vs water activity in methanol (+), *tert*-butyl alcohol ( $\times$ ), MTBE ( $\square$ ), and hexane ( $\triangle$ ). Standard error estimates were based on three replica simulations which were started from different initial velocities, and data points are joined by lines to guide the eye.

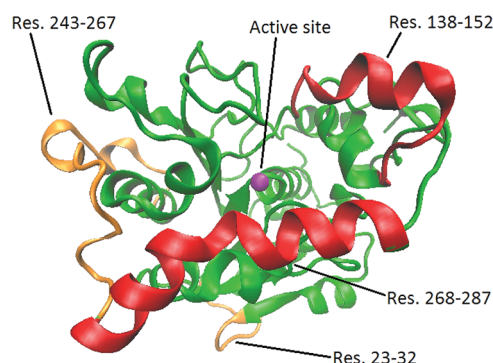
higher than the values obtained in the other solvents. For *tert*-butyl alcohol,  $P_{\text{slow}}$  was 79% at the lowest water activity and decreased with increasing  $a_w$ .

The residence times in *tert*-butyl alcohol are substantially longer than in the other solvents. Possibly, this is caused by the relatively high viscosity of *tert*-butyl alcohol, which is 1 order of magnitude higher than that of the other solvents. This is shown in Table 2 which lists viscosities taken from the DIPPR database.<sup>88</sup>

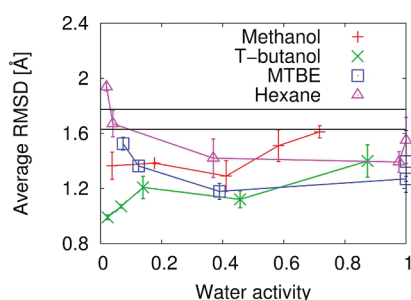
The results demonstrate that direct interactions between CALB and organic solvent molecules can be different for different solvents. These interactions may be modulated by the water activity, as here was the case for *tert*-butyl alcohol, and will depend on the protein surface characteristics (polar vs nonpolar).

**3.4. CALB Structure.** The root-mean-square deviation (rmsd) of CALB  $C_\alpha$  atoms with respect to the crystal structure coordinates (1TCA) was monitored for each simulation. The N- and C terminals (residues 1–20 and 308–317, respectively) were omitted from all rmsd calculations due to their flexibility. For a few simulations, additional flexible regions were omitted from the calculation. Depending on the simulation, these regions were the loops composed of residues 23–32 or 67–75, the helix  $\alpha 5$  and adjacent loop segment (residues 138–152), and the helix  $\alpha 10$  with preceding loop (residues 243–292; Figure 8). This analysis was carried out for all simulations, and stable rmsd curves were obtained in each case. Representative rmsd plots are shown in Figure S2 (Supporting Information).

Figure 9 shows the total rmsd averaged over the last 10 ns of the simulation. In the polar solvents and as well in MTBE, the rmsd was generally lower than in pure water (indicated by the horizontal lines). In *tert*-butyl alcohol, the rmsd increased with increasing water activity. The lowest average rmsd value ( $0.99 \pm 0.02$  Å) was observed at low  $a_w$ . The same trend was observed in methanol, but with a smaller slope and with



**Figure 8.** Image of CALB based on crystal structure (1TCA),<sup>62</sup> generated by VMD.<sup>63</sup> Marked regions are composed of residues 23–32, 138–152 (including helix  $\alpha 5$ ), 243–267, and 268–287 (including helix  $\alpha 10$ ). The location of the active site serine (Ser105) is indicated by a van der Waals sphere.



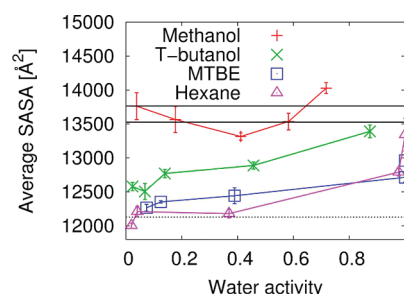
**Figure 9.** Average rmsd from crystal structure (1TCA) vs water activity in methanol (+), *tert*-butyl alcohol (x), MTBE (□), and hexane (Δ). Residues 1–20 (N-terminal) and 308–317 (C-terminal) were omitted from all rmsd calculations. Standard error estimates were based on three replica simulations which were started from different initial velocities. Error bar limits for the average rmsd in pure water are marked by horizontal lines.

generally higher rmsd values. For both MTBE and hexane, the average rmsd seemed to have U-shaped dependence on  $a_w$  with a minimum somewhere between 0.4 and 1. This was especially pronounced in hexane, where the rmsd became as high as  $1.94 \pm 0.03$  Å at low water activity. The results are qualitatively similar to the corresponding results for cutinase reported by Micaelo and Soares.<sup>37</sup> In that study, the minimum rmsd was in hexane attained at a hydration level of 7.5% (w/w). The minima observed here for CALB in MTBE and hexane (Figure 9) correspond to a hydration level of 10–15% (w/w).

Average rmsd per residue data were calculated from the final 10 ns of each simulation. The results are briefly summarized below. For most individual residues and secondary structure elements, the rmsd seemed to be uncorrelated with solvent and water activity (data not shown). The active site, defined as the residues within 5 Å of the catalytic Ser105 in the crystal structure, was well preserved in all simulations with average rmsd values of 0.4–0.8 Å. The regions for which the local rmsd for some system exceeded 2 Å were a loop composed of residues 23–32, the helix  $\alpha 5$  (residues 138–152), another loop composed of residues 243–267, and the helix  $\alpha 10$  (268–287; Figure 8). These regions are further discussed in the section S1 “CALB Structure–Average RMSD per Residue” in the Supporting Information.

The average solvent-accessible surface area (SASA) of CALB was evaluated based on the final 10 ns of each simulation, using

VMD.<sup>63</sup> Only non-hydrogen CALB atoms were considered, and a ball radius of 1.4 Å, roughly corresponding to the radius of a water molecule, was used. The average total SASA of CALB is shown in Figure 10. For all solvents except methanol,



**Figure 10.** Average total SASA of CALB vs water activity in methanol (+), *tert*-butyl alcohol (x), MTBE (□), and hexane (Δ). Standard error estimates were based on three replica simulations which were started from different initial velocities. Error bar limits for the average SASA in pure water are marked by horizontal black lines, and the SASA of the crystal structure<sup>62</sup> (1TCA) is marked with a horizontal dotted line.

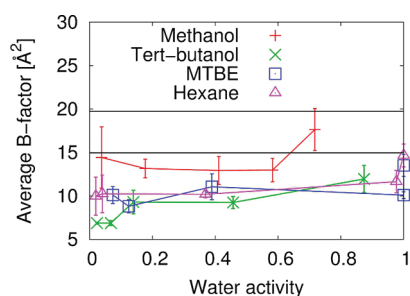
the SASA clearly increased with increasing water activity. At a fixed  $a_w$ , the values obtained in the different solvents were correlated with increasing solvent polarity, and the SASA was generally lower in the organic solvents than in pure water. Waterlike SASA values were nevertheless obtained in methanol at all  $a_w$ . The total SASA was consistently higher than that of the crystal structure (1TCA), except for hexane at low  $a_w$  (H43).

The SASA of hydrophobic and hydrophilic residues were as well evaluated for each system (data not shown). For methanol, neither the hydrophobic or hydrophilic SASA correlated with water activity. For *tert*-butyl alcohol and MTBE, both the hydrophilic and hydrophobic SASA increased with increasing  $a_w$  with approximately equal slopes. In hexane, a similar trend was observed although the hydrophilic SASA increased twice as fast with increasing  $a_w$  as the hydrophobic. Generally, the hydrophobic SASA exceeded that of the crystal structure (1TCA). For H43, the hydrophilic SASA was  $6500 \pm 90$  Å<sup>2</sup>, which is substantially lower than the corresponding value for the crystal structure, 6980 Å<sup>2</sup>.<sup>62</sup> In all other systems, the hydrophilic SASA was at least as large as the crystal structure value. This demonstrates that, particularly in hexane, there was a tendency for the hydrophilic surface area to be reduced as the water activity decreased. This accounts for the low total SASA and accordingly for the high rmsd values seen in hexane at low  $a_w$  (Figure 9 and Figure 10).

**3.5. CALB Flexibility.** The flexibility of CALB was assessed by calculating the B-factor of each  $C_\alpha$  atom. The B-factor profiles were in qualitative agreement with crystal structure data for CALB<sup>62,89</sup> and simulation results of Trodler and Pleiss and Skjot and co-workers.<sup>40,90</sup> Representative profiles are given in Figure S4 (Supporting Information).

The overall flexibility of CALB was characterized by averaging B-factors over all  $C_\alpha$  atoms excluding the N- and C-termini (residues 1–20 and 308–317) (Figure 11). The average B-factor generally increased with increasing water activity and approached the corresponding value of pure water at high  $a_w$ . In methanol most of the increase occurred when  $a_w$  exceeded 0.6, while in MTBE and hexane, most of the increase took place just as  $a_w$  approached 1. Water clusters at the surface





**Figure 11.** Average B-factor of CALB (omitting residues 1–20 and 308–317) vs water activity in methanol (+), *tert*-butyl alcohol (x), MTBE (□), and hexane (△). Standard error estimates were based on three replica simulations which were started from different initial velocities. Error bar limits for the average B-factor in pure water are marked by horizontal lines.

of CALB were also seen to percolate at these activities (Figure 5). In these solvents, the flexibility was approximately constant at lower water activities. A possible interpretation is that single water molecules at the surface have limited effects on the flexibility and that it is first when clusters containing several water molecules are formed that the lubricating effect emerges. In *tert*-butyl alcohol, the average B-factor increased significantly with increasing  $a_w$ , also when  $a_w$  was below 0.4 (Figure 11). This may be related to the exceptionally slow dynamics of *tert*-butyl alcohol molecules around CALB at low water activities (Figure 7).

These results are consistent with simulation studies of cutinase<sup>35</sup> and *Trypanosoma cruzi* triosephosphate isomerase<sup>38</sup> in which flexibility increased with increasing hydration and with previous finding that the flexibility depends on the organic solvent.<sup>40</sup> The role of the solvent for modulating flexibility seems, however, to be more complex than that of merely being a medium more or less able to dissolve the water layer around the protein. It is here demonstrated that the overall flexibility of CALB is different in different solvents, even if compared at similar water activities. Important for the flexibility is that the organic solvent molecules modulate the dynamical behavior of the water molecules bound to the protein. Rapidly exchanged water molecules promote protein flexibility while slowly exchanged water molecules counteract it, as suggested by Trodler and Pleiss.<sup>40</sup> The fraction and residence time of slowly exchanged water molecules is respectively larger and longer in hexane, MTBE, and *tert*-butyl alcohol than in methanol. This corresponds well with the lower flexibility observed in the three former solvents. The exceptionally low flexibility of CALB in *tert*-butyl alcohol may be due to the slow dynamics of *tert*-butyl alcohol molecules around the surface, which seems to be caused by the high viscosity of this solvent.

#### 4. CONCLUSION

MD simulations of CALB in pure water and the organic solvents methanol, *tert*-butyl alcohol, MTBE, and hexane at several water activities have been performed. The bulk water activity was evaluated by determining the concentration of water in bulk and using activity coefficients obtained from free energy perturbation and from binary water/organic solvent simulations analyzed by Kirkwood–Buff theory.

At similar water activity, the number of water molecules in the first solvation shell, i.e., the hydration level, was approximately the same in the four solvents although differences were observed at low and high activity. Analysis of

the water clusters on the surface of CALB showed that at low  $a_w$ , water molecules bind to the surface individually or in small clusters. At  $a_w \approx 0.4$ – $0.5$ , water molecules bind to existing water clusters leading to percolation of water clusters with increasing water activity. Calculation of residence times of water molecules at individual protein residues revealed that hydration water molecules became more volatile with increasing water activity and with increasing solvent polarity.

Structural and dynamical properties of CALB such as rmsd, SASA, and B-factors were sensitive to the water activity. Changing the water activity had different impacts in the different solvents. The average total rmsd from the crystal structure had a U-shaped dependence on the water activity in MTBE and hexane but increased monotonically with increasing water activity in *tert*-butyl alcohol and methanol. The SASA of CALB increased with increasing water activity, and at similar activities, it increased with solvent polarity. The overall flexibility of CALB increased as well with increasing water activity. At similar activities, the flexibility increased with increasing solvent polarity, with the exception that the lowest flexibility was observed in *tert*-butyl alcohol. This could possibly be attributed to the presence of more immobile water molecules at the CALB surface in the nonpolar solvents, and the relatively long residence times at the CALB surface for *tert*-butyl alcohol molecules.

While not yet realistic today, gains in understanding the hydration process may form the basis of studies aiming at improvements of not only enzyme activity but also stability. In practice stability is as much of concern as activity. If correlations between water cluster distributions and enzyme stability can be determined, this could suggest possible mutations on the protein surface to optimize protein stability. The challenge for protein engineering is to suggest the right mutations that increase stability and do not affect enzyme activity/selectivity. As we have shown here, the dynamical properties of the surface water not only depend on the water activity but also on the nature of the organic solvent suggesting that solvent selection may impact enzyme stability in addition to activity and reaction equilibria. This study is the first attempt to gain an understanding of the differences in microscopic hydration effects of organic solvent at similar water activities. Further studies in that direction could be used to create a database that compiles solvent effects on enzyme activity/stability.

#### ■ ASSOCIATED CONTENT

##### Supporting Information

Tables S1–S5 listing force field parameters used in this study; Table S6 listing parameters for the modified Margules models for water/methanol and water/*tert*-butyl alcohol mixtures used to obtain activity coefficients; Figure S1 comparing these models with simulations of the corresponding mixtures; further figures showing the rmsd of CALB from the crystal structure (1TCA) as a function of time for representative simulations (S2), the corresponding rmsd of helix  $\alpha 5$  as a function of organic solvent and water activity (S3), and representative B-factor profiles for CALB in water and each of the organic solvents studied (S4); section “CALB Structure–Average RMSD per Residue” discussing how the structure of specific CALB segments depends on the organic solvent and water activity. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: ja@kt.dtu.dk (J.A.); ghp@kemi.dtu.dk (G.H.P.).

### Present Address

<sup>†</sup>FOI—Swedish Defense Research Agency, Division of Defense and Security, Systems and Technology, 147 25 Tumba, Sweden

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge access to the Danish Center of Scientific Computing at the University of Southern Denmark, Odense, Denmark. R.W. acknowledges financial support from IP Bioproduction, European Union Sixth Framework Programme. G.H.P. acknowledges financial support from the Danish National Research Foundation via a grant to MEMPHYS—Center for Biomembrane Physics.

## REFERENCES

- (1) Carrea, G.; Riva, S. *Angew. Chem., Int. Ed.* **2000**, *39*, 2226–2254.
- (2) Klivanov, A. M. *Trends Biotechnol.* **1997**, *15*, 97–101.
- (3) Laane, C.; Boeren, S.; Vos, K.; Veeger, C. *Biotechnol. Bioeng.* **1987**, *30*, 81–87.
- (4) Valivety, R. H.; Johnston, G. A.; Suckling, C. J.; Halling, P. J. *Biotechnol. Bioeng.* **1991**, *38*, 1137–1143.
- (5) Nordblad, M.; Adlercreutz, P. J. *Biotechnol.* **2008**, *133*, 127–133.
- (6) Zaks, A.; Klivanov, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 3192–3196.
- (7) Carrea, G.; Ottolina, G.; Riva, S. *Trends Biotechnol.* **1995**, *13*, 63–70.
- (8) Zaks, A.; Klivanov, A. M. *Science* **1984**, *224*, 1249–1251.
- (9) Zaks, A.; Klivanov, A. M. *J. Biol. Chem.* **1988**, *263*, 8017–8021.
- (10) Affleck, R.; Xu, Z.; Suzawa, V.; Focht, K.; Clark, D. S.; Dordick, J. S. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1100–1104.
- (11) Watanabe, K.; Yoshida, T.; Ueki, S. *Bioorg. Chem.* **2004**, *32*, 504–515.
- (12) Broos, J.; Visser, A. J. W. G.; Engbersen, J. F. J.; Verboom, W.; van Hoek, A.; Reinhoudt, D. N. *J. Am. Chem. Soc.* **1995**, *117*, 12657–12663.
- (13) Valivety, R. H.; Halling, P. J.; Macrae, A. R. *Biotechnol. Lett.* **1993**, *15*, 1133–1138.
- (14) Bovara, R.; Carrea, G.; Ottolina, G.; Riva, S. *Biotechnol. Lett.* **1993**, *15*, 937–942.
- (15) Martinelle, M.; Holmquist, M.; Hult, K. *Biochim. Biophys. Acta* **1995**, *1258*, 272–276.
- (16) Graber, M.; Bousquet-Dubouch, M.; Lamare, S.; Legoy, M. *Biochim. Biophys. Acta* **2003**, *1648*, 24–32.
- (17) Graber, M.; Irague, R.; Rosenfeld, E.; Lamare, S.; Franson, L.; Hult, K. *Biochim. Biophys. Acta* **2007**, *1774*, 1052–1057.
- (18) Foresti, M. L.; Galle, M.; Ferreira, M. L.; Briand, L. E. *J. Chem. Technol. Biotechnol.* **2009**, *84*, 1461–1473.
- (19) Xu, Z.; Affleck, R.; Wangikar, P. W.; Suzawa, V.; Dordick, J. S.; Clark, D. S. *Biotechnol. Bioeng.* **1994**, *43*, 515–520.
- (20) Kim, J.; Clark, D. S.; Dordick, J. S. *Biotechnol. Bioeng.* **2000**, *67*, 112–116.
- (21) Wescott, C. R.; Klivanov, A. M. *J. Am. Chem. Soc.* **1993**, *115*, 1629–1631.
- (22) Wescott, C. R.; Klivanov, A. M. *J. Am. Chem. Soc.* **1993**, *115*, 10362–10363.
- (23) Ke, T.; Wescott, C. R.; Klivanov, A. M. *J. Am. Chem. Soc.* **1996**, *118*, 3366–3374.
- (24) Colombo, G.; Ottolina, G.; Carrea, G.; Bernardi, A.; Scolastico, C. *Tetrahedron: Asymmetry* **1998**, *9*, 1205–1214.
- (25) Peters, G. H. In *Enzyme Functionality*; Svendsen, A., Ed.; Marcel Dekker, Inc.: New York, 2004; pp 97–148.
- (26) Hartsough, D. S.; Merz, K. M. *J. Am. Chem. Soc.* **1992**, *114*, 10113–10116.
- (27) Hartsough, D. S.; Merz, K. M. *J. Am. Chem. Soc.* **1993**, *115*, 6529–6537.
- (28) Norin, M.; Haeffner, F.; Hult, K.; Edholm, O. *Biophys. J.* **1994**, *67*, 548–559.
- (29) Zheng, Y.; Ornstein, R. L. *J. Am. Chem. Soc.* **1996**, *118*, 4175–4180.
- (30) Zheng, Y.; Ornstein, R. L. *Biopolymers* **1996**, *38*, 791–799.
- (31) Zheng, Y.; Ornstein, R. L. *Protein Eng.* **1996**, *9*, 485–492.
- (32) Toba, S.; Merz, K. M. *J. Am. Chem. Soc.* **1996**, *118*, 6490–6498.
- (33) Toba, S.; Merz, K. M. *J. Am. Chem. Soc.* **1997**, *119*, 9939–9948.
- (34) Yang, L.; Dordick, J. S.; Garde, S. *Biophys. J.* **2004**, *87*, 812–821.
- (35) Soares, C. M.; Teixeira, V. H.; Baptista, A. M. *Biophys. J.* **2003**, *84*, 1628–1641.
- (36) Micaêlo, N. M.; Teixeira, V. H.; Baptista, A. M.; Soares, C. M. *Biophys. J.* **2005**, *89*, 999–1008.
- (37) Micaêlo, N. M.; Soares, C. M. *FEBS J.* **2007**, *274*, 2424–2436.
- (38) Díaz-Vergara, N.; Piñeiro, Á. *J. Phys. Chem. B* **2008**, *112*, 3529–3539.
- (39) Cruz, A.; Ramirez, E.; Santana, A.; Barletta, G.; López, G. E. *Mol. Simul.* **2009**, *35*, 205–212.
- (40) Trodler, P.; Pleiss, J. *BMC Struct. Biol.* **2008**, *8*, 9–1–10.
- (41) Halling, P. J. *Trends Biotechnol.* **1989**, *7*, 50–52.
- (42) Halling, P. J. *Biochim. Biophys. Acta* **1990**, *1040*, 225–228.
- (43) Valivety, R. H.; Halling, P. J.; Peilow, A. D.; Macrae, A. R. *Biochim. Biophys. Acta* **1992**, *1122*, 143–146.
- (44) Valivety, R. H.; Halling, P. J.; Macrae, A. R. *Biochim. Biophys. Acta* **1992**, *1118*, 218–222.
- (45) Halling, P. J. *Enzyme Microb. Technol.* **1994**, *16*, 178–206.
- (46) Bell, G.; Janssen, A. E. M.; Halling, P. J. *Enzyme Microb. Technol.* **1997**, *20*, 471–477.
- (47) Branco, R. J. F.; Graber, M.; Denis, V.; Pleiss, J. *ChemBioChem* **2009**, *10*, 2913–2919.
- (48) Anderson, E. M.; Larsson, K. M.; Kirk, O. *Biocatal. Biotransform.* **1998**, *16*, 181–204.
- (49) Damstrup, M.; Abildskov, J.; Kiil, S.; Jensen, A. D.; Sparsø, F. V.; Xu, X. J. *Agric. Food Chem.* **2006**, *54*, 7113–7119.
- (50) Wang, L.; Du, W.; Liu, D.; Dai, N. *J. Mol. Catal. B: Enzym.* **2006**, *43*, 29–32.
- (51) Su, E.; Wei, D. *J. Mol. Catal. B: Enzym.* **2008**, *55*, 118–125.
- (52) Fjerbaek, L.; Christensen, K. V.; Norddahl, B. *Biotechnol. Bioeng.* **2009**, *102*, 1298–1315.
- (53) Kaieda, M.; Samukawa, T.; Kondo, A.; Fukuda, H. *J. Biosci. Bioeng.* **2001**, *91*, 12–15.
- (54) Humeau, C.; Girardin, M.; Rovel, B.; Miclo, A. J. *Biotechnol.* **1998**, *63*, 1–8.
- (55) Chamouleau, F.; Coulon, D.; Girardin, M.; Ghoul, M. *J. Mol. Catal. B: Enzym.* **2001**, *11*, 949–954.
- (56) Petersson, A. E. V.; Adlercreutz, P.; Mattiasson, B. *Biotechnol. Bioeng.* **2006**, *97*, 235–241.
- (57) Mora-Pale, J. M.; Pérez-Munguía, S.; González-Mejía, J. C.; Dordick, J. S.; Bárzana, E. *Biotechnol. Bioeng.* **2007**, *98*, 535–542.
- (58) Leonard-Nevers, V.; Martona, Z.; Lamarea, S.; Hult, K.; Graber, M. *J. Mol. Catal. B: Enzym.* **2009**, *59*, 90–95.
- (59) Nordblad, M.; Adlercreutz, P. *Biotechnol. Bioeng.* **2007**, *99*, 1518–1524.
- (60) Pepin, P.; Lortie, R. *Biotechnol. Bioeng.* **1999**, *63*, 502–505.
- (61) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (62) Uppenberg, J.; Hansen, M. T.; Patkar, S.; Jones, T. A. *Structure* **1994**, *2*, 293–308.
- (63) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (64) Hansen, H. K.; Rasmussen, P.; Fredenslund, A.; Schiller, M.; Gmehling, J. *Ind. Eng. Chem. Res.* **1991**, *30*, 2352–2355.
- (65) Peters, G. H.; van Aalten, D. M. F.; Edholm, O.; Toxværd, S.; Bywater, R. *Biophys. J.* **1996**, *71*, 2245–2255.

- (66) Peters, G. H.; Bywater, R. P. *Biophys. J.* **2001**, *81*, 3052–3065.
- (67) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, *61*, 704–721.
- (68) Bas, D. C.; Rogers, D. M.; Jensen, J. H. *Proteins* **2008**, *73*, 765–783.
- (69) MacKerell, A. D. Jr. *J. Phys. Chem.* **1998**, *102*, 3586–3616.
- (70) MacKerell, A. D. Jr.; Feig, M.; Brooks, C. III. *J. Comput. Chem.* **2004**, *25*, 1400–1415.
- (71) Vorobyov, I.; Anisimov, V.; Greene, S.; Venable, R.; Moser, A.; Pastor, R.; MacKerell, A. Jr. *J. Chem. Theory Comput.* **2007**, *3*, 1120–1133.
- (72) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (73) Dixit, S. D.; Chipot, C. *J. Phys. Chem. A* **2001**, *105*, 9795–9799.
- (74) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025–9031.
- (75) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: New York, 2002.
- (76) Kirkwood, J. G.; Buff, F. P. *J. Chem. Phys.* **1951**, *19*, 774–777.
- (77) O'Connell, J. P. *Mol. Phys.* **1971**, *20*, 27–33.
- (78) Christensen, S.; Peters, G. H.; Hansen, F. Y.; Abildskov, J. *Fluid Phase Equilib.* **2007**, *261*, 185–190.
- (79) Christensen, S.; Peters, G. H.; Hansen, F. Y.; O'Connell, J. P.; Abildskov, J. *Fluid Phase Equilib.* **2007**, *260*, 169–176.
- (80) Christensen, S.; Peters, G. H.; Hansen, F. Y.; O'Connell, J. P.; Abildskov, J. *Mol. Simul.* **2007**, *33*, 449–457.
- (81) Wedberg, R.; Peters, G. H.; Abildskov, J. *Fluid Phase Equilib.* **2008**, *273*, 1–10.
- (82) Wedberg, R.; O'Connell, J. P.; Peters, G. H.; Abildskov, J. *Mol. Simul.* **2010**, *36*, 1243–1252.
- (83) Wedberg, R.; O'Connell, J. P.; Peters, G. H.; Abildskov, J. *Fluid Phase Equilib.* **2011**, *302*, 32–42.
- (84) Wedberg, R.; O'Connell, J. P.; Peters, G. H.; Abildskov, J. *J. Chem. Phys.* **2011**, *135*, 084113–1–9.
- (85) Schröder, C.; Rudas, T.; Boresch, S.; Steinhauser, O. *J. Chem. Phys.* **2006**, *124*, 234907–1–18.
- (86) Sirotkin, V. A. *Biochim. Biophys. Acta* **2005**, *1750*, 17–29.
- (87) Makarov, V. A.; Andrews, B. K.; Smith, P. E.; Pettitt, B. M. *Biophys. J.* **2000**, *79*, 2966–2974.
- (88) DIPPR® Project No. 801 Database. <http://dippr.byu.edu/>, 2011.
- (89) Uppenberg, J.; Öhrner, N.; Norin, M.; Hult, K.; Kleywegt, G. J.; Patkar, S.; Waagen, V.; Anthonsen, T.; Jones, T. A. *Biochemistry* **1995**, *34*, 16838–16851.
- (90) Skjot, M.; De Maria, L.; Chatterjee, R.; Svendsen, A.; Patkar, S. A.; Østergaard, P. R.; Brask, J. *ChemBioChem* **2009**, *10*, 520–527.