Molecular mechanism of the hydration of Candida antarctica lipase B in gas phase: water adsorption isotherms and molecular dynamics simulations.

R. J. F. Branco, Marianne Graber, Vinciane Denis, Juergen Pleiss

Molecular mechanism of the hydration of *Candida antarctica* lipase B in gas phase: water adsorption isotherms and molecular dynamics simulations

Ricardo J. F. Branco[‡], *Marianne Graber*[†], *Vinciane Denis*[†] *and Jürgen Pleiss*[‡*]

[‡]Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany.

[†]Université de la Rochelle, Pôle Sciences et Technologie -Bât. Marie Curie, UMR6250 LIENSs CNRS-ULR, Avenue Michel Crépeau, 17042 La Rochelle, Cedex 01, France

Dedicated to Professor Karl Hult on the occasion of his 65th birthday

[*] Communicating author:

Prof. Dr. Jürgen Pleiss
Institute of Technical Biochemistry
University of Stuttgart
Allmandring 31
D-70569 Stuttgart
Germany
E-mail: Juergen.Pleiss@itb.uni-stuttgart.de
Phone: (+49) 711-68563191

ABSTRACT

Hydration is a major determinant of activity and selectivity of enzymes in organic solvents or in gas phase. The molecular mechanism of the hydration of *Candida antarctica* lipase B (CALB) and its dependence on the thermodynamic activity of water a_w was studied by molecular dynamics simulations and compared to experimentally determined water sorption isotherms. Hydration occurred in two phases. At low water activity, single water molecules bound to specific water binding sites at the protein surface. As the water activity increased, water networks gradually developed. The number of protein-bound water increased linearly with a_w, until at a_w = 0.5 a spanning water network was formed consisting of 311 water molecules which covered the hydrophilic surface of CALB, with the exception of the hydrophobic substrate binding site. At higher water activity, the thickness of the hydration shell increased up to 10 Å close to a_w = 1. Above a limit of 1600 protein-bound water molecules the hydration shell becomes unstable and the formation of pure water droplets occurs in this oversaturated simulation conditions. While the structure and the overall flexibility of CALB was independent of the hydration state, the flexibility of individual loops was sensitive to hydration: some loops such as part of the substrate binding site became more flexible, while other parts of the protein became more rigid upon hydration. However, the molecular mechanism of how flexibility is related to activity and selectivity is still elusive.

Keywords: thermodynamic activity of water • hydration mechanism • molecular modeling • • deglycosylation

INTRODUCTION

Candida antarctica lipase B (CALB) is an efficient catalyst for hydrolysis of ester substrates in water and esterification in organic solvents. It is used in many industrial applications because of its high enantioselectivity, wide range of substrates, thermal stability, and stability in organic solvents.[1] CALB belongs to the α/β hydrolase fold family with a conserved catalytic triad consisting of Ser, His, and Asp/Glu.[2] In contrast to most lipases, CALB has no lid which shields the active site, but the hydrophobic substrate binding site is solvent exposed.[3] It has been shown that like many other enzymes, CALB retains activity in organic solvents and has interesting catalytic properties such as higher thermostability and altered stereoselectivity.[4] However, it is essential to add small amounts of water to maintain stability and flexibility of enzymes in organic solvent. Thus protein-bound water is essential for catalysis and serves as a lubricant for the enzyme.[5] In contrast, fully dry enzymes are inactive and enzymes in organic solvents with high amounts of water show denaturation.[6]

The influence of water on stability, biocatalytic activity, and selectivity of enzymes has been studied extensively. The best way to quantify the availability of water to the biocatalyst in non-conventional media is to use the parameter "thermodynamic water activity". Most studies were performed in organic solvents at varying water activity. A similar dependence of the biocatalytic activity of enzymes on the thermodynamic activity of water a_w was found: at very low water activity, the biocatalytic activity was low, and increased with increasing a_w to an optimum. For a higher a_w, the biocatalytic activity decreased. The optimal a_w depends on the enzyme under investigation, the reaction, and the solvent. In contrast, no general trend has been found for the dependence of selectivity, especially enantioselectivity, on a_w.[7, 8]

The dependence of flexibility, activity, and selectivity on the protein hydration level has been extensively investigated in condensed phase, mainly for proteins suspended in co-solvent mixtures of water and organic solvents.[5, 9-20] Water is known to have a crucial role in protein structure, flexibility, and activity.[21-25] Water molecules bind via hydrogen bonds to the side chain and backbone atoms of proteins,[26] to polar atoms of substrates and ligands,[27] as well as to other water molecules. This enables water to mediate protein-protein and protein-ligand contacts, and to take part in enzyme catalysis. X-ray crystallography has long been used to analyze water at protein surfaces, since crystal structures determined at high resolution provide a detailed picture of protein hydration.[28] Comparative studies of crystal structures

showed that there are water binding sites on the surface and the interior of a protein which are occupied by water molecules in different crystal structures.[29-31]

Computer simulations of enzymes in different solvents are a valuable tool to investigate the effect of solvent on the structure and dynamics of proteins.[21, 32, 33] While in most simulations proteins showed no significant structural changes in different solvents, the molecular flexibility decreased when simulated in organic solvents,[33, 34] which has been confirmed by different experimental techniques such as time-resolved fluorescence anisotropy,[35] ESR,[36] and dielectric relaxation spectroscopy.[37] It has been suggested that in water polar side chains orient toward the surface, thus increasing the hydrophilic surface and decreasing the polar intra-molecular interactions that mediate the rigidity of the protein,[38] while in organic solvents the surface area is reduced which leads to improved packing and increased stability.[36] In addition, a spanning water network was observed in simulations of CALB in a water/organic solvent mixture.[33] The spanning water network was formed in the presence of a non-polar solvent [34, 39] and resulted from a slow exchange of water molecules at the protein surface.

Previously, the catalytic activity of CALB at different water activities was determined in a continuous solid/gas reactor.[40, 41] In this type of reactor, a solid packed enzymatic sample is percolated by a carrier gas, which simultaneously carries gaseous substrates to the enzyme and removes gaseous products. The water activity in the reaction medium can be fully controlled and fixed independently of the other reaction medium components, by incorporating water in the inlet gas at the correct molar fraction. As a consequence, the sole role of water on biocatalysis can be precisely assessed.[42] In this system, it was shown that there is an optimal water activity in respect to alcoholysis (a_w=0.05-0.1) and enantioselectivity (a_w=0.2).[40] The water effect on enantioselectivity has been interpreted by binding of a water molecule in a well-defined binding site, the "specificity pocket" which is located near the active site of the enzyme. In this paper, our goal is to study the hydration state of CALB in gas phase both by measuring water sorption isotherms and by molecular modeling. The experiments and the molecular modeling studies are performed with a deglycosylated lipase. Because the commercially available CALB is a glycosylated heterogeneous protein expressed in *Aspergillus oryzae*, it was deglycosylated and desalted to avoid water binding by salts.

EXPERIMENTAL SECTION AND COMPUTATIONAL METHODS

Enzymes and chemicals

Chirazyme L-2, lyo (CALB) was purchased from Roche Diagnostics (Penzberg, Germany), as a crude lyophilized material. Peptide-N-glycosidase F (PNGase F) lyophilized and endoglycosidases F1, F2, and F3 were obtained from Sigma-Aldrich (Saint-Quentin Fallavier, France) as "Native Protein Deglycosylation Kit". All other chemicals used were analytical grade and obtained from Sigma-Aldrich (Saint-Quentin Fallavier, France).

Deglycosylation

Deglycosylation of CALB with endoglycosidases F1, F2, and F3 was performed with 3 µL of the Sigma enzymatic kit solution for 0.2 mg of native CALB in 37.5 µL de-ionized water in 10 µL of 5x reaction buffer. The mixture was then incubated 1 hour at 37°C (see details in Supporting Information).

Desalting

Desalting was performed using a PD-10 column (Amersham Biosciences, Pantin, France). PD-10 desalting columns are pre-packed, disposable columns containing Sephadex™ G-25 medium for group separation of high (Mr > 5000) from low molecular weight substances (Mr < 1000). After desalting, CALB was freeze dried. It was checked that these different treatments did not cause loss of enzyme activity.

Measurement of CALB activity

The CALB enzymatic activity was determined with a pH-stat system which measures the liberation of butyric acid from tributyrin by titrating with 100 mM sodium hydroxide. 24.25 mL 10 mM potassium phosphate buffer, pH 7.0 are incubated in a thermostated vessel at 25°C, equipped with a magnetic or propeller stirrer (the stirring speed is set to a value that a further enhancement of the stirrer speed did not increase the base consumption per min). After addition of 735 µL tributyrin, the pH-stat system was started to keep the pH at 7.0. When the pH stabilized, 5-10 U of enzyme was added. The consumption of 100 mM sodium hydroxide was monitored for 0 to 15 min. The specific activity was calculated from the base consumption at the linear part of the graph.

Electrophoresis analysis

Enzyme purity and deglycosylation was evaluated with sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 12% separation gel under the condition developed by Laemmli.[43] A Mini-Protean 3 (Bio-Rad Laboratories, Marnes la Coquette, France) was used for electrophoresis. As reference proteins molecular weight marker from 6 kDa to 66 kDa was purchased from Sigma-Aldrich (Saint-Quentin Fallavier, France). For the detection of proteins, gels were stained with Coomassie brilliant blue R 250 after the electrophoresis.

Measurement of biocatalyst water isotherms

Water adsorption isotherms of lyophilized lipase were measured at 45°C by using a DVS advantage automated moisture sorption analyzer (Surface Measurement Systems Ltd., London, UK) between 0 and 98% relative humidity. The samples were initially dried for 10 h at 45°C under totally dried flowing air in the DVS moisture sorption analyzer to obtain the starting dry sample mass. The required humidities are generated by mixing dry and saturated vapor gas flows in the correct proportions using two mass flow controllers and one vapor humidifier. The instrument was run with a *dm/dt* (*m*, mass; *t*, time) set at 0.002% min^{-1} to reach the equilibrium. This means that the relative humidity will remain constant until the mass variation falls consistently below this threshold. Once this condition has been reached, the relative humidity of the system will be raised or lowered to the next level.

The water adsorption curve was modeled using the Brunauer-Emmet-Teller (BET)[44] gas adsorption equation (eq 1):

$$\frac{a_w}{(1-a_w)\times m} = \frac{1}{m_0 \times C} + \frac{a_w \times (C-1)}{m_0 \times C} \qquad \text{(eq 1)}$$

where m is the mass of bound water (in mg per mg of dry enzyme), m_0 is the monolayer water coverage, a_W the water activity and C a constant. From the linear regression of the plot of $\frac{a_w}{(1-a_w)\times m}$ versus a_W, the theoretical monolayer water coverage can be calculated:

$m_0 = \frac{1}{S+I}$, where S is the slope and I the *y*-intercept.[45]

Molecular dynamics simulation

The structure of CALB was taken from the Protein Data Bank (entry 1TCA[2]), and the crystallographic water molecules as well as the N-acetyl-D-glucosamine moiety at Asn 74 were removed. The protein was placed in a cubic box with a side length between 170 and 370

Å filled with 981-1000 argon atoms as a model carrier gas and different numbers of water molecules. 8 protein boxes were created with 27, 64, 125, 338, 720, 1708, 2195, and 4091 water molecules. As a reference for the evaluation of the thermodynamic activity of water, a simulation with 4095 water molecules and 1000 argon atoms was performed.

Molecular dynamics (MD) simulations were carried out using the GROMACS 4.0.3 program package[46] and the united-atom force field GROMOS 43a1.[47] The system was neutralized, assuming the catalytic His 224 as positively charged. All simulations were run under isothermal-isobaric (NPT) ensemble conditions, coupled to the Berendsen barostat (reference pressure of 1.0 bar and coupling time constant of 0.6 ps) and thermostat (reference temperature 318.15 K and coupling time constant of 0.1 ps).[48] The compressibility was adjusted to 1.27×10^{-2} bar^{-1}[49] to reproduce the experimental density of argon in gas phase. The simulation time step was 2 fs and all H-bonds were constrained using the LINKS algorithm[50] as implemented in GROMACS.

After energy minimization of 2000 steps (steepest descent and conjugated gradient) the system was equilibrated in three subsequent simulations of 100 ps each, applying a decreasing force constant of 1000, 100, and 10 kJ mol^{-1} to all protein atoms. Finally, all constraints were removed and each system was simulated up to 20 ns.

The last 5 ns were analyzed to determine for each box the number of water molecules bound to the protein and the thermodynamic activity of water. The size distribution of water clusters was determined by the program g-clustsize as implemented in the GROMACS software package. All protein systems took about 10 ns to equilibrate as confirmed by the maximum cluster size and the number of clusters in the box reaching a plateau after 10 ns of simulation (Supporting Figure S3, S5 and S6).

In all simulations, one large cluster of N_{prot} water molecules attached to CALB was identified, as well as a number of single water molecules N_{gas}. In simulations with less than 1700 water molecules, no isolated pure water droplets were formed during simulation time. Because the thermodynamic activity of water is given by the vapor pressure of water above a sample divided by the vapor pressure of pure water (reference value), it can be derived from N_{gas} for 5 systems (27, 64, 125, 338, and 720 water molecules, Supporting Table S2). For the mixture

of argon and water, the total pressure is $p_{tot} = p_{arg} + p_{wat} = 1$ bar. Thus, the partial pressure of water, p_{wat}, can be calculated from the number of water molecules in the gas phase:

$$p_{wat} = p_{tot} * N_{gas} / N_{tot} \quad \text{(eq 2)}$$

with the number of water molecules in the gas phase, N_{gas}, and the number of particles in the gas phase (water plus argon), N_{tot}. In the pure water simulation with 4095 water molecules, a gas phase of 60±4 isolated water molecules was in equilibrium with water droplets after 8 ns of simulation, thus the partial vapor pressure of water in this fully water-saturated system is taken as a reference, and the water activity of a system with N water molecules a_w (N) can be determined as:

$$a_w (N) = p_{wat} (N) / p_{wat} (\text{reference}) \quad \text{(eq 3)}$$

Inserting (eq 2) into (eq 3) results in (eq 4):

$$a_w (N) = [N_{gas} (N) / N_{tot} (N)] / [(N_{gas} (\text{reference}) / N_{tot} (\text{reference})] \quad \text{(eq 4)}$$

RESULTS

CALB deglycosylation with endoglycosidase F1

Enzymatic methods are much more restrictive than chemical methods for protein deglycosylation, because the enzymatic activity generally depends on structural elements of the N-glycans chains. However, enzymes are very effective for deglycosylation of small glycoprotein quantities or when it is necessary to retain the protein conformation. Complete deglycosylation of CALB was achieved by using endoglycosidase F1 in high quantity (4.2 µL of commercial preparation for 0.2 mg of native lipase) (Figure 1).

Water sorption isotherm measurements

It was checked by enzyme activity measurement that neither deglycosylation nor desalting or freeze drying prior to water sorption isotherm measurement led to a loss of activity. Moreover enzyme activity was retained after isotherm measurement itself. Therefore it can be assumed that our results describe water sorption by native CALB.

The water adsorption and desorption isotherm measurements were carried out for a 0.1 incremental step in a_W, between 0.00 and 0.98. The experiment was repeated twice, showing very good reproducibility. Desalted and deglycosylated CALB showed a low adsorption of water until a_W=0.40, corresponding to 120 bound water molecules. Then, the isotherm became exponential until at a_W=0.98 a maximum number of protein-bound water of 1720 and 1790 molecules for two experimental runs was reached (**Figure 2 and Supporting Table S1**). The isotherm resembles the BET isotherm, whose model assumes a fixed number of independent binding sites, which accommodate a monolayer of adsorb molecules and allow subsequent layers to bind more weakly. Using the linear regression of the BET equation in the range of a_W from 0.1 to 0.4, the water corresponding to the monolayer on the enzyme was found to be equal to 0.063 mg of water per mg of dry protein. This bound water quantity is obtained for a_W between 0.3 and 0.4. The amount of water adsorbed when the equilibrium is approached from the “dry” side was identical to the amount of water adsorbed when the equilibrium is approached from the “wet” side, corresponding to a very low hysteresis between water adsorption and desorption (Supporting Figure S8).

Protein-bound water calculation

For each box the thermodynamic activity of water was calculated from (eq. 3) and the number of protein-bound water molecules was derived from the size of the largest water cluster (Supporting Table S2). Upon increasing the water activity, water molecules gradually bind to the surface of CALB. The analysis of the MD simulations revealed two distinct phases in the protein hydration: 1) at very low water activity, the water binding sites on the protein surface were occupied by single water molecules (

Figure 3**a**) corresponding to a molar fraction of 54 mol water/mol protein (3% w/w at a_w = 0.2). The size of these patches increased with water activity and spanning the hydrophilic surface of the protein, corresponding to a molar fraction of 311 mol water/mol protein (17% w/w at a_w = 0.5). However, the hydrophobic substrate binding site was not covered by the spanning water network **(**

Figure 3**b)**; 2) Finally, at high water activity $a_w > 0.5$, the number of protein-bound water molecules rapidly increased, and a continuous multilayer of water molecules on the hydrophilic surface was formed **(**

Figure 3**c)** with a thickness of up to 10 Å and a molar fraction of 677 mol water/mol protein (37% w/w at a_w = 0.7). Near the active site an isolated water cluster was formed without hydrogen bonding to the spanning water network or the hydrophobic substrate binding site. Simulations with 1708, 2195, and 4091 water molecules showed oversaturation and the formation of pure water droplets. Because the experimental investigations were performed under non-saturating conditions ($a_w < 1$), these simulations were excluded from analysis.

In hydration phase I ($a_w < 0.5$) the number of protein-bound waters was proportional to the water activity, since hydration occurs at isolated water binding sites and then gradually a spanning water network was formed which was in agreement with the experimental water sorption isotherms. In hydration phase II ($a_w > 0.5$) the number of protein-bound water molecules became exponential, corresponding to the observed multilayer hydration shell formation.

Protein structure and flexibility

The root mean square deviation (RMSD) of the CALB backbone atoms from the initial crystallographic structure ranged between 2.2 Å (simulation with 64 water molecules, corresponding to a_w = 0.23) and 2.7 Å (simulation with 125 water molecules, corresponding to a_w = 0.26). No correlation between a_w and RMSD was found, which indicates that there is no systematic conformational change upon increasing the hydration level. Interestingly, the

simulation with the highest number of water molecules showed one of the lowest RMSD values (Supporting Figure S3).

Similarly, the global flexibility measured as root mean square fluctuation (RMSF) averaged over all residues seems not to be affected by the hydration level (1.1 - 1.2 Å). However, four regions were identified which were influenced by protein hydration (Supporting Figure S4). For the N-terminus (residues 10-20) and the loop between helices α9 and α10 (residues 248-256) the flexibility decreased with increasing hydration by 1.3 and 0.2 Å, respectively. In contrast, the loop between sheet β5 and helix α5 (residues 133-150), and helix α10 (residue 277-290) showed an increased flexibility with increasing hydration, by 0.7 and 0.6 Å, respectively. The averaged RMSF in these five regions ranges between 0.9 and 2.2 Å. The flexibility of the catalytic amino acids Ser 105, Asp 187, and His 224 was low and did not change upon hydration.

DISCUSSION

The role of water in stability, catalytic activity, and selectivity is still not fully understood. Though it has been observed that thermodynamic activity is a major determinant of these properties,[34, 51-53] it is difficult to assess its role in mixtures of organic solvent and water, since the physico-chemical properties of the organic solvent is expected to influence the result. Therefore, working under solvent-free conditions using a gas-phase reactor has become a promising method to directly address the effect of water to the properties of enzymes. CALB is a widely used, robust, and highly selective biocatalyst, which has been shown to be stable in a wide range of organic solvent – water mixtures, although its enantioselectivity is sensitive to the solvent and water activity.[40, 54] Previously, binding of water to glycosylated, immobilized CALB was determined, and two phases of protein hydration were observed: up to a thermodynamics water activity a_w of 0.5, the amount of water binding to the protein was proportional to a_w, while at higher water activities protein hydration increased exponentially.[53] This dependence of CALB hydration on a_w was confirmed in the present study on hydration of free, deglycosylated and desalted CALB, although the amount of bound water was reduced by 30 to 50 %. This decrease shows that salt (from suspension buffer before enzyme drying) as well as glycosylation strongly influence the water adsorption, as already observed with other proteins.[55] The shape of experimental water sorption and desorption curves were very similar to water sorption curves previously obtained for other proteins, following the pattern of the BET model.[55-57] Moreover, the value of 0.063 mg of water per mg of dry protein found for BET monolayer water was in the range 0.05-0.07 found for other proteins.[56, 57] The absence of hysteresis between adsorption and desorption indicated that there was no irreversibly bound water on the protein. Recently, a very marked hysteresis between water adsorption and desorption in gas phase, was observed for alcohol dehydrogenase deposited onto non-porous glass carriers together with the co-factor NADPH, sucrose as stabilizing agents and salts resulting from the buffers. This hysteresis was mainly attributed to additives and leaching of biocatalyst away from the carrier surface, instead of being attributed to the protein itself.[55]

When trying to find correlation between biocatalyst hydration state and CALB catalytic activity at different water activity in gas phase, it appears that the optimal water activity in respect to alcoholysis (a_w=0.05-0.1) and enantioselectivity (a_w=0.2) both correspond to a water activity level at which the water is highly structured. The overlaying phenomena of

water adsorption and enzyme activity and selectivity will be investigated more precisely in the future.

In conclusion to the experimental results, it has to be noticed that the present results concerning water sorption by CALB in gas phase should also give information about the water sorption in water - organic solvent mixtures. Indeed, it was shown for several proteins in different organic solvents that the water sorption isotherms in water - organic solvent mixtures were approximately superimposable to the water vapor sorption isotherms obtained in gas phase, provided that data points were plotted both in gas and liquid phase, as the amount of water bound to protein versus water activity.[56]

MD simulation results help us to understand this similarity between the behavior of the protein in gas phase and in organic co-solvent mixtures. In both cases, the exchange rate between the spanning water network at the protein surface and the bulk solvent is much lower than in liquid water.[33] In gas phase and in organic solvent, an increase in the global protein flexibility is not expected, since water exchange rate is not affected *per se* by the increase of hydration of the protein surface. This does not exclude that water hydration leads to local changes of flexibility, either a decrease caused by the spanning water network or an increase upon formation of water layers, as it was shown by simulation results. Most remarkable in regard to catalytic activity is the increase of flexibility of helix α10 located at the entrance of the active site. Although CALB, in contrast to other lipases, has been shown to have no lid which opens upon interaction with a hydrophobic substrate interface, this helix might play a role in interaction with the substrate. In addition, the observed lack of hydration of the hydrophobic patch contacting the hydrophobic substrate interface seems to be essential for substrate recognition and protein orientation at the interface.

To assess the mechanism of CALB hydration in molecular detail, MD simulations at different water activities were performed which led to a similar dependence of the number of protein-bound water on a_w, though the total number of protein-water was substantially higher in the simulation and similar to the amount of water bound to the immobilized, glycosylated CALB. Three major conclusions could be drawn from the simulation:

1. Hydration occurs in two phases: for water activities below a critical water activity (a_w = 0.5 for CALB), water initially binds to well-defined water binding sites at the protein surface,[25,

31] and then gradually a spanning water network is formed.[33, 39, 58] The amount of protein-bound water increased linearly with water activity, depending on the number of water binding sites on the surface and their affinity toward water. The increase of the number of bound water molecules with water activity as well as the critical water activity are expected to be specific for each protein. At water activities above the critical water activity the amount of protein-bound water molecules increased exponentially.

2. For oversaturated conditions (systems with more than 1600 water molecules) the multilayer hydration shell further grew and became unstable which led to the formation of pure water droplets and a smaller protein hydration shell, with a maximum of 10 Å of thickness at the hydrophilic side of the protein, while the hydrophobic substrate binding site is not hydrated at all, but an isolated water cluster attached to the active site residues is formed at the center of this hydrophobic patch. This observation is also supported by previous studies in liquid phase for other enzymes,[9, 11] where a water monolayer completely covering the protein was never observed. Thus, the structure and physico-chemical properties of the protein surface not only lead to the formation of tight water binding sites, but also to a protein-specific shape of the hydration shell at high water activity.

3. The effect of hydration to protein flexibility cannot be generalized; while the flexibility of some parts of the protein surface increased upon hydration, for other parts of the surface the formation of a water networks led to a decrease of flexibility. It has been shown that catalytic activity and enantioselectivity of some enzymes are strongly correlated to the thermodynamic activity of water (a_w).[51] By carefully measuring the contribution of enthalpy and entropy to enantioselectivity of CALB,[59] the group of Karl Hult pointed to the pivotal role of the flexibility of a substrate in the binding site which in their model was interpreted as the contribution of entropy. Although the molecular mechanism of how flexibility is related to activity and selectivity is still elusive, there are strong indications that flexibility and dynamics of the enzyme-substrate complex are crucial to understand the effect of substrate structure, organic solvent, and water activity to selectivity and enzymatic activity.

Acknowledgements

We thank the German Science Foundation DFG (Sonderforschungsbereich 716) and the French National Agency for Research ANR (Chimie et procédés pour le développement durable) for financial support. R.B. also wants to thanks Christian Gruber for fruitful discussions and help.

LEGENDS TO FIGURES

Figure 1: SDS-PAGE showing CALB deglycosylation with endoglycosidase F1 after 1h30 reaction. Lanes 1 and 2: Chirazyme (200µg/mL) 2 and 1 µL, respectively. Lane 3: MW Markers. Lanes 4 and 5: Chirazyme treated with endoglycosidase F1, 1 and 3 µL, respectively.

Figure 2 Water adsorption isotherm at 45 °C of desalted and deglycosylated *Candida antarctica* lipase B (CALB). The open diamonds correspond to the mean value of two different cycles of water adsorption performed on 0.5mg of the same enzyme sample. The filled diamonds correspond to the calculation of maximum cluster size of protein-bound water molecules, derived from MD simulations.

Figure 3**a-c:** Dynamics of protein hydration. MD simulation snapshot after 20 ns of simulation of the protein with a solvent box with 125 (phase I), 338 (transition between phase I and II), and 720 water molecules (phase II), corresponding to water activities of 0.26, 0.49 and 0.75, respectively. Each view corresponds to the same orientation of the enzyme, with the amino acids of the catalytic triad (Ser 105, Asp 187 and His 224) coloured on the surface in green. Pictures rendered in Pymol visualization software.

FIGURES

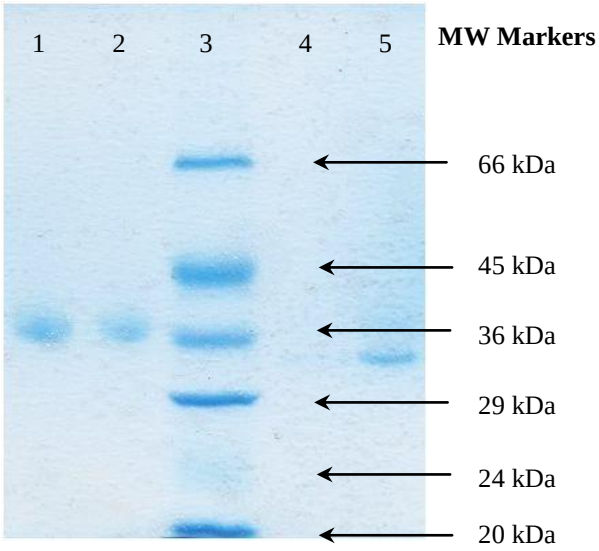


Figure 1

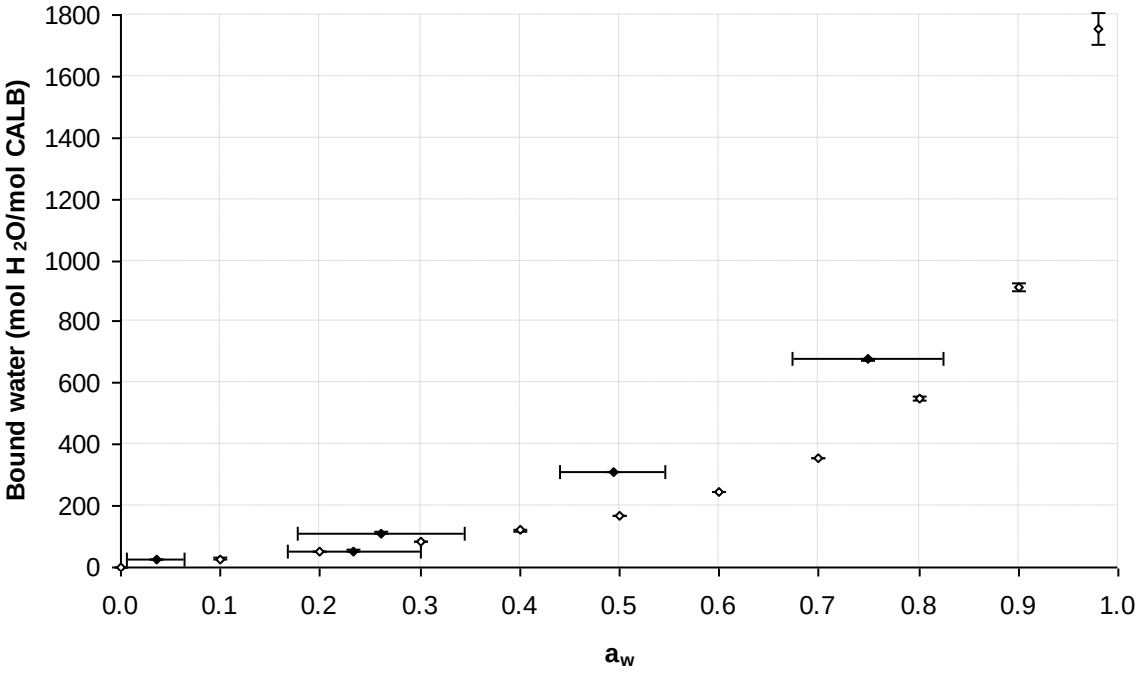

Bound water (mol H 2O/mol CALB)
1800
1600
1400
1200
1000
800
600
400
200
0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
aw

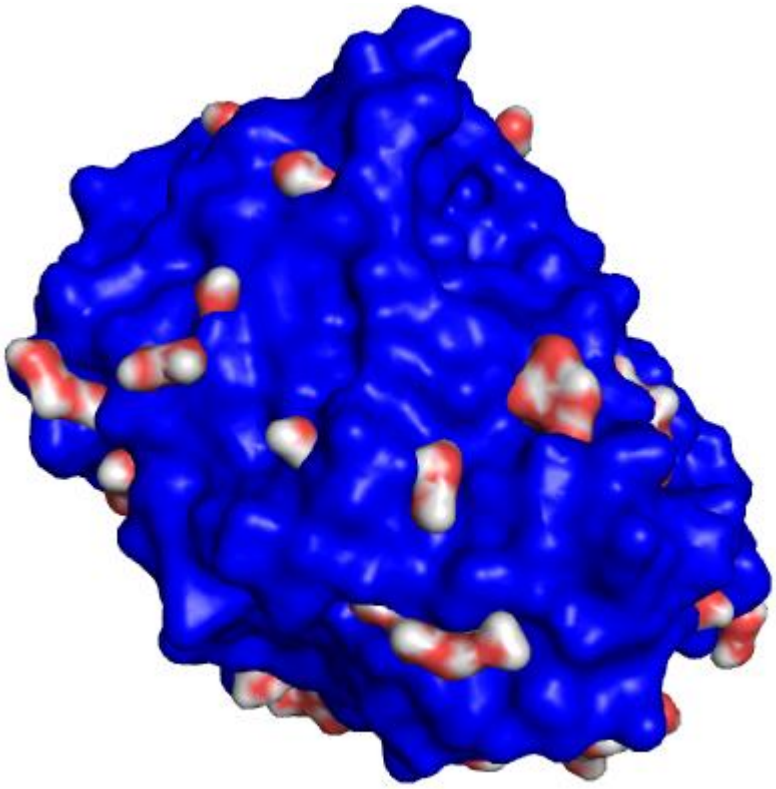
Figure 2

471

472

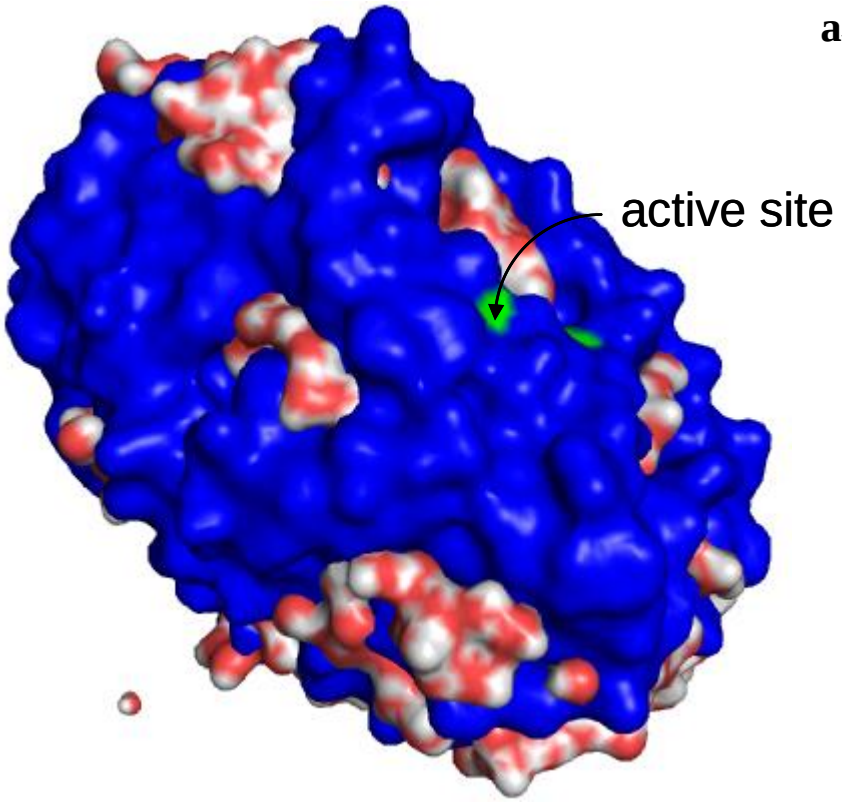
a)

a_w = 0.26



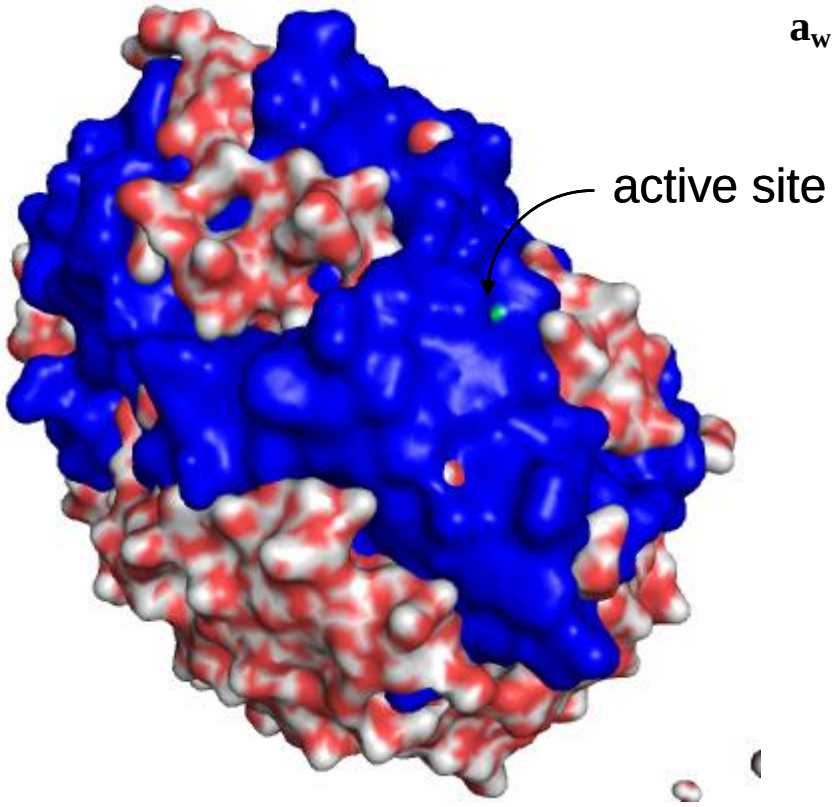
b)

a_w = 0.49



c)

a_w = 0.75



473

474 **Figure 3a-c**

475

476 REFERENCES

477

478 [1] D. Rotticci, J. C. Rotticci-Mulder, S. Denman, T. Norin and K. Hult, *Chembiochem* **2001**,
479 *2*, 766-770.

480 [2] J. Uppenberg, M. T. Hansen, S. Patkar and T. A. Jones, *Structure* **1994**, *2*, 293-308.

481 [3] M. Martinelle, M. Holmquist and K. Hult, *Biochim Biophys Acta* **1995**, *1258*, 272-276.

482 [4] O. Kirk and M. Würtz Christensen, *Org Process Res Dev* **2002**, *6*, 446-451.

483 [5] A. Zaks and A. M. Klibanov, *J Biol Chem* **1988**, *263*, 3194-3201.

484 [6] K. Griebenow and A. M. Klibanov, *J. Am. Chem. Soc.* **1996**, *118*, 11695-11700.

485 [7] Jonsson, E. Wehtje, P. Adlercreutz and B. Mattiasson, *Biochim Biophys Acta* **1999**, *1430*,
486 313-322.

487 [8] M. Persson, E. Wehtje and P. Adlercreutz, *Chembiochem* **2002**, *3*, 566-571.

488 [9] N. M. Micaelo and C. M. Soares, *Febs J.* **2007**, *274*, 2424-2436.

489 [10] N. M. Micaelo, V. H. Teixeira, A. M. Baptista and C. M. Soares, *Biophys. J.* **2005**, *89*,
490 999-1008.

491 [11] M. C. Parker, B. D. Moore and A. J. Blacker, *Biotechnol. Bioeng.* **1995**, *46*, 452-458.

492 [12] H. Wu, M. H. Zong and X. Y. Chen, *Biotechnol. Appl. Biochem.* **2009**, *53*, 201-207.

493 [13] F. Secundo, G. Barletta and G. Mazzola, *Biotechnol. Bioeng.* **2008**, *101*, 255-262.

494 [14] X. Y. Chen, M. H. Zong, W. Y. Lou and H. Wu, *Appl. Biochem. Biotechnol.* **2008**, *151*,
495 21-28.

496 [15] N. Diaz-Vergara and A. Pineiro, *J. Phys. Chem. B* **2008**, *112*, 3529-3539.

497 [16] J. M. Mora-Pale, S. Perez-Munguia, J. C. Gonzalez-Mejia, J. S. Dordick and E. Barzana,
498 *Biotechnol. Bioeng.* **2007**, *98*, 535-542.

499 [17] M. Graber, R. Irague, E. Rosenfeld, S. Lamare, L. Franson and K. Hult, *BBA-Proteins*
500 *Proteomics* **2007**, *1774*, 1052-1057.

501 [18] G. Bell, A. E. M. Janssen and P. J. Halling, *Enzyme Microb. Technol.* **1997**, *20*, 471-477.

502 [19] G. Bell, P. J. Halling, B. D. Moore, J. Partridge and D. G. Rees, *Trends Biotechnol.*
503 **1995**, *13*, 468-473.

504 [20] A. Zaks and A. M. Klibanov, *J. Biol. Chem.* **1988**, *263*, 8017-8021.

505 [21] P. Trodler, R. D. Schmid and J. Pleiss, *BMC Struct Biol* **2009**, *9*, 38.

506 [22] C. Mattos, *Trends Biochem Sci* **2002**, *27*, 203-208.

507 [23] F. Despa, *Ann N Y Acad Sci* **2005**, *1066*, 1-11.

508 [24] N. Prabhu and K. Sharp, *Chem. Rev.* **2006**, *106*, 1616-1623.

509 [25] V. Helms, *ChemPhysChem* **2007**, *8*, 23-33.

510 [26] S. Park and J. G. Saven, *Proteins* **2005**, *60*, 450-463.

511 [27] Y. Lu, R. Wang, C. Y. Yang and S. Wang, *J Chem Inf Model* **2007**, *47*, 668-675.

512 [28] M. Nakasako, *Philos Trans R Soc Lond B Biol Sci* **2004**, *359*, 1191-1204; discussion
513 1204-1196.

514 [29] P. C. Sanschagrin and L. A. Kuhn, *Protein Sci* **1998**, *7*, 2054-2064.

515 [30] C. A. Bottoms, T. A. White and J. J. Tanner, *Proteins* **2006**, *64*, 404-421.

516 [31] F. Bos and J. Pleiss, *Antimicrob. Agents Chemother.* **2008**, *52*, 1072-1079.

517 [32] G. Colombo and G. Carrea, *J Biotechnol* **2002**, *96*, 23-33.

518 [33] P. Trodler and J. Pleiss, *BMC Struct Biol* **2008**, *8*, 9.

519 [34] C. M. Soares, V. H. Teixeira and A. M. Baptista, *Biophys. J.* **2003**, *84*, 1628-1641.

520 [35] J. Broos, A. Visser, J. F. J. Engbersen, W. Verboom, A. vanHoek and D. N. Reinhoudt,
521 *J. Am. Chem. Soc.* **1995**, *117*, 12657-12663.

522 [36] S. Toba and K. M. Merz, *J. Am. Chem. Soc.* **1997**, *119*, 9939-9948.

523 [37] J. Mijovic, Y. Bian, R. A. Gross and B. Chen, *Macromolecules* **2005**, *38*, 10812-10819.

524 [38] D. S. Hartsough and K. M. Merz, *J. Am. Chem. Soc.* **1993**, *115*, 6529-6537.

525 [39] N. Smolin, A. Oleinikova, I. Brovchenko, A. Geiger and R. Winter, *J. Phys. Chem. B*
526 **2005**, *109*, 10995-11005.

527 [40] V. Leonard, L. Fransson, S. Lamare, K. Hult and M. Graber, *Chembiochem* **2007**, *8*, 662-
528 667.

529 [41] M. Graber, M. P. Bousquet-Dubouch, S. Lamare and M. D. Legoy, *BBA-Proteins*
530 *Proteomics* **2003**, *1648*, 24-32.

531 [42] S. Lamare, M. D. Legoy and M. Graber, *Green Chemistry* **2004**, *6*, 445-458.

532 [43] U. K. Laemmli, *Nature* **1970**, *227*, 680-685.

533 [44] S. Brunauer, P. H. Emmett and E. Teller, *J. Am. Chem. Soc.* **1938**, *60*, 309-319.

534 [45] H. R. Costantino, J. G. Curley and C. C. Hsu, *Journal of Pharmaceutical Sciences* **1997**,
535 *86*, 1390-1393.

536 [46] B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*,
537 435-447.

538 [47] X. Daura, A. E. Mark and W. F. van Gunsteren, *J. Comput. Chem.* **1998**, *19*, 535-547.

539 [48] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola and J. R. Haak, *J.*
540 *Chem. Phys.* **1984**, *81*, 3684-3690.

541 [49] G. P. Baxter and H. W. Starkweather, *Proc. Natl. Acad. Sci. U. S. A.* **1929**, *15*, 441-444.

542 [50] B. Hess, H. Bekker, H. J. C. Berendsen and J. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463-
543 1472.

544 [51] N. Fontes, M. C. Almeida, C. Peres, S. Garcia, J. Grave, M. R. Aires-Barros, C. M.
545 Soares, J. M. S. Cabral, C. D. Maycock and S. Barreiros, *Ind. Eng. Chem. Res.* **1998**, *37*,
546 3189-3194.

547 [52] P. Vidinha, N. Harper, N. M. Micaelo, N. M. T. Lourengo, M. da Silva, J. M. S. Cabral,
548 C. A. M. Afonso, C. M. Soares and S. Barreiros, *Biotechnol. Bioeng.* **2004**, *85*, 442-449.

549 [53] V. Leonard-Nevers, Z. Marton, S. Lamare, K. Hult and M. Graber, *Journal of Molecular*
550 *Catalysis B-Enzymatic* **2009**, *59*, 90-95.

551 [54] J. Ottosson, L. Fransson, J. W. King and K. Hult, *Biochim. Biophys. Acta-Protein Struct.*
552 *Molec. Enzym.* **2002**, *1594*, 325-334.

553 [55] K. Dimoula, M. Pohl, J. Buechs and A. C. Spiess, *Biotechnology Journal* **2009**, *4*, 712-
554 721.

555 [56] S. B. Lee and K. J. Kim, *J. Ferment. Bioeng.* **1995**, *79*, 473-478.

556 [57] Y. Caro, M. Pina, F. Turon, S. Guilbert, E. Mougeot, D. V. Fetsch, P. Attwool and J.
557 Graille, *Biotechnol. Bioeng.* **2002**, *77*, 693-703.

558 [58] A. Oleinikova, N. Smolin, I. Brovchenko, A. Geiger and R. Winter, *J. Phys. Chem. B*
559 **2005**, *109*, 1988-1998.

560 [59] J. Ottosson, L. Fransson and K. Hult, *Protein Sci.* **2002**, *11*, 1462-1471.

561

562