Title

Solvent Entropy Contributions to Catalytic Activity in Designed and Optimized Kemp Eliminases

Journal

The Journal of Physical Chemistry B, 122(21)

ISSN

1520-6106

Authors

Belsare, Saurabh

Pattni, Viren

Heyden, Matthias

et al.

Publication Date

2018-05-31

DOI

10.1021/acs.jpcb.7b07526

Peer reviewed

Solvent Entropy Contributions to Catalytic Activity in Designed and Optimized Kemp Eliminases

Saurabh Belsare[1†], Viren Pattni[7†], Matthias Heyden[7*], Teresa Head-Gordon[1-6*]

[1]The UC Berkeley-UCSF Graduate Program in Bioengineering
[2]Kenneth S. Pitzer Center for Theoretical Chemistry, [3]Department of Chemistry, [4]Department of Bioengineering, [5]Department of Chemical and Biomolecular Engineering, University of California, [6]*Chemical Sciences Division, Lawrence Berkeley National Labs*
Berkeley, CA, 94720

[7]Max-Planck-Institut fur Kohlenforschung
Mülheim an der Ruhr, Germany

ABSTRACT

We analyze the role of solvation for enzymatic catalysis in two distinct, artificially designed Kemp Eliminases, KE07 and KE70, and mutated variants that were optimized by laboratory directed evolution. Using a spatially resolved analysis of hydration patterns, intermolecular vibrations and local solvent entropies, we identify distinct classes of hydration water and follow their changes upon substrate binding and transition state formation for the designed KE07 and KE70 enzymes and their evolved variants. We observe that differences in hydration of the enzymatic systems are concentrated in the active site and undergo significant changes during substrate recruitment. For KE07, directed evolution reduces variations in the hydration of the polar catalytic center upon substrate binding, preserving strong protein-water interactions, while the evolved enzyme variant of KE70 features a more hydrophobic reaction center for which the expulsion of low-entropy water molecules upon substrate binding is substantially enhanced. While our analysis indicates a system-dependent role of solvation for the substrate binding process, we identify more subtle changes in solvation for the transition state formation, which are less affected by directed evolution.

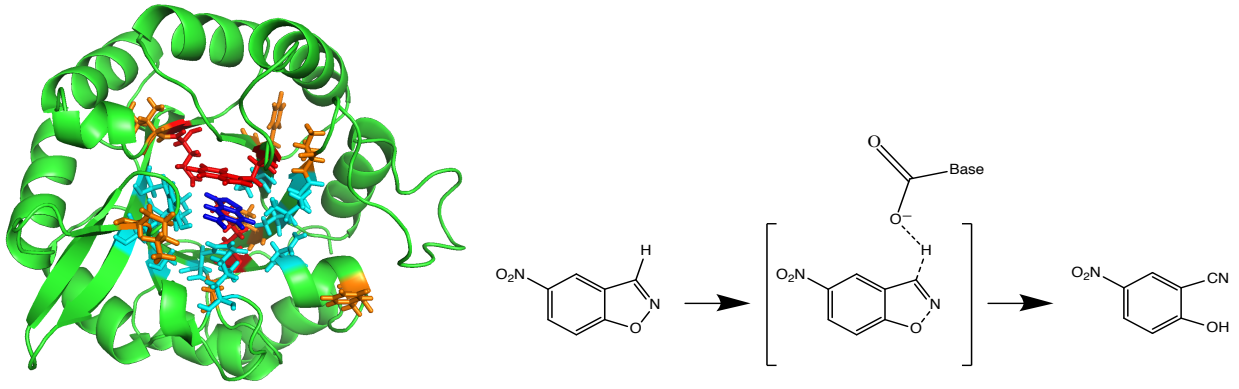
[†]authors contributed equally
[*]corresponding authors: heyden@kofo.mpg.de, thg@berkeley.edu

INTRODUCTION

Enzymes are molecular catalysts that have been highly optimized by natural evolution to increase the rate of a targeted biochemical reaction anywhere between ~10-20 orders of magnitude compared to the uncatalyzed reaction in solvent (typically water). In addition to steric positioning effects by residues in the catalytic pocket of an enzyme,[1] a large component of their catalytic power also arises from the pre-organization of stabilizing electric fields for the transition state (TS) due to the enzyme scaffold, that is absent in bulk water environments.[2-3] However, in addition to the protein itself, the solvent dipoles and their fluctuations can, in principle, also contribute to the enzyme catalytic activity.[4-5] This might arise especially from the integrated effect of a hydration layer which behaves differently from the bulk in terms of structure, dipole orientation, and self-diffusion, due to the varied chemistry and topography immediate to the enzyme surface,[6-7] and the possibility that these solvent alterations persist from the surface for more than one solvation layer.[8-9]

It also arises from features of water in the active site that can effect the catalytic process: 1) the binding of the substrate to the enzyme typically involves the expulsion of water from the catalytic center and a partial dehydration of the substrate itself; 2) conformational changes of the enzyme-substrate complex may alter the pattern of solvent-exposed sidechains; 3) changes in the electron density within the reactants, e.g. upon formation of the TS, will affect electrostatic interactions with the solvent environment; 4) product formation and product release will result in the resolvation of functional groups previously buried within the catalytically active complex. Both protein surface and active site perturbations from the aqueous solvent are thus expected to have significant enthalpic and entropic contributions to the free energy that will impact the quantitative values of the Michaelis Menten constants k_{cat} and K_M.

In this work, we employ atomistic molecular dynamics simulations with a polarizable force field to analyze solvent entropy contributions to distinct steps of an enzymatic reaction catalyzed by the engineered KE07 (Figure 1A) and KE70 (Figure S1) enzymes. These contain active sites that were engineered into TIM barrel scaffolds in a centrally located position in the β-barrel to carry out the Kemp Elimination reaction, *i.e.* the deprotonation of the 5-nitrobenzisoxazole substrate by a catalytic base leading to the formation of a cyanophenol product (Figure 1B).



(A) (B)

Figure 1. (A) *Kemp Eliminase KE07 enzyme and* (B) *the targeted catalytic reaction*. The catalytic (red) and additional stabilizing residues (cyan) create the designed KE07 enzyme. The mutations introduced by laboratory directed evolution (orange) result in an improved variant R7.2 with an increased catalytic turnover rate. The substrate is shown in dark blue.

Previous work has shown how improvements in the catalytic activity of enzymes in the Kemp Eliminase family through directed evolution can be traced back to changes in side chain entropy[10] and protein electrostatics[3] through the analysis of molecular simulations. There has also been some analysis of the role of the solvent in mediating activity[11]. However, in the present study a novel spatially resolved variant of the two-phase thermodynamics approach[12-14] allows us to analyze the entropy of solvent molecules in great detail in the vicinity of the catalytic center and on the protein surface. We compare results obtained for the designed Kemp Eliminase enzymes and their mutated counterparts obtained from laboratory directed evolution (LDE). The hydration shell of both designed enzymes and their evolved variants show entropic signatures that differ from bulk water as one might expect due to steric hindrances and interactions with the protein surface. However, our main focus lies on quantifying changes in solvation patterns during distinct stages of the catalytic process in the active site, i.e. the free enzyme (apo), the substrate-bound complex of the enzyme (EL), and the TS of the catalyzed reaction (EL*), to determine how they correlate with improvements in catalytic performance after LDE. We demonstrate not only that there are notable differences in the hydration of the catalytic centers for the two distinct enzymatic systems, but also that directed evolution has essentially opposite effects on these hydration patterns for KE07 vs. KE70. Our results indicate that hydration likely plays an important role for the mechanism of the enzymatically catalyzed reactions, but also that this role is highly context dependent and does not follow a single general rule.

THEORY

Thermodynamic properties of liquids cannot be accurately computed from equilibrium ensemble simulations using, for example, harmonic or quasi-harmonic expressions. However, the two-phase thermodynamic (2PT) provides an alternative approach, which utilizes the spectrum of vibrational modes described by the vibrational density of states (VDOS), $\Omega(\omega)$.[12-13] The latter can be obtained in a straight-forward manner from equilibrium simulations by means of time correlation functions (TCF) of atomic velocities **v**. For atomic liquids one obtains, for example:

$$\Omega(\omega) = \frac{2m}{k_B T} \int e^{i\omega\tau} \langle \mathbf{v}(t)\mathbf{v}(t+\tau)\rangle_t \, d\tau \qquad (1)$$

Here, m is the particle mass, $\langle ... \rangle_t$ denotes an ensemble average, k_B is Boltzmann's constant and T the system temperature. Within the 2PT approach, the VDOS is described as a superposition of an abstract solid consisting of harmonic oscillators (HO) and a gas-like hard sphere (HS) fluid, whose thermodynamic quantities can both be described analytically based on their VDOS.[12] The key feature of the 2PT approach is to separate the total VDOS of the liquid into contributions belonging to these analytical models, $\Omega^{HS}(\omega)$ and $\Omega^{HO}(\omega)$, which represent distinct subsets of the degrees of freedom of the simulated liquid. Diffusive motions, which determine the zero-frequency intensity of the VDOS through the Green-Kubo relation of the velocity time-correlation function, are described entirely by the HS fluid model. The specifics of the abstract HS model are defined by its number density and particle size, which within the 2PT approach are determined by a fixed procedure based on the number density of the simulated liquid and its diffusion coefficient. This in turn describes the HS model contribution to the liquid VDOS, which describes a fraction f of the total number of degrees of freedom (N_DOF) in the liquid and includes the full liquid VDOS at zero-frequency, $\Omega(0) = \Omega^{HS}(0)$, and relaxation contributions at non-zero frequencies. The VDOS of the remaining degrees of freedom, $(1-f)\, N_\mathrm{DOF}$, which are described within the HO model, is obtained simply as:

$$\Omega^{HO}(\omega) = \Omega(\omega) - \Omega^{HS}(\omega) \qquad (2)$$

The thermodynamic properties of both models, including the absolute entropy, are available from analytical expressions. The entropy of the HS model for $f N_\mathrm{DOF}$ degrees of freedom is determined by the Carnahan-Starling equation of state,[15] the entropy of the HO model is obtained by treating $\Omega^{\mathrm{HO}}(\omega)$ as a continuous distribution of harmonic oscillators, for which the absolute entropy of each degree of freedom is described by

$$S^{HO}(\omega) = \frac{\beta\hbar\omega}{\exp(\beta\hbar\omega)-1} - \ln[1-\exp(-\beta\hbar\omega)] \quad (3)$$

with $\beta = 1/(k_B T)$. Variations of the method exist for the treatment of polyatomic molecules, in which the degrees of freedom describing molecular translation, rotation and intramolecular vibration are treated separately, resulting in VDOS expressions $\Omega_{trans}(\omega)$, $\Omega_{rot}(\omega)$ and $\Omega_{intra}(\omega)$.[13, 16] Diffusive rotational motion is then described by a rigid rotor approximation. The latter approach is employed here, however, contributions of intramolecular vibrations are deliberately ignored as they were shown to contribute only a miniscule contribution to the entropy at room temperature.[16] In the quantum HO model expressions, the intramolecular vibrational modes i with frequency ω_i^0 of water are essentially restricted to the ground state in thermal equilibrium due to $\hbar\omega_i^0 \gg k_B T$.

In this study, we utilized a spatially resolved version of the 2PT method, which allows studying contributions of individual hydration sites to the solvation entropy of a solute (3D–2PT).[14] This approach decomposes the simulation box into a grid of voxels of specified dimensions, which is used to define a local water VDOS computed from velocity time correlation functions up to a maximum correlation time of 1.6 ps. Employing the 2PT equations to this local VDOS results in local entropy expressions per water molecule found in a given site **r**, from which a solvation entropy contribution can be obtained by comparison to bulk water:

$$S_{solv}(\mathbf{r}) = S(\mathbf{r}) - S^{bulk} \quad (4)$$

The voxels of the analysis grid in this work uniformly exhibit a cubic volume, V, with an edge length of 2.25 Å. Local entropies were calculated for each voxel in which a minimum number of water molecules, i.e. local water number density $\rho_W(\mathbf{r})$, was observed during a given simulation trajectory. This threshold was set to 30% of the bulk water density, ρ_W^{bulk}. This filter eliminated low occupancy hydration sites, *i.e.* hydrations sites with very unfavorable local excess chemical potentials ($\mu_W^{ex}(\mathbf{r})/k_B T = -\ln\left[\frac{\rho_W(\mathbf{r})}{\rho_W^{bulk}}\right] > 1.2$). This is the case, for example, for voxels occupied by protein atoms where essentially no water molecules are observed. These low water occupancy voxels do not contribute significantly to the solvation of the studied enzyme or enzyme-substrate complex. Further, the absence of sufficient numbers of water molecules impedes the sampling of a statistically converged VDOS and the corresponding entropy estimation. Notably, the number density for a given voxel in an analysis grid can fluctuate from simulation to simulation due to slow conformational fluctuations of the protein. Results from

independent simulations were combined by computing weighted averages for local entropies from equivalent positions of the analysis grid in each simulation. The local water density $\rho_{\mathrm{W}}^{i}(\mathbf{r})$, observed in simulation *i* was used as a weighting factor, unless it fell short of the threshold value mentioned above, in which case 0 was used to eliminate contributions from this trajectory to the local average.

$$\bar{S}(\mathbf{r}) = \frac{\sum_i w_i(\mathbf{r}) S_i(\mathbf{r})}{w_i(\mathbf{r})} \quad \text{with} \quad w_i(\mathbf{r}) = \begin{cases} 0 & \text{if } \rho_{\mathrm{W}}^{i}(\mathbf{r}) < 0.2\, \rho_{W}^{\mathrm{bulk}} \\ \rho_{\mathrm{W}}^{i}(\mathbf{r}) & \text{otherwise} \end{cases} \tag{5}$$

With this definition, we obtain contributions to the solvation entropy from a given region in the hydration environment of the enzyme/enzyme-substrate complex as a sum of averaged voxel contributions times the product of the local average water number density $\bar{\rho}_W(\mathbf{r})$ and the volume *V*:

$$\Delta S_{\mathrm{solv}}^{\mathrm{region}} = \sum_j^{\mathrm{region}} \bar{\rho}_W(\mathbf{r}_j) \cdot V \cdot \left[\bar{S}(\mathbf{r}_j) - S^{\mathrm{bulk}}\right] \tag{6}$$

Statistical error bars for the reported entropy values were estimated based on the variations of local water entropies and number densities from a set of independent simulations.

SIMULATION METHODS

The KE07 and KE70 designed models for the apo state were obtained from the PDB database (PDB ID: 2RKX and 3NPU, respectively). The EL structures for KE07 and KE70 were obtained from the original KE07 and KE70 publications.[17] The LDE evolved structures, KE07-R7.2 and KE70-R6, for both the apo and reactant states were modeled via homology modeling using MODELLER[18].[19] The corresponding site mutations obtained from LDE are described in Table S1 and S2 in the Supplementary Information. The substrate geometry and position of the active site for the EL* state was the same as in the EL state with only the charges changed to reflect the TS. All 4 distinct proteins in the apo, substrate-bound complex EL and EL* TS states (12 systems in total) were solvated in water boxes with water extending to a minimum of ~10 Å from the surface of the protein.

The solvated systems were subjected to an energy minimization and were then equilibrated for 500ps in the NPT ensemble at 300K and 1atm, followed by a subsequent 3 ns NVT trajectory at 300K, in which structures and velocities were saved every 100ps. The resulting collection of 30 independent snapshots was then used as starting points for separate NVE trajectories of 100ps length each, for which coordinates, velocities and dipole moments

were saved every 8 fs. The resulting 30 x 100 ps trajectories per system were then used to calculate the local translational and rotational VDOS of water molecules in voxels of the analysis grid, which was defined relative to the position and orientation of the enzyme and the enzyme-substrate complex. Bulk water properties, *i.e.* the bulk water number density and entropy, were determined for water molecules with a minimum distance of 10 Å to the closest protein atom.

All simulations were carried out using the AMOEBA force field with the OpenMM[20] software package on NVIDIA GPUs; details of the AMOEBA force field are reported elsewhere.[21-22] The parameters for the ligand in the EL and EL* states were generated using the standard AMOEBA parameterization protocol[22] and have been reported in Ref. [3].

RESULTS

The 2PT entropy of bulk water has been calculated for the AMOEBA model as 57.9 J mol^{-1} K^{-1}, which is underestimated with respect to the experimental value of 69.9 J mol^{-1} K^{-1}.[16] This variation can be attributed to the approximate nature of the 2PT entropy expression and is in line with the previously reported consistent underestimation of the bulk water entropy by 2PT for the majority of established empirical water models as well as water in *ab initio* MD simulations.[13, 16] Nonetheless, a systematic underestimation of the absolute entropy does not impede our analysis of entropy variations relative to bulk water (described by the same model) within the hydration shell of our studied systems, which we carried out using the spatially resolved 3D-2PT method as described in the Methods section.

The spatial extent of the solvation layer (or equivalently the number of hydration waters, N_{water}) for a biomolecule can be defined based on a multitude of hydration water properties. In addition to structural properties, describing *e.g.* the local density[23] or the tetrahedrality[24] of the water hydrogen bond network, dynamical retardation factors are often used to quantify the impact of interactions with the solvated molecule surface and to estimate the range of solute-induced effects.[7, 9, 25-29] Single particle dynamics, as observed for example by NMR experiments, identify a dynamical slowdown that is restricted to the first solvation shell.[25] On the other hand, for collective motions on the sub-picosecond timescale, as probed by THz spectroscopy or extended depolarized light scattering, non-bulk properties were reported to be extending up to ~15-20 Å from the surface of solvated proteins.[9, 29] Other recent studies identify perturbed sub-diffusive hydration water dynamics extending ~5-6 Å or roughly 2 hydration layers from the protein-water interface.[26] In summary, it can therefore be concluded that estimates for the extent

of thickness of effective solvation layers depend on the experimental observable utilized as a probe. Here, we utilize the local solvent entropy based on estimates obtained by the 3D-2PT approach, which essentially reports on the average local structure of the liquid through the spectrum of intermolecular vibrations and local diffusion properties.

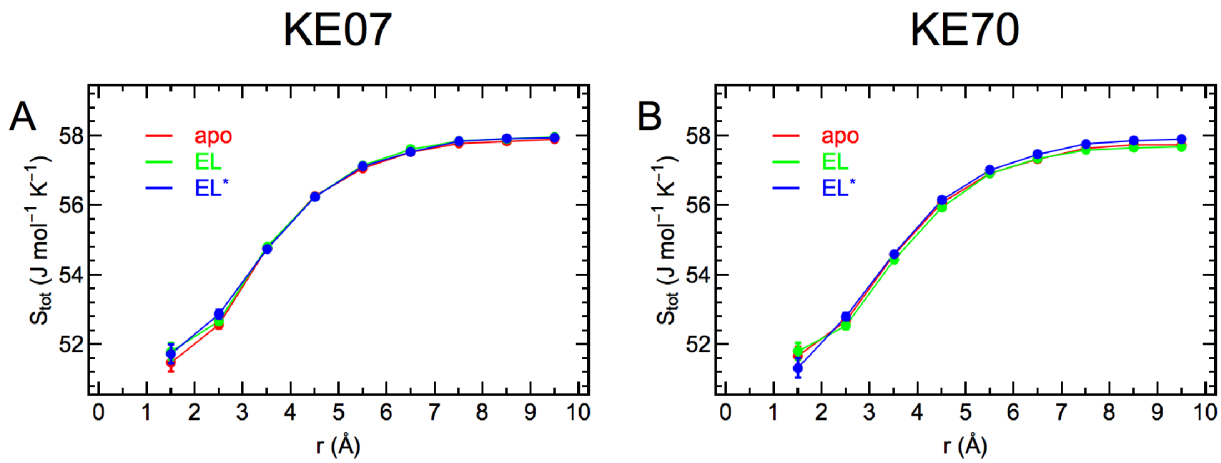


Figure 2. Solvent entropy per water molecule as a function of distance from the nearest protein surface for KE07 (A) and KE70 (B) enzymes during distinct stages of catalysis.

Figure 2 depicts the average local solvent entropy as a function of distance from the closest protein surface for the enzymes KE07 and KE70. We observe a significant reduction of the solvent entropy primarily within the first hydration layer (2-3 Å), which amounts to -6 J mol^{-1} K^{-1} per hydration water molecule and shows only minor variations between both enzymes, upon binding of the substrate, or formation of the TS. A lower-than-bulk entropy can be observed up to distances of ~6-7 Å from the protein surface, while essentially bulk-like properties are retained for larger distances. Equivalent observations are made for the enzyme variants obtained after several rounds of LDE, KE07-R7.2 and KE70-R6, which are not explicitly shown for clarity. This result matches previous studies on sub-diffusive properties for hydration water,[26] which may indicate a close correlation between local diffusion dynamics and the local entropy. A direct relation between diffusion and entropy has already been established previously for bulk atomic liquids.[30]

Next, we analyzed the distribution of local solvent entropies at varying hydration sites in close proximity of the solvated enzyme surface. For this purpose, we selected voxels of the analysis grid within 6 Å of the closest protein atom for the apo KE07 and KE70 enzymes, i.e. within a distance for which a significantly reduced entropy is observed in Fig. 2. The

corresponding probability distributions of local solvent entropies exhibit a pronounced asymmetry (Fig. 3) featuring a long tail for low entropies, which indicates the presence of distinct types of hydration water species in the hydration shell, similar to a previous report on an enzyme of the matrix metalloproteinase family.[14]

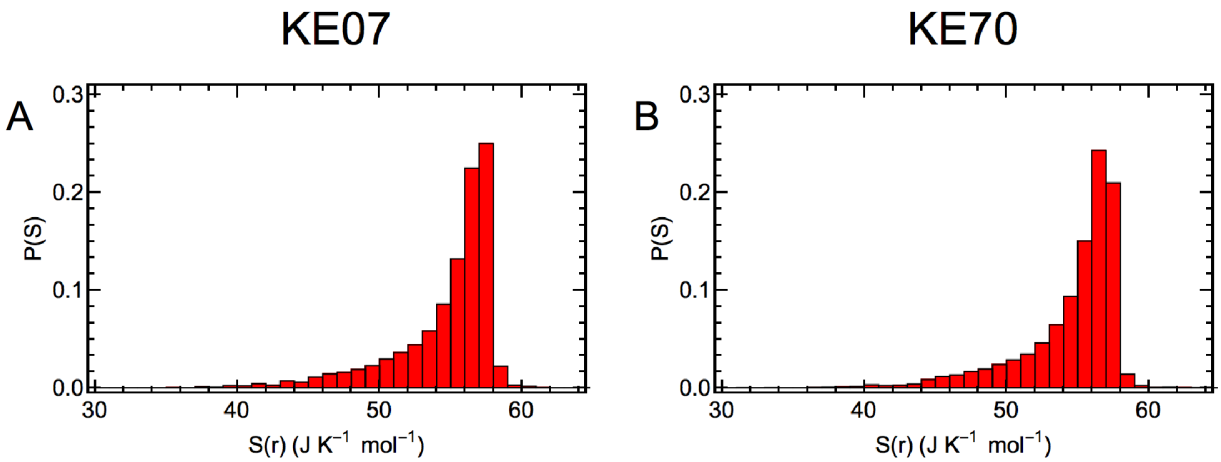


Figure 3: Histogram of local solvent entropies obtained for all hydration sites within 6 Å of the surfaces of the KE07 (A) and KE70 (B) enzymes prior to binding of the substrate (apo). Equivalent distributions were observed in presence of the substrates and for the LDE-derived enzyme variants, which are omitted here for brevity.

In order to analyze the distinct species of hydration water, we refine a procedure employed in a previous study[14] and identify 5 distinct parameters that describe the properties and the local environment of water molecules in their respective hydration sites: 1) the local number density (if $\frac{\rho_W(\mathbf{r})}{\rho_W^{\text{bulk}}} > 0.3$); 2) the local entropy as obtained from 3D-2PT; 3) the local diffusivity as characterized by the zero-frequency intensity of the VDOS; 4) the integral of the translational VDOS between 200 and 300 cm^{-1}; 5) the integral of the rotational VDOS between 200 and 400 cm^{-1}. The latter two parameters describe the number of vibrational modes involved in high frequency hydrogen bond stretch vibrations and low-frequency librations, respectively. The translational VDOS is obtained from velocity TCF's of water molecule centers of mass, while the rotational VDOS is defined by angular velocity TCF's around the three moments of inertia as described in Refs. [13] and [14]. Distinct hydration water species were then identified via a *k*-means clustering algorithm within the 5-dimensional space spanned by these parameters. This allowed us to distinguish k=4 distinct hydration water species, whose properties we characterize as "bound", "weakly-bound", "unbound" and "bulk-like". The chosen names are based on indications for the strength of binding to the protein surface, as discussed in the following. The

number of distinct species *k* is an input parameter for the algorithm and is chosen to identify the maximum number of statistically distinct hydration water species. In comparison to previous work,[14] the inclusion of low occupancy hydration sites allowed us to identify hydration sites with higher-than-bulk entropies that are specific to the KE07 ad KE70 artificial enzymes.

The average VDOS for each water species, the separate contributions of translational and rotational degrees of freedom, and the underlying distributions of local hydration water entropies are shown in Figure 4 for hydration water of the KE07 enzyme in the apo state. Strong binding to the protein is typically indicated in particular by a pronounced reduction of the average local hydration water entropy, a reduced diffusivity, *i.e.* $\Omega_{\mathrm{trans}}(0)$ and $\Omega_{\mathrm{total}}(0)$, enhanced numbers of high frequency hydrogen bond stretch vibrations ($\Omega_{\mathrm{trans}}(\omega)$ between 200 and 300 cm^{-1}) and a reduced number of low-frequency librations ($\Omega_{\mathrm{rot}}(\omega)$ between 200 and 400 cm^{-1}).

The VDOS of the "bound" hydration water species exhibits the lowest diffusivity (Figs. 4A and 4C), a pronounced high-frequency shoulder in the translational VDOS with higher than bulk intensities beyond 200 cm^{-1} (Fig. 4A), a significant blue-shift of librational modes in the rotational VDOS (Fig. 4B) and an overall reduced local hydration water entropy (distribution average of 48.0 J K^{-1} mol^{-1} in Fig. 4D), justifying the classification as bound water. "Weakly bound" hydration is characterized by a less pronounced decrease in diffusivity and the local entropy (distribution average of 54.8 J K^{-1} mol^{-1}), which go along with a small increase of the translational VDOS between 100 and 200 cm^{-1} relative to bulk, while the signatures of hydrogen bond stretch vibrations and librations are essentially bulk-like. The "unbound" hydration water species features a reduced diffusivity, while at the same time its vibrational signatures exhibits properties characteristic of a highly-strained hydrogen bond network environment. The high-frequency shoulder in the translational VDOS, which describes hydrogen bond stretch vibrations, exhibits a pronounced red-shift and compared to bulk water reduced intensities beyond 200 cm^{-1}. Likewise, the libration band in the rotational VDOS is distorted and features an increase of low-frequency modes between 200 and 400 cm^{-1}. The corresponding distribution of the local entropies shows an overall reduction relative to bulk water (distribution average of 55.3 J K^{-1} mol^{-1}), however, we also note a high entropy tail with higher-than-bulk entropies, which contributes to this species of hydration water molecules. Overall, these properties indicate geometrical constraints for translational motion (reduced diffusivity), but a reduced of hydrogen bond formation, either to the protein or surrounding water molecules (red-shift of intermolecular

vibrations, high local entropies). "Bulk-like" water, as the name suggests, exhibits average properties that essentially match the properties of bulk water at distances >10 Å from the protein. These water molecules are primarily found at distances far away from the protein surface (c*f.* Fig. 2) or contribute only a miniscule amount to the more immediate hydration of the catalytically active center analyzed in detail in the following.

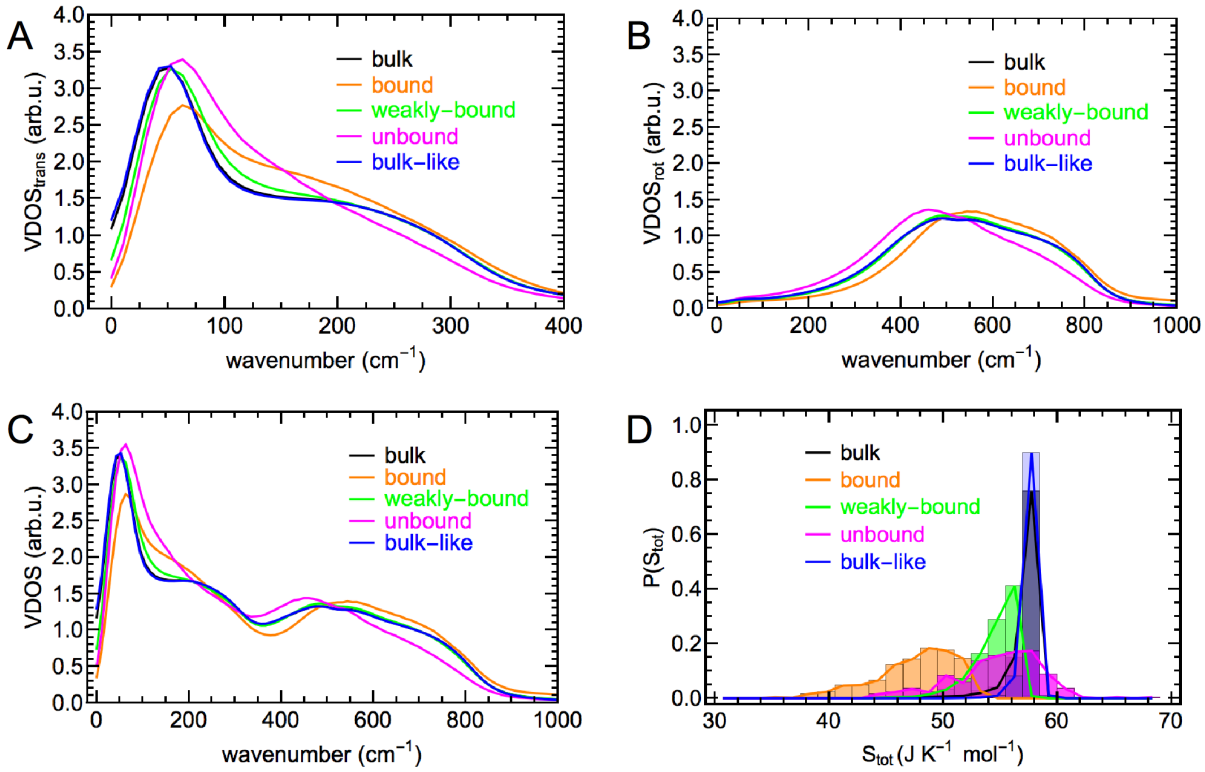


Figure 4: Thermodynamic and vibrational properties of the distinct water species obtained from *k*-means clustering. A) Translational, B) rotational and C) total VDOS of the individual water species in comparison to the bulk. D) Probability distribution of the solvent entropies of individual water species.

Equivalent hydration water species with identical average properties were observed in simulations of all 12 distinct systems, however, we observed significant variations in the composition (*i.e.,* percentage of each type of water species), total number of water molecules and average solvation entropy contributions for the hydration environment of the catalytic site between both enzyme variants, KE07 and KE70, before and after LDE, and for the analyzed stages of the catalytic process (Figure 5 and 6). Figure 5 shows a detailed comparison of the composition of hydration water within a 10 Å spherical volume around the catalytic center for the KE07 and KE07-R7.2 enzyme variants during the simulated stages of the catalytic process.

Shown are the explicit numbers of water molecules within the selected spherical volume, assigned to each hydration species described in Fig. 4, together with the total entropic solvation free energy contribution, $-T\Delta S_{\mathrm{solv}}$, obtained from summation over the selected spherical volume as described by Eq. 6.

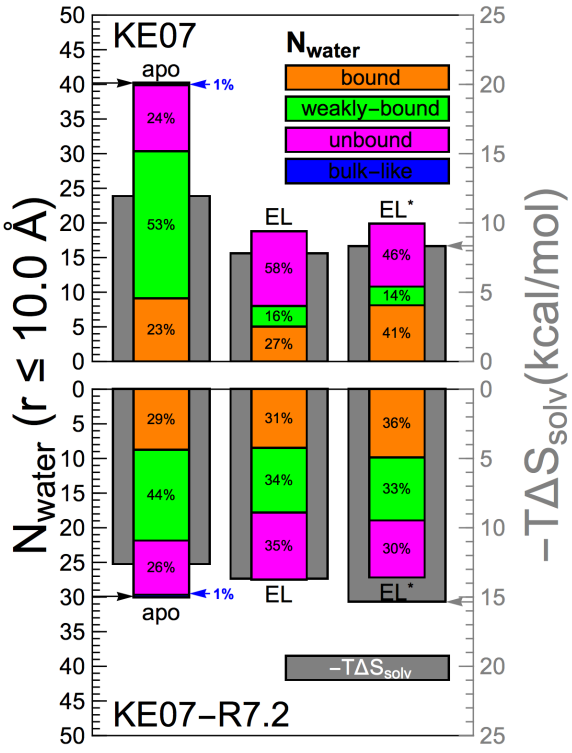


Figure 5: Comparison of hydration patterns of the catalytic center for the KE07 and KE07-R7.2 enzyme variants. Left axis and colored bars: number of hydration water molecules within a 10 Å spherical radius around the catalytic center and decomposition into hydration water species. The upper and lower halves of the figure show results for KE07 and KE07-R7.2, respectively. Right axis and gray bars: entropic solvation free energy contribution of all water molecules within the spherical volume according to Eq. 6.

For the apo-form of the original KE07 enzyme (upper half of Fig. 5), ~40 water molecules are observed in the vicinity of the catalytic site. During substrate binding (*i.e.,* formation of the EL state), roughly 50% of these water molecules are expelled from the catalytic site and are released into the bulk. The released water consists primarily of "bound" and "weakly bound" waters. The release of these low entropy water molecules reduces the unfavorable, negative solvation entropy of the catalytic site, resulting in an apparent free energy change of $-T\Delta S_{\mathrm{solv}} \approx$ -4 kcal/mol. This observation is in line with the idea that strongly bound, low-entropy hydration water molecules at ligand-binding sites can promote binding.[31] However, we note that changes in the hydration water entropy are insufficient to predict solvation

contributions to the free energy, which must include also the solvation enthalpy. Hence, we cannot claim that this change in solvation entropy plays a relevant role for the substrate-binding process. However, it does clearly highlight a change in the solvation pattern of the catalytic site upon substrate binding, which is in the focus of the present study. The formation of the EL* state results in no significant change in the number of water molecules in the selected volume (relative to EL), while however the fraction of the "bound" water species increases in the TS at the expense of "unbound" water. Using simulation to directly calculate the substrate molecular dipole moment in the EL and EL* states (Figure S2), the dipole moment of the EL* state is much larger (3.2 D vs 1.8 D for the EL state), which likely explains the increase in bound waters for the TS.

Comparison to the optimized KE07-R7.2 enzyme after LDE (lower half of Fig. 5) reveals significant changes in the hydration of the catalytic site, as well as its evolution during the catalytic process. Only ~30 water molecules are found within 10 Å of the catalytic center for the apo KE07-R7.2 enzyme. Interestingly, the difference to apo KE07 lies primarily in the number of "weakly-bound" water molecules, while the amount of "bound" and "unbound" water remains approximately the same. Overall, the unfavorable entropic contribution to the solvation free energy of these 30 water molecules is even larger in magnitude than for the designed KE07 enzyme, despite the lower number of water molecules involved. Hence, the average entropy of water molecules hydrating the catalytic site must be significantly more reduced than one might conclude based on our decomposition into discrete hydration water species.

Table S1 shows that 3 out of 8 mutations introduce a new polar site in KE07-R7.2, while the remaining mutations are neutral in terms of polarity. An increased overall polarity of the catalytic center is likely to result in additional attractive protein-water interactions, which would be compensated by a reduced entropy as observed in our analysis. Interestingly, the number of water molecules within the analyzed volume around the catalytic site remains almost constant upon substrate binding for KE07-R7.2: only ~3 water molecules are released to provide room for the substrate. This is a significant change relative to the designed KE07, for which ligand binding is accompanied by the release of ~20 water molecules.

Our decomposition scheme indicates that substrate binding to KE07-R7.2 results in the conversion of "weakly-bound" into "unbound" water. Based on average properties of these hydration water species this would translate into a small increase of the solvation entropy, i.e. a

favorable contribution to a change in the solvation free energy. However, the explicitly calculated change in solvation entropy reports an opposite effect. This seemingly contradicting observation is caused by the remaining width of the entropy distribution reported for the individual hydration water species in Fig. 4D. The explicitly computed change in entropy indicates that the average hydration water entropies decrease upon substrate binding, *e.g*. due to enhanced binding or steric factors. Overall, this entropic signature of the change in solvation is inverted qualitatively compared to the designed KE07 enzyme, for which substrate binding resulted in a favorable change of the solvation entropy. Despite the apparent differences in hydration of the KE07 and KE07-R7.2 enzyme variants, the formation of the TS invokes comparable changes in the hydration pattern for both enzyme variants. Similar to our observation for KE07, the number of water molecules remains approximately constant, the fraction of "bound" water increases and the solvation entropy decreases upon TS formation for KE07-R7.2, resulting in an unfavorable contribution to a change in the solvation free energy of +2 kcal/mol. This is consistent with previous studies which have shown that the improved catalytic activity of the evolved KE07 enzyme relative to its design is for the trivial reason that the EL state is destabilized[10, 32-33].

Our analysis for the KE70 enzyme and its LDE derived variant KE70-R6 indicates a very different role of hydration for an optimized catalytic activity, as shown in Fig. 6. In the apo form, the catalytic center of the KE70 designed enzyme is significantly less hydrated than KE07, with just ~18 water molecules within the analyzed spherical volume (upper half of Fig. 6). Upon substrate binding, i.e. formation of the EL complex, 3-4 water molecules are expelled, which roughly corresponds to the volume required by the entering substrate. The composition analysis indicates that the expelled water molecules can be mainly categorized as "unbound" and correspondingly the change in solvation entropy is essentially neutral. The formation of the TS, similar to the previously studied KE07 enzyme variants, does not change the number of hydration water molecules. However, the composition analysis, as well as the change in the solvation entropy, indicate an increase in the number of "bound", low-entropy water molecules, indicating a similar destabilization of the EL state for KE70-R6 as observed for KE07-R7.2.

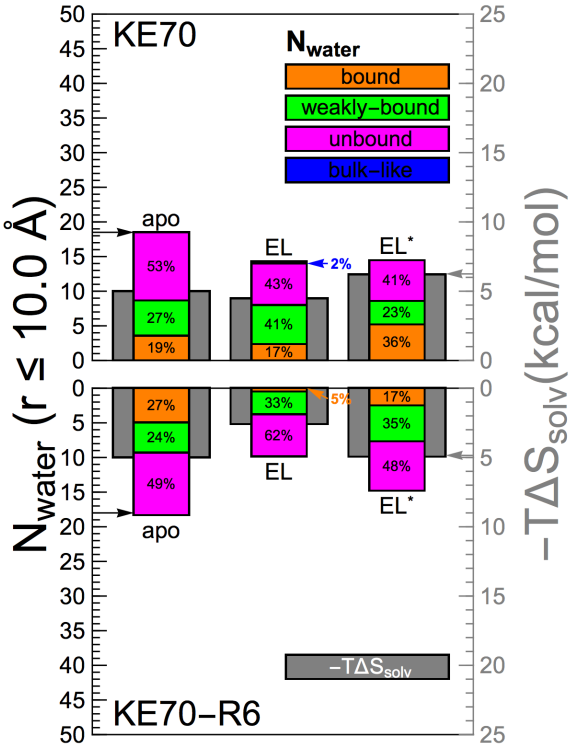


Figure 6: Comparison of hydration patterns of the catalytic center for the KE70 and KE70-R6 enzyme variants. Left axis and colored bars: number of hydration water molecules within a 10 Å spherical radius around the catalytic center and decomposition into hydration water species. The upper and lower halves of the figure show results for KE70 and KE70-R6, respectively. Right axis and gray bars: entropic solvation free energy contribution of all water molecules within the spherical volume according to Eq. 6.

The LDE-optimized KE70-R6 enzyme features a decreased number of charged polar amino acids relative to KE70 as indicated in Table S2. Despite this, the hydration of the apo-state of the enzyme is almost unchanged compared to KE70 (Fig. 6, lower half), while the fraction of bound water molecules is even increased. Even the solvation entropy contribution of water within 10 Å of the catalytic center is essentially identical to the KE70 design. Upon ligand binding ~8 primarily "bound" water molecules are expelled from the catalytic center, in contrast to a minor change in hydration upon substrate binding for the designed KE70 enzyme, such that "bound" water molecules are now almost absent in the EL complex. Consequently, the solvation entropy shows a sizable, favorable change for the EL formation in KE70-R6, which by itself would amount to a free energy contribution of –2.5 kcal/mol. Only after TS formation, EL*, a fraction of the expelled "bound" water molecules re-enters the volume around the catalytic center and essentially restores the solvation entropy observed for the apo-state, albeit with a reduced number of water molecules. Thus hydration would contribute to an increase in the free

energy difference between the EL and EL* state, slowing down catalysis, but apparently this is compensated for with improved binding of the hydrophobic substrate with the more greasy active site cavity, consistent with the nearly order of magnitude improvement in K_M for the evolved KE70 enzyme.

DISCUSSION AND CONCLUSIONS

Catalytic site hydration can play a significant role in enzymatic reaction mechanisms. We have used a two-phase thermodynamic (2PT) theory, which utilizes the spectrum of vibrational modes described by the vibrational density of states[12-13], to analyze the changes in hydration for two designed Kemp Eliminases and their evolution using LDE to improve their performance. Overall, our analysis shows that LDE affects the hydration of the catalytic center for two Kemp Eliminases in distinct ways towards improved enzymatic activity, but rendered through changes mostly in the hydration entropy destabilization of the EL state.

The catalytic site of the improved KE07-R7.2 enzyme is hydrated by a smaller number of water molecules, which are, however, more strongly bound in average as indicated by the more negative solvation entropy, consistent with the increase in polarity of the catalytic site. Upon substrate-binding, only a minimal amount of water is expelled from the catalytic center, just enough to make room for the substrate, while in case of the original KE07 enzyme, a major change in the hydration is observed. In contrast to KE07, LDE applied to the KE70 enzyme rendered the catalytic site more hydrophobic, expelling a large number of "bound" low entropy water molecules, in line with a potential role of low-entropy water molecules to promote ligand-binding. As mentioned above, changes in the solvation enthalpy also need to be considered, before speculating on a potential role of solvation in stabilizing the substrate-bound (EL) state. However, it is intriguing that directed evolution results in opposite changes in the evolution of hydration during substrate binding for the two studied KE designs, but with a net effect of destabilizing the EL state in both cases.

A potential role of solvation for the formation of the transition state is less apparent from our analysis, *i.e.* the effects of LDE on changes in catalytic site hydration are significantly less pronounced for the TS formation for all designed and evolve enzyme systems. The number of water molecules near the reaction center increases slightly or remains constant, while the solvation entropy decreases. The latter effect is more pronounced for the KE70 enzymes than for KE07, but in both cases changes introduced by directed evolution are much less apparent than

for the formation of the initial enzyme-substrate complex, EL. While counterintuitive, it is consistent with the general feature of LDE when applied to designed Kemp Eliminase enzymes, where mutations are found to reduce the free energy difference between the EL and EL* state by destabilizing the EL state, as opposed to stabilizing the EL* state as natural enzymes do.[10, 32-33]

Supporting Information Available: KE70 Mutations, Ligand dipole moment for EL vs EL*, Changes in residue nature due to mutations.

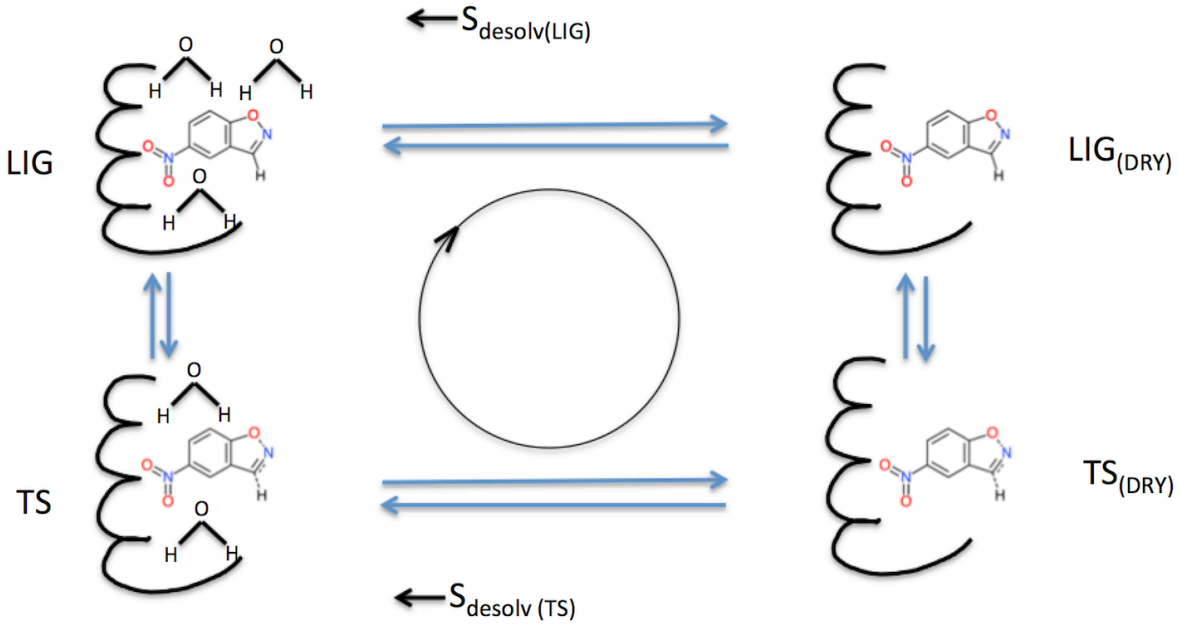
ACKNOWLEDGEMENTS. The hydration entropy analysis is supported by the Cluster of Excellence RESOLV (EXC 169) funded by the Deutsche Forschungsgemeinschaft (DFG), the RESOLV Graduate School "Solvation Science" (GSS), and UC Berkeley CalSolv center. THG also acknowledges support by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 for applications to catalysis.

REFERENCES

1. Khanjin, N. A.; Synder, J. P.; Menger, F. M. Mechanism of Chromismate Mutase: Contribution of Conformational Restriction to Catalysis in the Klaisen Rearrangement. *J. Am. Chem. Soc.* **1999,** *121* (50), 11831-11846.
2. Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006,** *106*, 3210-3235.
3. Bhowmick, A.; Sharma, S. C.; Head-Gordon, T. The Importance of the Scaffold for de Novo Enzymes: A Case Study with Kemp Eliminase. *J. Am. Chem. Soc.* **2017,** *139* (16), 5793-5800.
4. Fenimore, P. W.; Frauenfelder, H.; McMahon, B. H.; Young, R. D. Bulk-Solvent and Hydration-Shell Fluctuations, Similar to Alpha- and Beta-Fluctuations in Glasses, Control Protein Motions and Functions. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101* (40), 14408-14413.
5. Frauenfelder, H.; Fenimore, P. W.; Chen, G.; McMahon, B. H. Protein Folding is Slaved to Solvent Motions. *Proc. Nat. Acad. Sci. U. S. A.* **2006,** *103(42)*, 15469-15472.
6. Johnson, M. E.; Malardier-Jugroot, C.; Murarka, R. K.; Head-Gordon, T. Hydration Water Dynamics Near Biological Interfaces. *J. Phys. Chem. B* **2009,** *113*, 4082-4092.
7. Russo, D.; Hura, G.; Head-Gordon, T. Hydration Dynamics near a Model Protein Surface. *Biophys. J.* **2004,** *86* (3), 1852-1862.
8. DiTucci, M. J.; Heiles, S.; Williams, E. R. Role of Water in Stabilizing Ferricyanide Trianion and Ion-Induced Effects to Hydrogen-Bonding Water Network at Long Distance. *J. Am. Chem. Soc.* **2015,** *137(4)*, 1650-1657.
9. Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M. An Extended Dynamical Hydration Shell around Proteins. *Proc. Nat. Acad. Sci. U. S. A.* **2007,** *104(52)*, 20749-20752.

10. Bhowmick, A.; Sharma, S. C.; Honma, H.; Head-Gordon, T. The Role of Side Chain Entropy and Mutual Information for Improving the de Novo Design of Kemp Eliminases KE07 and KE70. *Phys. Chem. Chem. Phys.* **2016,** *18*, 19386-19396.
11. Kiss, G.; Rothlisberger, D.; Baker, D.; Houk, K. N. Evaluation and Ranking of Enzyme Designs. *Protein Science* **2010,** *19* (9), 1760-1773.
12. Lin, S.-T.; Blanco, M.; Goddard, W. A. I. I. I. The Two-Phase Model for Calculating Thermodynamic Properties of Liquids from Molecular Dynamics: Validation for the Phase Diagram of Lennard-Jones Fluids. *J. Chem. Phys.* **2003,** *119(22)*, 11792-11805.
13. Lin, S.-T.; Maiti, P. K.; Goddard, W. A. I. I. I. Two-Phase Thermodynamic Model for Efficient and Accurate Absolute Entropy of Water from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2010,** *114(24)*, 8191-8198.
14. Pattni, V.; Vasilevskaya, T.; Thiel, W.; Heyden, M. Distinct Protein Hydration Water Species Defined by Spatially Resolved Spectra of Intermolecular Vibrations. *J. Phys. Chem. B* **2017,** *121(31)*, 7431-7442.
15. Carnahan, N. F.; Starling, K. E. Thermodynamic Properties of a Rigid-Sphere Fluid. *J. Chem. Phys.* **1970,** *53* (2), 600-603.
16. Pascal, T. A.; Scharf, D.; Jung, Y.; Kühne, T. D. On the Absolute Thermodynamics of Water from Computer Simulations: A Comparison of First-Principles Molecular Dynamics, Reactive and Empirical Force Fields. *J. Chem. Phys.* **2012,** *137* (24), 244507.
17. Röthlisberger, D.; Khersonsky, A. M.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; et. al. Kemp Elimination Catalysis by Computational Enzyme Design. *Nature* **2008,** *453*, 190-195.
18. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using Modeller. *Curr. Protoc. Bioinform.* **2014,** *5.6*, 1-32.
19. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinform.* **2016,** *54*, 5.6.1-5.6.37.
20. Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.; Beauchamp, K. A.; Lane, T. J.; Wang, L. P.; Shukla, D.; et. al. OpenMM 4: A Reusable, Extensible, Hardware Independent Library for High Performance Molecular Simulation. *J. Chem. Theor. Comput.* **2013,** *9(1)*, 461-469.
21. Ren, P.; Ponder, J. P. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003,** *107*, 5933-5947.
22. Ren, P.; Wu, C.; Ponder, J. W. Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. *J. Chem. Theo. Comp.* **2011,** *7*, 3143-3161.
23. Merzel, F.; Smith, J. C. Is the First Hydration Shell of Lysozyme of Higher Density than Bulk Water? *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99* (8), 5378-5383.
24. Duboue-Dijon, E.; Laage, D. Characterization of the Local Structure in Liquid Water by Various Order Parameters. *J. Phys. Chem. B* **2015,** *119* (26), 8406-8418.
25. Denisov, V. P.; Halle, B. Protein Hydration Dynamics in Aqueous Solution. *Faraday Discuss* **1996,** *103*, 227-244.
26. Perticaroli, S.; Ehlers, G.; Stanley, C. B.; Mamontov, E.; O'Neill, H.; Zhang, Q.; Cheng, X. L.; Myles, D. A. A.; Katsaras, J.; Nickels, J. D. Description of Hydration Water in Protein (Green Fluorescent Protein) Solution. *J. Am. Chem. Soc.* **2017,** *139* (3), 1098-1105.
27. Pal, S. K.; Peon, J.; Zewail, A. H. Biological Water at the Protein Surface: Dynamical Solvation Probed Directly with Femtosecond Resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99* (4), 1763-1768.

28. Nickels, J. D.; Sakai, V. G.; Sokolov, A. P. Dynamics in Protein Powders on the Nanosecond-Picosecond Time Scale Are Dominated by Localized Motions. *J. Phys. Chem. B* **2013,** *117* (39), 11548-11555.
29. Comez, L.; Paolantoni, M.; Sassi, P.; Corezzi, S.; Morresi, A.; Fioretto, D. Molecular Properties of Aqueous Solutions: A Focus on the Collective Dynamics of Hydration Water. *Soft Matter* **2016,** *12* (25), 5501-5514.
30. Dzugutov, M. A Universal Scaling Law for Atomic Diffusion in Condensed Matter. *Nature* **1996,** *381* (6578), 137-139.
31. Nucci, N. V.; Pometun, M. S.; Wand, A. J. Site-Resolved Measurement of Water-Protein Interactions by Solution NMR. *Nat. Struct. Mol. Biol.* **2011,** *18* (2), 245-249.
32. Frushicheva, M. P.; Cao, J.; Chu, Z. T.; Warshel, A. Exploring Challenges in Rational Enzyme Design by Simulating the Catalysis in Artificial Kemp Eliminase. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107*, 16869-74.
33. Frushicheva, M. P.; Cao, J.; Warshel, A. Challenges and Advances in Validating Enzyme Design Proposals : The Case of Kemp Eliminase Catalysis. *Biochemistry* **2011,** *50* (18), 3849-3858.



TOC Graphic