"Toward improved non-aqueous biocatalysis: Mapping relationships between enzyme structure, interfacial solvent dynamics, and enzyme dynamics"

Katie R. Mitchell-Koch, Wichita State University, Department of Chemistry

The use of enzymes in organic solvent expands the scope of reactions catalyzed by a given enzyme, provided that the enzyme retains its structure and activity in solution. Hydrophobic substrates, such as those derived from petroleum, are solubilized in organic solvent, and competing reactions, such as hydrolysis, are suppressed in organic media. One of the drawbacks of biocatalysis in organic solvent is that enzymes generally have hindered dynamics and suppressed catalytic rates, relative to their function in water. Our grant work has focused on uncovering whether these hindered protein dynamics can be attributed to hindered dynamics of solvent in the protein solvent shell. Studies of proteins in their native solvent, water, indicate that protein and solvent dynamics are intimately connected. Furthermore, a phenomenon in non-aqueous enzymology called "salt activation" has been observed. When enzymes are lyophilized in the presence of a salt, then used in organic solvent, they display higher kinetics, on the order of $10\text{-}10^4$, depending on the enzyme, solvent, and reaction. Measurements of the dynamics of residual waters in the enzyme solvation shell, done by NMR spectroscopy, indicated faster moving water molecules around salt-activated enzymes, correlating with faster enzymatic reaction rates.

The PRF funded work is based on the following hypotheses: 1) Dynamics of solvent molecules at the enzyme surface are heterogeneous (as has been observed in the hydration dynamics of many proteins); at different regions of the protein, solvent will be slowed to a different extent, relative to its dynamics in the bulk liquid; 2) The regions of fast and slow solvent will vary by the nature of the organic solvent, and will likely be different from the pattern of heterogeneous dynamics observed for the enzyme in water; 3) Locally slower solvent shell dynamics provide a higher-friction environment and effectively a higher local viscosity, hindering the local motions of the protein.

Molecular dynamics (MD) simulations have been used to characterize the dynamics of the solvation shell by region around a solvent-compatible enzyme, *Candida antarctica* lipase B, in water, acetonitrile, *n*-butanol, *tert*-butanol, and cyclohexane.

As described in last year's report, we first determined sound methodology for MD simulations of enzymes in organic and mixed organic-aqueous solutions. Namely, we determined which crystallographic waters to keep (how many, and which, based on their dynamics and location) in order to have the most quickly equilibrated protein structure. The rationale for using structural equilibration as a criterion for analysis is to maximize computational efficiency, and also based on the well-founded postulate that a rapidly-equilibrated protein has started the simulation at a lower point on the potential energy surface and is more likely to find a valid, commonly-sampled, stable equilibrium structure. We used our recommended methods for simulations of the solvent-compatible enzyme *Candida antarctica* lipase B (CALB) in water and organic solvents. The focus of simulations analysis has been protein dynamics and solvent dynamics. The dynamics of solvent in the first and second solvation shell have been evaluated regionally around the surface of the protein, characterizing solvent properties around each secondary structure (α-helix, β-sheet, loop region). The dynamics of the protein in aqueous solvent serves as a comparison to the dynamics of CALB in organic solvent.

In evaluating the hydration dynamics, several correlations were observed: fastest solvation layer dynamics are observed around regions of the enzyme having highest flexibility (indicated by root mean square fluctuation values); and the most dense regions of solvent (indicated by radial distribution function around protein atoms) exhibit the fastest solvation layer dynamics. We calculated the excess (pairwise) entropy of the solvation layer water molecules, and showed that the regions of high density and fast diffusion have the highest pairwise entropy (lower tetrahedral order). This finding for the biomolecular solvation layer echoes what is seen for many bulk liquids: a relationship between solvent density, diffusion, and entropy that is called Rosenfeld scaling. For simple liquids, diffusion decreases as density increases (and excess entropy decreases), while for bulk water and other tetrahedral liquids, the density-diffusion relationship is reverse: diffusion increases as density and excess entropy increase concomitantly. We are currently evaluating whether Rosenfeld scaling is valid for organic solvents. Our data describing the solvation layer structure-dynamics around different regions of CALB indicates there is a density-diffusion relationship for organic solvents as well.

Another goal of the work has been to determine how protein structure influences solvation layer dynamics. For the water data, we used multi-regression analysis, and determined that both hydrophobicity and surface curvature were correlated with water dynamics, with curvature contributing more to the correlation (~80%). That is to say, hydrophilic and convex topologies exhibit faster hydration dynamics and hydrophobic and concave topologies exhibit a hindered hydration layer. One might expect some correlation between hydrophobicity and

topology- after all, proteins fold to bury hydrophobic residues, and convex pockets are often found in protein interiors. However, in this case, we limited our analysis to the exterior of the protein, where the majority of the solvent-accessible surface area is located. Furthermore, we examined the correlation between hydrophobicity and topology, and found poor correlation ($R^2 \sim 50\%$). Thus, in studying CALB, the two effects can be separated. We are using the same approach (multi-regression analysis) to determine how organic solvents with different polarity, shape, size, and hydrogen-bonding properties are influenced by the protein interface, but are considering other descriptors of the protein surface (following our hypothesis that protein surface properties affect different solvents in different ways).

In comparing solvent dynamics for a given region of CALB across multiple solvents, it is seen that the dynamics vary considerably. Some solvents will be dramatically slowed relative to their dynamics in bulk, while others will be only mildly affected by interfacial interactions. This supported our working hypothesis. We also compared the flexibility of the protein across different solvents, and saw correlations between protein and solvent mobility. We compared side chain mobility and $C\alpha$ (protein backbone) flexibility separately, and together, and saw consistent correlations between all forms of protein dynamics and the local viscosity of the solvent in its regional solvation shell (where local viscosity, η_{local} is approximated by $\eta_{bulk} D_{bulk}/D_{solvent\ shell}$). This supported our hypothesis that protein dynamics in organic solvent are hindered by slower solvation layer dynamics. Furthermore, we saw that in certain regions where organic solvent moves faster than water, the flexibility of the protein is higher than in water. Figure 1 shows the flexibility of the $C\alpha$ atoms for 3 different regions of CALB versus the local solvation shell viscosity in water, acetonitrile, *n*-butanol, *tert*-butanol, and cyclohexane.

Next, we used Markov state modeling to quantitatively compare the rates of conformational transitions in different solvents. What we found was that conformational sampling shifts by solvent, relative to the native conformational dynamics of CALB, such that the states are different, and one-to-one comparisons cannot be made directly. However, we can see that opening and closing rates of the $\alpha5$ helix that gates the substrate-binding pocket are substantially slower in *n*-butanol and *tert*-butanol, which move slowly around the $\alpha5$ helix. This is true even though the $\alpha5$ helix is displaced to a smaller extent in butanols than in water. The results of the Markov state modeling have led to a new hypothesis that protein conformational transitions in organic solvent can be modeled (quantified) by a modified Kramers' equation, considering local solvation layer viscosity and changes to the protein's potential energy surface due to solvent effects. Work is underway for this.

Figure 1: $C\alpha$ backbone flexibility (relative to water, from integrated RMSF) versus local solvent shell viscosity ($1/\eta_{local}$) for helix $\alpha10$ (top) loop 23-32 (center) and loop 246-267 (bottom). *Published in* Mitchell-Koch and Dahanayake, *Frontiers in Molecular Biosciences* **2018**, 5:65.

Over the past year, the grant has supported two graduate students, both of whom defended their doctoral dissertations. One graduated in December 2017, and the other in May 2018. One of the graduate students has continued on doing postdoctoral work supported by the PRF grant. The other has taken a postdoctoral position at National Institute of Environmental Health Sciences in North Carolina. Both have gained valuable experience in molecular dynamics simulations, data analysis, and coding that have translated into success across multiple research endeavors. The initial results obtained from this work were used in a successful grant proposal to the National Science Foundation with Co-PI Prof. Vinh Nguyen from the Department of Physics at Virginia Tech (dielectric spectroscopy of solvation layer dynamics and protein collective motions). Therefore, the research directions started with this Doctoral New Investigator grant will be continuing as a funded research effort for at least another two years. We anticipate the results having a broad impact in understanding the role of solvent in biomolecular function and in expanding the use of non-aqueous biocatalysis. The next steps are to investigate the role of selective mutations in altering solvent shell dynamics, and potentially altering protein dynamics. If our hypothesis is valid, this will provide a new strategy in protein design and engineering, providing a rational means to improve catalytic efficiency of enzymes in organic solvent. This in turn will provide new strategies for energy- and atom-efficient biocatalytic transformations of petroleum-deriv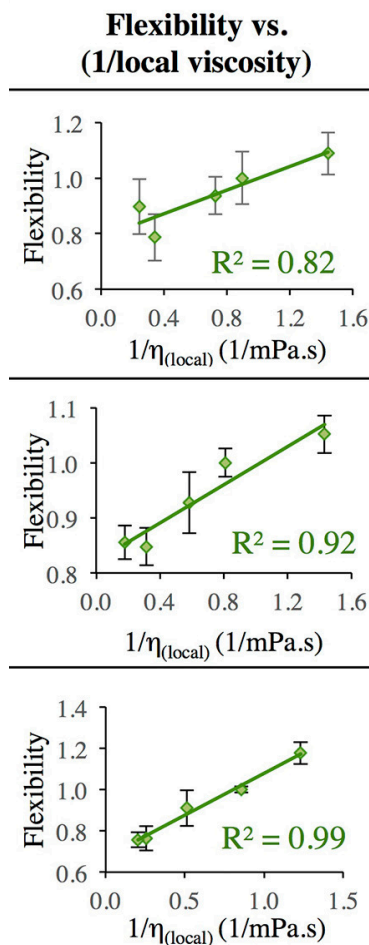ed substrates into commodity and fine chemicals.