PAPER

Check for updates

Universal kinetic solvent effects in acid-catalyzed reactions of biomass-derived oxygenates†

Theodore W. Walker, ‡[a] Alex K. Chew, ‡[a] Huixiang Li,[bc] Benginur Demir,[ad]
Z. Conrad Zhang,[b] George W. Huber,[a] Reid C. Van Lehn *[a] and
James A. Dumesic*[ad]

The rates of Brønsted-acid-catalyzed reactions of ethyl *tert*-butyl ether, *tert*-butanol, levoglucosan, 1,2-propanediol, fructose, cellobiose, and xylitol were measured in solvent mixtures of water with three polar aprotic cosolvents: γ-valerolactone; 1,4-dioxane; and tetrahydrofuran. As the water content of the solvent environment decreases, reactants with more hydroxyl groups have higher catalytic turnover rates for both hydrolysis and dehydration reactions. We present classical molecular dynamics simulations to explain these solvent effects in terms of three simulation-derived observables: (1) the extent of water enrichment in the local solvent domain of the reactant; (2) the average hydrogen bonding lifetime between water molecules and the reactant; and (3) the fraction of the reactant accessible surface area occupied by hydroxyl groups, all as a function of solvent composition. We develop a model, constituted by linear combinations of these three observables, that predicts experimentally determined rate constants as a function of solvent composition for the entire set of acid-catalyzed reactions.

Broader context

This article presents a model-predictive tool for the computational design of liquid-phase biomass conversion processes over Brønsted-acid catalysts. Catalytic conversion of lignocellulosic biomass is a promising method to produce liquid fuels and chemicals from a renewable carbon resource. Liquid-phase, acid-catalyzed reactions of biomass-derived oxygenates are of particular interest in this context, being characterized by rapid catalytic turnovers and low catalyst costs. One simple strategy to improve product yields for such processes is to change the solvent composition. However, broadly-applicable design rules are not available to guide the solvent selection process for new liquid phase biomass conversion processes. Herein, we demonstrate how the reactivity of biomass-derived oxygenates in mixtures of water and polar aprotic cosolvents can be understood in terms of the hydrophilicity of the reactant. We explain this behavior by using classical molecular dynamics simulations to quantify the size and properties of water-enriched local domains that nucleate in the immediate vicinity of the reactants in mixed solvent environments, effectively collocating reactants and catalytic protons in solution. By correlating experimentally determined reaction rates with the properties of these water-rich domains, we provide a method for estimating rates of acid-catalyzed reactions of biomass-derived oxygenates in mixed solvent environments. The development of this modeling tool represents an important step toward a general understanding of solvent effects in acid-catalyzed biomass conversion processes, and toward the rational design of new liquid phase processes.

Introduction

Acid-catalyzed reactions in the liquid phase are ubiquitous in the production and upgrading of renewable biomass-derived oxygenates,[1–5] which have garnered interest as sources of organic carbon for the production of renewable liquid fuels.[5–10] An important variable to control the reactivity and selectivity for these catalytic processes is the solvent composition, which affects reaction rates,[11,12] product selectivities,[13,14] the stability of desired products,[1,15–17] and the economics of downstream separations. Prior studies have shown that aqueous mixtures containing polar aprotic cosolvents (*i.e.*, mixed-solvent environments) are of particular interest in this context; a minimum amount of water is often required to facilitate the solvation of

[a] *Department of Chemical and Biological Engineering, University of Wisconsin – Madison, Madison, WI, 53706, USA. E-mail: vanlehn@wisc.edu, jdumesic@wisc.edu*

[b] *Dalian National Laboratory for Clean Energy, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, 457 Zhongshan Road, Dalian 116023, Liaoning, China*

[c] *University of Chinese Academy of Sciences, Beijing 100049, China*

[d] *DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI, 53706, USA*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7ee03432f

‡ These authors contributed equally.

This journal is © The Royal Society of Chemistry 2018

Energy Environ. Sci., 2018, **11**, 617–628 | 617

biomass-derived materials, while the cosolvent can enhance reaction performance.[11,18,19] For example, the rate of xylose dehydration to furfural in aqueous mixtures of γ-valerolactone increases 30-fold compared to the same reaction in pure water, while the formation of undesired humins *via* degradation of the reactant and product is suppressed,[12] improving the furfural selectivity by over 50%.

While extensive knowledge has been accumulated regarding optimal solvent compositions for key applications,[12,13,15,20,21] it is not generally possible to anticipate how mixed-solvent environments will perform in new processes *a priori*, because the mechanistic details underlying kinetic solvent effects in multi-component systems are not well understood.[22–24] Computational efforts in the past decade have relied on *ab initio* quantum chemical methods to quantify the effects of solvent on the barriers to elementary reaction steps.[25] These studies have provided detailed insights in important case studies,[26–30] but broadly applicable design rules have not been developed due (in part) to the limitation of *ab initio* techniques to capture the slow intermolecular re-organizations that constitute solvent–solute interactions.[31–37]

In contrast to *ab initio* methods, classical molecular dynamics (MD) simulations can access longer time scales ($\sim\mu$s) and larger length scales ($\sim$nm), and at a lower computational cost. MD simulations therefore permit characterization of the solvent environment in the immediate vicinity of a reactant molecule (*i.e.*, the local solvent domain of the reactant), which can be compared to the solvent environment far from the reactant (*i.e.*, the bulk solvent domain).[21,38,39] For example, MD simulations have examined the preferred configurations of solvent molecules at biomass-relevant reactant surfaces.[21,27,38] Based on key reaction rate measurements in this study and insights from our recent work,[40] we hypothesize that trends in acid-catalyzed reaction rates as a function of solvent environment can be related to the formation and properties of a water-rich or -deficient local domain near the reactant. These properties can be quantified *via* classical MD simulations, and moreover can be determined for a large series of possible reactants and solvent compositions given the computational efficiency of classical MD simulations.

Herein, we report the effects of mixed-solvent environments, consisting of water mixed with a polar aprotic cosolvent, on experimentally determined rates of Brønsted-acid-catalyzed reactions for seven biomass-derived model compounds. Irrespective of the selected cosolvent, consisting of 1,4-dioxane, γ-valerolactone, and tetrahydrofuran, we find that reactants with more available hydroxyl groups become more reactive as the water content of the solvent environment is decreased, and this behavior is true for both hydrolysis and dehydration reactions considered in this study. By contrasting properties between local and bulk solvent domains using MD simulations, we develop a computational approach to predict reaction rates as a function of solvent composition. We predict experimentally determined reaction rates using three computationally determined observables: (1) the extent of water enrichment in the local solvent domain of the reactant (Γ); (2) the average hydrogen bonding lifetime between water molecules and the reactant (τ); and (3) the fraction of the reactant accessible surface area occupied by hydroxyl groups (δ).

To our knowledge, the approach developed in this study provides the first tool of its kind in the context of biomass conversion in multicomponent solvent environments. As such, this study represents a step toward the model-predictive design of liquid-phase biomass conversion technologies. Moreover, this approach demonstrates that contrasting properties between local and bulk solvent domains using MD simulations can provide insight into the kinetics of acid-catalyzed reactions in mixed-solvent environments.

Results and discussion

Universal effects of reactant and cosolvent properties on reactivity

Scheme 1 shows the seven Brønsted-acid catalyzed reactions considered in this study, including those reported in our prior work.[40] For brevity, the reactant abbreviations in Scheme 1 are used throughout this text. Reactions were carried out in pure water, and in aqueous mixtures of the three polar aprotic cosolvents: 1,4-dioxane (DIO); γ-valerolactone (GVL); and tetra-hydrofuran (THF). Reactants in Scheme 1 are organized according to decreasing hydrophilicity, as estimated by the accessible hydroxyl fraction (δ), which we have defined as the accessible surface area (ASA) occupied by the (N) hydroxyl groups in a reactant molecule ($\text{ASA}_{\text{OH},k}$) normalized by the ASA occupied by the (M) total atoms in the molecule, shown in eqn (1).

$$\delta = \frac{\sum_{k=1}^{N} \text{ASA}_{\text{OH},k}}{\sum_{l=1}^{M} \text{ASA}_l} \qquad (1)$$

We interpret this dimensionless quantity as a qualitative molecular descriptor of the reactant hydrophilicity, normalized by its molecular size; larger values of δ indicate more hydrophilic molecules. See ESI† for more details on all molecular descriptors.

The forward rates of all reactions (r_i) in Scheme 1 are described by apparent first-order kinetics with respect to the concentrations of the reactants (C_i) and acidic protons in solution (C_{H^+}).[40–43] Accordingly, we measured apparent rate constants in each solvent environment by fitting the experimental reactions kinetics data (obtained in batch reactors) to expressions in the form of eqn (2), where ($k_{\text{org},j}^i$) is the apparent rate constant for the ith reaction, and the subscript denotes the identity and composition (in jth mass fraction) of the organic cosolvent.

$$r_i = -\frac{\text{d}C_i}{\text{d}t} = k_{\text{org},j}^i C_i C_{\text{H}^+} \qquad (2)$$

Fig. 1 shows the measured apparent rate constants for xylitol (XYL) dehydration in DIO/water mixtures ($k_{\text{DIO},j}^{\text{XYL}}$), normalized by the rate constant in pure water ($k_{\text{H}_2\text{O}}^{\text{XYL}}$), as a function of the mass fraction of DIO in the solvent environment (m_{DIO}). As the mass fraction of DIO increases, the value of the rate constant for XYL dehydration increases by nearly two orders of magnitude.
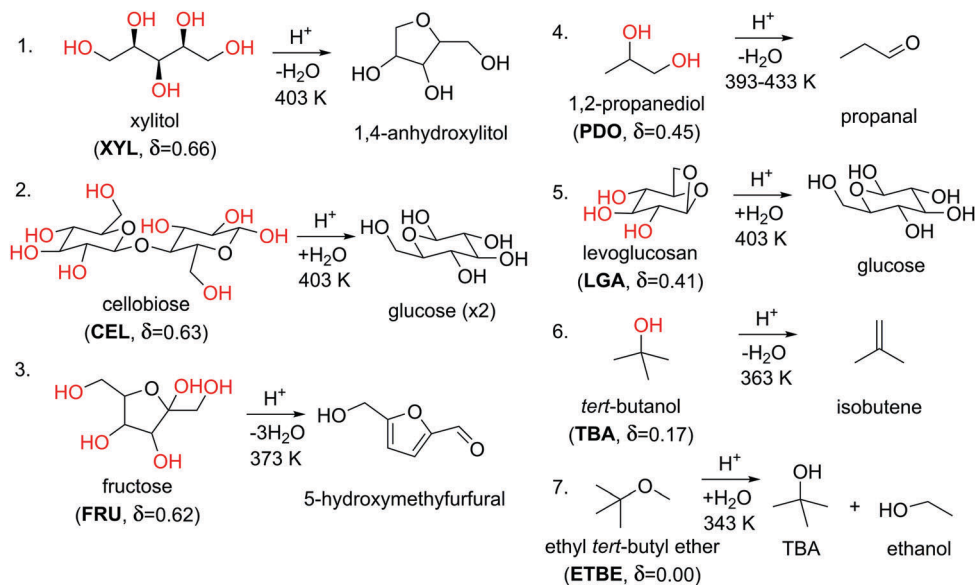
Scheme 1 Brønsted acid-catalyzed reactions of seven model compounds. Rate constants associated with reactions 3, 4, and 6 were taken from prior work.[40] Hydroxyl groups are highlighted in red for emphasis. Reactants are arranged according to decreasing hydrophilicity, as estimated by the δ parameter.
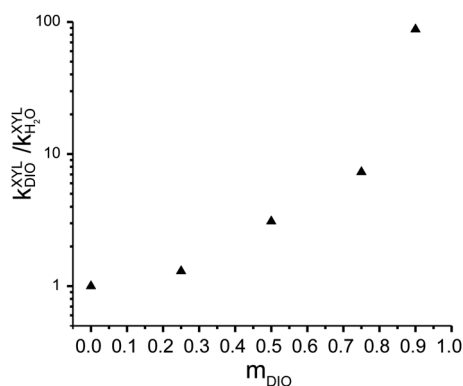
Fig. 1 Apparent rate constant for XYL dehydration normalized by the rate constant in pure water *versus* the mass fraction of the organic cosolvent in DIO/water mixtures. Reaction conditions: 75–200 mM XYL; 0.03–1.3 M trifluoromethane sulfonic (triflic) acid; 403 K.

Table 1 Kinetic solvent parameters for the Brønsted-acid-catalyzed dehydration of XYL in mixtures of water with three organic cosolvents. Reaction conditions: 403 K; 0.015–1.3 M triflic acid. Confidence intervals were calculated at the 95% confidence level (see ESI)

$k_{H_2O}^{XYL} = 1.04 \times 10^{-4} \pm 9.3 \times 10^{-6}$ L mol^{-1} s^{-1}

Mass fraction of the organic phase	$\sigma_{GVL,j}^{XYL}$	$\sigma_{DIO,j}^{XYL}$	$\sigma_{THF,j}^{XYL}$
0.90	2.05 ± 0.07	1.80 ± 0.08	1.85 ± 0.10
0.75	1.02 ± 0.08	1.02 ± 0.08	0.74 ± 0.13
0.50	0.41 ± 0.09	0.50 ± 0.08	N/A[a]
0.25	0.11 ± 0.10	0.18 ± 0.08	0.23 ± 0.07

[a] Not available due to phase separation between water and organic cosolvent.

In general, the values of the rate constants measured in this study are a strong function of solvent composition, and this phenomenon has previously been noted elsewhere.[11,12,40]

To compare reactivity trends across different reactions and solvent environments, the rate constants associated with each of the reactions in Scheme 1 were described in terms of a kinetic solvent parameter ($\sigma_{org,j}^i$) as shown in eqn (3).

$$\sigma_{org,j}^i = \log_{10}\left(\frac{k_{org,j}^i}{k_{H_2O}^i}\right) \quad (3)$$

Positive kinetic solvent parameters indicate an increase in the rate of reaction in a particular solvent environment compared to the same reaction in pure water, while negative values have a converse implication. Table 1 presents the kinetic solvent parameters for the rate of XYL dehydration in aqueous mixtures

of up to 90 wt% DIO, GVL, and THF. These results demonstrate a general trend of increasing reactivity with decreasing water content for each of the three cosolvents. A complete list of the kinetic solvent parameters collected in this study is presented in the ESI.†

Fig. 2(a) presents the kinetic solvent parameters for XYL and *tert*-butanol (TBA) dehydration in aqueous mixtures of DIO, GVL, and THF as a function of the mass fraction of the organic cosolvent (m_{org}). These two reactions represent upper and lower limits in our dataset with respect to reactant hydrophilicity and kinetic behavior. XYL has an accessible hydroxyl fraction (δ) of 0.66 and becomes monotonically more reactive with decreasing water content of the solvent environment. In contrast, TBA has an accessible hydroxyl fraction of 0.17, and its reactivity is a non-monotonic function of solvent composition.[40] In general, the extent to which the rate of TBA dehydration is affected by the increasing addition of organic cosolvent is smaller than that of XYL dehydration.

The comparison of XYL and TBA dehydration suggests that reactivity trends with respect to the water content of the solvent
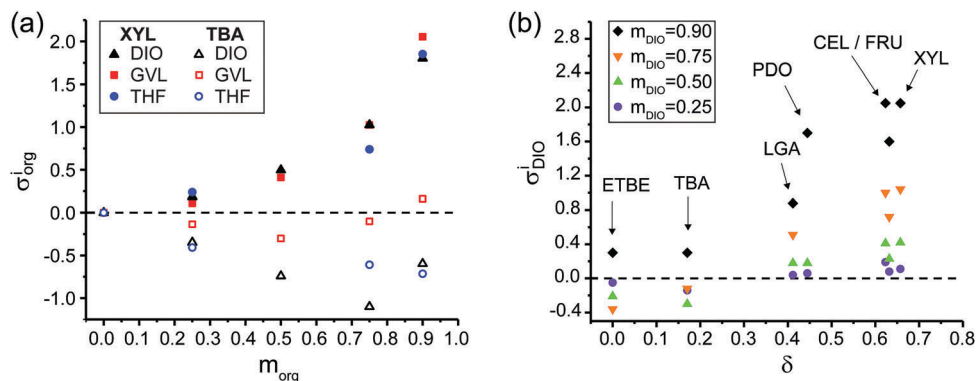
This journal is © The Royal Society of Chemistry 2018

Energy Environ. Sci., 2018, **11**, 617–628 | **619**

Fig. 2 Kinetic solvent parameters (σ_{org}^i) as a function of: (a) solvent composition in aqueous mixtures of GVL, XYL and THF (open symbols = TBA, closed symbols = XYL), and; (b) the accessible hydroxyl fraction (δ) in DIO/water mixtures. Reaction conditions: see Table S2 (ESI†).

environment may depend on the hydrophilicity of the reactant, as we have postulated elsewhere.[40] Fig. 2(b) presents the kinetic solvent parameters as a function of the accessible hydroxyl fraction for all seven reactions in aqueous DIO mixtures. At a fixed solvent composition, the kinetic solvent parameters generally become more positive, indicating increasing rates of reaction compared to pure water, with an increasing accessible hydroxyl fraction. Similar behavior was observed in GVL and THF mixtures (see ESI†). Note that this behavior is observed for both the dehydrations and hydrolysis reactions considered in this study.

Proposed mechanism: reaction rates correlate with formation of water-enriched local solvent domain

To understand the aforementioned trends in reactivity, we note that rates of chemical reactions are controlled by the thermodynamic properties of the reactants, catalyst, and transition states in the elementary steps.[24] A quantitative understanding of solvent effects therefore requires knowledge of the reaction mechanism, and a rigorous characterization of the transition-state-solvent interactions in each of the kinetically relevant steps.[30,35,40]

In many cases, however, catalytic reactions occur *via* a sequence of elementary steps where a single step is rate-limiting, and the rate is then controlled by the thermodynamic properties of the transition state for this step.[44–46] In these cases,[47,48] the reaction mechanism may be analyzed *via* a sequence of quasi-equilibrated steps, such that the reactant and proton are treated as being in equilibrium with the transition state. Furthermore, transition states in acid-catalyzed reactions typically display strong carbocation-like character.[46,49] Accordingly, as shown in Fig. 3(b), we consider the acid-catalyzed reactions in this study as being composed of two generalized steps, both of which can be impacted by the solvent:

(i) the transfer of a proton from the bulk domain to the reactant, and;

(ii) the formation of a carbocation-like transition state.

For step (i), we have shown that water-enriched local solvent domains form around hydroxyl groups in the presence of polar aprotic cosolvents.[40] In mixed-solvent environments, hydrophilic reactants thus drive the formation of local solvent domains in which the local density of water molecules near the reactant is greater than the density of water molecules in the bulk solvent, as
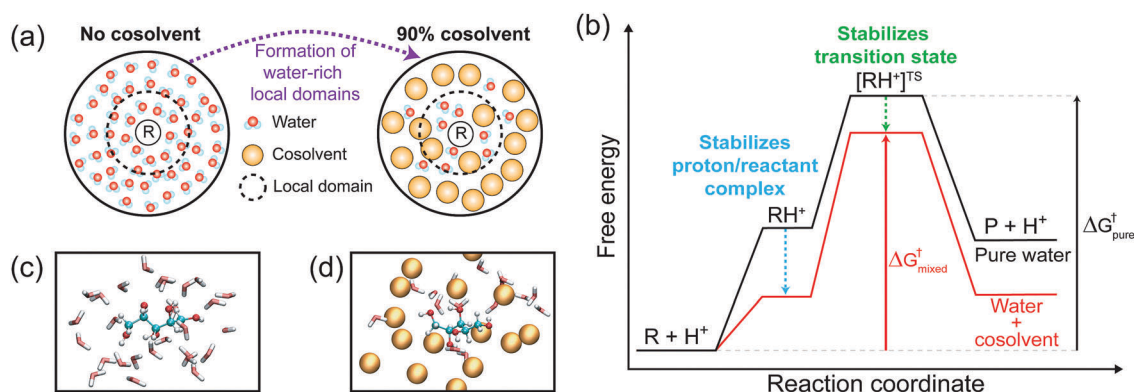

Fig. 3 (a) Role of cosolvent molecules on the distribution of solvent molecules. Favorable interactions with hydrophilic reactants in mixed-solvent environments drive the formation of water-rich local domains around the reactant. While there are fewer water molecules in the local domain relative to pure water, the local water density is enriched relative to the bulk density in the solvent mixture. (b) Proposed effect of cosolvent molecules on a reaction free energy landscape. Stabilization of the proton and transition state in the water-rich local domain, relative to the bulk domain, lowers the apparent free energy barrier for the reaction in a mixed-solvent environment. (c and d) MD simulation snapshots of XYL in (c) pure water and (d) 90 wt% DIO, which is drawn as a single representative bead to match the schematics in (a).

shown in Fig. 3(a). A proton is therefore destabilized in the bulk solvent relative to the local domain, <mark>because of its higher affinity for water than for the organic phase, leading to a thermodynamic driving force for transferring the proton to the reactant.</mark>[12,50–52] Vlachos and co-workers have postulated a similar mechanism for fructose dehydration in DMSO–water mixtures.[27]

For step (ii), we hypothesize that the stability of the reactant, proton, and transition state are correlated in these local solvent domains, because water molecules that bind strongly to the reactant are preorganized into configurations that stabilize charged transition states.[35,53] Similar solvent preorganization is thought to contribute to enzyme catalytic efficiency,[54] has been speculated as a key effect in the acid-catalyzed glucose to fructose isomerization reaction,[35] and may contribute to the enhanced reactivity of hydronium ions in confined environments (zeolites).[55,56]

We now explore whether reactant–solvent–cosolvent interactions can be tuned to deliberately drive the formation of water-enriched local domains near hydrophilic reactants. Water and the hydrophilic reactant are then characterized as being confined to a local domain, with confinement enhancing

reactivity by increasing reactant–proton association, and stabilizing the carbocation-like transition states common to acid-catalyzed reaction mechanisms. With MD simulations, we probe our hypotheses by analyzing the local solvent domain near the reactant and deriving simulation measurables that can be used to predict experimental reaction rates. Simulation snapshots for XYL in pure water and 90 wt% DIO are shown in Fig. 3(c) and (d), respectively.

MD simulations: formation of water-enriched local domains in solvent mixtures

Fig. 4 shows the radial distribution function (RDF), which quantifies the density of water molecules at a radius r away from a central molecule normalized by the bulk water density. The RDF is schematically depicted in Fig. 4(a) for a mixed-solvent environment in which XYL is the central molecule. <mark>The RDF for XYL in a 90 wt% DIO/water mixture and in pure water (0 wt% DIO/water) is shown in Fig. 4(b). We define the cutoff, r_{cutoff}, between the local and bulk solvent domains as the distance at which the RDF reaches unity.</mark>

RDFs between TBA–water and XYL–water in various DIO/water mixtures are shown in Fig. 4(c) and (d), respectively.
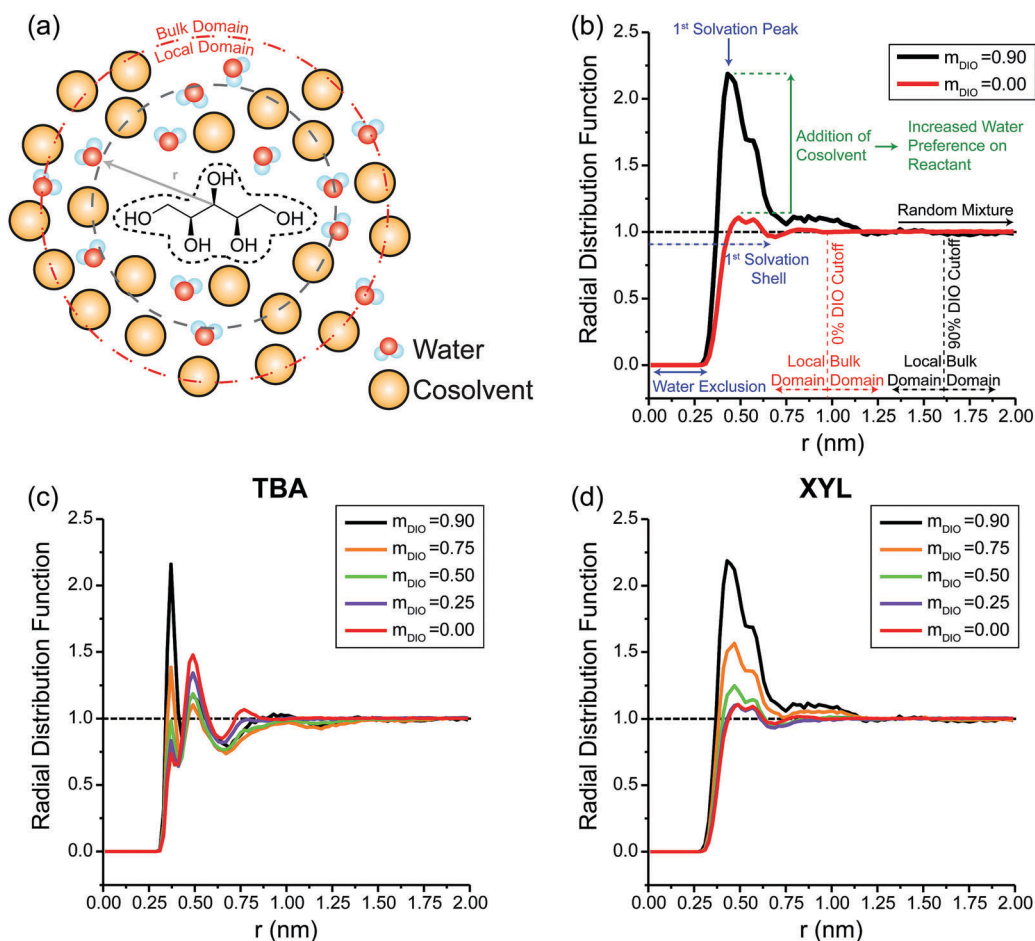
Fig. 4 (a) Schematic depiction of the radial distribution function for XYL. The distance, r, is calculated between the center of mass of the reactant and the oxygen atom of each water molecule. (b) Radial distribution function for XYL in 90 wt% DIO and pure water (m_{DIO} = 0). The cutoff between local and bulk domain is defined as the distance when the RDF between the reactant and water reaches unity (i.e. a random mixture). (c and d) Radial distribution function for TBA (c) and XYL (d) for various wt% of organic solvent.

This journal is © The Royal Society of Chemistry 2018

Energy Environ. Sci., 2018, **11**, 617–628 | 621

From the RDFs, we find that the water content in the local solvent domain of each reactant increases compared to the bulk solvent domain when a high concentration of cosolvent is present in the solvent environment. This behavior is apparent from the increase in the magnitude of the first solvation peak for systems containing large concentrations of DIO relative to its magnitude for a pure water system, indicating that water preferentially partitions to the local solvent domain around the reactant. The RDF for XYL has a broader first solvation peak than TBA, which indicates greater water enrichment in the local solvent domain when the reactant has more hydroxyl groups.

### Quantifying water enrichment within the local domain of the reactant

Because RDFs are difficult to compare across reactants with different concentrations of cosolvents, we calculated the preferential exclusion coefficient (Γ) as a molecular descriptor to quantify the local domain composition around the reactant. We define Γ as the excess number of cosolvent molecules within the local solvent domain of the reactant relative to the bulk solvent domain. Preferential exclusion coefficients are calculated from the MD simulations according to eqn (4), where n_C and n_W are the total number of cosolvent and water

molecules, and the superscripts L and B indicate molecules within the local and bulk domains, respectively.[57–60]

$$\Gamma = -\left\langle n_C^L - n_W^L \left(\frac{n_C^B}{n_W^B} \right) \right\rangle \qquad (4)$$

Positive values of Γ indicate that the concentration of cosolvent is lower in the local solvent domain of the reactant than in the bulk solvent domain. Positive Γ is also referred to as preferential hydration,[60] because the exclusion of cosolvent indicates that the reactant has a higher affinity for water. Negative values of Γ indicate that the concentration of cosolvent is higher in the local solvent domain of the reactant than in the bulk solvent domain and that the reactant has a higher affinity for the cosolvent.

Preferential exclusion coefficients (Γ) calculated for TBA and XYL in DIO/water and GVL/water mixtures at various cosolvent concentrations are shown as filled lines in Fig. 5(a) and (b). Experimentally determined kinetic solvent parameters (σ) are depicted as dashed lines for comparison. For TBA, Γ is negative with the exception of 90% GVL/water. Conversely, Γ is positive for XYL across the range of cosolvent compositions, which means that XYL preferentially excludes cosolvent and has a higher affinity for water. We find that Γ correlates well with σ, even capturing the non-monotonic behavior in TBA.
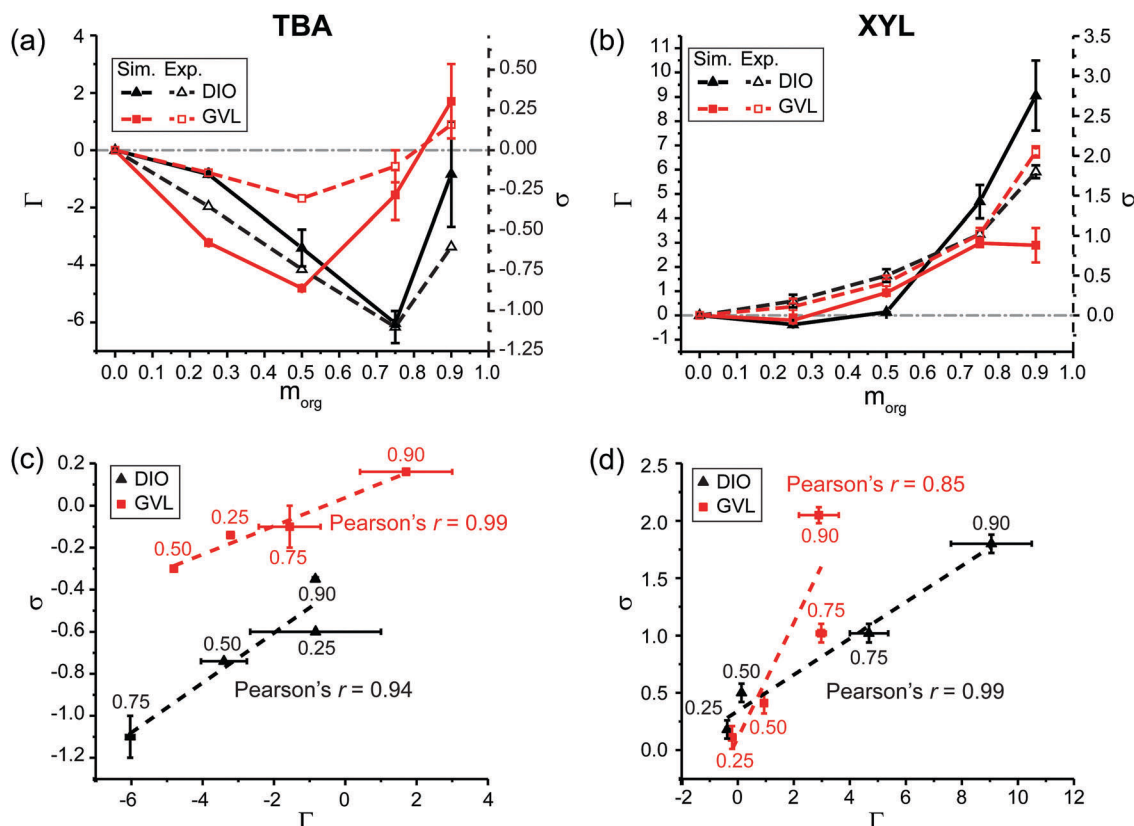
Fig. 5 Relationship between experimentally determined kinetic solvent parameters (σ) and simulated preferential exclusion coefficient (Γ) for (a) TBA and (b) XYL for various wt% of organic cosolvent in GVL/water and DIO/water mixtures. The gray dotted line denotes when σ and Γ are zero. Kinetic solvent parameters are also plotted against Γ for (c) TBA and (d) XYL. The dashed lines in (c) and (d) represent the best-fit line. Data points are labeled with the wt% of the organic cosolvent. Pure water systems have been omitted from parts (c) and (d) because σ and Γ will always be zero.

622 | *Energy Environ. Sci.*, 2018, **11**, 617–628

This journal is © The Royal Society of Chemistry 2018

Table 2 Pearson correlation coefficient between the kinetic solvent parameters (σ) and the preferential exclusion coefficient (Γ) for various reactants and cosolvents[a]

Cosolvent	Reactant[c]							All[b]
	ETBE	TBA	LGA	PDO	FRU	CEL	XYL	
DIO	**−0.08**	0.94	0.84	0.99	0.96	0.98	0.99	0.84
GVL	0.80	0.99	0.91	0.92	0.93	0.80	0.85	0.72
THF	**0.49**	**−0.68**	**0.76**	**0.27**	**0.45**	**0.55**	**0.26**	0.60

[a] Uses data from solvent mixtures containing 25 wt%, 50 wt%, 75 wt%, and 90 wt% of the cosolvent. [b] Uses data from all reactants and all four cosolvent wt%. [c] Data in bold have a Pearson's r less than 0.80.

Fig. 5(c) and (d) show a strong linear correlation between simulation-derived Γ and the experimentally determined σ, as indicated by Pearson correlation coefficients (Pearson's r). Pearson correlation coefficients close to 1 indicate a total positive linear correlation, whereas values near −1 indicate a linear negative correlation, and zero indicates no linear correlation. Pearson correlation coefficients for all reactants in each of the three cosolvent mixtures are summarized in Table 2. We find that Γ and σ are highly correlated (Pearson's $r \geq 0.80$) for the majority of the systems, with the exception of THF/water systems.

The agreement between simulations and experiments suggests that higher reaction rates correspond to the enrichment of water near the reactant despite differences in reactant hydrophilicity and reaction mechanisms. However, the poor correlation between Γ and σ in some reactant/cosolvent environments suggests that additional parameters that characterize the local solvent domain will improve our understanding of the solvent effects that contribute to experimentally determined reaction rates.

Quantifying reactant–water hydrogen bonding strength

Following the hypothesis that charged transition states may be stabilized by water molecules in the local solvent domain that are preorganized into favorable binding configurations, we next calculated the average reactant–water hydrogen bonding lifetime (τ_{HB}) as a molecular descriptor to quantify the strength of water binding to the reactant in the mixed-solvent systems.[61] We expect that hydrogen bonds between the reactant and water are stronger (*i.e.*, longer-lasting) in mixed-solvent environments that have large cosolvent concentrations, because water–water hydrogen bonds are unable to form, increasing the preference of reactant–water hydrogen bonds. Stronger interactions between water and reactant may translate to a lower transition state free energy and thus, an increase in reaction rate.[62,63]

Hydrogen bonding lifetimes for TBA and XYL are shown in Fig. 6. When increasing the concentration of the organic cosolvent, hydrogen bonding lifetimes between the reactant and the surrounding water molecules increase monotonically across different cosolvent environments, indicating stronger reactant–water hydrogen bonds. XYL has a lower hydrogen bonding lifetime compared to TBA, possibly due to a higher reaction temperature for XYL (403 K *vs.* 363 K); higher system temperatures often result in lower hydrogen bonding lifetimes. To remove temperature effects, we define the hydrogen bonding lifetime
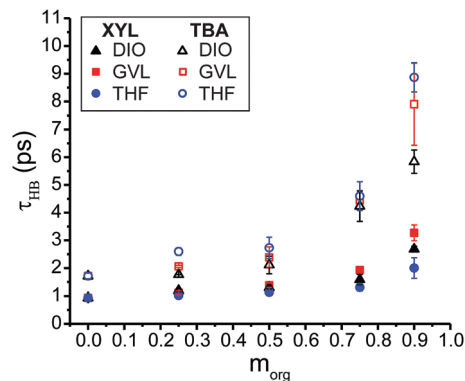
Fig. 6 Average reactant–water hydrogen bond lifetimes (in picoseconds) for TBA and XYL as a function of mass fraction of organic cosolvent in DIO/water, GVL/water, and THF/water mixtures.

ratio, τ, by normalizing the hydrogen bonding lifetime in the organic cosolvent mixtures ($\tau_{HB,org}$) by the hydrogen bonding lifetime of the same reactant in pure water (τ_{HB,H_2O}), as shown in eqn (5).

$$\tau = \frac{\tau_{HB,org}}{\tau_{HB,H_2O}} \tag{5}$$

A monotonic increase in hydrogen bonding lifetime ratio with respect to cosolvent fraction was observed for reactants across all cosolvent environments (Table S5, ESI†), confirming that reactant–water hydrogen bonding strength generally increases as the amount of available water in the mixture decreases.

Multidescriptor correlation between experimental and simulation results

Since both Γ and τ measure contributions to experimentally determined reaction rates, we explored the use of these descriptors in combination to improve the correlations in Table 2, as shown in eqn (6).

$$\sigma_{pred} = A + Bb + Cc + \ldots \tag{6}$$

A, B, and C are coefficients that quantify the relationship between the simulated molecular descriptors b and c and experimental data (σ). Coefficients for the multidescriptor correlation are obtained by linear regression between the simulated descriptors and experimental kinetic solvent parameters. The model accuracy is then assessed by using the coefficients to calculate a predicted kinetic solvent parameter, σ_{pred}, and plotting it against the experimentally determined kinetic solvent parameter σ_{exp}.

To determine if hydrogen bonding strength improves the correlation between simulation results and experimental models, we define two models: eqn (7) uses a single descriptor, Γ and eqn (8) uses two descriptors, Γ and τ.

$$\sigma_{pred} = A + B(\Gamma) \tag{7}$$

$$\sigma_{pred} = A + B(\Gamma) + C(\tau) \tag{8}$$

As shown in Table 3, the slope between σ_{pred} and σ_{exp} for most reactant/cosolvent systems is close to unity and the root-mean-square error (RMSE) between the predicted and experimental

This journal is © The Royal Society of Chemistry 2018

Energy Environ. Sci., 2018, **11**, 617–628 | 623

Table 3 Best-fit slope and root-mean-square error (RMSE) between predicted kinetic solvent parameters (σ_{pred}) and experimental kinetic solvent parameters (σ_{exp})

Cosolvent	Reactant	1 descriptor fit Eqn (7)		2 descriptor fit Eqn (8)	
		Slope[a]	RMSE[b]	Slope[a]	RMSE[b]
DIO	ETBE	**0.01**	*0.11*	**0.17**	*0.10*
	TBA	**0.89**	0.09	0.99	0.03
	LGA	**0.70**	*0.11*	0.98	0.03
	PDO	0.97	0.05	N/A[c]	N/A[c]
	FRU	0.92	*0.16*	1.00	0.03
	CEL	0.97	0.08	0.97	0.07
	XYL	0.98	0.09	0.98	0.08
GVL	ETBE	**0.63**	*0.14*	**0.63**	*0.14*
	TBA	0.97	0.03	0.98	0.02
	LGA	**0.82**	*0.14*	0.95	0.07
	PDO	**0.85**	*0.29*	N/A[c]	N/A[c]
	FRU	**0.86**	*0.27*	1.00	0.00
	CEL	**0.64**	*0.35*	0.97	*0.11*
	XYL	**0.73**	*0.39*	1.00	0.00

[a] Slopes in bold have a value less than 0.90. [b] RMSEs in italics have values greater than the experimental error of ≈ 0.10. [c] N/A: omitted values since only three experimental values are available (see Table S2, ESI†), resulting in an exact solution to eqn (8). THF cosolvent mixture results are omitted for the same reason.

values is lowered when using the two-descriptor model in eqn (8) compared to only fitting a single descriptor in eqn (7). Therefore, the addition of the hydrogen bonding lifetime ratio, which characterizes the binding strength between the reactant and water within the local domain, improves the overall correlation between σ_{pred} and σ_{exp}.

The results in Table 3 suggest that the two-descriptor correlation model can accurately reproduce reaction rates as a function of cosolvent concentration for a single reactant in a single solvent mixture. To probe if Γ and τ can be used to predict reaction performance for a series of various reactants, we calculated best-fit parameters for eqn (8) using the combined data for all seven reactants in DIO/water mixtures and used the resulting two-descriptor correlation model to calculate values of σ. Comparing the calculated and experimental values of σ resulted in a slope of 0.71 and RMSE of 0.36 (Fig. S7, ESI†). We further found that incorporating a reactant-specific descriptor, the accessible hydroxyl fraction (δ), led to the improved multidescriptor correlation model described in eqn (9).

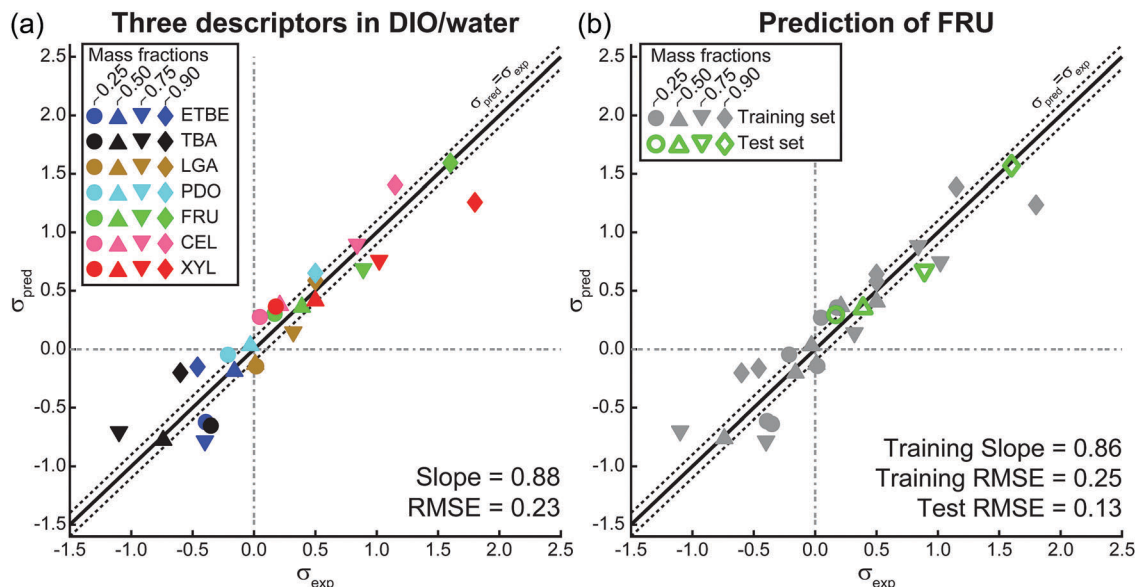$$\sigma_{pred} = A + B(\Gamma) + C(\tau) + D(\delta) \qquad (9)$$

The three molecular descriptors in eqn (9) are statistically uncorrelated (Fig. S9, ESI†) with one another, and moreover are physically motivated vis-à-vis the generalized reaction mechanism proposed above. The preferential exclusion coefficient captures the extent to which the reactant accumulates excess water in its immediate vicinity, which creates a thermodynamic driving force for the transfer of a proton from the bulk phase to the reactant, thereby initiating the acid-catalyzed reaction mechanism. The hydrogen bonding lifetime ratio captures information regarding the binding strength of water to the reactant, which

we interpret as a measure of the ability of the encapsulated water cluster to stabilize the carbocation-like transition states common to acid-catalyzed reactions, as shown in Fig. 3. The accessible hydroxyl fraction captures information relating to the percentage of the surface area of the reactant molecule occupied by hydroxyl groups, which we interpret as a qualitative descriptor of the reactant's hydrogen bonding capacity, normalized by its molecular size. To further validate the addition of each parameter, we compared several simpler models (see Table S8, ESI†) using the Akaike Information Criterion (AIC).[64] Comparison of the AIC's associated with each of the competing models indicated that eqn (9) affords a sufficient improvement in the model's ability to predict experimental rate constants so that its greater complexity is statistically justified.

Fig. 7(a) compares kinetic solvent parameters calculated using the multidescriptor correlation model in eqn (9) to experimentally determined values for DIO/water mixtures, affording a slope of 0.88 and RMSE of 0.24 with few false positive/negatives. As the correlation is maintained across several different reactants, we suggest that acid-catalyzed reactions behave similarly in each of these cosolvent mixtures, where knowledge of simulation-derived Γ, τ, and δ can predict experimental σ_{exp} with accuracy. To demonstrate the predictive power of the model in eqn (9), we selected one reactant (the test set), fit the parameters in eqn (9) using the remaining six reactants (the training set), and then calculated kinetic solvent parameters for the test set reactant using the parameters derived from the training set. This procedure assesses the ability of the multidescriptor correlation model to predict kinetic solvent parameters for reactants that are not used to determine model parameters. Fig. 7(b) shows the results of this procedure for DIO/water mixtures using FRU as the test set reactant. We find that the best-fit slope and RMSE between experimental and predicted kinetic solvent parameters are nearly the same when using the six-reactant training set compared to using all seven reactants. The test set RMSE is 0.13, which is lower than the average RMSE of the training set. These results confirm the robustness and predictability of the multidescriptor correlation model. Table S9 (ESI†) further extends these results to all reactants in all solvent systems.

Table 4 summarizes the model coefficients and error for the different cosolvent environments. Simulated parameters were re-scaled to values between 0 and 1 so that the values of the coefficients in eqn (9) can be compared (see ESI† for re-scaling details). DIO/water mixtures were found to have the best correlation when fitting. GVL/water and THF/water cosolvent environments have the highest error. The coefficients differ across cosolvents, suggesting that the simulated parameters contribute to a different degree when changing cosolvent environments. In general, Γ has the largest weight in the multidescriptor correlation model for DIO/water and GVL/water systems, indicating that the formation of a water-rich local domain is the dominant contributor to reactivity, in agreement with the correlations in Table 2. Γ has a smaller weight in THF/water system further supported by low correlations between Γ and σ in Table 2. The low correlation implies that there may be other descriptors that could be included to improve the model, such as additional

624 | *Energy Environ. Sci.*, 2018, **11**, 617–628

This journal is © The Royal Society of Chemistry 2018

Fig. 7 (a) Comparison of kinetic solvent parameters calculated using the multidescriptor correlation model (σ_{pred}) to experimentally determined values (σ_{exp}) for all seven reactants in DIO/water mixtures. Each reactant has four data points for 0.25, 0.50, 0.75, and 0.90 mass fractions of DIO, with the exception of PDO (see Table S1, ESI†). The slope of the best-fit line for all the data points and the average root-mean-squared error (RMSE) between the values of σ_{pred} and σ_{exp} are shown at the bottom right. The solid black line indicates a perfect correlation ($\sigma_{pred} = \sigma_{exp}$) and dotted lines are drawn at $\sigma_{exp} = 0$ and $\sigma_{pred} = 0$ to help visualize false positive/negative predicted values. Lines above and below the $\sigma_{pred} = \sigma_{exp}$ line are shifted by ±0.10, denoting the approximate experimental error. (b) Prediction of kinetic solvent parameters using FRU as a test set with all other reactants taken as a training set. The slope of the best-fit line and RMSE between the values of σ_{pred} and σ_{exp} for the training and test sets are shown at the bottom right.

Table 4 Coefficients of the multidescriptor correlation model describing the rates of all seven reactions in each of the three cosolvent mixtures with the best-fit slope and root-mean-square error (RMSE) between predicted (σ_{pred}) and experimentally determined (σ_{exp}) kinetic solvent parameters. The descriptors $\tilde{\Gamma}$, $\tilde{\tau}$, and $\tilde{\delta}$ are equal to Γ, τ, and δ but re-scaled to values between 0 and 1 so the coefficients are comparable. Errors for the coefficients are shown in Table S7 (ESI)

| | $\sigma_{pred} = A + B(\tilde{\Gamma}) + C(\tilde{\tau}) + D(\tilde{\delta})$ | | | | | |
Cosolvent	A	B	C	D	Slope	RMSE
DIO	−1.497	1.489	1.270	1.025	0.88	0.23
GVL	−1.389	1.542	0.851	1.077	0.70	0.37
THF	−0.825	0.501	1.510	0.302	0.53	0.58

cosolvent-specific descriptors (e.g., dielectric constants, size, dipole moment, etc.).[65]

Conclusions

We have analyzed the effects of three polar aprotic cosolvents, in aqueous mixtures of varying composition, on the acid-catalyzed reactions of seven biomass-derived model compounds. General trends in reactivity, as expressed by changes in apparent rate constants as a function of solvent composition, were correlated to simulation-derived observables obtained from classical molecular dynamics simulations. We find that the presence of organic cosolvents in the solvent mixture leads to the formation of water-enriched local solvent domains near hydrophilic reactants and increases the strength of hydrogen bonding between reactants and local water molecules. These effects are

quantified by molecular descriptors describing: (1) the local density of water near the reactant (Γ); (2) the average hydrogen-bond lifetime between the reactant and neighboring water molecules (τ); and (3) the accessible surface area occupied by hydroxyl groups on the reactant (δ). By combining these three observables in a multiple linear regression scheme, we have developed a multidescriptor correlation model that predicts rate constants as a function of solvent composition. This development represents an important step toward the computational design of new liquid-phase biomass conversion processes, informed by a first principles approach.

Methods

Reaction kinetics studies

Reactions were carried out in closed, 10 mL thick-walled glass reactors. In a typical experiment, an appropriate amount of reactant (e.g., XYL), acid catalyst, water and organic cosolvent (e.g., DIO) were charged to the reactors, which were then sealed and placed in an oil bath at the appropriate temperature. Reactors were removed at times corresponding to the desired reaction time, and quenched in an ice bath at 273 K. Reaction products were analyzed using high-performance liquid chromatographs equipped with differential refractometers and photo diode array detectors, or gas chromatographs equipped with flame ionization detectors. All products were quantified using calibration curves with external standards. Conditions for each reaction (temperature, fractional conversion, etc.) were chosen so that each reaction was selective (>90%) to a single product.

This procedure allowed for reliable measurements of rate constants based on the rate reactant consumption in a MATLAB-based optimization routine as detailed in the ESI.† Trifluoromethane sulfonic (triflic) acid (pK_{a,H_2O} = −14.7, $pK_{a,DMSO}$ = −14.3, $pK_{a,MeCN}$ = 0.7) was used as catalyst in all experiments, which has been shown to behave as a strong acid even in pure polar aprotic solvents.[40,66,67] We thus assume complete dissociation of the acidic proton in all mixed solvent environments, allowing for normalization of the apparent rate constants on a per-proton basis.

Molecular dynamics simulations

Classical molecular dynamics simulations were performed using Gromacs version 2016.[68] Reactants and cosolvents were parameterized using the CGenFF/CHARMM36 force fields[69–71] while water was modeled using the SPC/E[72] model. Solvent mixtures were initially equilibrated in an *NPT* ensemble using a Berendsen barostat at 1 bar and velocity-rescale thermostat at 300 K. The initial simulation box size containing the cosolvent and water was set to (6 nm)^3 in all simulations. A single reactant molecule was then added to the system, equilibrated with the same barostat and thermostat at the temperature of the reaction for 500 ps, followed by a *NPT* MD simulation with Parrinello–Rahman and Nose–Hoover as the barostat and thermostat, respectively, for 200 ns. A schematic of the system setup is shown in Fig. S4 (ESI†). The accessible surface area, radial distribution functions, and preferential exclusion coefficients were calculated using the final 190 ns of simulation data. For preferential exclusion coefficients, the simulation trajectory was partitioned into two separate trajectories of 95 ns to obtain the error in the calculations as the standard deviation of the calculated values. Simulation analysis was performed using the MDTraj tool box[73] and analysis tools developed in-house. A total of 91 MD simulations (∼18.2 μs simulation time) was performed. Details of the simulation parameters are available in the ESI,† as well as methods for computing simulation observables.

Preferential exclusion coefficient

Preferential exclusion coefficients were calculated using eqn (4) for each configuration in the final 190 ns of simulation data. n_W^L and n_C^L are calculated by counting the total number of molecules within a cutoff distance of the reactant, where the cutoff between local and bulk domains is determined by the RDF for each system composition (Fig. 4(b)). Distances are computed between the center of mass of each molecule. Cutoff radii are tabulated in Table S5 (ESI†).

Hydrogen bonding lifetime

Hydrogen bonding lifetimes were calculated based on the Luzar and Chandler approach.[74] The kinetics of hydrogen bond formation and breaking is described in eqn (10).

$$\frac{dc(t)}{dt} = -kc(t) + k'n(t) \tag{10}$$

where k and k' are rate constants for breaking and making hydrogen bonds, respectively. $c(t)$ is the hydrogen bond correlation function that outputs the probability a hydrogen bond is intact at time t given that it was intact at $t = 0$. $n(t)$ is the probability that a hydrogen bond is broken but the hydrogen bonding groups are still within hydrogen bonding distance. The average bonding lifetime (τ_{HB}) could be found by the reciprocal of the forward rate constant ($1/k$) and calculated using Gromacs 5.0.1.[75–79] Additional details of hydrogen bonding lifetime calculations are available in the ESI.†

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 J. N. Chheda, G. W. Huber and J. A. Dumesic, *Angew. Chem., Int. Ed.*, 2007, **46**, 7164–7183.

2 O. Bobleter, in *Polysaccharides*, ed. S. Dumitriu, Marcel Dekker, New York, 2nd edn, 2005, pp. 893–937.

3 S. R. D. C. E. Wyman, M. E. Himmel, J. W. Brady, C. E. Skopec and L. Viikari, in *Polysaccharides*, ed. S. Dumitriu, Marcel Dekker, New York, 2nd edn, 2005, pp. 995–1033.

4 A. Corma, S. Iborra and A. Velty, *Chem. Rev.*, 2007, **107**, 2411–2502.

626 | *Energy Environ. Sci.*, 2018, **11**, 617–628

This journal is © The Royal Society of Chemistry 2018

5 G. W. Huber, S. Iborra and A. Corma, *Chem. Rev.*, 2006, **106**, 4044–4098.

6 Y. Roman-Leshkov, C. J. Barrett, Z. Y. Liu and J. A. Dumesic, *Nature*, 2007, **447**, 982–985.

7 M. Stöcker, *Angew. Chem., Int. Ed.*, 2008, **47**, 9200–9211.

8 L. Tock, M. Gassner and F. Maréchal, *Biomass Bioenergy*, 2010, **34**, 1838–1854.

9 D. A. Simonetti and J. A. Dumesic, *Catal. Rev.*, 2009, **51**, 441–484.

10 T. Y. Nguyen, C. M. Cai, R. Kumar and C. E. Wyman, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 11673–11678.

11 M. A. Mellmer, D. Martin Alonso, J. S. Luterbacher, J. M. R. Gallo and J. A. Dumesic, *Green Chem.*, 2014, **16**, 4659–4662.

12 M. A. Mellmer, C. Sener, J. M. R. Gallo, J. S. Luterbacher, D. M. Alonso and J. A. Dumesic, *Angew. Chem., Int. Ed.*, 2014, **53**, 11872–11875.

13 Y. Román-Leshkov, J. N. Chheda and J. A. Dumesic, *Science*, 2006, **312**, 1933–1937.

14 Z. Wei, Y. Li, D. Thushara, Y. Liu and Q. Ren, *J. Taiwan Inst. Chem. Eng.*, 2011, **42**, 363–370.

15 M. A. Mellmer, J. M. R. Gallo, D. Martin Alonso and J. A. Dumesic, *ACS Catal.*, 2015, **5**, 3354–3359.

16 Y. J. Pagan-Torres, T. Wang, J. M. R. Gallo, B. H. Shanks and J. A. Dumesic, *ACS Catal.*, 2012, **2**, 930–934.

17 T. Zhang, R. Kumar and C. E. Wyman, *RSC Adv.*, 2013, **3**, 9809–9819.

18 B. Mostofian, C. M. Cai, M. D. Smith, L. Petridis, X. Cheng, C. E. Wyman and J. C. Smith, *J. Am. Chem. Soc.*, 2016, **138**, 10869–10878.

19 T. Y. Nguyen, C. M. Cai, R. Kumar and C. E. Wyman, *ChemSusChem*, 2015, **8**, 1716–1725.

20 J. He, M. Liu, K. Huang, T. W. Walker, C. T. Maravelias, J. A. Dumesic and G. W. Huber, *Green Chem.*, 2017, **19**, 3642–3653.

21 V. Vasudevan and S. H. Mushrif, *RSC Adv.*, 2015, **5**, 20756–20763.

22 E. Cesarotti, R. Ugo and L. Kaplan, *Coord. Chem. Rev.*, 1982, **43**, 275–298.

23 P. J. Dyson and P. G. Jessop, *Catal. Sci. Technol.*, 2016, **6**, 3302–3316.

24 R. J. Madon and E. Iglesia, *J. Mol. Catal. A: Chem.*, 2000, **163**, 189–204.

25 L. Shuai and J. Luterbacher, *ChemSusChem*, 2016, **9**, 133–155.

26 S. Behtash, J. Lu, E. Walker, O. Mamun and A. Heyden, *J. Catal.*, 2016, **333**, 171–183.

27 W. J. Cheong and P. W. Carr, *Anal. Chem.*, 1988, **60**, 820–826.

28 S. Caratzoulas and D. G. Vlachos, *Carbohydr. Res.*, 2011, **346**, 664–672.

29 D. Y. Murzin, *Catal. Sci. Technol.*, 2016, **6**, 5700–5713.

30 J. Zhang, A. Das, R. S. Assary, L. A. Curtiss and E. Weitz, *Appl. Catal., B*, 2016, **181**, 874–887.

31 R. S. Assary, T. Kim, J. J. Low, J. Greeley and L. A. Curtiss, *Phys. Chem. Chem. Phys.*, 2012, **14**, 16603–16611.

32 V. Choudhary, S. H. Mushrif, C. Ho, A. Anderko, V. Nikolakis, N. S. Marinkovic, A. I. Frenkel, S. I. Sandler and D. G. Vlachos, *J. Am. Chem. Soc.*, 2013, **135**, 3997–4006.

33 A. Curioni, M. Sprik, W. Andreoni, H. Schiffer, J. Hutter and M. Parrinello, *J. Am. Chem. Soc.*, 1997, **119**, 7218–7229.

34 E. J. Meijer and M. Sprik, *J. Am. Chem. Soc.*, 1998, **120**, 6345–6355.

35 S. H. Mushrif, J. J. Varghese and C. B. Krishnamurthy, *Phys. Chem. Chem. Phys.*, 2015, **17**, 4961–4969.

36 X. H. Qian and X. F. Wei, *J. Phys. Chem. B*, 2012, **116**, 10898–10904.

37 J. Zhang, A. Das, R. S. Assary, L. A. Curtiss and E. Weitz, *Appl. Catal., B*, 2016, **181**, 874–887.

38 B. Mostofian, C. M. Cai, M. D. Smith, L. Petridis, X. L. Cheng, C. E. Wyman and J. C. Smith, *J. Am. Chem. Soc.*, 2016, **138**, 10869–10878.

39 S. H. Mushrif, S. Caratzoulas and D. G. Vlachos, *Phys. Chem. Chem. Phys.*, 2012, **14**, 2637–2644.

40 M. A. Mellmer, C. Sanpitakseree, B. Demir, P. Bai, K. Ma, M. Neurock and J. A. Dumesic, *Nature Catalysis*, 2018, DOI: 10.1038/s41929-018-0027-3.

41 M. Sasaki, M. Furukawa, K. Minami, T. Adschiri and K. Arai, *Ind. Eng. Chem. Res.*, 2002, **41**, 6642–6649.

42 J. U. Oltmanns, S. Palkovits and R. Palkovits, *Appl. Catal., A*, 2013, **456**, 168–173.

43 K. T. O'Reilly, M. E. Moir, C. D. Taylor, C. A. Smith and M. R. Hyman, *Environ. Sci. Technol.*, 2001, **35**, 3954–3961.

44 T. D. Swift, C. Bagia, V. Choudhary, G. Peklaris, V. Nikolakis and D. G. Vlachos, *ACS Catal.*, 2013, **4**, 259–267.

45 G. R. Akien, L. Qi and I. T. Horváth, *Chem. Commun.*, 2012, **48**, 5850–5852.

46 A. J. Bennet and M. L. Sinnott, *J. Am. Chem. Soc.*, 1986, **108**, 7287–7294.

47 J. A. Dumesic, *J. Catal.*, 1999, **185**, 496–505.

48 J. A. Dumesic, *J. Catal.*, 2001, **204**, 525–529.

49 J. K. Lee, A. D. Bain and P. J. Berti, *J. Am. Chem. Soc.*, 2004, **126**, 3769–3776.

50 C. P. Kelly, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2007, **111**, 408–422.

51 M. D. Tissandier, K. A. Cowen, W. Y. Feng, E. Gundlach, M. H. Cohen, A. D. Earhart, J. V. Coe and T. R. Tuttle, *J. Phys. Chem. A*, 1998, **102**, 7787–7794.

52 C. Kalidas, G. Hefter and Y. Marcus, *Chem. Rev.*, 2000, **100**, 819–852.

53 A. Warshel, *Proc. Natl. Acad. Sci. U. S. A.*, 1978, **75**, 5250–5254.

54 A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. B. Liu and M. H. M. Olsson, *Chem. Rev.*, 2006, **106**, 3210–3235.

55 T. C. Bruice and F. C. Lightstone, *Acc. Chem. Res.*, 1999, **32**, 127–136.

56 Y. S. Liu, A. Vjunov, H. Shi, S. Eckstein, D. M. Camaioni, D. H. Mei, E. Barath and J. A. Lercher, *Nat. Commun.*, 2017, **8**, DOI: 10.1038/ncomms14113.

57 M. Kang and P. E. Smith, *Fluid Phase Equilib.*, 2007, **256**, 14–19.

58 C. P. Schneider and B. L. Trout, *J. Phys. Chem. B*, 2009, **113**, 2050–2058.

59 D. Shukla and B. L. Trout, *J. Phys. Chem. B*, 2011, **115**, 1243–1253.

This journal is © The Royal Society of Chemistry 2018

Energy Environ. Sci., 2018, **11**, 617–628 | **627**

60 I. L. Shulgin and E. Ruckenstein, *J. Phys. Chem. B*, 2007, **111**, 3990–3998.

61 N. Smolin and R. Winter, *J. Phys. Chem. B*, 2008, **112**, 997–1006.

62 J. E. Sutton and D. G. Vlachos, *ACS Catal.*, 2012, **2**, 1624–1634.

63 S. Wang, V. Petzold, V. Tripkovic, J. Kleis, J. G. Howalt, E. Skulason, E. M. Fernandez, B. Hvolbaek, G. Jones, A. Toftelund, H. Falsig, M. Bjorketun, F. Studt, F. Abild-Pedersen, J. Rossmeisl, J. K. Norskov and T. Bligaard, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20760–20765.

64 H. Akaike, *IEEE Trans. Autom. Control*, 1974, **19**, 716–723.

65 M. J. Kamlet, J. L. M. Abboud, M. H. Abraham and R. Taft, *J. Org. Chem.*, 1983, **48**, 2877–2887.

66 A. Trummal, L. Lipping, I. Kaljurand, I. A. Koppel and I. Leito, *J. Phys. Chem. A*, 2016, **120**, 3663–3669.

67 E. Raamat, K. Kaupmees, G. Ovsjannikov, A. Trummal, A. Kütt, J. Saame, I. Koppel, I. Kaljurand, L. Lipping and T. Rodima, *J. Phys. Org. Chem.*, 2013, **26**, 162–170.

68 S. Pall, M. J. Abraham, C. Kutzner, B. Hess and E. Lindahl, *Lect. Notes Comput. Sci. Eng.*, 2015, **8759**, 3–27.

69 R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and A. D. MacKernell Jr, *J. Chem. Theory Comput.*, 2013, **8**, 3257–3273.

70 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. MacKernell Jr, *J. Comput. Chem.*, 2009, **31**, 671–690.

71 W. Yu, X. He, K. Vanommeslaeghe and A. D. MacKernell Jr, *J. Comput. Chem.*, 2012, **33**, 2451–2468.

72 H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Phys. Chem.*, 1987, **91**, 6269–6271.

73 R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernández, C. R. Schwantes, L. P. Wang, T. J. Lane and V. S. Pande, *Biophys. J.*, 2015, **109**, 1528–1532.

74 A. Luzar and D. Chandler, *Nature*, 1996, **379**, 55–57.

75 H. J. C. Berendsen, D. Vanderspoel and R. Vandrunen, *Comput. Phys. Commun.*, 1995, **91**, 43–56.

76 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.

77 E. Lindahl, B. Hess and D. van der Spoel, *J. Mol. Model.*, 2001, **7**, 306–317.

78 D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.

79 D. van der Spoel, P. J. van Maaren, P. Larsson and N. Timneanu, *J. Phys. Chem. B*, 2006, **110**, 4393–4398.

628 | *Energy Environ. Sci.*, 2018, **11**, 617–628

This journal is © The Royal Society of Chemistry 2018