

# Boosting Performance by Water Solvation Shell with Hydrogen Bonds on Protonic Ionic Liquid: Insights into the Acid Catalysis of Glycosidic Bond

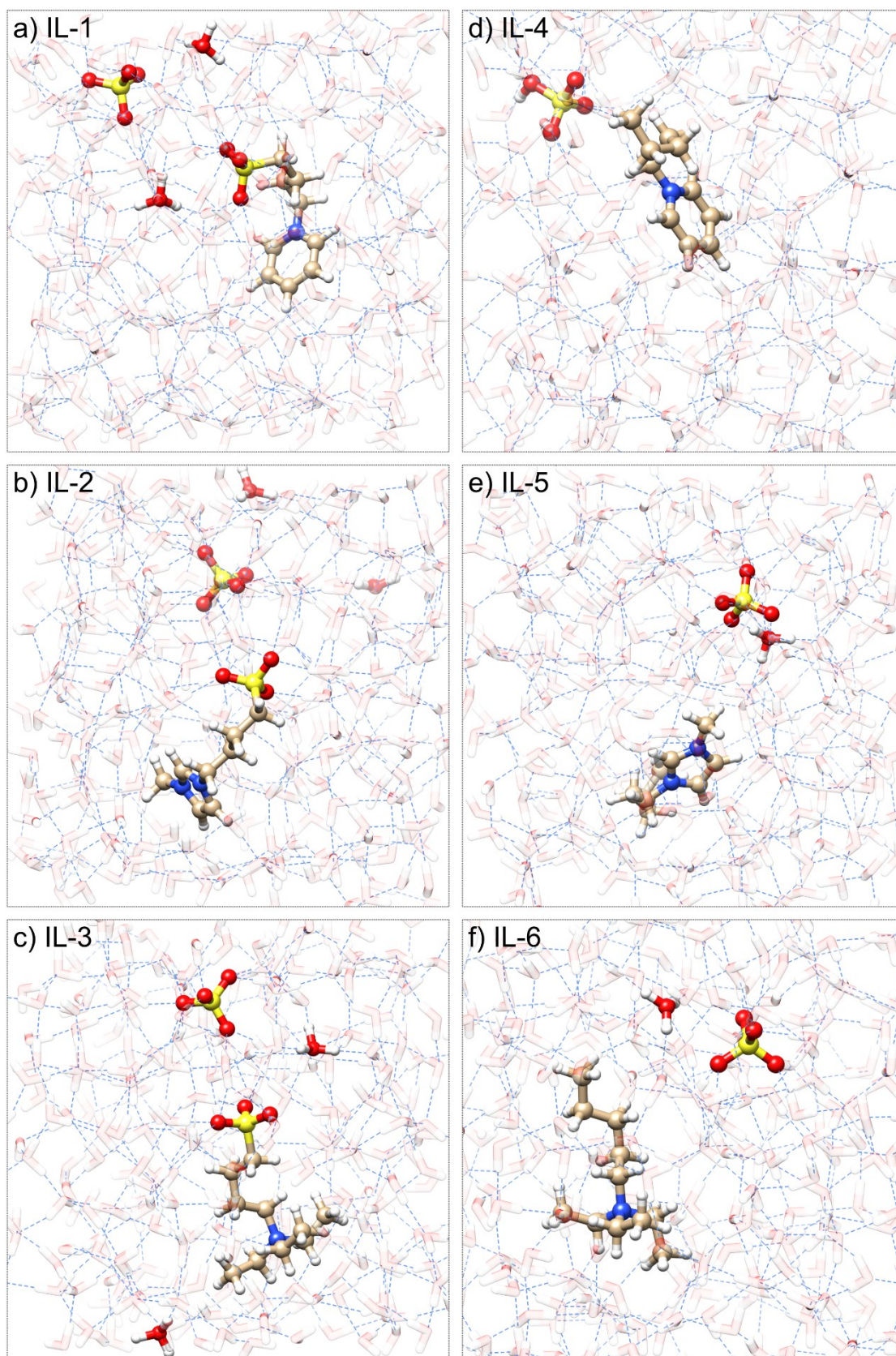
Kaixin Li<sup>a</sup>, Limin Deng<sup>a</sup>, Shun Yi<sup>a</sup>, Yabo Wu<sup>a</sup>, Guangjie Xia<sup>b\*</sup>, Jun Zhao<sup>c\*</sup>, Dong LU<sup>d</sup>, Yonggang Min<sup>a\*</sup>

*<sup>a</sup>School of Materials and Energy, Center of Emerging Material and Technology, Guangdong University of Technology, Guangzhou 510006, China*

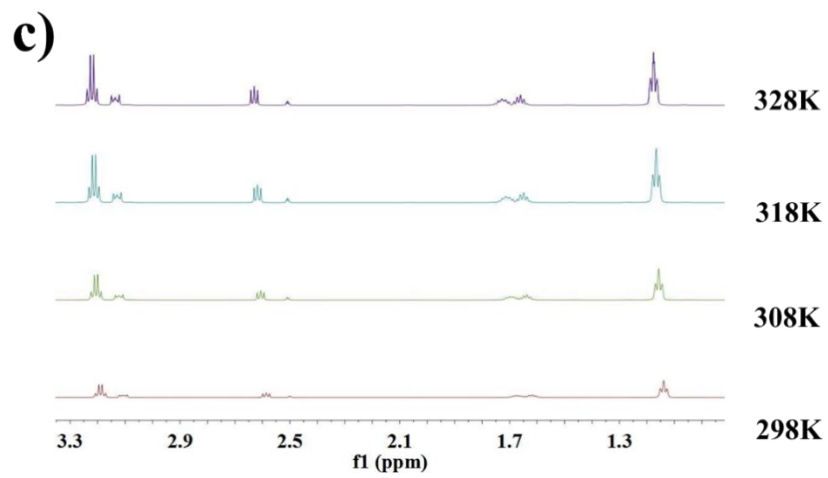
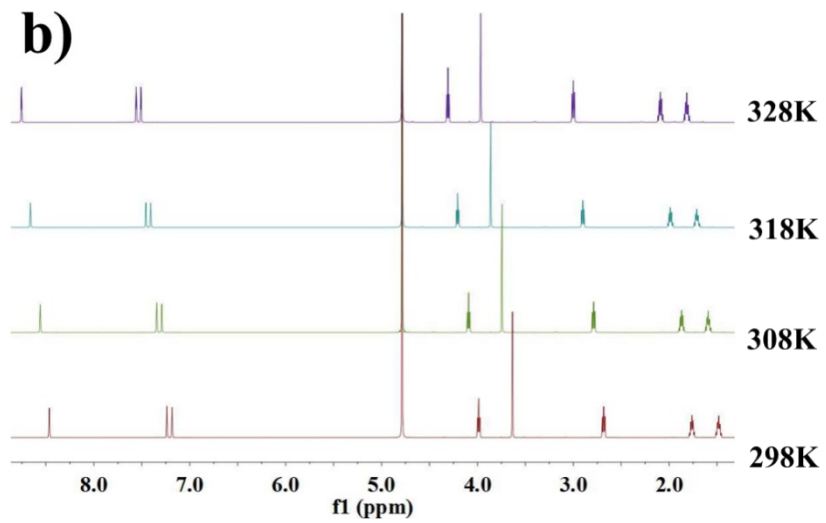
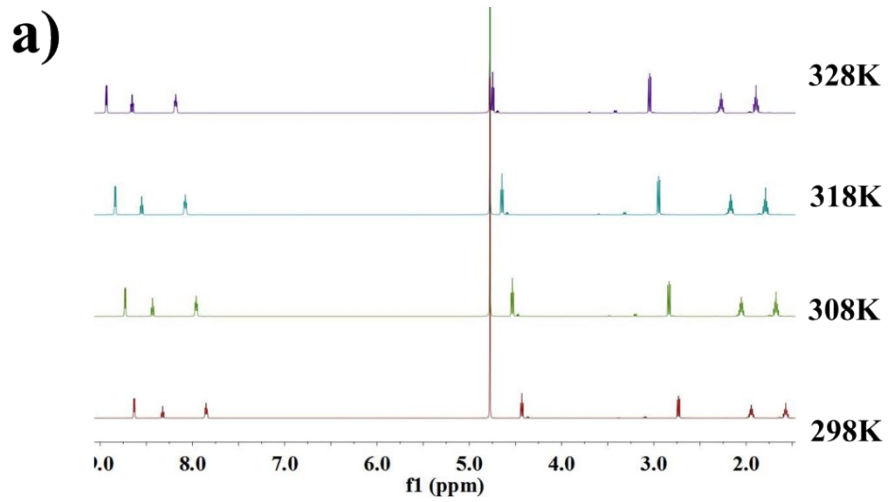
*<sup>b</sup>Department of Chemistry, Southern University of Science and Technology, Shenzhen, Guangdong, 518055, China*

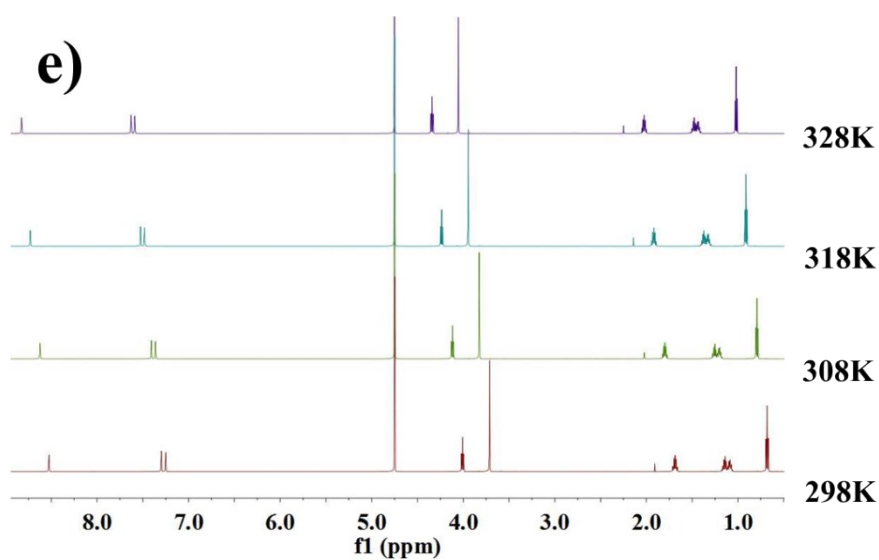
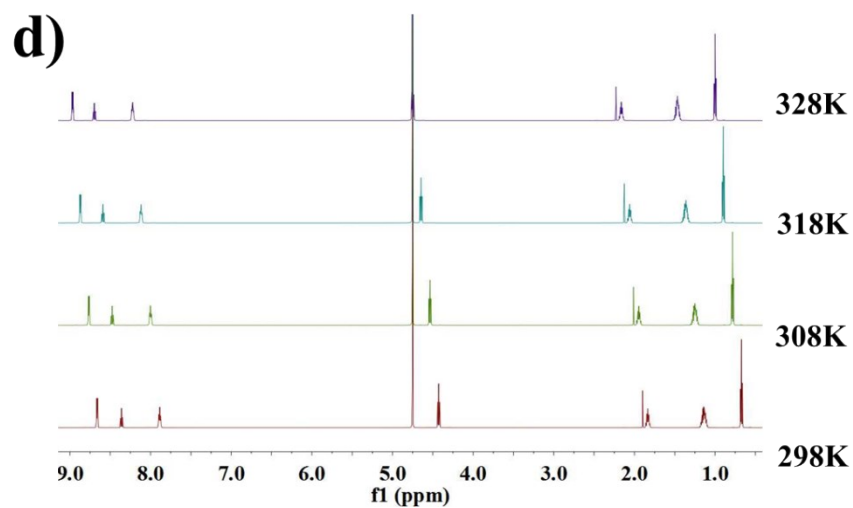
*<sup>c</sup>Institute of Bioresource and Agriculture, Hong Kong Baptist University, Hong Kong SAR*

*<sup>d</sup>Center for Engineering Materials and Reliability, Guangzhou HKUST Fok Ying Tung Research Institute, Guangzhou, Guangdong 511458, China*



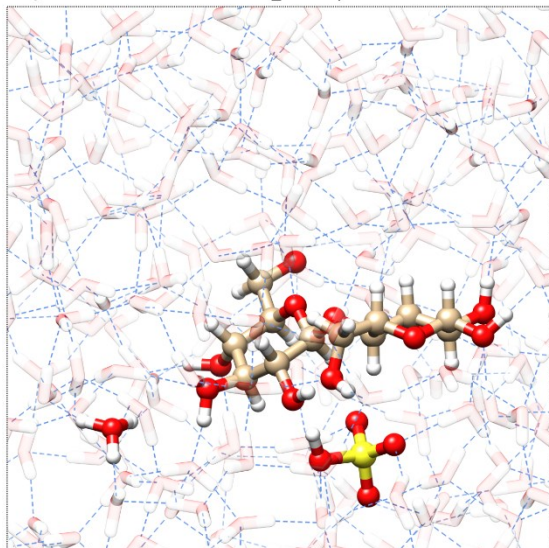
**Fig. S1.** The six IL models implemented with a periodic cubic box of 20 Å filled with explicit water molecules. The color are the same with those in Fig. 4. The solvent waters are presented in transparency. The configurations are stable ones, which are obtained by AIMD of 20 ps and annealing of 3 ps followed by further optimization.



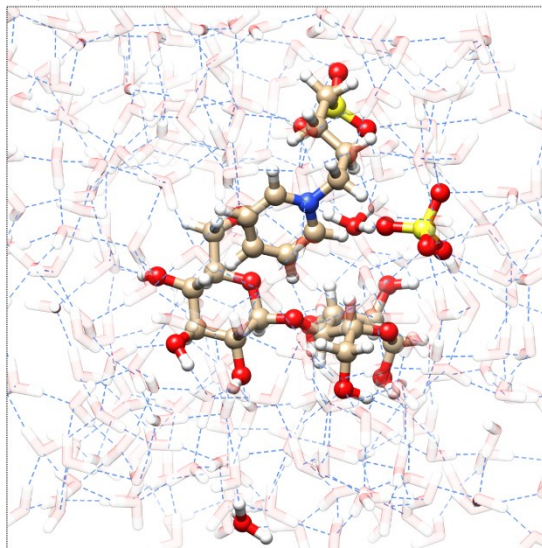


**Fig. S2.** The temperature-dependent changes of C-H chemical shifts for a) IL-1; b) IL-2; c) IL-3; d) IL-4; e) IL-5 in H<sub>2</sub>O at the concentration of 0.5 mol L<sup>-1</sup>.

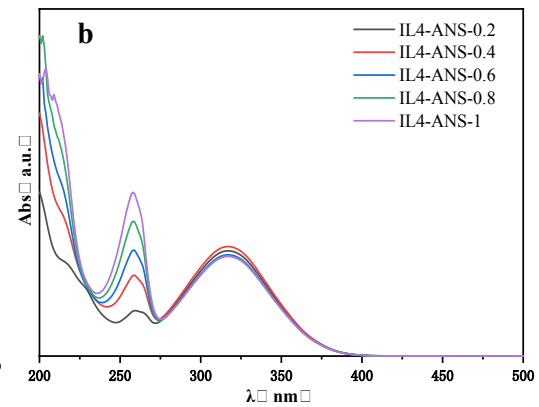
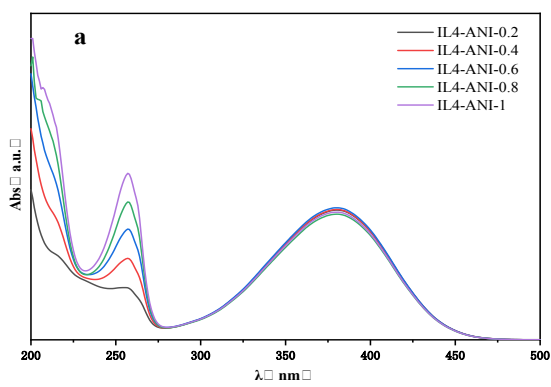
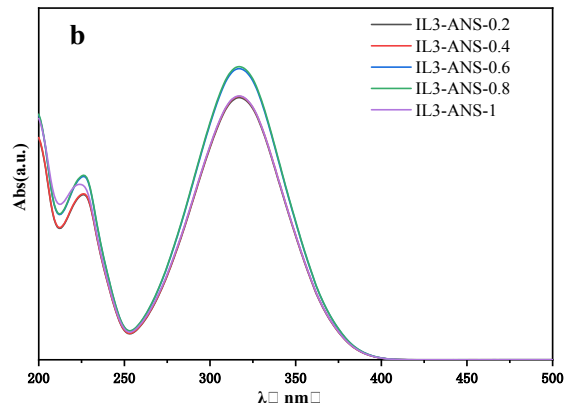
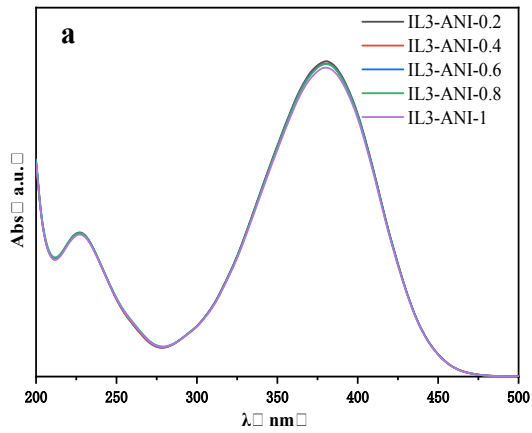
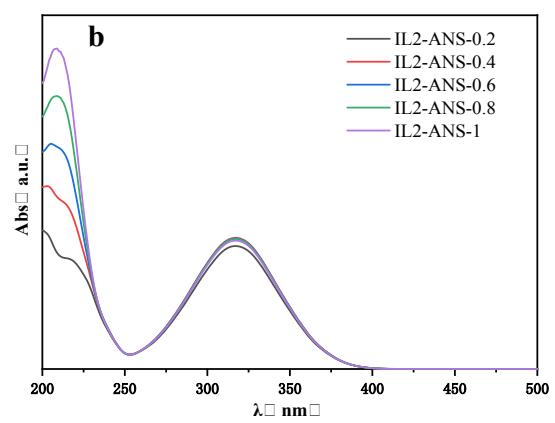
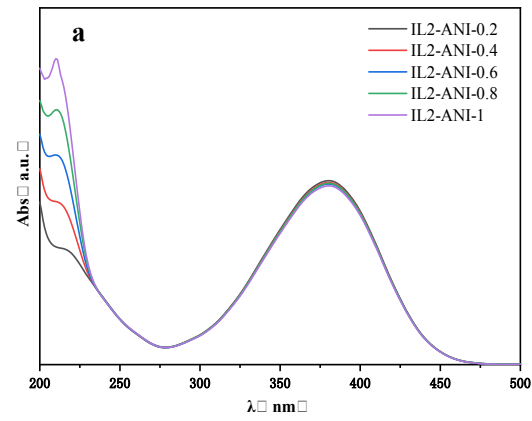
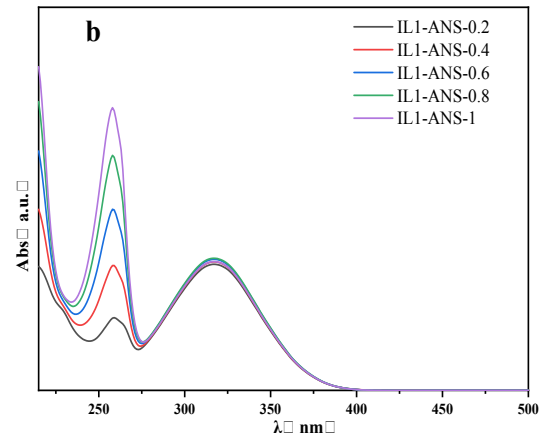
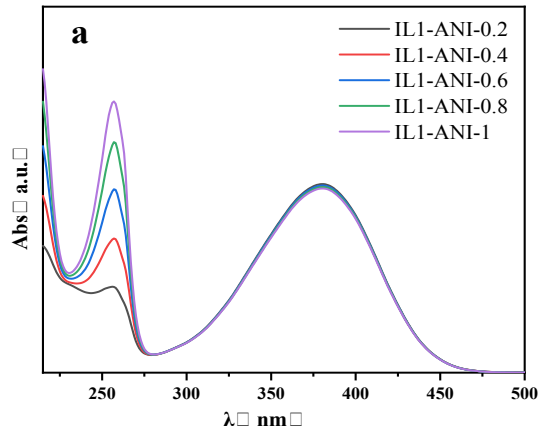
a) cellobiose + H<sub>2</sub>SO<sub>4</sub>

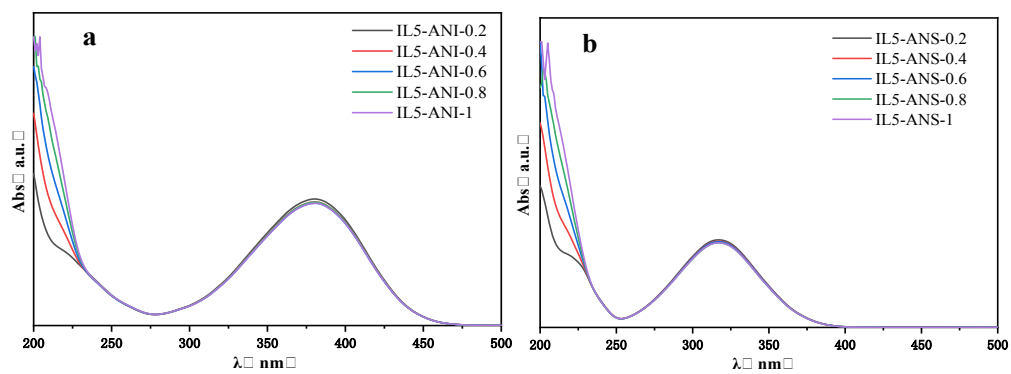


b) cellobiose + IL-1

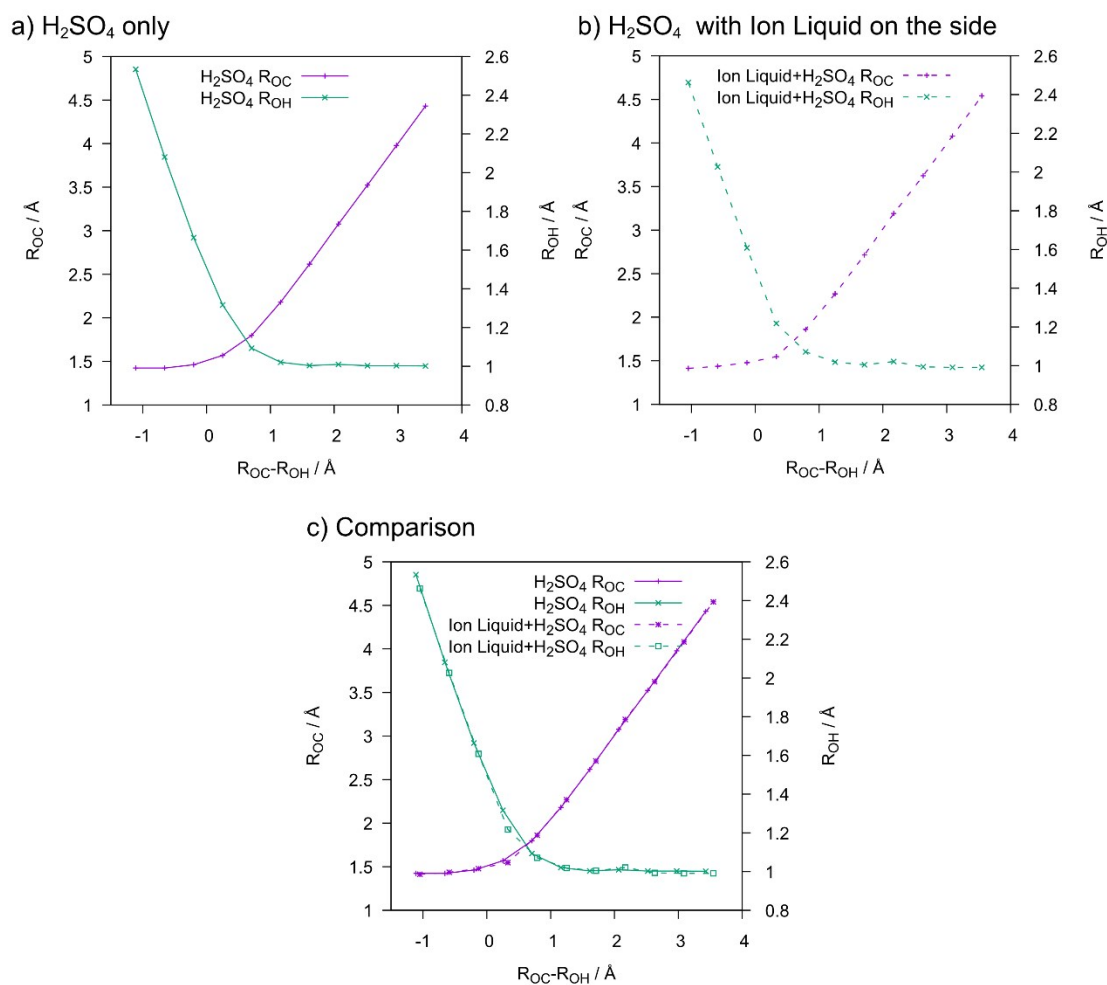


**Fig. S3.** The models involving one molecule of cellobiose and one molecule of catalyst (H<sub>2</sub>SO<sub>4</sub> or IL-1) are simulated within a periodic cubic box of 20 Å filled with explicit water molecules. The colours are the same as those in Fig. 4. The solvent waters are presented in transparency. The configurations are stable ones, which are obtained by AIMD of 20 ps and annealing of 3 ps followed by further optimization.

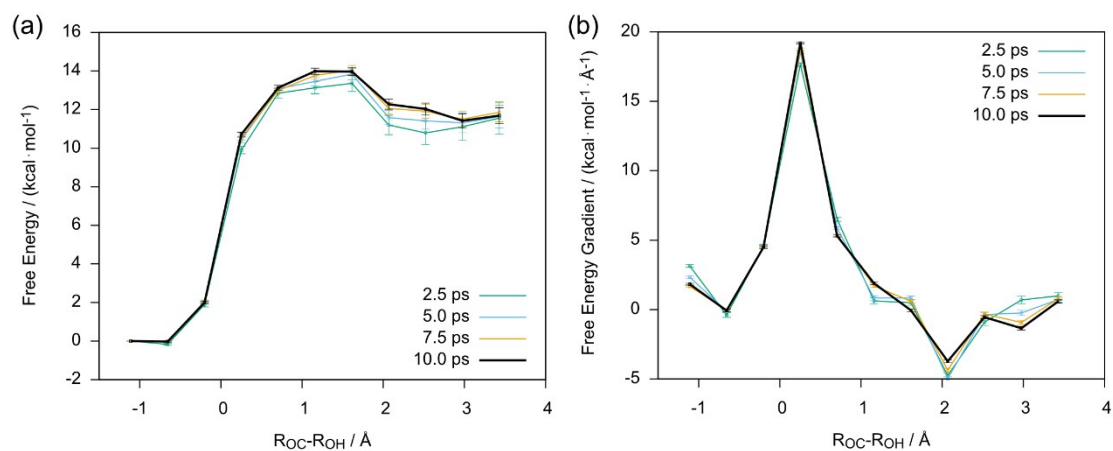




**Fig. S4.** The UV-vis spectrum of IL1-IL 5 in aqueous probe solution.



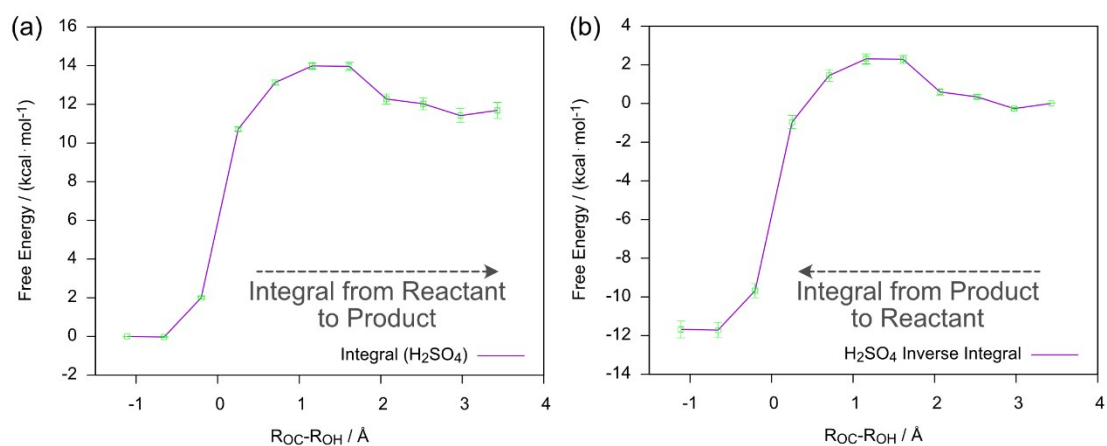
**Fig. S5.** The average O-H distance ( $d_{OH}$ ) and C-O bond distance ( $d_{OC}$ ) during the AIMD calculation with constrain of different value of CV.



**Figure S6.** Convergence Test of the statistic time on (a) free energy and (b) free energy gradient. The statistic times to get the free energy gradient and free energy curve increase from 2.5 ps to 10.0 ps. The trajectories of the first 2 ps before the statistic not used to make

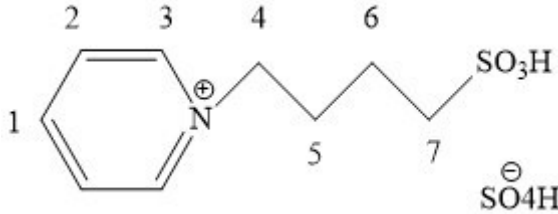


the system to reach the equilibrium. The error bar represents the standard error of the mean (SEM) of each point.

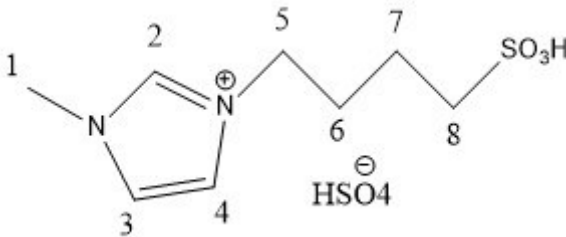


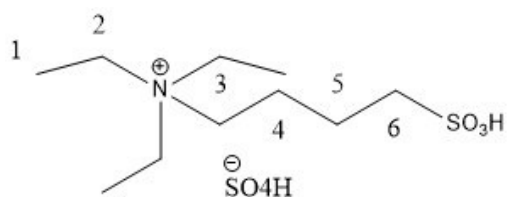
**Figure S7.** The comparison on the integral directions during the calculation of free energy curve. (a) Integrate from reactant to product; (b) Integrate from product to reactant. The error bar represents the standard error of the mean (SEM) of each point.

**Table S1.** The  $^1\text{H}$  NMR parameters for the six ILs

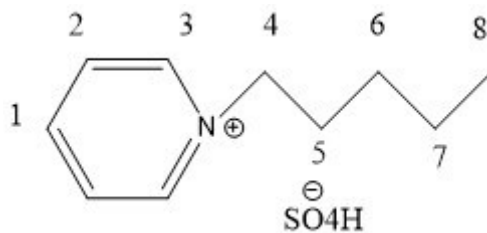
<b>IL-1:</b>									
									
<b>IL-1</b>	T	1	2	3	4	5	6	7	
J(Hz)	298K	7.1	7.9	5.6	7.5	7.7	7.8	7.6	
	308K	7.1	7.9	5.5	7.4	7.7	7.7	7.6	
	318K	7.1	7.9	5.5	7.4	7.6	7.7	7.6	
	328K	7.0	7.9	5.5	7.1	7.5	7.6	7.5	
$\delta$ (ppm)	298K	7.88	8.35	8.65	4.45	1.97	1.60	2.76	
	308K	7.96	8.43	8.73	4.53	2.05	1.68	2.84	
	318K	8.08	8.55	8.84	4.65	2.17	1.79	2.95	
	328K	8.18	8.65	8.93	4.75	2.27	1.89	3.05	
$\Delta\delta/\Delta T$ (ppd/K)		10.2	10.2	9.5	10.2	10.2	9.8	9.8	

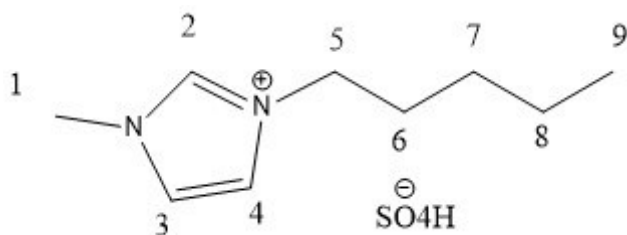
<b>IL-2:</b>									
									
<b>IL-2</b>	T	1	2	3	4	5	6	7	8
J(Hz)	298K	-	-	1.5	1.4	7.1	7.05	7.92	7.50
	308K	-	-	1.4	1.4	7.1	7.12	8.04	7.27
	318K	-	-	1.4	1.3	7.1	7.16	8.59	7.83
	328K	-	-	1.4	1.3	7.1	7.3	7.93	7.65
$\delta$ (ppm)	298K	3.63	8.46	7.24	7.18	3.99	1.76	1.48	2.68
	308K	3.74	8.56	7.34	7.29	4.09	1.87	1.59	2.79
	318K	3.86	8.66	7.46	7.41	4.21	1.99	1.71	2.90
	328K	3.97	8.75	7.56	7.51	4.31	2.09	1.82	3.00
$\Delta\delta/\Delta T$ (ppd/K)		11.4	9.7	10.8	11.1	10.8	11.1	11.4	10.7

**IL-3:**

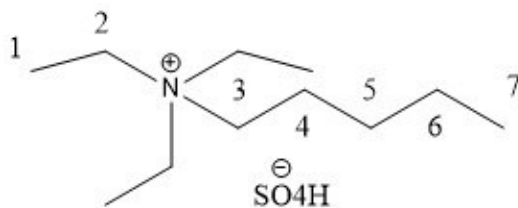
<b>IL-3</b>	T	1	2	3	4	5	6
J(Hz)	298	7.1	7.2	7.9	7.0	6.8	7.5
	308	7.2	7.2	8.0	7.0	6.5	7.45
	318	7.2	7.2	8.0	7.1	6.6	7.4
	328	7.2	7.2	8.1	7.2	6.7	7.35
$\delta$ (ppm)	298	1.14	3.19	3.11	1.68	1.62	2.59
	308	1.16	3.21	3.12	1.70	1.64	2.61
	318	1.17	3.22	3.13	1.71	1.65	2.62
	328	1.18	3.23	3.14	1.73	1.66	2.63
$\Delta\delta/\Delta T$ (ppd/K)		1.3	1.3	1	1.6	1.3	1.3

**IL-4:**

<b>IL-4</b>	T	1	2	3	4	5	6	7	8
J(Hz)	298	7.9	7.0	5.5	7.4	2.3	2.7	7.5	7.1
	308	7.9	7.0	5.5	7.4	2.3	1.9	7.5	7.1
	318	7.9	6.9	5.6	7.4	2.4	3.2	7.5	7.1
	328	7.8	6.8	5.8	7.4	2.5	3.2	7.5	7.1
$\delta$ (ppm)	298	8.36	7.89	8.66	4.43	1.16	1.12	1.83	0.67
	308	8.48	8.00	8.77	4.54	1.27	1.23	1.95	0.78
	318	8.58	8.12	8.88	4.65	1.39	1.34	2.06	0.90
	328	8.70	8.22	8.97	4.75	1.49	1.45	2.16	1.00
$\Delta\delta/\Delta T$ (ppd/K)		11.2	11.1	10.4	10.7	11.1	11	11	11.1

**IL-5:**

<b>IL-5</b>	T	1	2	3	4	5	6	7	8	9
J(Hz)	298	-	-	1.8	1.7	7.1	5.8	5.3	7.2	7.2
	308	-	-	1.7	1.7	7.2	8.1	8.2	7.2	7.2
	318	-	-	1.8	1.7	7.2	8.2	7.3	7.3	7.2
	328	-	-	1.7	1.6	7.2	6.4	7.1	7.3	7.2
$\delta$ (ppm)	298	3.71	8.52	7.30	7.25	4.01	1.15	1.10	1.69	0.68
	308	3.83	8.62	7.41	7.36	4.12	1.25	1.20	1.80	0.79
	318	3.95	8.73	7.52	7.48	4.24	1.37	1.32	1.92	0.91
	328	4.05	8.82	7.63	7.59	4.34	1.49	1.43	2.02	1.02
$\Delta\delta/\Delta T$ (ppd/K)		11.4	10.1	11	11.4	11.1	11.4	11.1	11.1	11.4

**IL-6:**

<b>IL-6</b>	T	1	2	3	4	5	6	7
J(Hz)	298	7.3	7.3	8.2	8.4	7.0	7.8	7.2
	308	7.3	7.3	9.0	8.4	7.1	7.8	7.2
	318	7.3	7.3	9.1	8.4	6.8	7.7	7.2
	328	7.3	7.3	8.2	8.3	7.0	7.7	7.2
$\delta$ (ppm)	298	1.21	3.22	3.07	1.36	1.30	1.61	0.91
	308	1.22	3.22	3.07	1.37	1.31	1.62	0.91
	318	1.23	3.23	3.08	1.38	1.32	1.63	0.92
	328	1.27	3.27	3.12	1.41	1.36	1.67	0.95
$\Delta\delta/\Delta T$ (ppd/K)		1.9	1.6	1.6	1.6	1.9	1.9	1.3

**Table S2.** The values of CV, Free Energy Gradient and Free Energy, together with the standard derivation and standard error of the mean.

<b>Reaction with IL-1</b>					
CV: $d_{OC}-d_{OH}$ (Å)	Free Energy Gradient (dA/dCV) / (kcal mol <sup>-1</sup> Å <sup>-1</sup> )	S.D. <sup>†</sup> of (dA/dCV) / (kcal mol <sup>-1</sup> Å <sup>-1</sup> )	SEM <sup>‡</sup> of (dA/dCV) / (kcal mol <sup>-1</sup> Å <sup>-1</sup> )	Free Energy (A) / (kcal mol <sup>-1</sup> )	SEM of Free Energy /(kcal mol <sup>-1</sup> )
-1.05	0.7281	7.9854	0.0799	0.0000	0.0000
-0.59	0.2994	9.4805	0.0948	0.1377	0.0367
-0.13	7.0698	8.1512	0.0815	3.3898	0.0803
0.33	11.8500	4.3167	0.0432	8.8408	0.1178
0.79	2.9083	7.9422	0.0794	10.1786	0.1377
1.25	0.6242	10.3528	0.1035	10.4658	0.1742
1.71	-0.5243	11.3452	0.1135	10.2246	0.2219
2.17	-2.2049	10.3412	0.1034	9.2104	0.2740
2.63	0.3369	12.1284	0.1213	9.3653	0.3216
3.09	0.4851	13.1157	0.1312	9.5885	0.3774
3.55	-0.3228	11.8799	0.1188	9.4400	0.4377
<b>Reaction with H<sub>2</sub>SO<sub>4</sub></b>					
CV: $d_{OC}-d_{OH}$ (Å)	Free Energy Gradient (dA/dCV) / (kcal mol <sup>-1</sup> Å <sup>-1</sup> )	S.D. <sup>†</sup> of (dA/dCV) / (kcal mol <sup>-1</sup> Å <sup>-1</sup> )	SEM <sup>‡</sup> of (dA/dCV) / (kcal mol <sup>-1</sup> Å <sup>-1</sup> )	Free Energy (A) / (kcal mol <sup>-1</sup> )	SEM of Free Energy /(kcal mol <sup>-1</sup> )
-1.11	1.8253	6.3805	0.0638	0.0000	0.0000
-0.66	-0.0703	8.7512	0.0875	-0.0319	0.0290
-0.20	4.4861	8.0554	0.0806	2.0048	0.0687
0.25	19.2043	5.5844	0.0558	10.7235	0.1053
0.71	5.3013	7.2355	0.0724	13.1303	0.1306
1.16	1.8924	10.3071	0.1031	13.9894	0.1635
1.61	-0.0490	9.9994	0.1000	13.9672	0.2103
2.07	-3.7164	11.4453	0.1145	12.2799	0.2557
2.52	-0.5527	10.8724	0.1087	12.0290	0.3076
2.98	-1.3432	12.7005	0.1270	11.4192	0.3570
3.43	0.5901	11.7091	0.1171	11.6871	0.4146

<sup>†</sup> Standard derivation (S.D.) of the free energy gradient, which is caused by the endogenic fluctuation from the thermal movement of all the molecules, especially solvent waters.

<sup>‡</sup> Standard error of the mean (SEM) of the free energy gradient, which is calculated from the S.D. divided by the square root of the number of samples. The longer the simulation time, the smaller SEM, indicating the more accurate free energy gradient.

**Table S3.** Convergence Test of the planewave cutoff ( $E_{\text{cutoff}}$ , unit in Ry) and relative cutoff ( $E_{\text{relcut}}$ , unit in Ry) in CP2k Package on the reaction energy difference of cellobiose hydrolysis (CB. hydro.) catalyzed by pure sulfuric acid ( $\Delta E$ ). The  $E_{\text{cutoff}}$  defines the planewave cutoff for the finest level of the multi-grid; the  $E_{\text{relcut}}$  controls which product Gaussians are mapped onto which level of the multi-grid. The  $\Delta E$  is directly obtained by subtracting the relative energy of the two typical stable configurations before and after the reaction. These stable configurations are obtained by the annealing of 3 ps from the AIMD trajectory followed by geometry optimization. The typical 3D configurations of reactant and product are shown in Fig. 4e.

$E_{\text{cutoff}}$ Test ( $E_{\text{relcut}} = 60$ Ry)					
CB. hydro.	$E_{\text{cutoff}}= 300$ Ry	$E_{\text{cutoff}}= 350$ Ry	<b><math>E_{\text{cutoff}}=400</math> Ry</b>	$E_{\text{cutoff}}= 500$ Ry	$E_{\text{cutoff}}= 600$ Ry
Reactant	-4775.573132	-4776.022207	-4776.797743	-4777.724001	-4778.511249
Product	-4775.554023	-4776.002702	-4776.777418	-4777.703727	-4778.490893
$\Delta E$	11.99	12.24	<b>12.75</b>	12.72	12.77
$E_{\text{relcut}}$ Test ( $E_{\text{cutoff}} = 400$ Ry)					
CB. hydro.	$E_{\text{relcut}}= 40$ Ry	$E_{\text{relcut}}= 50$ Ry	<b><math>E_{\text{relcut}}= 60</math> Ry</b>	$E_{\text{relcut}}= 70$ Ry	$E_{\text{relcut}}= 80$ Ry
Reactant	-4776.798215	-4776.797741	-4776.797743	-4776.797740	-4776.797740
Product	-4776.778318	-4776.777581	-4776.777418	-4776.777580	-4776.777580
$\Delta E$	12.49	12.65	<b>12.75</b>	12.65	12.65

**Table S4.** Convergence Test of the statistic time on free energy and free energy gradient.  $T_{\text{statistic}}$  is the statistic time to get the free energy gradient and free energy curve. The trajectories of the first 2 ps before the statistic not used to make the system to reach the equilibrium. The number follows the “ $\pm$ ” sign represents the standard error of the mean (SEM) of the free energy gradient or free energy, which decreases a lot with the increase of simulation time.

Free Energy Curve (kcal mol <sup>-1</sup> )				
CV: $d_{\text{OC}-d_{\text{OH}}}$ (Å)	$T_{\text{statistic}} = 2.5$ ps	$T_{\text{statistic}} = 5.0$ ps	$T_{\text{statistic}} = 7.5$ ps	$T_{\text{statistic}} = 10.0$ ps
-1.110	0.00±0.00	0.00±0.00	0.00±0.00	0.00±0.00
-0.656	-0.17±0.05	-0.05±0.04	-0.01±0.03	-0.03±0.03

-0.202	1.89±0.12	2.00±0.09	2.05±0.08	2.00±0.07
0.252	9.90±0.19	10.42±0.14	10.54±0.12	10.72±0.11
0.706	12.85±0.25	13.08±0.18	12.99±0.15	13.13±0.13
1.160	13.13±0.31	13.46±0.23	13.77±0.19	13.99±0.16
1.614	13.36±0.42	13.84±0.30	14.04±0.25	13.97±0.21
2.068	11.20±0.50	11.58±0.37	12.06±0.30	12.28±0.26
2.522	10.79±0.60	11.42±0.44	11.92±0.36	12.03±0.31
2.976	11.11±0.71	11.31±0.51	11.50±0.42	11.42±0.36
3.430	11.56±0.84	11.64±0.59	11.86±0.49	11.69±0.41

---

Free Energy Gradient (kcal mol<sup>-1</sup> Å<sup>-1</sup>)

---

CV: d <sub>OC</sub> -d <sub>OH</sub> (Å)	T <sub>statistic</sub> = 2.5 ps	T <sub>statistic</sub> = 5.0 ps	T <sub>statistic</sub> = 7.5 ps	T <sub>statistic</sub> = 10.0 ps
-1.110	3.14±0.10	2.33±0.08	1.67±0.07	1.83±0.06
-0.656	-0.39±0.17	-0.11±0.13	-0.03±0.10	-0.07±0.09
-0.202	4.55±0.15	4.51±0.10	4.54±0.09	4.49±0.08
0.252	17.64±0.13	18.54±0.08	18.72±0.07	19.20±0.06
0.706	6.50±0.15	5.86±0.11	5.39±0.09	5.30±0.07
1.160	0.62±0.23	0.84±0.16	1.71±0.12	1.89±0.10
1.614	0.50±0.19	0.84±0.15	0.60±0.12	-0.05±0.10
2.068	-4.75±0.22	-4.98±0.16	-4.36±0.13	-3.72±0.11
2.522	-0.90±0.24	-0.36±0.16	-0.31±0.13	-0.55±0.11
2.976	0.69±0.28	-0.24±0.19	-0.93±0.15	-1.34±0.13
3.430	0.99±0.24	0.72±0.18	0.79±0.14	0.59±0.12

---

**Table S5.** Test calculation on the commonly used pure GGA functional. D3 represents the Grimme's D3 correction on dispersion interaction. The reaction energy ( $\Delta E$ ) of cellobiose hydrolysis (CB. hydro.) catalyzed by pure sulfuric acid is calculated, which is obtained by directly subtracting the relative energy of the two stable configurations before and after the reaction. These stable configurations are obtained by the annealing of 3 ps from the AIMD

trajectory followed by geometry optimization. The typical 3D configurations of reactant and product are shown in Fig. 4e.

CB. hydro.	PBE	PBE+D3	revPBE+D3	BLYP+D3
Reactant	-4775.713253	-4776.797743	-4781.522921	-4771.964150
Product	-4775.698875	-4776.777418	-4781.501870	-4771.932198
$\Delta E$	9.02	12.75	13.21	20.05

**Table S6.** The comparison on the integral directions during the calculation of free energy curve. Integrate from reactant to product vs. integrate from product to reactant.

CV: $d_{OC-d_{OH}}$ (Å)	Integrate from Reactant to Product		Integrate from Product to Reactant	
	Relative Free Energy (kcal mol <sup>-1</sup> )	SEM (kcal mol <sup>-1</sup> )	Relative Free Energy (kcal mol <sup>-1</sup> )	SEM (kcal mol <sup>-1</sup> )
-1.11	0.0000	0.0000	-11.6871	0.4388
-0.66	-0.0319	0.0290	-11.7190	0.3991
-0.20	2.0048	0.0687	-9.6823	0.3625
0.25	10.7235	0.1053	-0.9636	0.3372
0.71	13.1303	0.1306	1.4432	0.3043
1.16	13.9894	0.1635	2.3023	0.2575
1.61	13.9672	0.2103	2.2801	0.2121
2.07	12.2799	0.2557	0.5929	0.1602
2.52	12.0290	0.3076	0.3419	0.1108
2.98	11.4192	0.3570	-0.2679	0.0532
3.43	11.6871	0.4146	0.0000	0.0000
<b>Barrier</b>	<b>13.9894</b>	<b>0.1635</b>	<b>13.9894</b>	<b>0.6963</b>



## Equations derived from Kamlet-Taft method for calculating the parameters of $\pi^*$ and $\beta$

Reference:

[1] Aryafard, M., et al., Experimental and Theoretical Studies of Preferential Solvation of 4-Nitroaniline and 4-Nitroanisole in an Amino Acid Ionic Liquid with Molecular Solvents. *Journal of Chemical & Engineering Data*, 2019. 64(12).

[2] Khupse, N.D. and A. Kumar, Delineating solute-solvent interactions in binary mixtures of ionic liquids in molecular solvents and preferential solvation approach. *Journal of Physical Chemistry B*, 2011. 115(4): p. 711-8.

Based on the references above, the equations for calculating the  $\pi^*$  and  $\beta$  are as followings:

$$v_0 = 34.17(kK) \quad (1)$$

$$s = -2.410 \quad (2)$$

$$v_{\max(ANS)} = v_0 + S\pi^* \quad (v = 1 / \lambda^* 10000) \quad (3)$$

$$\beta = (31.10 - 3.14\pi^* - v_{\max(ANI)}) / 2.79 \quad (4)$$