

Electronic Supplementary Information for

**An easy way to identify high performing covalent organic frameworks for  
hydrogen storage**

Minman Tong<sup>‡a</sup>, Weichen Zhu<sup>‡a</sup>, Jian Li<sup>b</sup>, Zhouyang Long<sup>a</sup>, Zhao Shuang<sup>a</sup>, Guojian Chen<sup>\*a</sup>, Youshi Lan<sup>\*c</sup>

## Methods

### COF structures

The CoRE COF database is updated to 449 structures in this work, including 49 3D-COFs and 400 2D-COFs (<https://core-cof.github.io/CoRE-COF-Database/>). Pore characterizations of largest cavity diameter (LCD), accessible surface area ( $S_{\text{acc}}$ ) and void fraction ( $\phi$ ) are performed using the Zeo++ version 0.3.<sup>1</sup> The  $S_{\text{acc}}$  is calculated using a probe molecule with size equal to the kinetic diameter of N<sub>2</sub> (3.68 Å). The  $\phi$  is computed with a probe size of 0.0 Å, which is the absolute amount unoccupied by the framework atoms.

**Simulation details.** Hydrogen adsorption in COFs is determined through Grand canonical Monte Carlo (GCMC) simulations using HT-CADSS code. H<sub>2</sub> is described as a rigid, three-site model (H–H bond length of 0.741 Å) with a Lennard-Jones (LJ) site ( $\sigma = 2.958$  Å and  $\varepsilon/k_{\text{B}} = 36.7$  K) and negative charge of  $-0.936e$  placed at the center of mass of H<sub>2</sub> and positive charge of  $0.468e$  (no LJ potentials) placed at the H nuclei.<sup>2,3</sup> The quadratic Feynman–Hibbs correction<sup>4</sup> is used to modify the Lennard-Jones potentials to account for quantum effects. The LJ parameters of the framework atoms of CoRE COFs are taken from DREIDING force field,<sup>5</sup> and the missing parameters in some COFs are taken from the Universal force field (UFF).<sup>6</sup> COFs are considered as robust ‘organic zeolites’ as the organic groups are connected via covalent bonds<sup>7</sup>. To the best of our knowledge, in the 449 CoRE COFs, only LZU-301<sup>8</sup> and COF-300<sup>9</sup> have been reported to show crystal-structure transformation induced by tetrahydrofuran or H<sub>2</sub>O molecules, and none of the CoRE COFs have been reported to exhibit gate-opening behavior for H<sub>2</sub> adsorption. Hence, the ‘gate-opening’ possibility is not considered in our simulation, and COFs are held as rigid frameworks in the simulations as others do.<sup>10-12</sup> All of the LJ cross potential parameters are described the Lorentz–Berthelot mixing rules. Partial point charges of the frameworks are determined by the charge equilibration method (QEq).<sup>13</sup> A cutoff radius is set to 14.0 Å for the simulations, and the long-range interactions are handled by the Ewald summation technique. The numbers of the unit cells contained in the simulation box are COF-dependent, and no finite-size effects existed by checking the simulations with larger boxes. Periodic boundary conditions are considered in all three dimensions. Peng-Robinson equation of state is used to convert the pressure to the corresponding fugacity. Each GCMC simulation consists of  $5 \times 10^6$  equilibration steps and  $5 \times 10^6$  production steps. The simulation process involves five types of trials: translation, rotation, random swapping, deletion of molecules, and identifying changes.

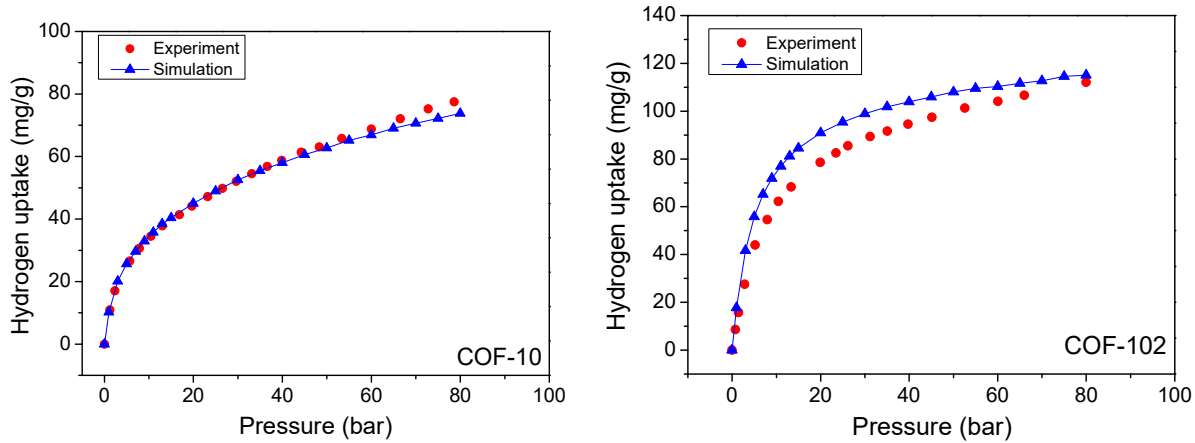
The accuracy of our simulation models are confirmed by comparing the simulated hydrogen uptake with the experimental results of some measured COFs, as shown in Fig. S1.

The isosteric heats of adsorption ( $Q_{st}$ ) under the adsorption conditions are calculated using the ensemble fluctuation method as follows,<sup>14</sup>

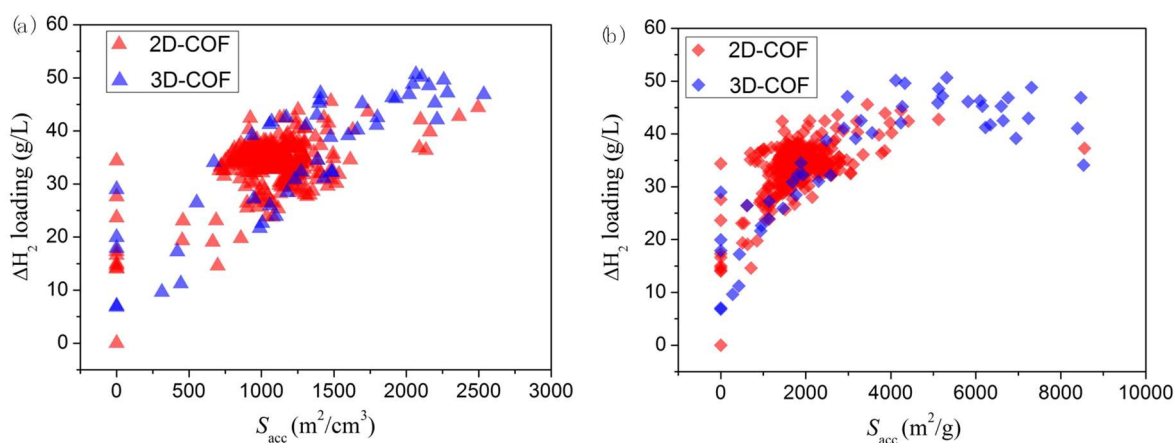
$$Q_{st} = RT - \frac{\langle U_{ff}N \rangle - \langle U_{ff} \rangle \langle N \rangle}{\langle N^2 \rangle - \langle N \rangle \langle N \rangle} - \frac{\langle U_{fm}N \rangle - \langle U_{fm} \rangle \langle N \rangle}{\langle N^2 \rangle - \langle N \rangle \langle N \rangle}$$

where the brackets  $\langle \dots \rangle$  denote the ensemble average,  $R$  is the gas constant, and  $N$  is the number of molecules adsorbed. The first and second terms are the contributions from the molecular thermal energy and adsorbate-adsorbate interaction energy  $U_{ff}$ , respectively, while the remaining term is the contribution from the adsorbate-adsorbent interaction energy  $U_{fm}$ . In the following, the contributions for the  $Q_{st}$  from the second and third terms are denoted as  $Q_{st,ff}$  and  $Q_{st,fm}$ , respectively.

To compute the free energy profiles, we performed NVT-ensemble Monte Carlo simulations using the histogram-sampling (HS) method.<sup>15, 16</sup> In the HS method, a histogram is made of the particle positions, mapped on the reaction coordinate  $q$ . Then the histogram can be converted into a free-energy profile using  $\beta F(q) = -\ln\langle P(q) \rangle$ , where  $P(q)$  denotes the probability of finding a molecule at a given position  $q$  according to the histogram.



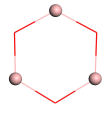
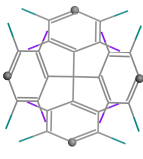
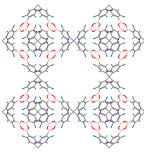
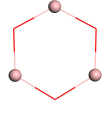
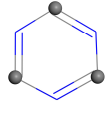
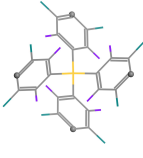
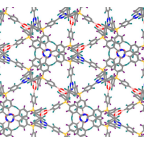
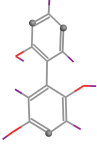
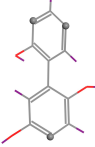
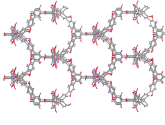
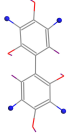
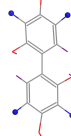
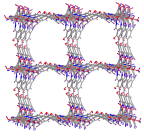
**Fig. S1** Comparison of the simulated adsorption isotherms of H<sub>2</sub> in COFs with the experimental measurements.



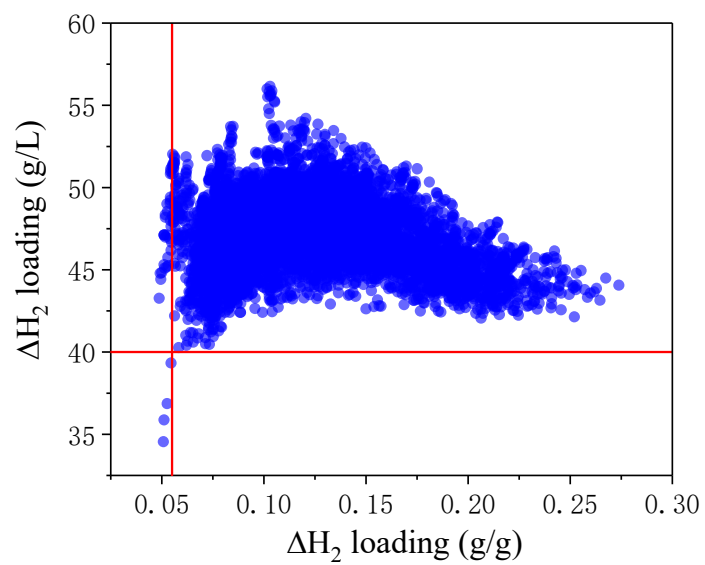
**Fig. S2** Relationship between the volumetric deliverable hydrogen capacity and (a) volumetric accessible surface area and (b) gravimetric accessible surface area.

The relationship between deliverable hydrogen capacities and accessible surface area are also investigated (Fig. S2). The optimum region of accessible surface area for top performing COFs are not determined because this region is wide and the structure-performance trend is steady and not clear enough.

**Table S1.** The structure features of the hypothetical COFs with the highest volumetric deliverable hydrogen capacity in each topology.

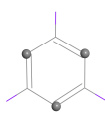
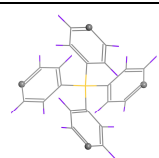
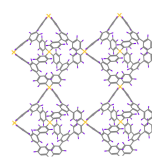
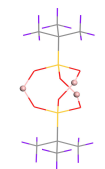
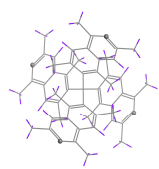
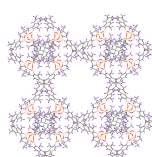
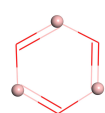
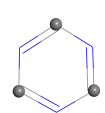
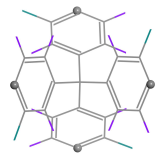
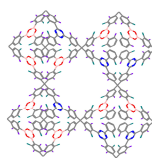
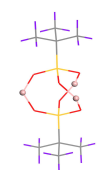
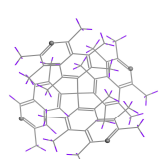
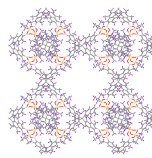
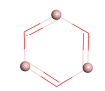
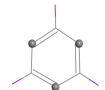
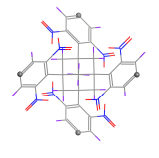
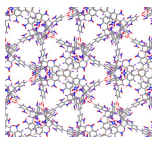
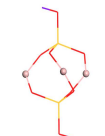
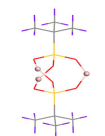
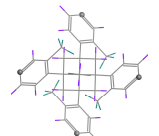
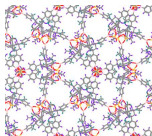
Type	Name	LCD (Å)	$\phi$	$\Delta H_2$ loading (g/L)	$\Delta H_2$ loading (g/g $\times$ 100%)	Linker1	Linker2	Center	Framework
<b>bor</b>	COF108_EX T4_F_MOD 2-108_B3O3 _No1	15.06	0.779	55.8	10.3		/		
<b>ctn</b>	COF105_F_ MOD2-102_ B3O3_P1-10 2_C3N3_P2 _No1	9.51	0.778	56.1	10.3				
<b>dia</b>	linker100_C _linker88_C _dia_relaxed	9.83	0.812	53.8	12.7		/		
<b>pts</b>	linker100_C H_linker104 _N_pts_relaxed	11.03	0.821	52.5	13.0		/		



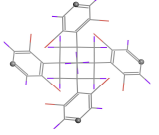
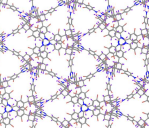
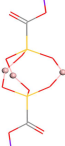
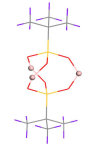
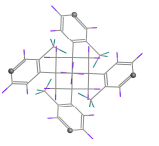
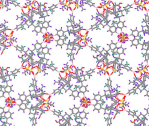


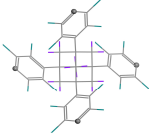
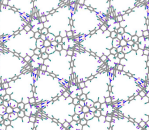
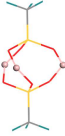
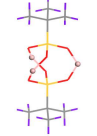
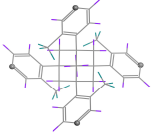
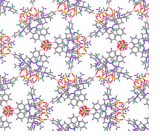


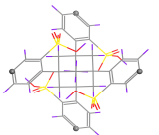
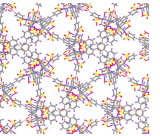
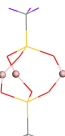
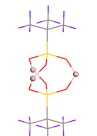
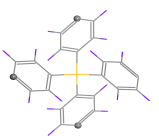
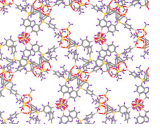
The highest volumetric deliverable hydrogen capacity achieved by each topology of COFs is present in Table S1. The LCD and  $\phi$  of the listed four hypothetical COFs are generally consistent with the optimum LCD of  $\sim 10$  Å and  $\phi$  of  $\sim 0.8$  analyzed from the CoRE COFs. But for the **bor**-type COFs, the optimum LCD is much larger to  $\sim 15$  Å because the **bor**-type COFs typically show two kinds of cages (one big and one small).




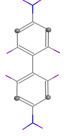
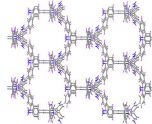

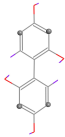
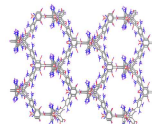
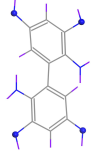
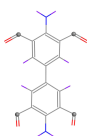
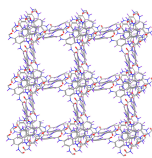
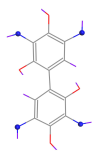
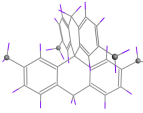
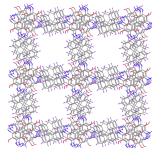
**Fig. S3** Volumetric and gravimetric deliverable hydrogen capacity of the 6893 hypothetical COFs. The red lines represent the DOE 2025 target of 40 g/L and 5.5 wt %.

**Table S2.** Textural and structural properties of the comparable COFs classified by topology (continued to Table 1).

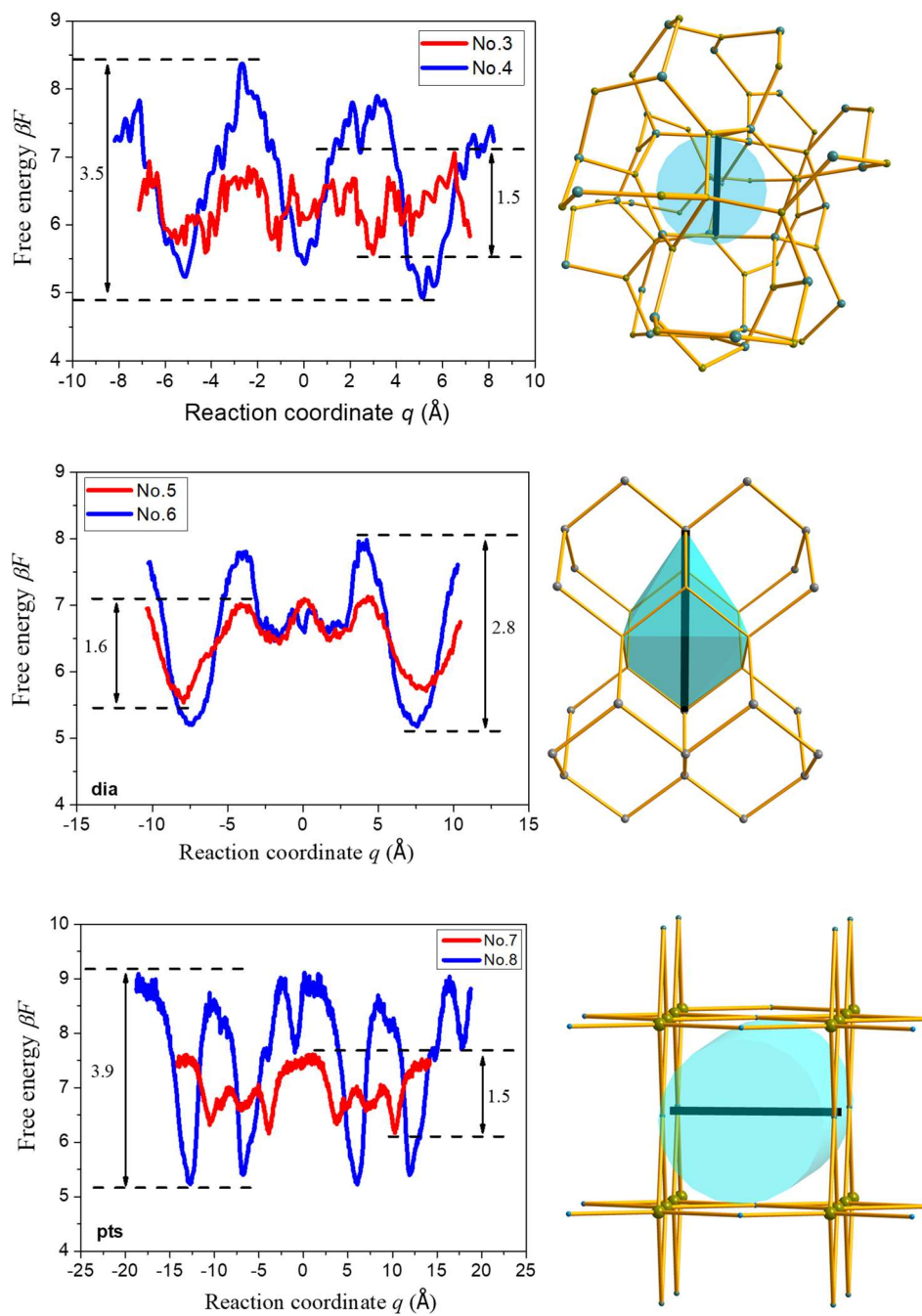
Topology	No. Name	LCD (Å)	$\phi$	$\Delta H_2$ loading (g/L)	$Q_{st}$ (kJ/mol)	Linker 1 <i>FC</i>	Linker2 <i>FC</i>	Center <i>FC</i>	Framework
<b>bor</b>	9. linker111_C_1 linker92_C_bo r_relaxed	14.23	0.800	51.2	3.9	 15.9	/	 140.1	
	10. COF108_EX T5_CH3_MO D1-108_SIO_ BUT_No1	<b>bor</b>	0.808	43.5	3.4	 137.7	/	 536.7	
<b>bor</b>	11. COF108_EX T4_F_MOD1- 108_B3O3-10 8_C3N3_No1	14.30	0.793	53.3	4.0	 8.7	 8.3	 135.8	
	12. COF108_EX T5_CH3_MO D2-108_SIO_ BUT_No1	14.28	0.798	43.9	3.6	 137.7	/	 536.7	
<b>ctn</b>	13. COF104_SPH _NO2_MOD2 -102_B3O3_P 1-102_PBB1_ P2_No1	10.90	0.790	52.1	4.3	 8.8	 16.6	 306.3	
	14. COF104_SPH _CF3_MOD1 -202_SIO_BU T_P1-202_SI O_OH_P2_N o1	10.90	0.803	47.5	3.5	 35.5	 136.5	 289.1	

ctn	15.								
	COF104_SPH _BR_MOD2- 102_C3N3_P 1-102_N3B3_ P2_No1	11.03	0.795	51.1	3.9				
					8.3	17.2	240.6		
ctn	16.								
	COF104_SPH _CF3_MOD1 -202_SIO_BU T_P1-202_SI O_COOH_P2 _No1	11.00	0.796	47.8	3.7				
					55.9	136.5	289.1		
ctn	17.								
	COF104_SPH _F_MOD3-10 2_N3B3_P1-1 02_PBB1_P2 _No1	11.15	0.807	53.0	3.5				
					17.1	16.6	242.3		
ctn	18.								
	COF104_SPH _CF3_MOD1 -202_SIO_BU T_P1-202_SI O_CF3_P2_N o1	11.14	0.795	47.0	3.7				
					53.3	136.5	289.1		
ctn	19.								
	COF104_SPH _SO3H_MOD 1-102_B3O3_ P1-102_PBB1 _P2_No1	11.07	0.792	51.9	3.9				
					8.8	16.6	311.5		
ctn	20.								
	COF105-202_ SIO_BUT_P1 -202_SIO_M ET_P2_No1	11.09	0.775	46.8	5.0				
					52.3	136.5	147.3		



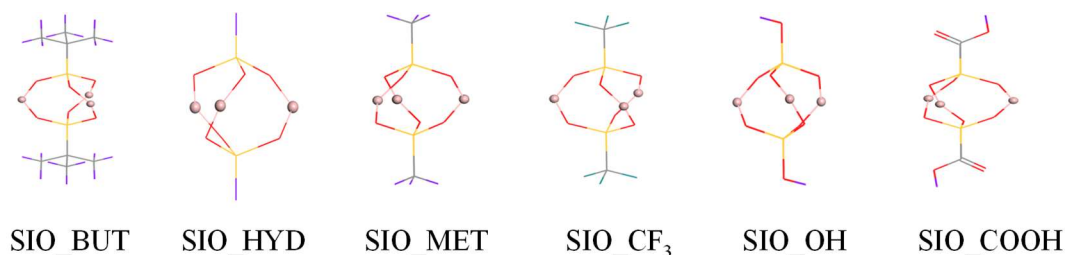
<b>dia</b>	21. linker107_C_1 linker87_C_di a_relaxed	9.83	0.798	52.2	4.2	 18.4	/	 67.4	
	22. linker104_C_1 linker11_C_di a_relaxed	9.81	0.787	46.9	4.1	 47.6	/	 67.1	
<b>pts</b>	23. linker101_NH _linker107_C O_pts_relaxed	10.70	0.795	51.9	4.1	 105.4	/	 110.7	
	24. linker102_CH 2_linker104_ NH_pts_relax ed	10.69	0.798	48.6	4.0	 101.8	/	 230.7	

In Table S2, some SBUs of good performing COFs (Nos. 13, 19, 21, 23) show higher *FC* values than that of the corresponding bad performing ones. Their *FC* values are highlighted with red color. By taking COF pair Nos. 13 and 14 as an example, the *FC* value of the center fragment in COF No. 13 is 306.3, higher than that of the center fragment in COF No. 14 which is 289.1, with the complexity ratio of  $306.3/289.1=1.1$ . However, the *FC* values of the linker fragments in COF No. 14 are much higher than that in COF No. 13, with the complexity ratios of  $35.5/8.8=4.0$  and  $136.5/16.6=8.2$ . Especially, **ctn** COF contains more linker fragment than center fragment (the molar ratio of linker/center in **ctn** COF is 4/3). Therefore, the structure of COF No. 14 is more complicated than COF No.13. The situation of COFs Nos. 19, 21, 23 are similar with COF No. 13.

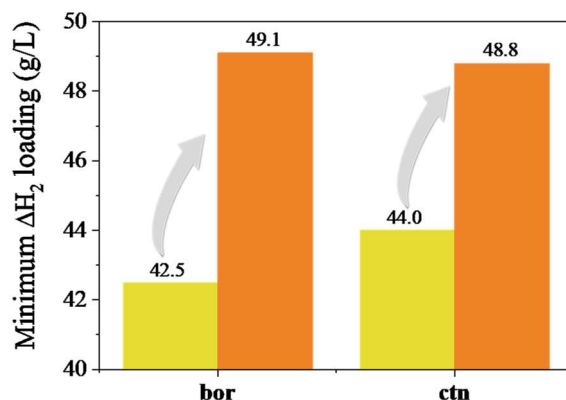


**Fig. S4** Free-energy profile for H<sub>2</sub> in dia-type COFs No. 5 and No. 6 and pts-type COFs No. 7 and No. 8 (in Table 1) along the midlines at 77 K and infinite dilution. The black lines in the channels represent the reaction path.

From Table 1 and Table S2, we can see that the bad performing **bor**- and **ctn**-type COFs usually have the stereo linker of borosilicate cluster (recorded as SIO in the GCOF name, also shown in Fig. S5), a building unit originated from the classical COF-202<sup>17</sup>. Consider that the structure composition is reflected in the GCOF name, we make a simple elimination of **bor**- and **ctn**-type GCOFs with SIO and polyatomic functional groups because excess modifications bring structure complexity. After filtration, the minimum volumetric deliverable hydrogen capacity in the corresponding region can be improved from 42.5 to 49.1 g/L for **bor**-GCOFs and from 44.0 to 48.8 g/L for **ctn**-GCOFs (Fig. S6).



**Fig. S5** A series of stereo linker of borosilicate clusters with three connection sites.



**Fig. S6** The minimum volumetric deliverable hydrogen capacity in the intersectant shadow region shown in Fig. 2 (a) and (b) for **bor**-GCOFs and Fig. 2 (c) and (d) for **ctn**-GCOFs after elimination of GCOFs with SIO and polyatomic functional groups ( $\text{NH}_2$ ,  $\text{NO}_2$ ,  $\text{CH}_3$ ,  $\text{CF}_3$ ,  $\text{SO}_3\text{H}$ ,  $\text{COOH}$  and  $\text{C}(\text{CH}_3)_3$ )

Knowing this principle, we seem able to understand why the reported best performing MOFs are always the classics with simple structures.<sup>18-21,22</sup> For example, recently, Ahmed et al.<sup>19</sup> computationally screened nearly half a million MOFs for  $\text{H}_2$  storage and experimentally assessed the most promising ones. At the same operation condition to us, the top 2 MOFs are MOF-5 and IRMOF-20 that show volumetric deliverable hydrogen capacities higher than 50 g/L (51.9 and 51.0 g/L respectively). MOF-5 and IRMOF-20 share the cubic **pcu** topology formed by the same

Zn<sub>4</sub>O<sub>13</sub> octahedral SBU and respectively the simple and flat benzene and thieno[3,2-b]thiophene links. These experimental results prove our finding that creating homogeneous energetic environment in nanoporous materials by using simple SBUs is significant for achieving high volumetric deliverable hydrogen capacity.

## References

1. D. Ongari, P. G. Boyd, S. Barthel, M. Witman, M. Haranczyk and B. Smit, *Langmuir*, 2017, **33**, 14529-14538.
2. F. Darkrim and D. Levesque, *J. Chem. Phys.*, 1998, **109**, 4981-4984.
3. A. Michels, W. de Graaff and C. A. Ten Seldam, *Physica*, 1960, **26**, 393-408
4. R. B. Getman, Y. Bae, C. E. Wilmer and R. Q. Snurr, *Chem. Rev.*, 2012, **112**, 703-723.
5. S. L. Mayo, B. D. Olafson and W. A. Goddard, *J. Phys. Chem.*, 1990, **94**, 8897-8909.
6. A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024-10035.
7. A. P. Cote, A. I. Benin, N. W. Ockwig, M. O'Keeffe, A. J. Matzger and O. M. Yaghi, *Science*, 2005, **310**, 1166-1170.
8. Y. X. Ma, Z. J. Li, L. Wei, S. Y. Ding, Y. B. Zhang and W. Wang, *J. Am. Chem. Soc.*, 2017, **139**, 4995-4998.
9. Y. Chen, Z. L. Shi, L. Wei, B. Zhou, J. Tan, H. L. Zhou and Y. B. Zhang, *J. Am. Chem. Soc.*, 2019, **141**, 3298-3303.
10. G. Garberoglio, *Langmuir* 2007, **23**, 12154-12158
11. J. L. Mendoza-Cortés, S. S. Han, H. Furukawa, O. M. Yaghi, W. A. Goddard III, *J. Phys. Chem. A* 2010, **114**, 10824-10833
12. K. Gopalsamy, C. Desgranges, J. Delhommelle, *J. Phys. Chem. C* 2017, **121**, 24692-24700.
13. B. A. Wells, C. De Bruindickason and A. L. Chaffee, *J. Phys. Chem. C*, 2015, **119**, 456-466.
14. F. Karavias and A. L. Myers, *Langmuir*, 1991, **7**, 3118-3126.
15. E. Beerdsen, D. Dubbeldam and B. Smit, *J. Phys. Chem. B*, 2006, **110**, 22754-22772.
16. B. Liu, Q. Yang, C. Xue, C. Zhong, B. Chen and B. Smit, *J. Phys. Chem. C*, 2008, **112**, 9854-9860.
17. J. R. Hunt, C. J. Doonan, J. D. Levangie, A. P. Cote and O. M. Yaghi, *J. Am. Chem. Soc.*, 2008, **130**, 11872-11873.
18. A. Ahmed, Y. Liu, J. Purewal, L. D. Tran, A. G. Wongfooy, M. Veenstra, A. J. Matzger and D. J. Siegel, *Energy Environ. Sci.*, 2017, **10**, 2459-2471.
19. A. Ahmed, S. Seth, J. Purewal, A. G. Wongfooy, M. Veenstra, A. J. Matzger and D. J. Siegel, *Nature Commun.*, 2019, **10**, 1568.
20. D. A. Gomezgualdron, T. C. Wang, P. Garciaholley, R. M. Sawelewa, E. Argueta, R. Q. Snurr, J. T. Hupp, T. Yildirim and O. K. Farha, *ACS Appl. Mater. Interfaces*, 2017, **9**, 33419-33428.
21. A. W. Thornton, C. M. Simon, J. Kim, O. Kwon, K. S. Deeg, K. Konstas, S. J. Pas, M. R. Hill, D. A. Winkler and M. Haranczyk, *Chem. Mater.*, 2017, **29**, 2844-2854.
22. P. Garciaholley, B. Schweitzer, T. Islamoglu, Y. Liu, L. Lin, S. Rodriguez, M. H. Weston, J. T. Hupp, D. A. Gomezgualdron and T. Yildirim, *ACS Energy Lett.*, 2018, **3**, 748-754.