Discovery of a molecular adsorbent for efficient CO_2/CH_4 separation using a Computation-ready Experimental database of porous molecular materials

Siyuan Yang,‡[ab] Qianqian Mao,‡[a] Heng Ji,‡[a] Dingyue Hu,[a] Jinjin Zhang,[a] Linjiang Chen,[c,d]* and Ming Liu[a,b]*

[a] Department of Chemistry, Zhejiang University, Hangzhou 310027.

[b] Hangzhou Global Scientific and Technological Innovation Center (HIC), Zhejiang University, Hangzhou 311200

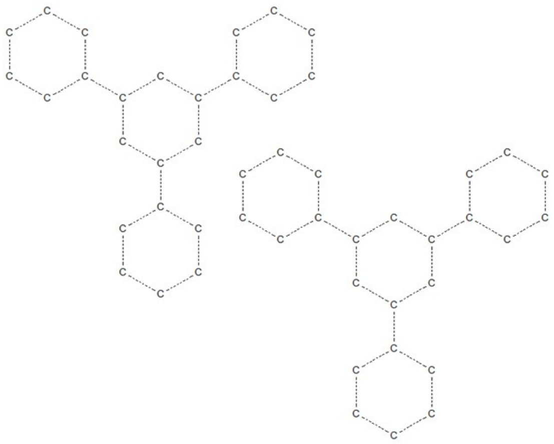
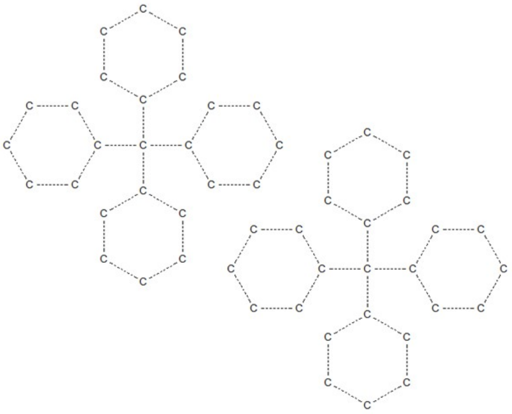
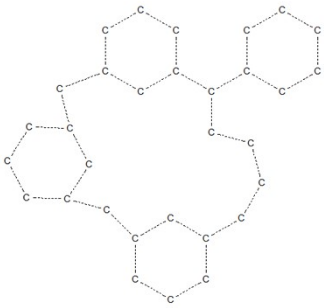
[c] Key Laboratory of Precision and Intelligent Chemistry, University of Science and Technology of China, Hefei, Anhui 230026.

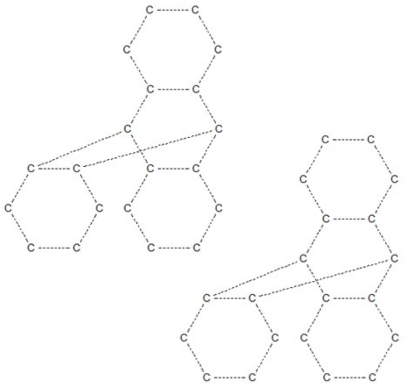
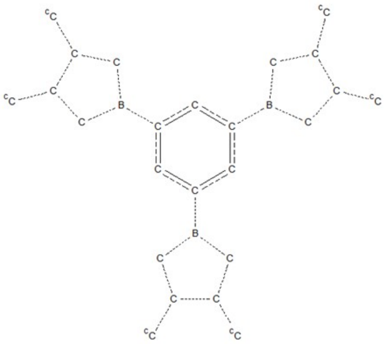
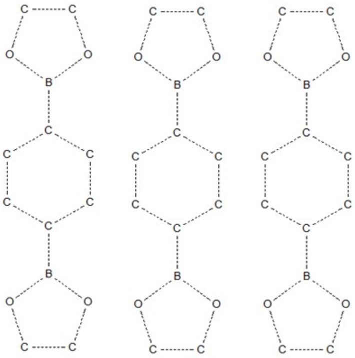
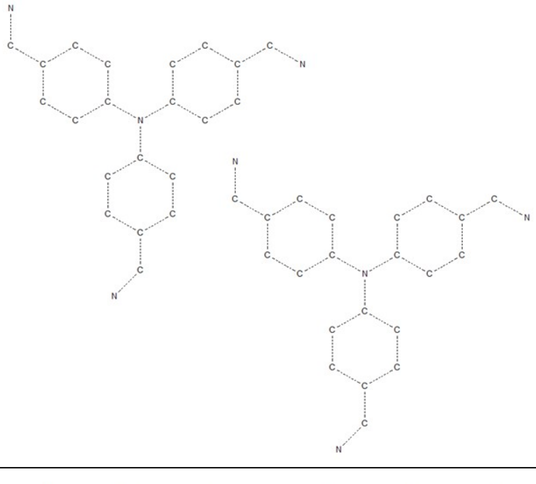
[d] School of Chemistry and School of Computer Science, University of Birmingham, Birmingham, B15 2TT

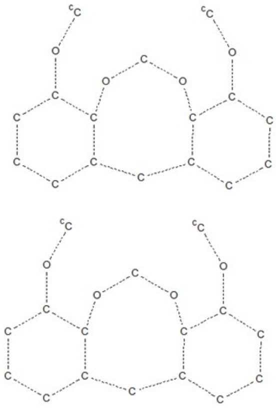
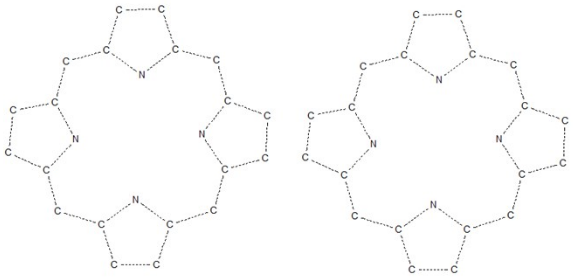
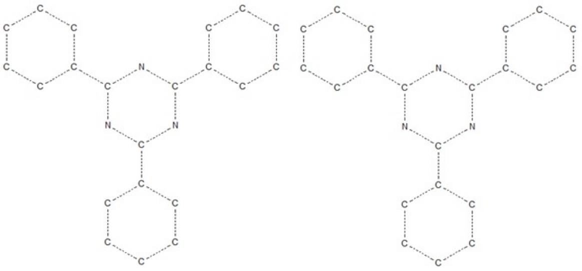
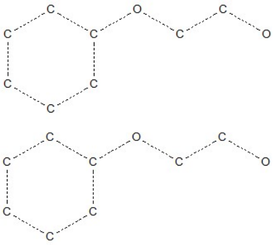
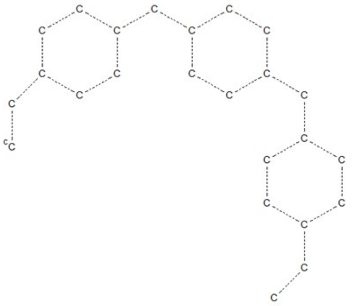
Contents

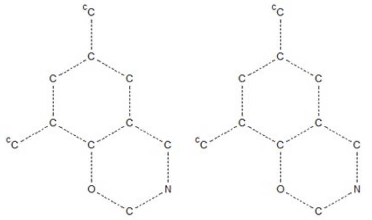
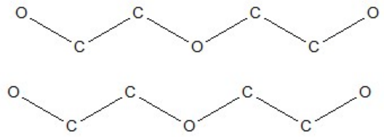
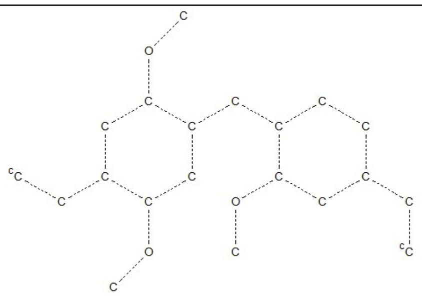
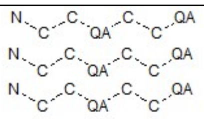
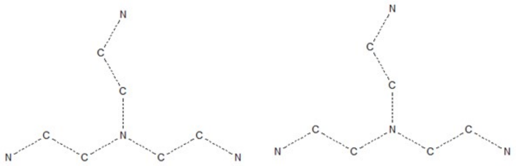
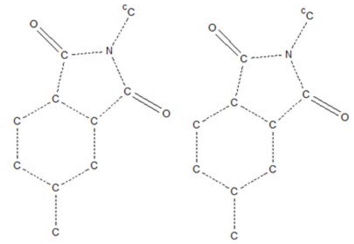
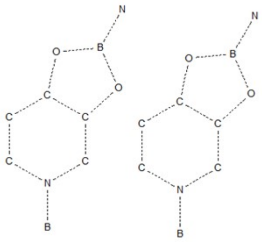
S1 ConQuest queries used for search space of macrocycles and cages in Cambridge Structural Database (CSD)

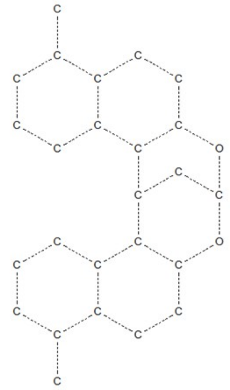
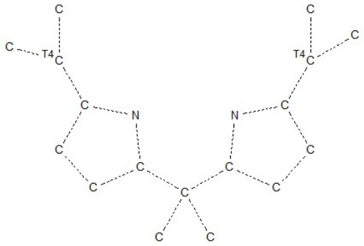
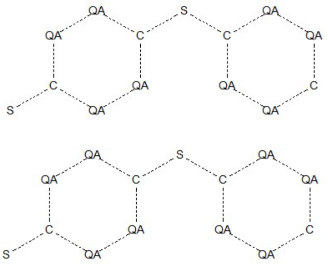
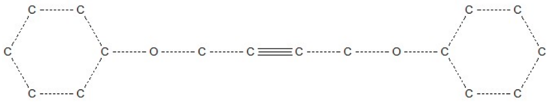
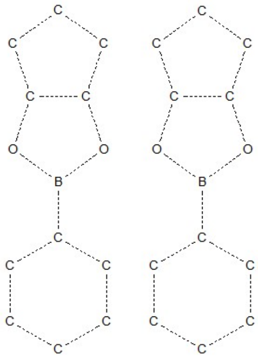
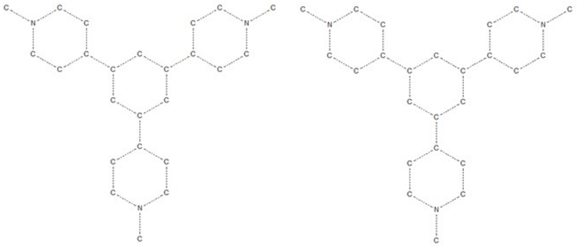
Table S1 Queries for searching macrocycles and cages.

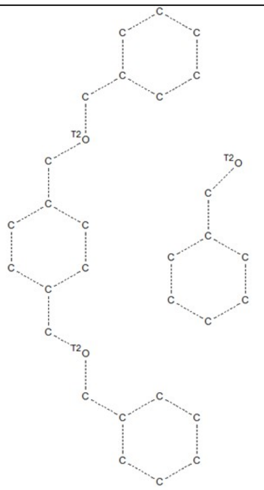
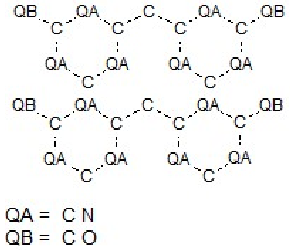
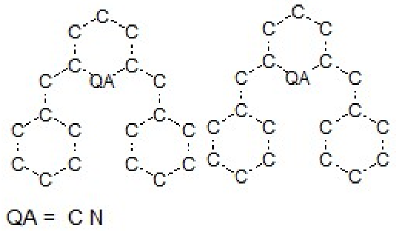
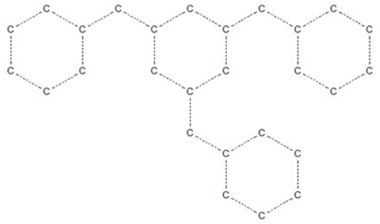
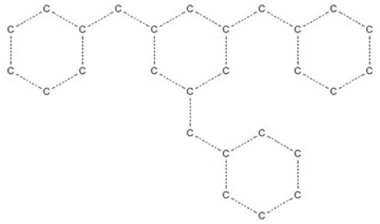
Query	Number of hits
	2054
	173
	941

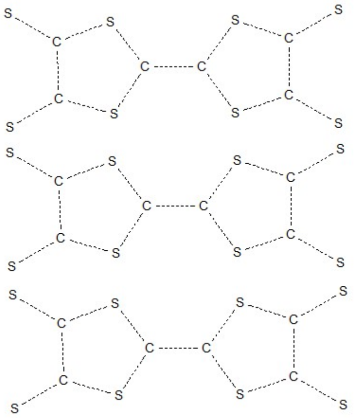
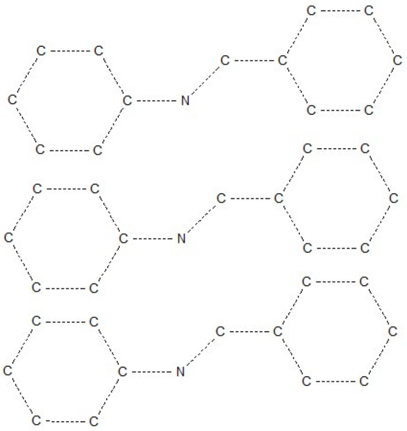
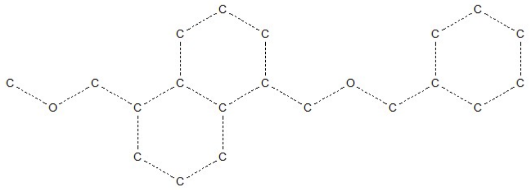
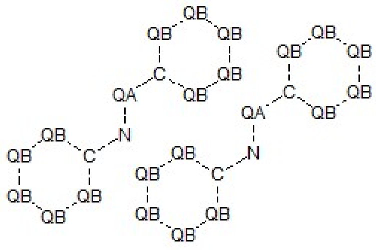
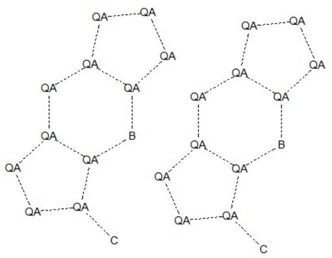
	271
	5
	11
	9
	725

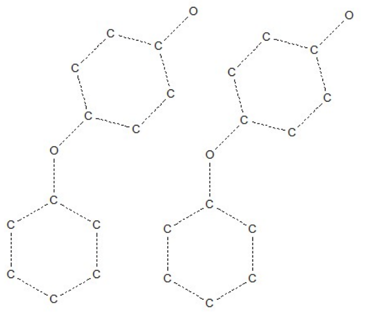
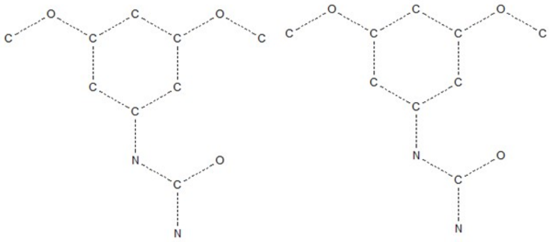
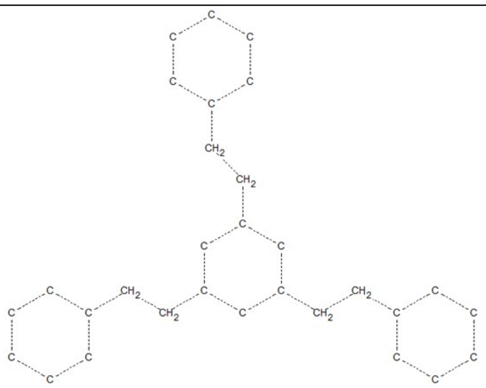
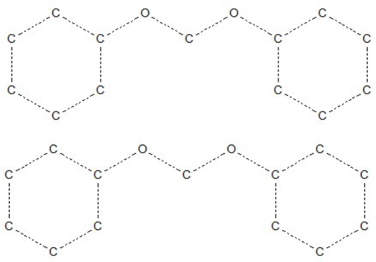
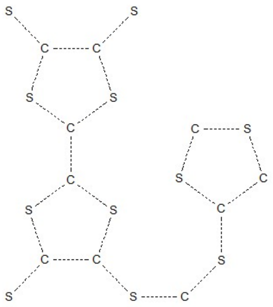
	28
	154
	49
	4245
	2438

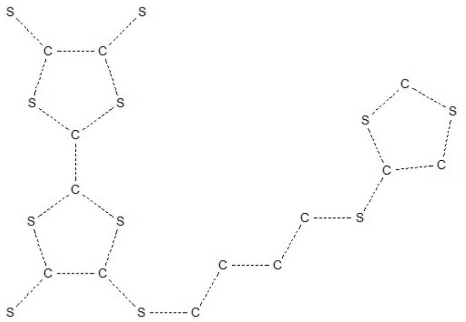
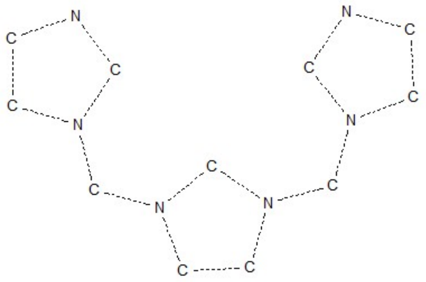
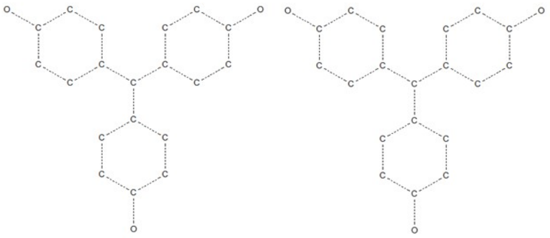
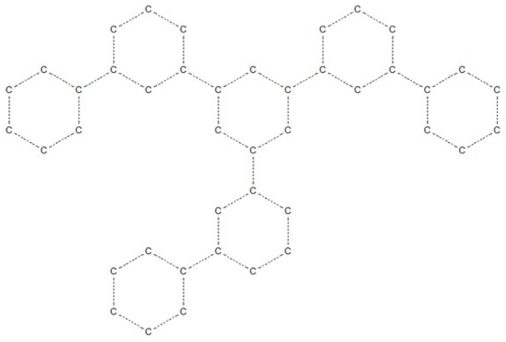
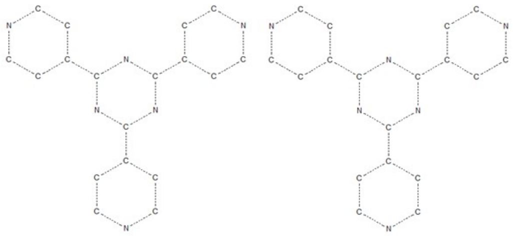
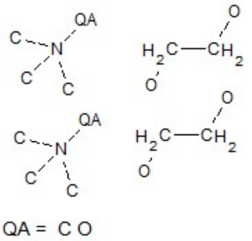
	17
	6567
	658
 QA = O N	2269
	418
	120
	4

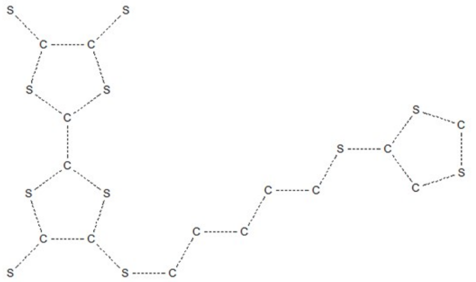
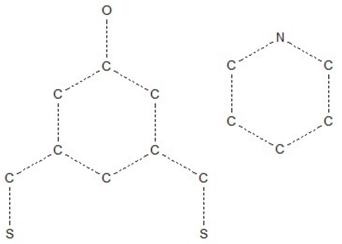
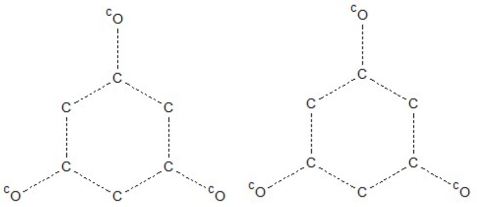
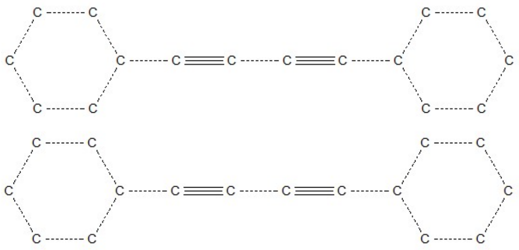
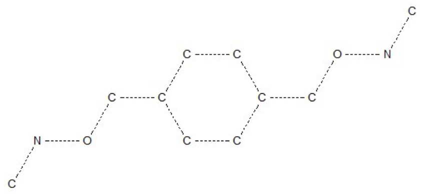
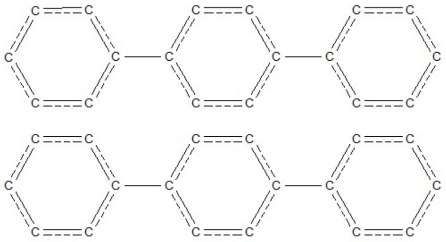
	72
	632
	94
	23
	32
	11

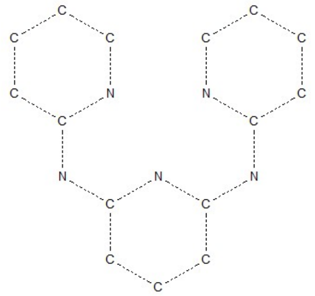
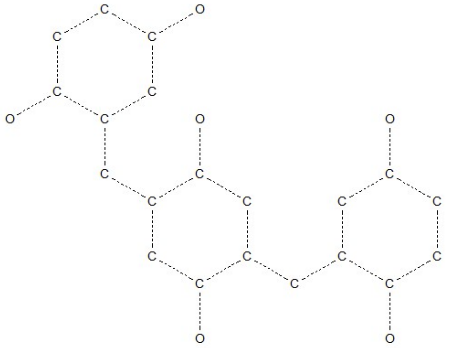
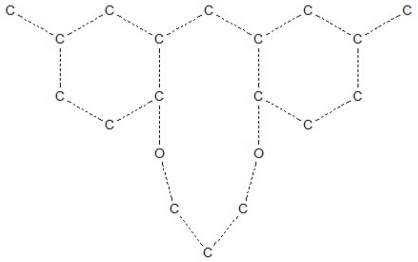
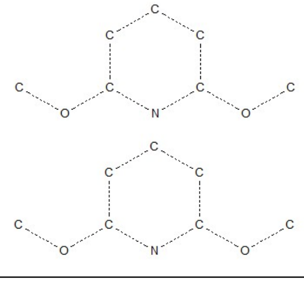
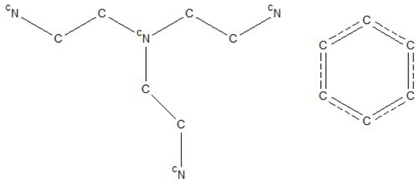
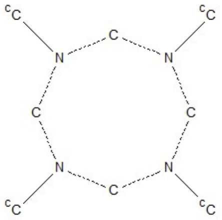
	38
	360
	3657
 	1795

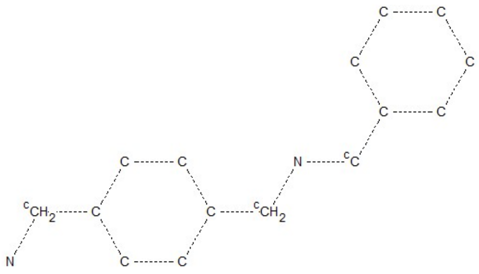
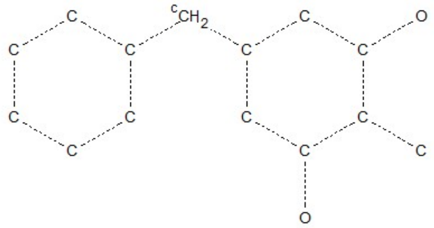
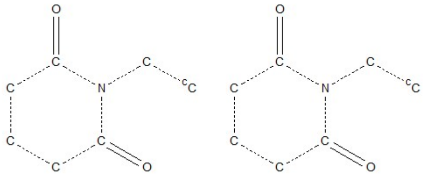
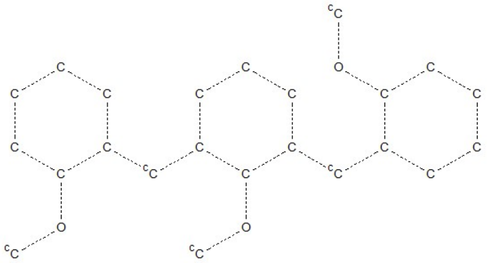
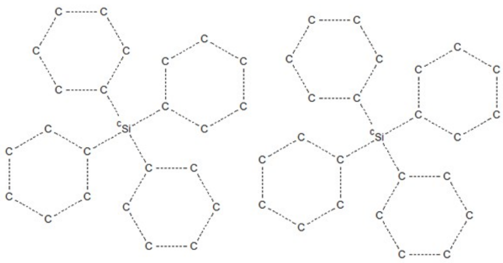
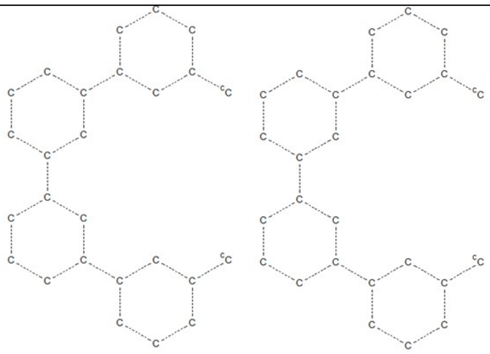
	10
	1000
	18
 QA = C S QB = C N	5625
	237

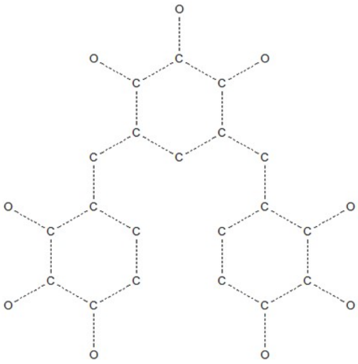
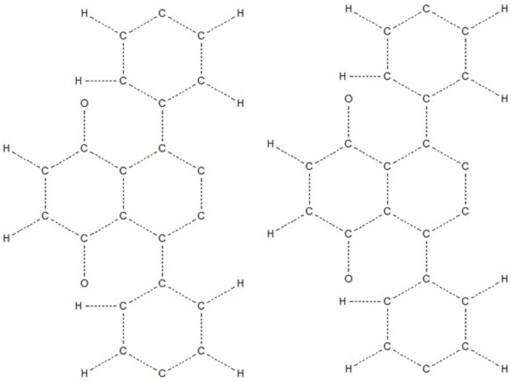
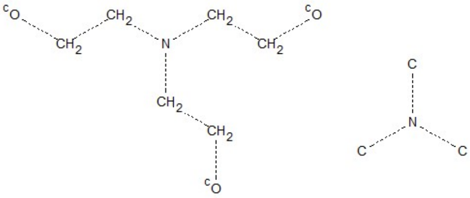
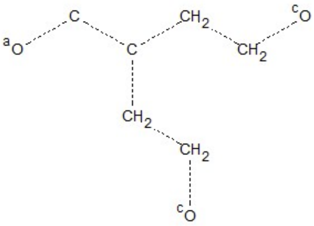
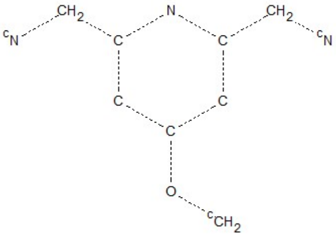
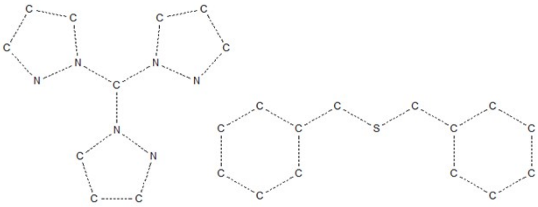
	98
	6
	15
	382
	0

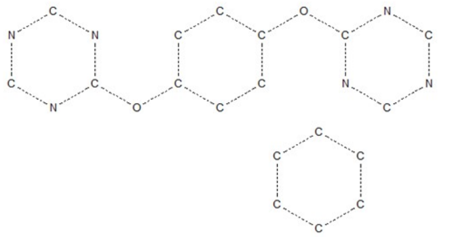
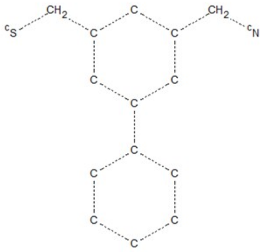
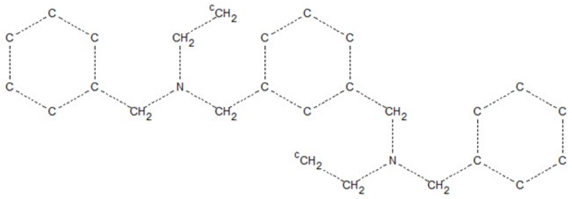
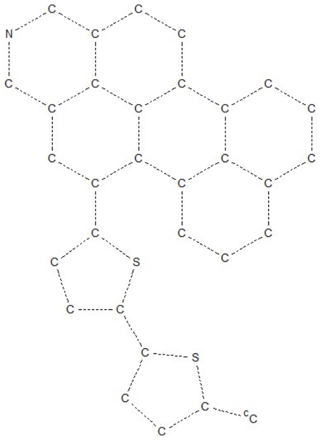
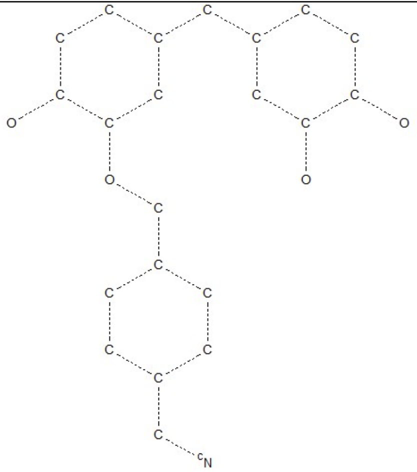
	3
	713
	74
	1930
	15
 QA = C O	15

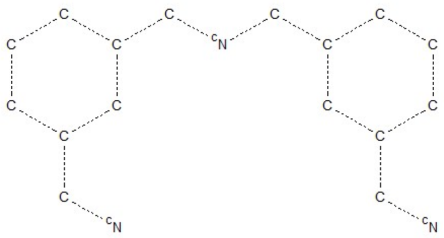
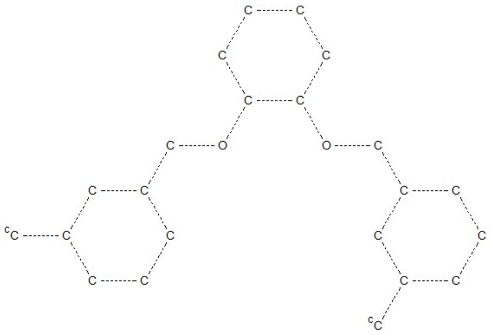
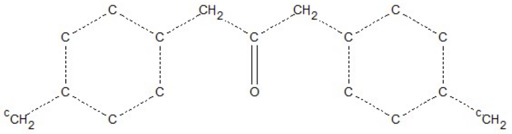
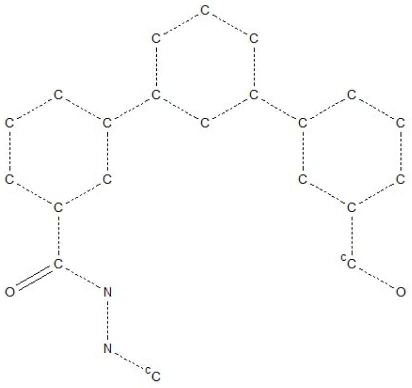
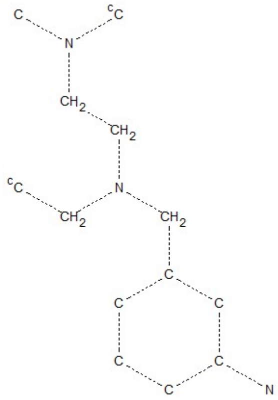
	5
	4
	91
	108
	12
	795

	93
	626
	8
	46
	467
	936

	231
	76
	606
	156
	7
	833

	271
	4
	584
	3
	1
	1

	1
	2
	7
	1
	2

	196
	18
	18
	1
	7

S2 Database Construction

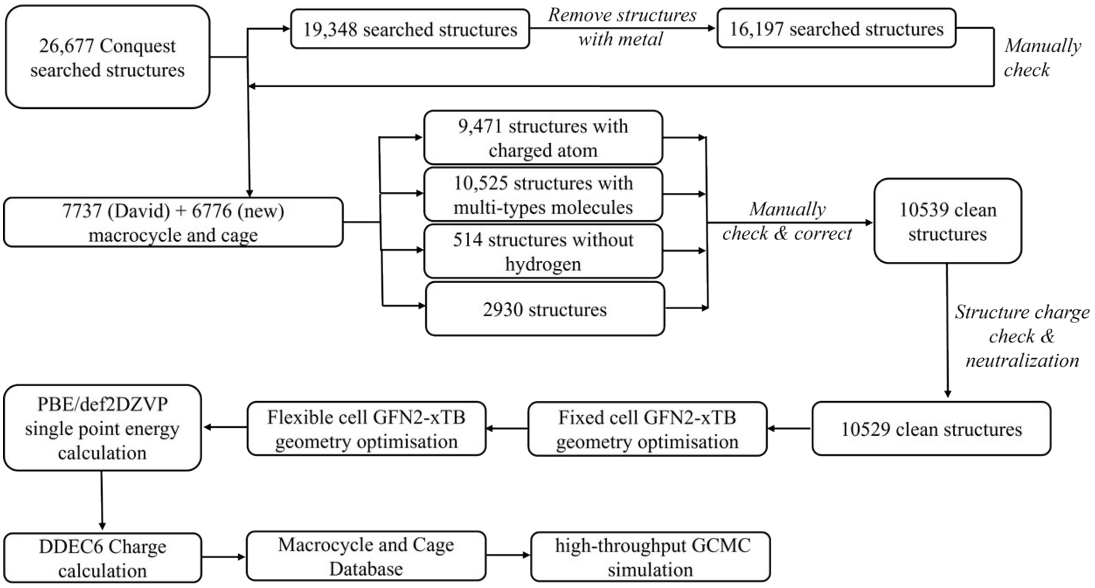


Fig. S1 Workflow for generating the macrocycles and cages database.

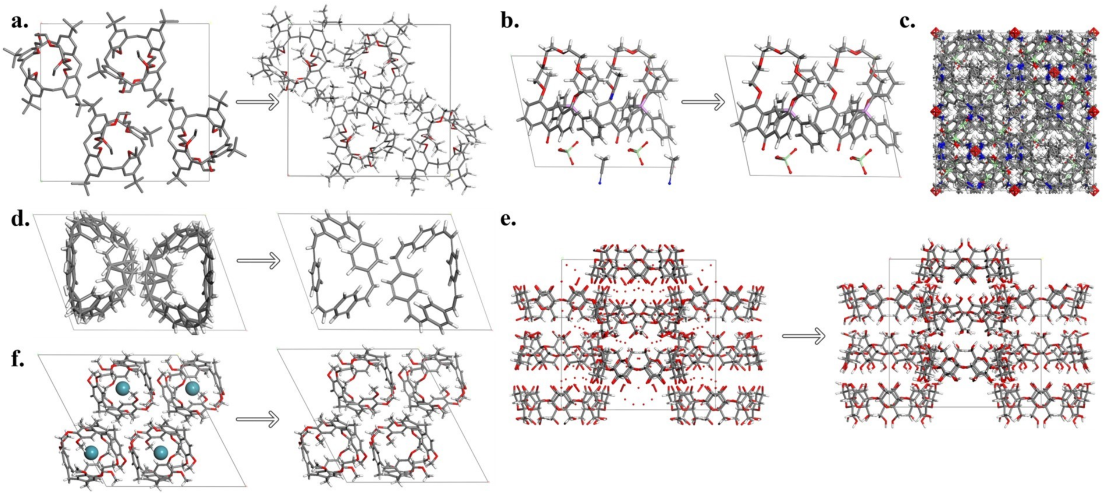


Fig. S2 Examples of structure cleaning: (a) Hydrogen atom addition of SIGDAN01, (b) solvent molecule removal of CAMVOE, (c) structure of IYEMAC is too messy to modify, (d) coordinate redundancy elimination of PCYPTR01, (e) add missing atom, removal of redundant coordinate and solvent molecule of NUNRIX, (f) guest molecule removal of OJITOR. Carbon: grey, hydrogen: white, oxygen: red, nitrogen: blue, chlorine: green. The guest molecule xenon is shown in dark cyan in the CPK model.

Structure charge neutralisation

Fast, high-throughput energy calculations were performed on the cleaned structures to verify charge neutrality. These calculations were performed using the Forcite module within *Materials Studio*[1] with the calculation accuracy set to "fine". The COMPASSIII force field was used for these evaluations, with charges automatically assigned according to the force field parameters. By examining the output files from the energy calculations, the charge neutrality of each structure could be easily determined.

For these net-charged structures, the host atoms were fixed and the number of counterions was modified. if deleting only a few counterions (less than 3) would neutralise the structure, we simply deleted the redundant ions. For the structures without enough counterions, we used the Fixed Loading Metropolis method in the Sorption module implemented in Materials Studio[1] to insert counterions, with the maximum loading steps set to 100,000 and both equilibrium and production steps set to 1,000,000. Rotation, translation, regrowth and conformer changes were allowed during the simulation. For these structures, a significant proportion of counterions must be deleted, we deleted all these ions and re-insert ions with the method described above to avoid the charge inequivalence caused by manual deletion.

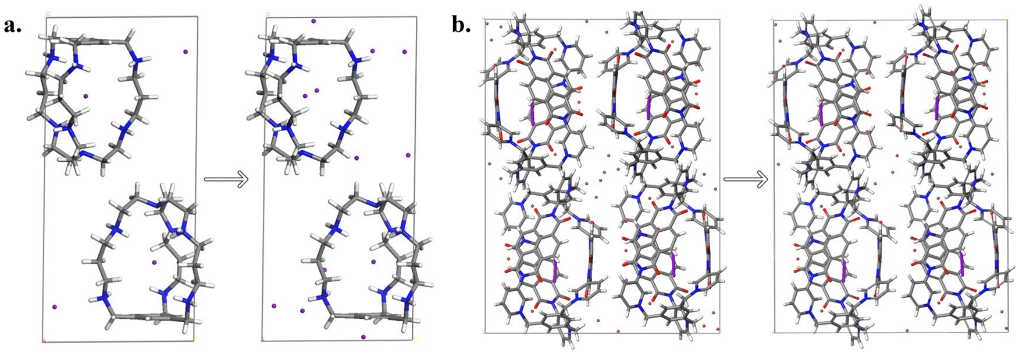
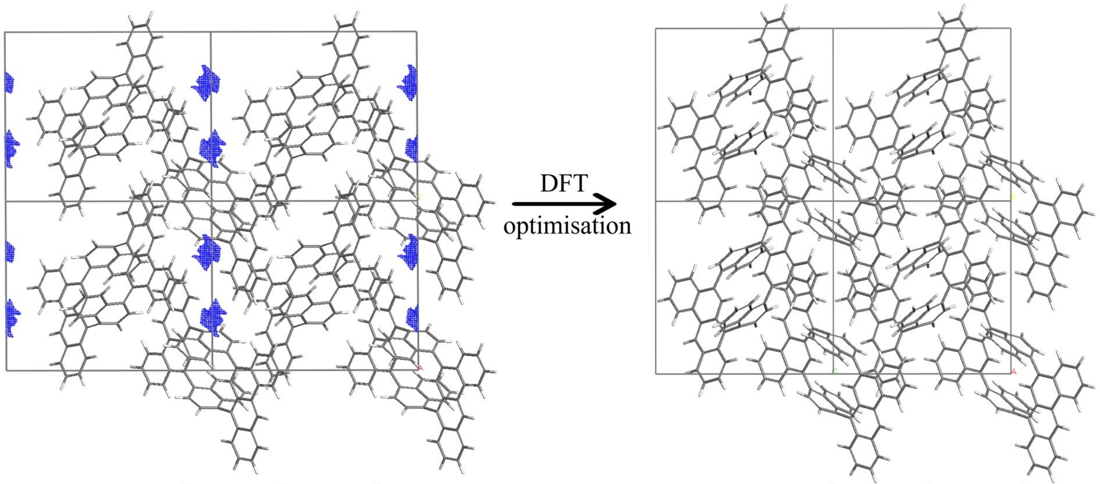


Fig. S3 Examples of structure charge neutralisation: (a) add 8 missing iodide ions of BIHQOZ, (b) delete 16 redundant bromide ions of EVIROT. Carbon: grey, hydrogen: white, oxygen: red, nitrogen: blue, bromine: brown, iodine: purple.

Table S2. Counts of operations in structure cleaning process.

Operations	Counts
add hydrogen	715
add missing atom	309
remove solvent	6741
modify atom position	70
delete redundant atoms	3130
charge neutralisation	163
delete guest molecule	267
correct element	12
no change needed	2386
bin: too messy to correct	688
bin: redundant structure	227
bin: not macrocycle or cage	661
bin: fullerene	336
bin: rotaxanes	493
bin: contain metal	385
Sum	16583

S3 REWKIQ, GFN2-xTB optimised structure vs PBE optimised structure



a, b, c = 9.70 Å, 13.31 Å, 16.35 Å
α, β, γ = 89.82°, 97.11°, 90.28°

a, b, c = 9.28 Å, 13.46 Å, 14.13 Å
α, β, γ = 89.97°, 78.71°, 89.95°

Fig. S4 Comparison between GFN2-xTB optimised structure and PBE optimised structure. The blue dots denote the accessible volume with probe radius of 1.65 Å.

S4 Experimental synthesis and characterization

All reagents were commercially available, used as supplied without further purification. ^{1}H Nuclear Magnetic Resonance (NMR) spectra were collected on a Bruker Advance spectrometer (operating at 600 MHz for ^{1}H NMR). Powder X-ray diffraction patterns (PXRD) measurements were conducted on a SmartLab SE instrument with Cu radiation source. Each PXRD pattern was obtained using a step size of 0.02˚ in a 2θ range of 2-60˚. Thermogravimetric analysis (TGA) were carried out under nitrogen atmosphere on a Discovery TGA 550 thermogravimeter with a heating rate of 5 K/min. CO_2 and CH_4 adsorption-desorption isotherms at 273 and 298 K up to 100 kPa were measured on a Advanced Gas Sorption and Micropore Analyzer (BSD-660M) after the samples were activated and degassed in vacuum at 353 K for 10 h. Breakthrough experiments were performed on a Micromeritics Instrument. Moreover, the heat of adsorption (Qst) curves are calculated through Clausius-Clapeyron equation.

Breakthrough Experiments

0.4662 g **NDI-Δ-CH_2Cl_2** was filled into the stainless-steel adsorption column (inner diameter 4.5 mm, length 15 cm), and the unfilled **NDI-Δ-CH_2Cl_2** powder part was filled with asbestos.

The fixed bed breakthrough experiments tests were carried out on the Breakthrough Analyzer BTA (Micromeritics Instrument Co, Ltd). Before the experiment, Ar gas was injected into the pipeline of the breakthrough system to ensure that the pipeline was fully cleaned, and the sample column was heated to 353 K for 4 h at the same time to fully active the **NDI-Δ-CH_2Cl_2** powder. After the sample column was lowered to room temperature (298 K), CO_2/CH_4 mixture (1:1, v:v) was injected at 1 bar at a constant flow rate (1 mL/min). The composition of the gas at the exit of the adsorption column was obtained by mass spectrometry CirrusTM 3 (MKS Co, Ltd), and the temperature of the adsorption column was controlled at 298 K. When the composition of the outgoing gas from the adsorption column was constant, the breakthrough experiment was completed. For the breakthrough cycling test, the Ar was injected into the adsorption column at the flow rate of 10 mL min^{-1} to purge the adsorbed substances in the adsorption column for 2 h. Meanwhile, the adsorption is rapidly heated to 253 K (heating rate is about 15-20 K/min) to ensure that all adsorbed substances in the adsorption column were desorbed. After that, carried out the breakthrough experiment, injected CO_2/CH_4 mixture (1:1, v:v) at 1 bar, room temperature and a constant flow rate (1 mL/min). Repeated the above procedures five times.

Synthesis and characteriazations

NDI-Δ was synthesized according to literature procedure[2]. A solution of *(RR)-trans*-1,2-cyclohexanediamine [0.855 g, 7.5 mmol] in anhydrous DMF (5 ml) was added quickly to a solution of naphthalenetetracarboxylic dianhydride (1.99 g, 7.4 mmol) in anhydrous DMF (100 ml) with vigorous stirring at 150 °C. The resulting dark red reaction mixture was stirred at 150 °C for 5 h and the DMF was evaporated under reduced pressure at 80 °C. The deep red residue was dissolved in CH_2Cl_2 and purified by flash column chromatography over neutral aluminum oxide (CH_2Cl_2/Me_2CO, 0–10% Me_2CO), followed by precipitation of the product with MeOH to afford pure **NDI-Δ** (0.77 g, 0.7 mmol) in 30 % yield as a slightly red solid. ^1H NMR (600 MHz, $CDCl_3$): δ (ppm): 8.48 (d, 12H), 6.23 (d, 6H), 2.46-2.51 (m, 6H), 2.00-1.93 (m, 12H), 1.63-1.69 (m, 12H).

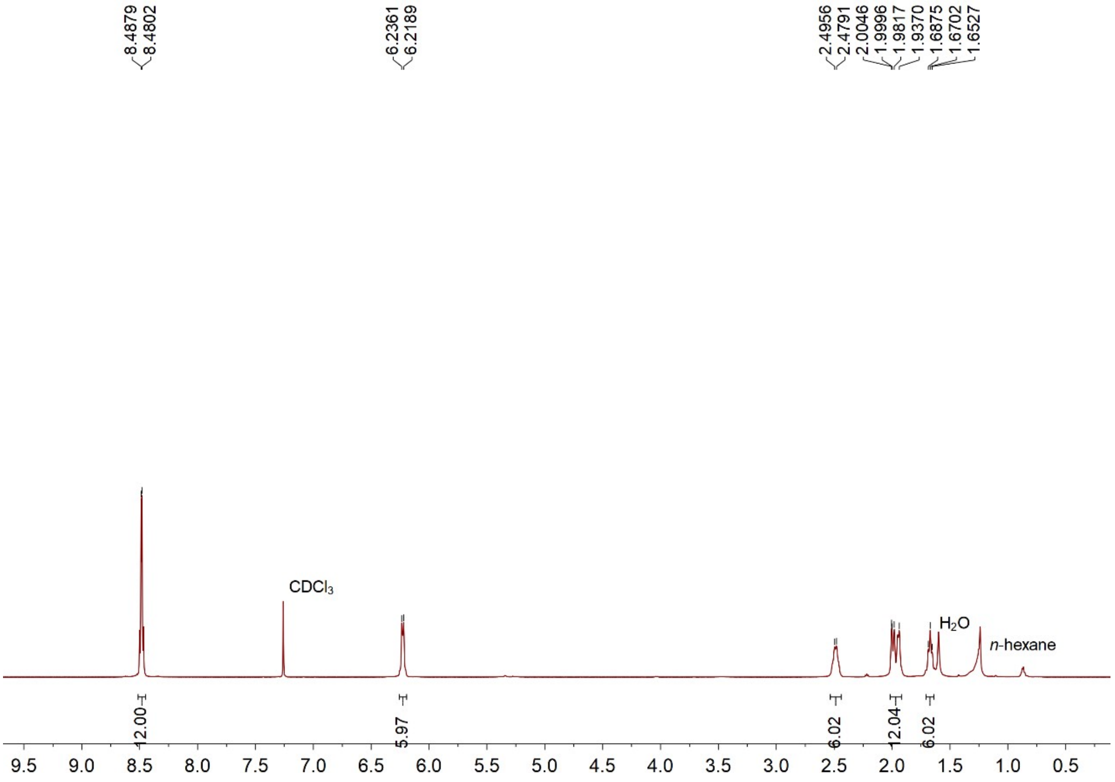


Fig. S5 ^1H NMR spectrum (600 MHz, $CDCl_3$, 298 K) of **NDI-Δ**

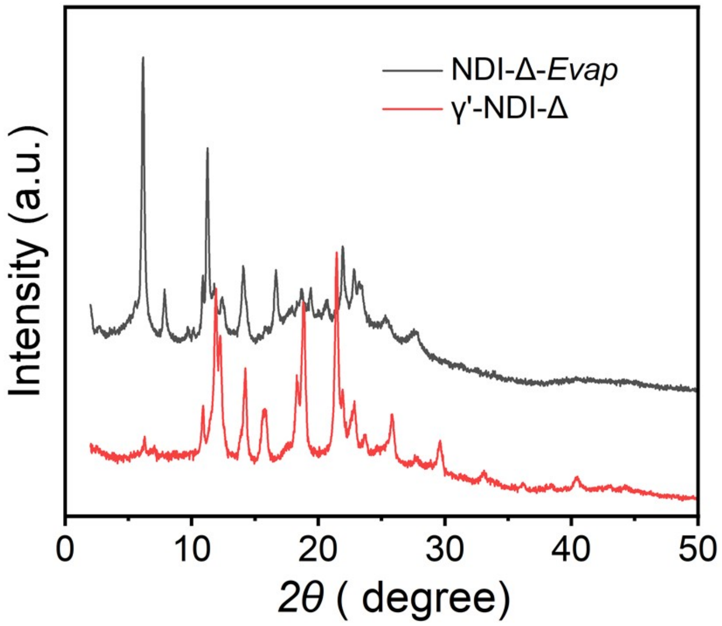


Fig.S6 Powder X-ray diffraction patterns for **γ'-NDI-Δ** and **NDI-Δ-*Evap***.

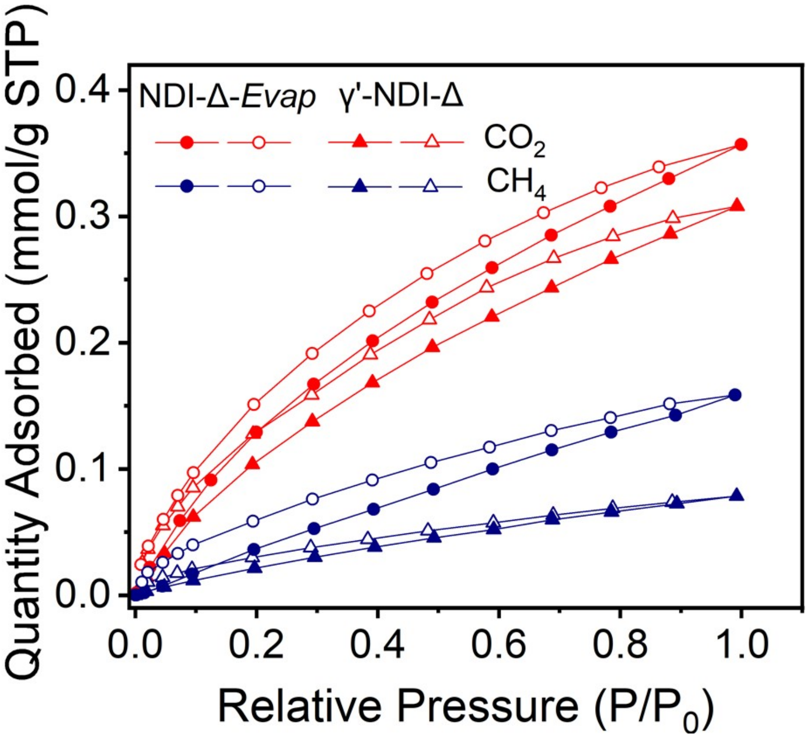


Fig. S7 298 K CO_2 and CH_4 sorption isotherms for **γ'-NDI-Δ** and **NDI-Δ-*Evap*** (closed symbols: adsorption; open symbols: desorption).

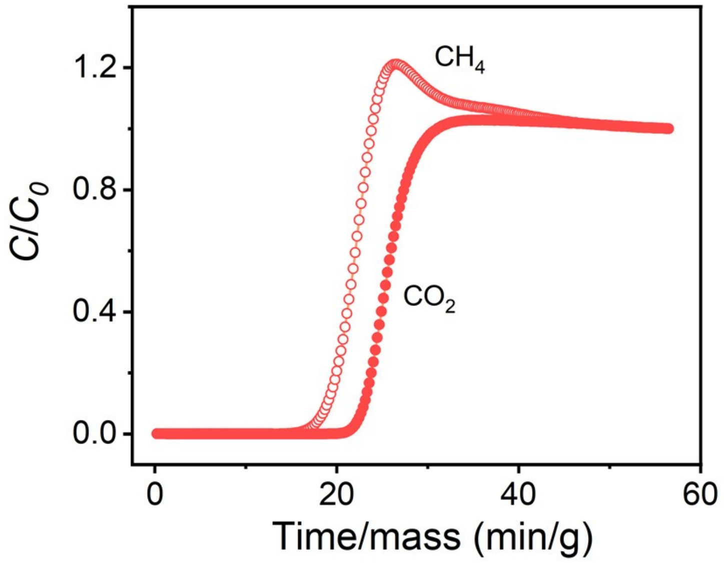


Fig. S8 Separation performances of **γ'-NDI-Δ**: Breakthrough curves for gas feed of equimolar binary mixture of CO_2/CH_4 with a flow rate of 1 mL · min^{-1} at 1 bar and 298 K.

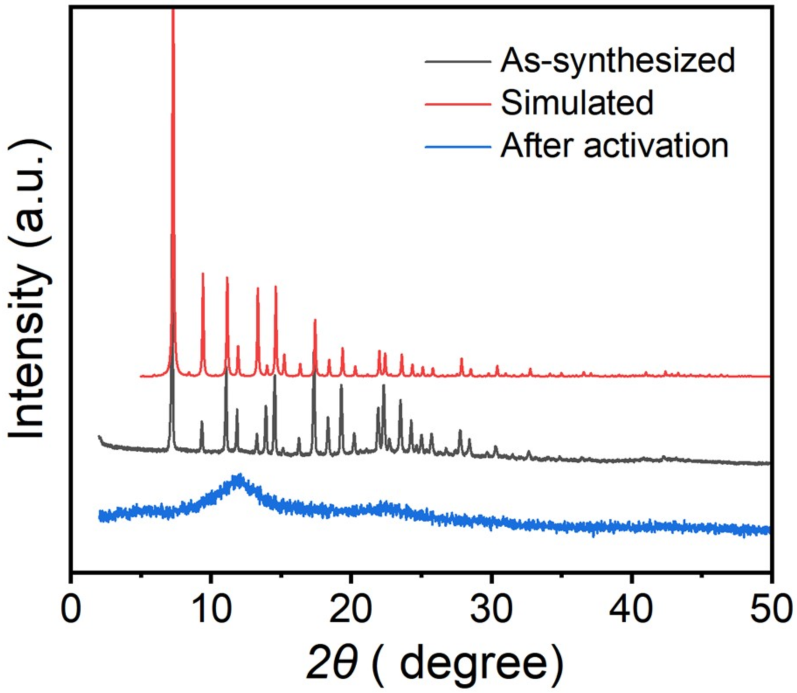


Fig. S9 Powder X-ray diffraction patterns for **NDI-Δ-CH_2Cl_2**

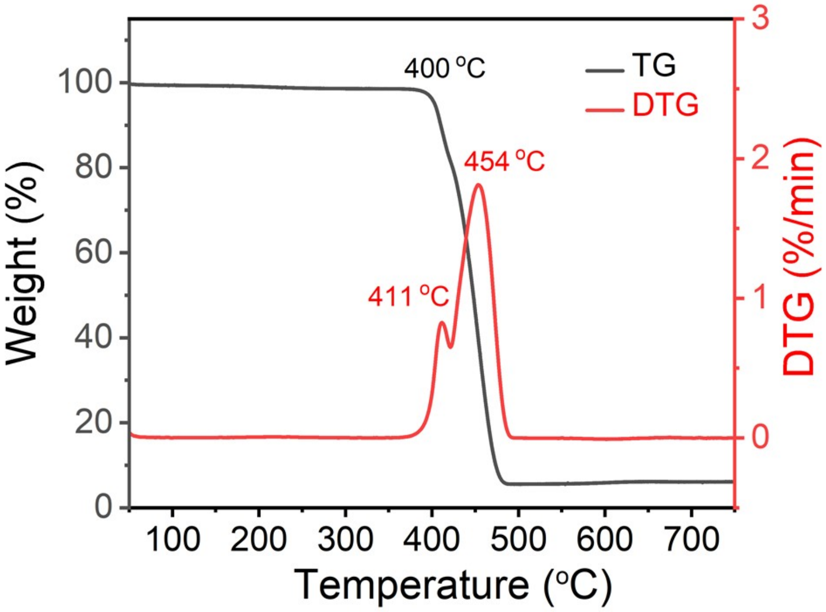


Fig. S10 TGA analysis for **NDI-Δ-CH_2Cl_2** after activation.

S5 The K-points settings by the "SupercellFolding" method in DFTB+

We used the SupercellFolding method in DFTB+ to set the k-points.[3] This method generates a sampling set containing all the special k-points in the Brillouin zone related to points that would occur in an enlarged supercell that repeats the current unit cell. it is set using the following format:

l_1 0 0
0 l_2 0
0 0 l_3
s_1 s_2 s_3

The l_1, l_2, l_3 indicate the uniform extension of lattice, and s_1, s_2, s_3 specify the point in the Brillouin-zone of the super lattice, in which the folding should occur. For cubic lattice, if l_i is odd, s_i is 0.0, and if l_i is even, s_i is 0.5. If one of the cell angles smaller or equal to 75º or greater or equal to 105º, the s_3 is set with the same way that described above. The s_1 and s_2 set to be 0.0 independent of the values of l_1 and l_2.

To set l_i, two criteria must match. The length of the supercell should be close and the shortest length of the supercell should be greater than 12. Here are some examples for the setting of

SupercellFolding method:

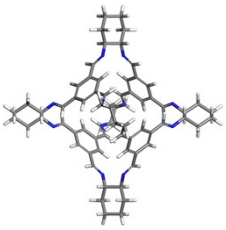
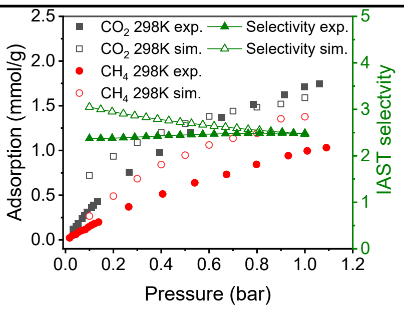
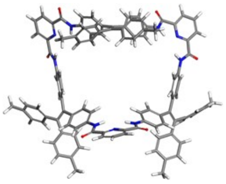
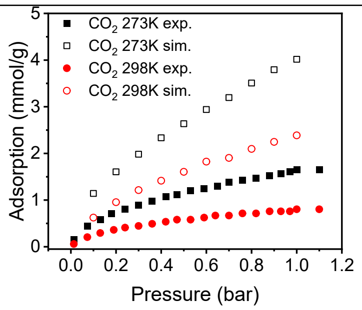
Table S3 Example of SupercellFolding method setting.

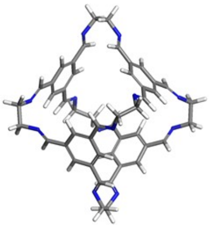
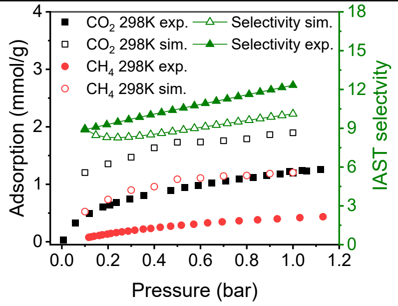
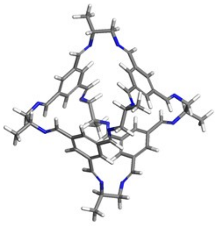
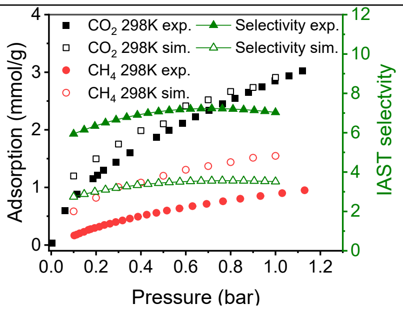
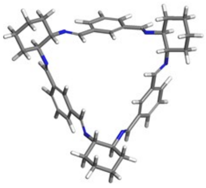
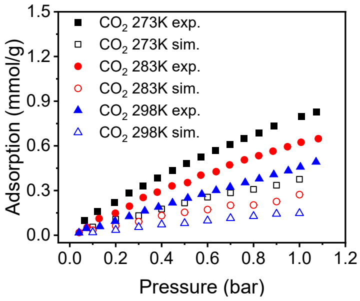
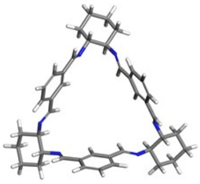
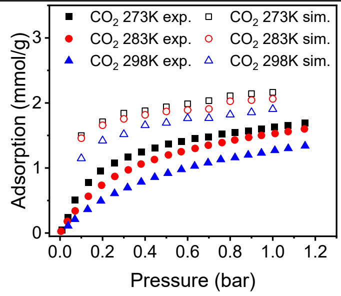
Structure Refcodes	Cell Parameters	$l_1, l_2, l_3, s_1, s_2, s_3$
AAGAGG10	9.274, 24.357, 9.168, 90, 90, 90	3, 1, 3, 0.0, 0.0, 0.0
EVUZED	18.734, 18.734, 18.405, 90, 90, 120	1, 1, 1, 0.0, 0.0, 0.0
PAHXIH01	40.809, 10.266, 19.311, 90, 117.97, 90	1, 4, 2, 0.0, 0.0, 0.5

S6 Comparison between simulated and experimental adsorption isotherms

The simulated isotherms for both CO_2 and CH_4 showed good agreement in shape with the experimental adsorption isotherms. For FOXLAG and PUDXAO, the CO_2 adsorption quantities were well reproduced. The adsorption mechanisms of these materials were captured by the GCMC simulations. The IAST selectivity ranking (FOXLAG < PUDXAO < PUDWUH), derived from both experimental and simulated isotherms, was also consistent. This suggests that GCMC simulations based on these materials can effectively guide the identification of promising candidates. The other three materials (MESGOJ, UTUJIF, UTUJUR) did not have experimental CH_4 adsorption isotherms in the original literature, so they were omitted from the IAST calculations.

Table S4 Comparison between simulated and experimental adsorption isotherms

Refcodes	Molecules	Adsorption isotherms	Ref.
FOXLAG*			4
MESGOJ			5

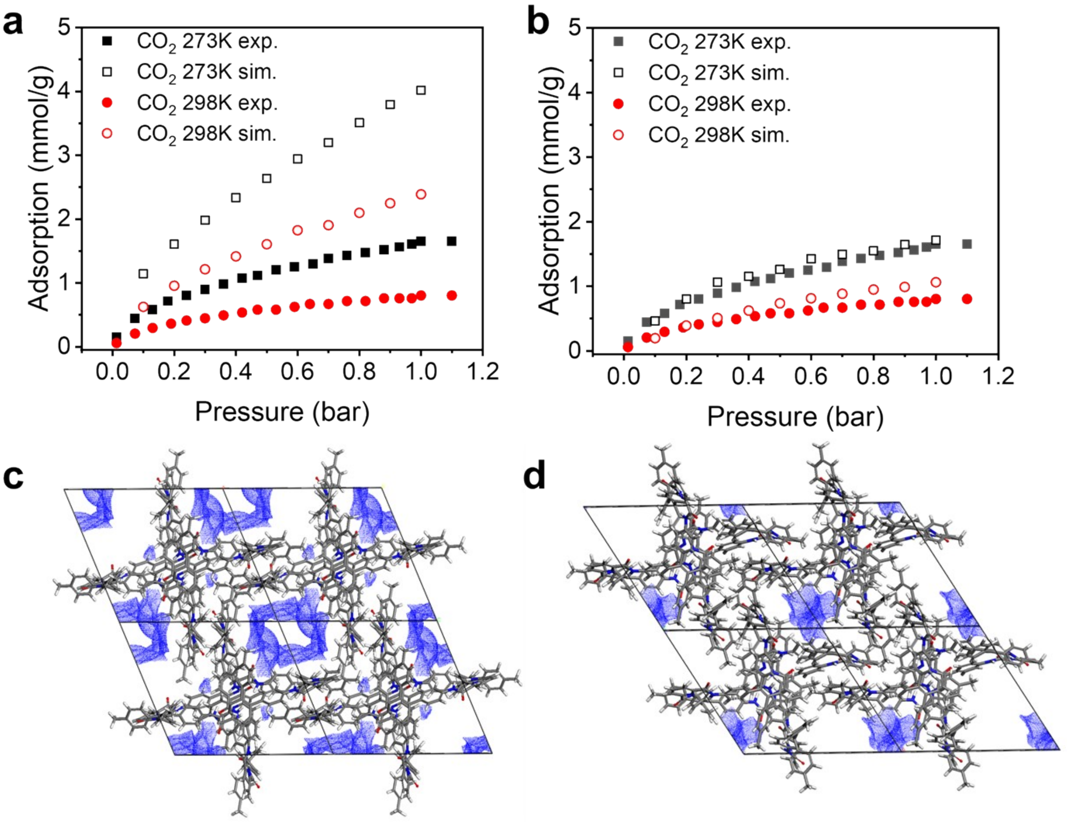
PUDWUH*			6
PUDXAO*			6
UTUJIF			7
UTUJUR			7

* The Ideal Adsorbed Solution Theory (IAST) calculations were performed using pyGAPS[8].

It is worth noting that GCMC simulations were performed on idealized and rigid structures from the database, they cannot precisely reproduce the exact adsorption quantities of all materials due to the loss of crystallinity after activation and the inherent flexibility of molecular materials.

The simulated and experimental isotherms of MESGOJ exhibit a relatively large discrepancy (Fig. S11a). After carefully reviewing the original paper, we suspect that the loss of crystallinity during activation may have compromised the experimental adsorption quantity. The BET surface area of this material after activation at 120 °C for 10 hours under vacuum was reported as 347 m²/g.[5] However, simply removing the solvent molecule from the structure deposited in the CCDC and calculating the surface area using nitrogen as the probe resulted in a value of 2133.87 m²/g. The structure then underwent two stages of optimization before being deposited in MCD, yielding a surface area of 1123.76 m²/g. Since the PXRD pattern of the activated structure was not reported in the original paper, we cannot directly assess its crystallinity.

To further demonstrate this phenomenon, a 100 ps NPT molecular dynamics (MD) simulation initiate from the MESGOJ structure in MCD was conducted at 120 °C and 1 bar using the Forcite module in *Materials Studio*[1]. The energy was described using the COMPASSIII force field, with a time step set to 0.5 fs. The final frame of the MD simulation was used for geometry optimization with the same force field, allowing the cell parameters to change during the optimization. After this process, the surface area of the material decreased to 413.495 m²/g, bringing it closer to the experimental value of 347 m²/g. We further conducted GCMC simulations based on the structure, using the same settings as in the high-throughput GCMC simulation. As shown in Fig. S11 b, the simulated isotherms at both temperatures exhibited good agreement with the experimental data. Additionally, we observed that the extrinsic 1D channel shrank after 100 ps of MD simulation (Fig. S11c, d).



a, b, c = 19.25, 19.94,18.05
α, β, γ = 104.87, 111.62, 89.76

a, b, c = 18.64, 18.28,18.29
α, β, γ = 108.76, 119.48, 91.90

Fig. S11 Comparison of simulated and experimental isotherms based on the MESGOJ structure in (a) the MCD structure and (b) the MCD structure after mimicking the activation process. Accessible surface area (blue dots) computed using *Materials Studio* with nitrogen as the probe for (c) the MESGOJ structure in MCD and (d) the MCD structure after mimicking the activation process.

In this case, the GCMC simulation based on the MCD structure can indicate the potential of this material. The experimental CO_2 capacity may be significantly underestimated, and higher adsorption quantities are expected under optimal activation conditions.

The limitations of the rigid-host assumption in GCMC simulations for accurately reproducing the adsorption quantities of molecular materials, due to their inherent flexibility, have been widely discussed in previous studies[9–11]. UTUJUR exhibited hysteresis during CO_2 desorption, while the CO_2 adsorption isotherms of UTUJIF displayed two successive plateaus,[7] both indicating the impact of host flexibility on adsorption behavior. To account for this effect, we employed a flexible-host hybrid GCMC/MD simulation approach implemented in RASPA[12]. The probability of an MD move was set to 2.4% during the hybrid GCMC/MD simulation, with each MD move consisting of 5 timesteps of 0.5 fs. The General Amber Force Field (GAFF)[13] was used to describe bond stretching, angle bending, torsion, improper torsion, and nonbonding interactions between host atoms. The nonbonding interaction settings between the host and guest molecules were kept consistent with those used in the high-throughput GCMC simulations. For both UTUJUR and UTUJIF, the simulated isotherms computed using the flexible-host assumption show improved agreement with experimental data (Fig. S12).

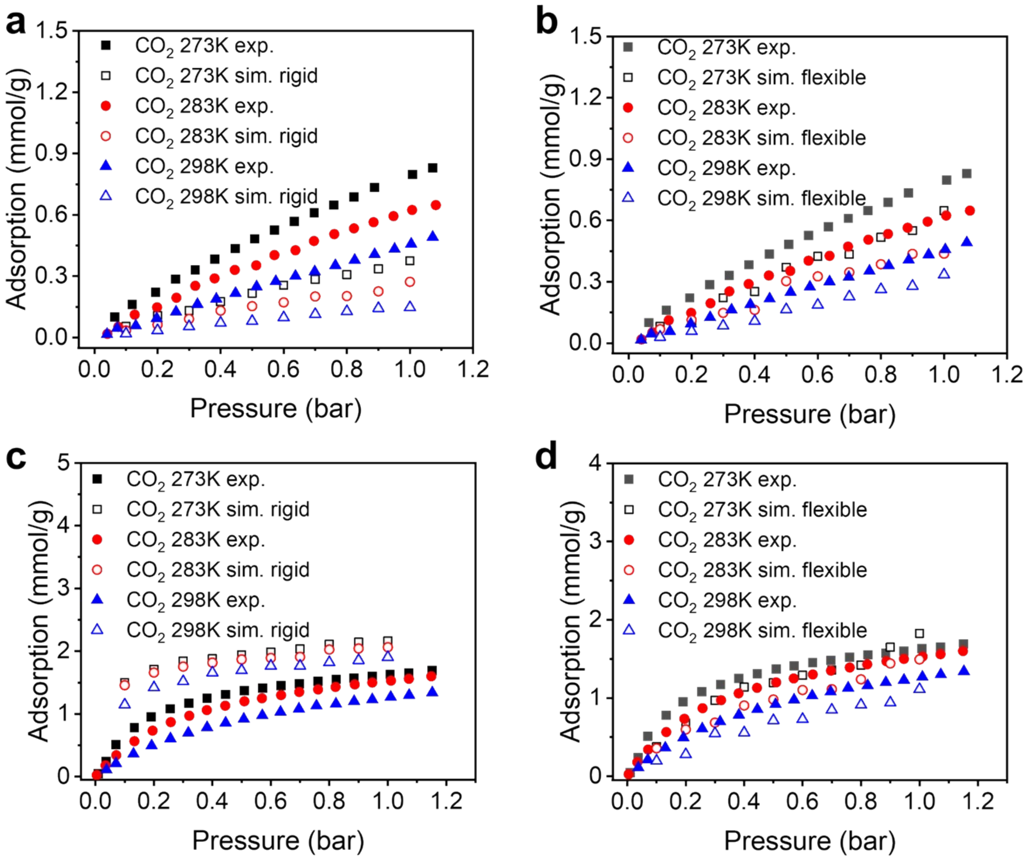


Fig. S12 Comparison of experimental isotherms for UTUJIF (a, c) and UTUJUR (b, d) with both rigid-host assumption GCMC simulations and flexible-host assumption hybrid GCMC/MD simulations.

However, since hybrid GCMC/MD simulations cannot run in parallel, with equilibrium and production cycles set to 100,000 and 200,000, respectively, a single simulation can take more than 10 days to complete, making it impractical for high-throughput screening. As mentioned earlier, the isotherms computed by GCMC showed good agreement in trend with experimental isotherms, and the IAST selectivity was well reproduced. Therefore, GCMC simulations were used for high-throughput screening to identify promising materials.

S7 Structure-property relationships between Selectivity/CO_2 adsorption quantities with common geometrical parameters

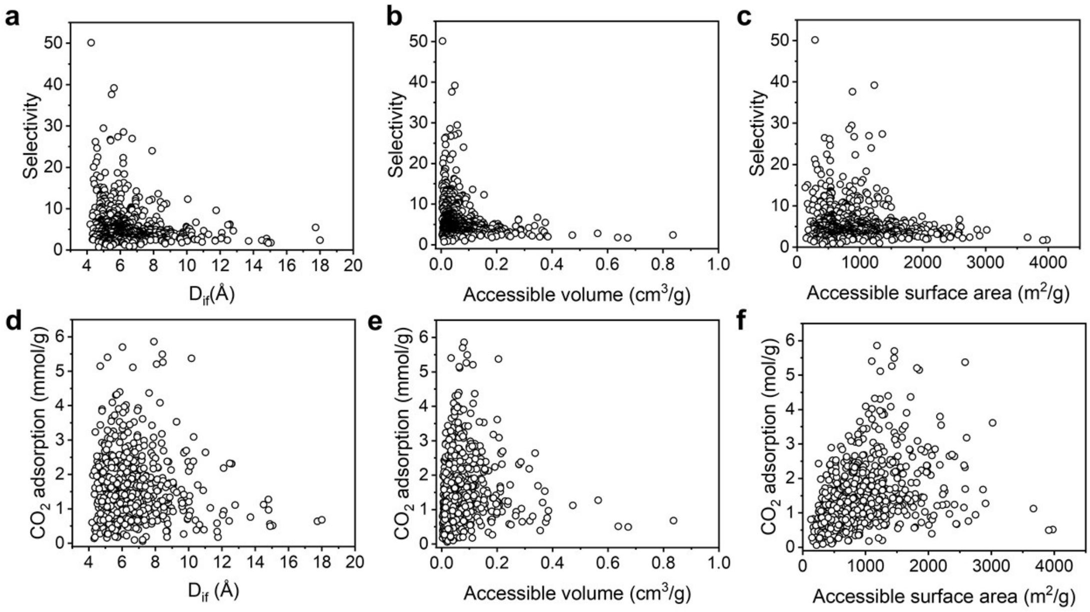


Fig. S13 Relationships of CO_2/CH_4 selectivity (a–c) and CO_2 adsorption quantities (d–f) with the diameter of the largest inclusion sphere along the free path (D_{if}) (a, d), accessible volume (b, e), and accessible surface area (c, f).

S8 Comparison between simulated and experimental CO_2 adsorption isotherms for NDI-Δ-CH_2Cl_2

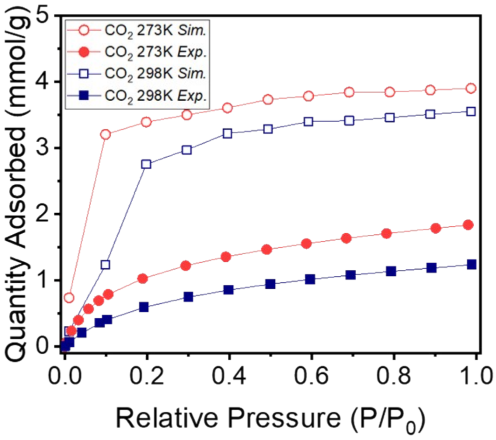


Fig. S14 Comparison between simulated and experimental CO_2 Adsorption Isotherms for **NDI-Δ-CH_2Cl_2**.

References

1 Materials Studio (version 2020) BIOVIA: Dassault Systèmes, San Diego 2020.

2 S. T. Schneebeli, M. Frasconi, Z. Liu, Y. Wu, D. M. Gardner, N. L. Strutt, C. Cheng, R. Carmieli, M. R. Wasielewski and J. F. Stoddart, *Angew. Chem., Int. Ed*, 2013, **52**, 13100–13104.

3 B. Hourahine, B. Aradi, V. Blum, F. Bonafé, A. Buccheri, C. Camacho, C. Cevallos, M. Y. Deshaye, T. Dumitric, A. Dominguez, S. Ehlert, M. Elstner, T. Van Der Heide, J. Hermann, S. Irle, J. J. Kranz, C. Köhler, T. Kowalczyk, T. Kubař, I. S. Lee, V. Lutsker, R. J. Maurer, S. K. Min, I. Mitchell, C. Negre, T. A. Niehaus, A. M. N. Niklasson, A. J. Page, A. Pecchia, G. Penazzi, M. P. Persson, J. Řezáč, C. G. Sánchez, M. Sternberg, M. Stöhr, F. Stuckenberg, A. Tkatchenko, V. W. Z. Yu and T. Frauenheim, *J. Chem. Phys.*, DOI:10.1063/1.5143190.

4 K. Krishnan, J. M. Crawford, P. K. Thallapally and M. A. Carreon, *Ind. Eng. Chem. Res.*, 2022, **61**, 10547–10553.

5 J.-B. Xiong, J.-H. Wang, B. Li, C. Zhang, B. Tan and Y.-S. Zheng, *Org. Lett.*, 2018, **20**, 321–324.

6 T. Tozawa, J. T. A. Jones, S. I. Swamy, S. Jiang, D. J. Adams, S. Shakespeare, R. Clowes, D. Bradshaw, T. Hasell, S. Y. Chong, C. Tang, S. Thompson, J. Parker, A. Trewin, J. Bacsa, A. M. Z. Slawin, A. Steiner and A. I. Cooper, *Nat. Mater.*, 2009, **8**, 973–978.

7 D. He, R. Clowes, M. A. Little, M. Liu and A. I. Cooper, *Chem. Commun.*, 2021, **57**, 6141–6144.

8 P. Iacomi and P. L. Llewellyn, *Adsorption*, 2019, **25**, 1533–1542.

9 L. Chen, P. S. Reiss, S. Y. Chong, D. Holden, K. E. Jelfs, T. Hasell, M. A. Little, A. Kewley, M. E. Briggs, A. Stephenson, K. M. Thomas, J. A. Armstrong, J. Bell, J. Busto, R. Noel, J. Liu, D. M. Strachan, P. K. Thallapally and A. I. Cooper, *Nat. Mater*, 2014, **13**, 954–960.

10 S. Yang, L. Chen, D. Holden, R. Wang, Y. Cheng, M. Wells, A. I. Cooper and L. Ding, *Z. Kristallogr. Cryst. Mater.*, 2019, **234**, 547–555.

11 L. Yang, H. Zheng, W. Mei, Z. Wang, Z. Zhang and K. Yu, *J. Phys. Chem. C*, 2024, **128**, 10997–11005.

12 D. Dubbeldam, S. Calero, D. E. Ellis and R. Q. Snurr, *Mol. Simul.*, 2016, **42**, 81–101.

13 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.