Electronic Supplementary Material (ESI) for Materials Horizons.

Supporting Information: Room-temperature stacking disorder in layered covalent-organic frameworks from machine-learning force fields

Ju Huang,[†] Seung-Jae Shin,[‡] Kasper Tolborg,[†] Alex M. Ganose,[¶] Gabriel Krenzer,[†] and Aron Walsh[*,†,§]

[†]*Department of Materials, Imperial College London, London SW7 2AZ, UK*

[‡]*Department of Materials Science and Engineering, Yonsei University, Seoul 03722, Republic of Korea*

[¶]*Department of Chemistry, Imperial College London, London W12 0BZ, UK*

[§]*Department of Physics, Ewha Womans University, Seoul 03760, Korea*

E-mail: a.walsh@imperial.ac.uk

Tab. S1. Optimised lattice parameters of **a.** eclipsed Tp-Azo, **b.** inclined Tp-Azo, **c.** eclipsed DAAQ-TFP and **d.** inclined DAAQ-TFP structures. 'E' and 'I' are the abbreviation of eclipsed and inclined, respectively.

Lattice parameters						
System	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
a. E-Tp-Azo	33.28	33.36	3.23	90	90	120.17
b. I-Tp-Azo	33.51	33.59	4.19	61.51	128.65	121.14
c. E-DAAQ-TFP	30.28	30.38	3.33	90	90	60.19
d. I-DAAQ-TFP	30.34	30.55	3.86	67.86	57.10	60.18

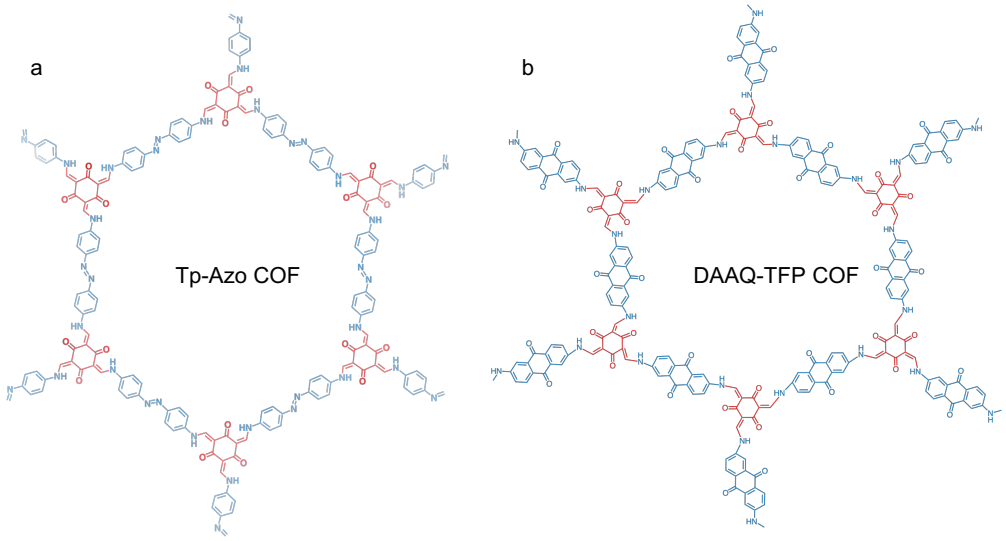

Fig. S1. Schematic representation of the skeletal structures of Tp-Azo[S1] and DAAQ-TFP[S2] COFs.

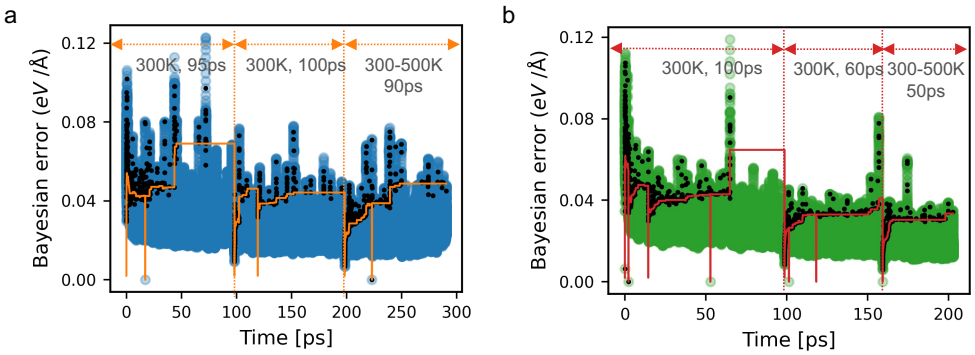

Fig. S2. Maximum Bayesian errors of the force for per atom of **a.** eclipsed and **b.** inclined DAAQ-TFP in the whole training process. The black dots represent where *ab initio* calculations performed. The orange and red lines in each plot correspond to the criterion of threshold in the training. The dips in the threshold curves correspond to restarts of the training runs, where the threshold was reset to its default value. The temperature and the time in the plot correspond to the training temperature and time in the first training, retraining and heating training.

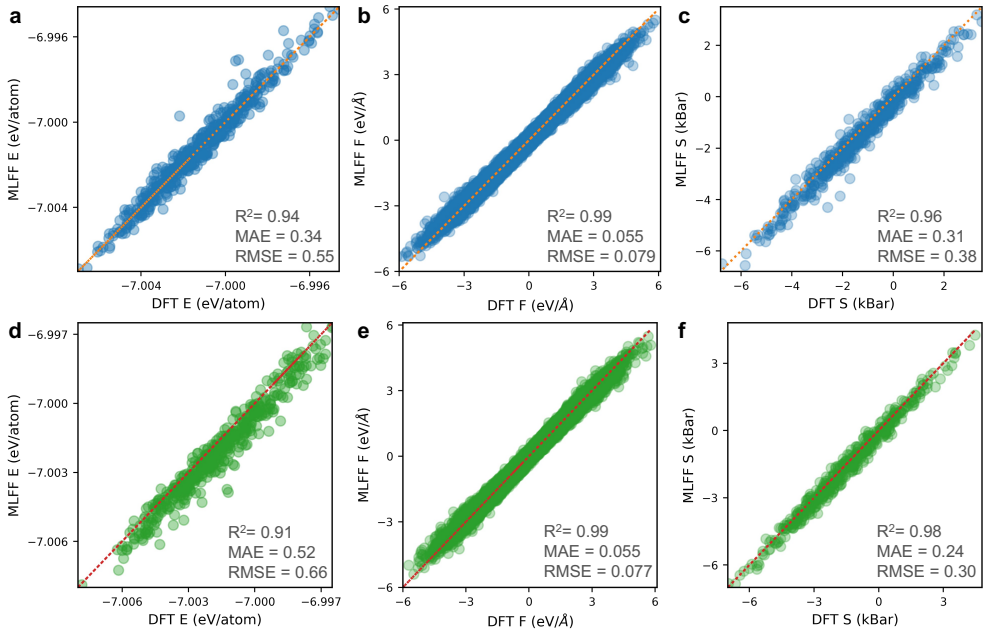

Fig. S3. Comparison of energy (E), force (F) and stress (S) of all sampled configurations calculated from MLFF and DFT in **a-c.** eclipsed and **d-f.** inclined Tp-Azo. The orange and red dotted line s correspond to lines of $x = y$. The accuracy score (R^2), mean absolute error (MAE), and root mean squared error (RMSE) of energy, force and stress are shown at the right bottom of each plot.

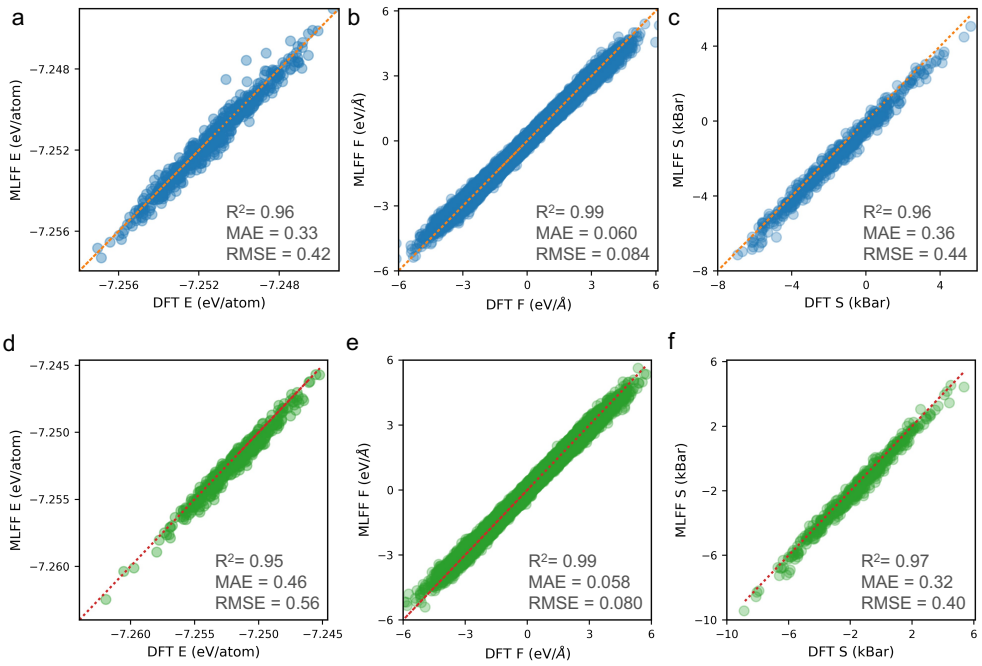

Fig. S4. The energies, forces and stress of all sampled configurations calculated from MLFF and DFT in **a-c.** eclipsed and **d-f.** inclined DAAQ-TFP. The red and orange dotted lines correspond to lines of $x = y$. The R^2, MAE, and RMSE of energy, force and stress are shown at the right bottom of each plot.

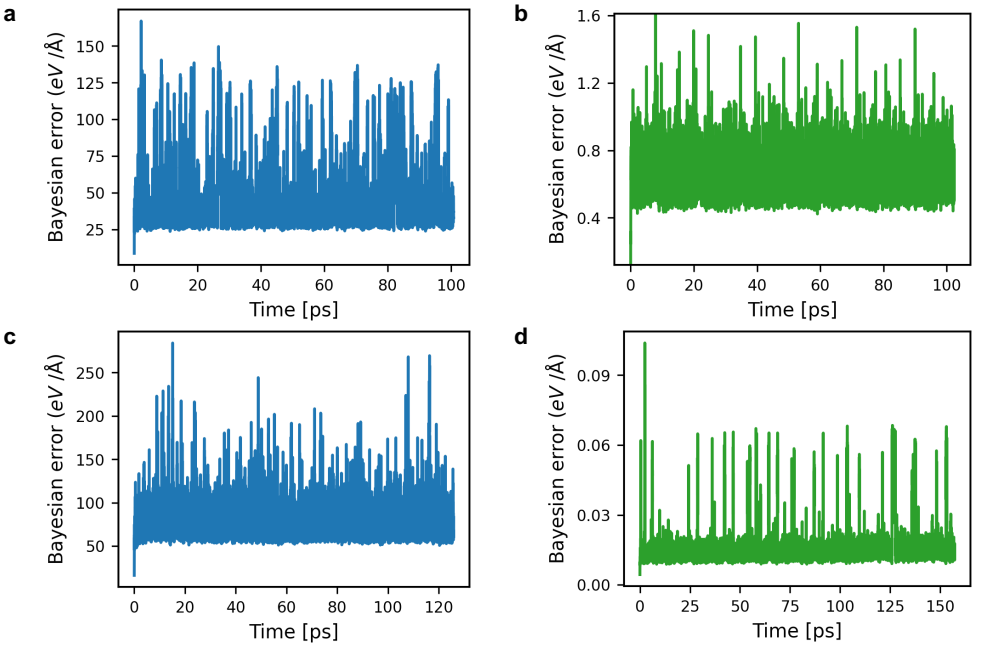

Fig. S5. The maximum Bayesian errors of force for per atom of **a.** eclipsed Tp-Azo, b. inclined Tp-Azo, **c.** eclipsed DAAQ-TFP and **d.** inclined DAAQ-TFP in MLFF-MD simulation.

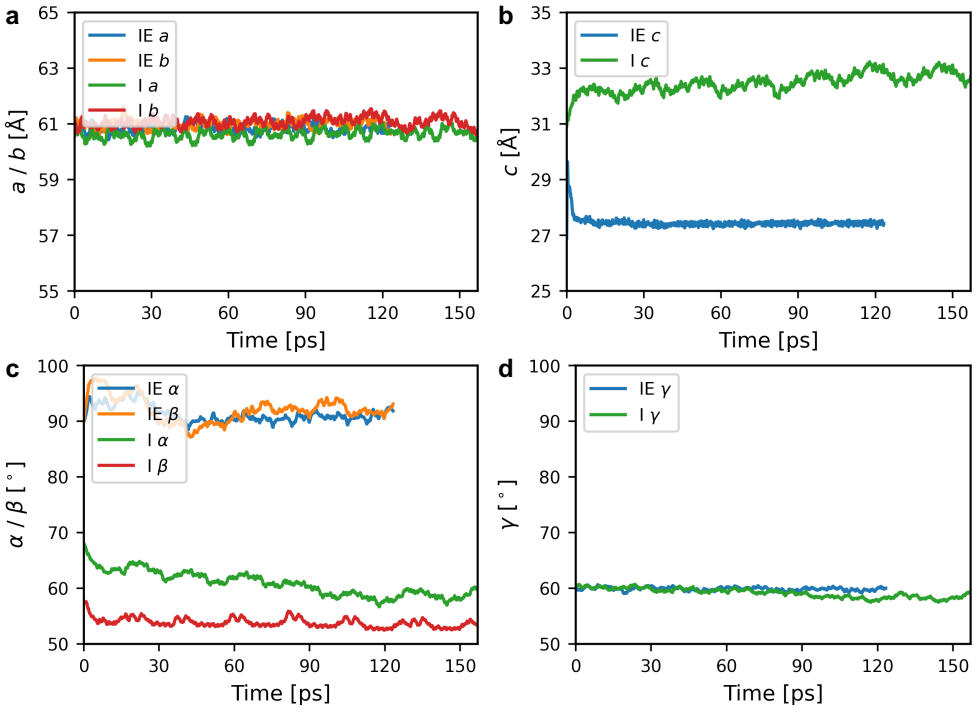

Fig. S6. Evolution of the lattice parameters of **a.** a and b, **b.** c, **c.** α and β, **d.** γ in initially eclipsed and inclined DAAQ-TFP in the trajectory. 'IE' and 'I' correspond to initially eclipsed and inclined configurations, respectively.

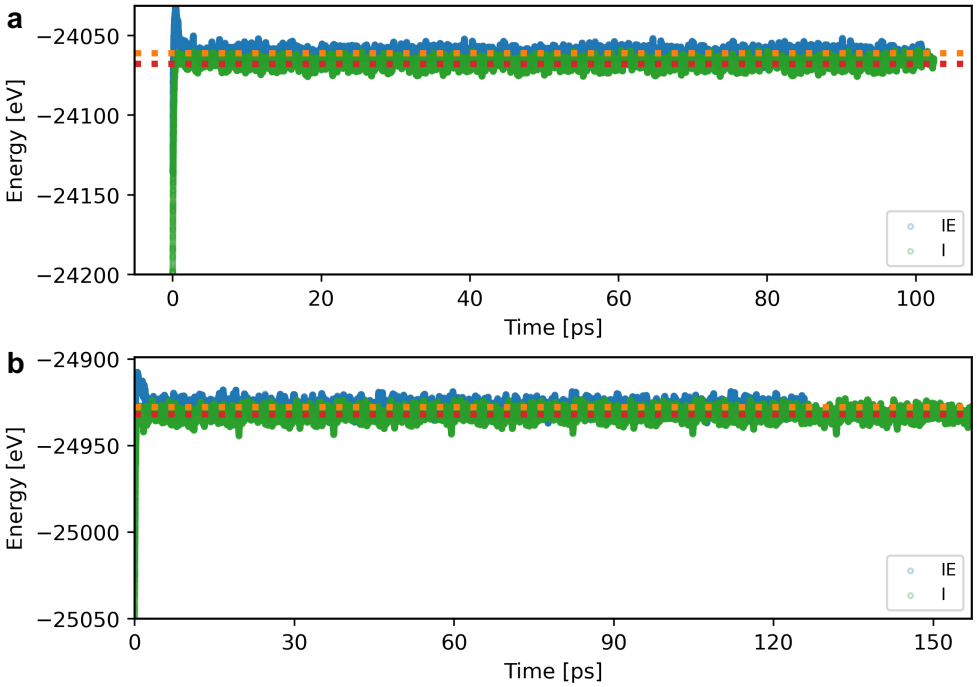

Fig. S7. The total internal energies of **a.** initially eclipsed and inclined Tp-Azo and **b.** initially eclipsed and inclined DAAQ-TFP along the MD. The dotted orange and red lines are the average energy lines of initially eclipsed and inclined COFs, respectively. 'IE' and 'I' correspond to the initially eclipsed and inclined configurations, respectively.

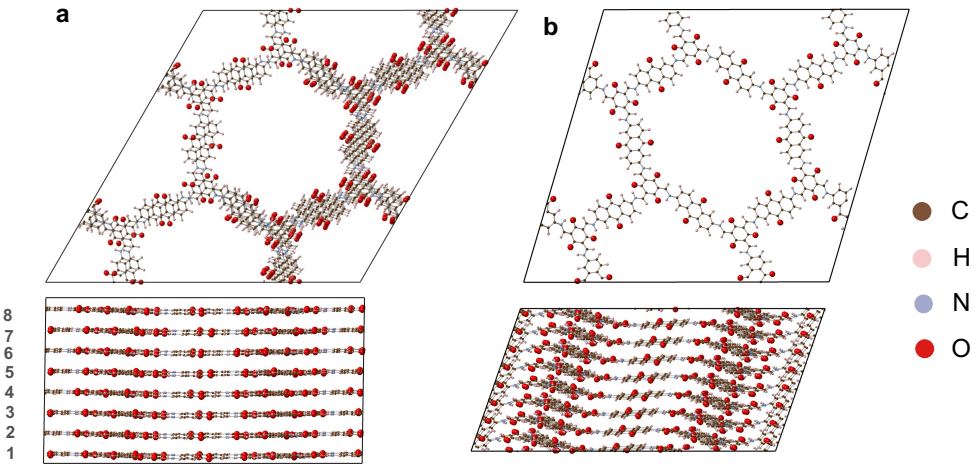

Fig. S8. Layered supercells from MLFF relaxation calculation of **a.** zigzag DAAQ-TFP, **b.** inclined DAAQ-TFP viewed from c direction (top) and ab plane (bottom). The structures are viewed along the c direction (top panel) and ab plane (bottom panel). Each layer is numbered from 1 to 8.

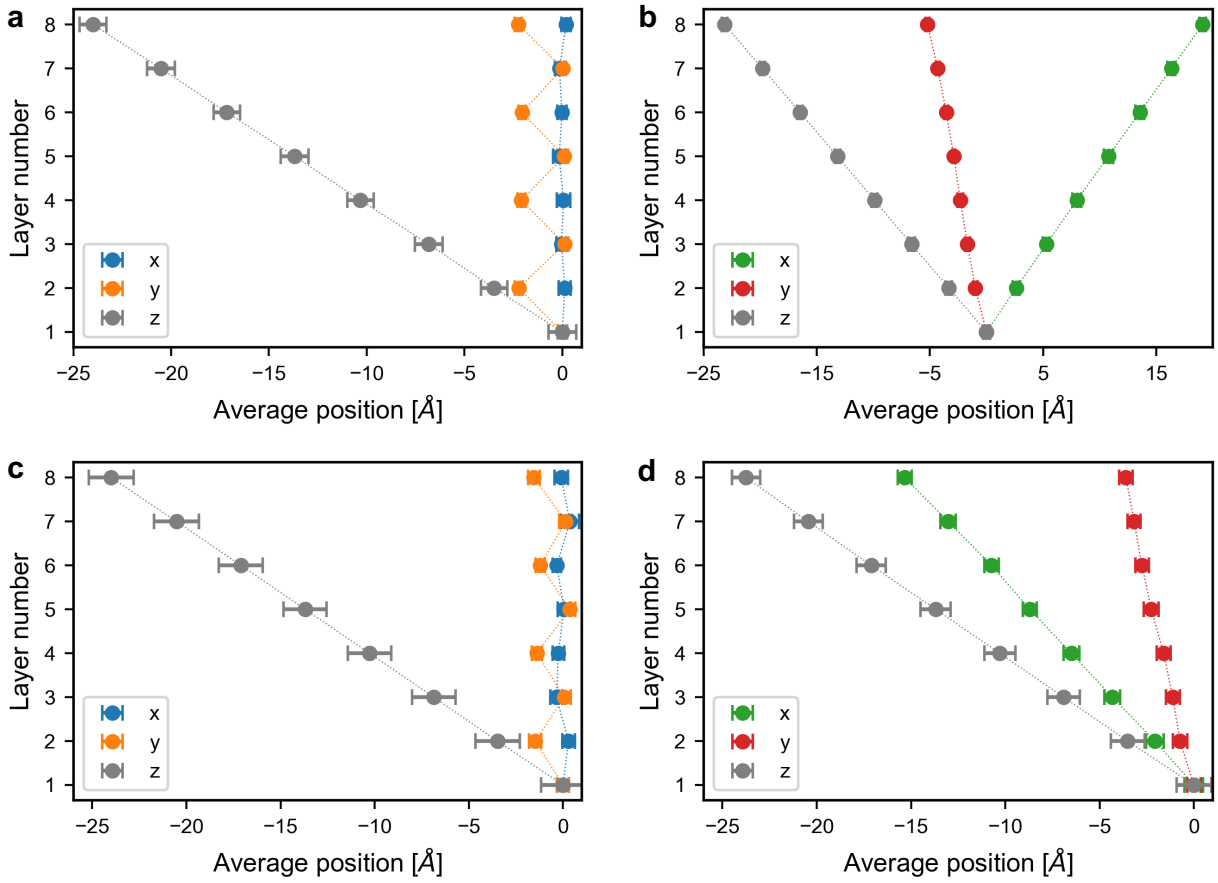

Fig. S9. Averaged positions of layers in x, y and z directions of **a.** zigzag Tp-Azo, **b.** inclined Tp-Azo, **c.** zigzag DAAQ-TFP, **d.** inclined DAAQ-TFP from the MD trajectory after the structures have been stable. The numbers from 1 to 8 corresponds to the number that labeled in Fig. 5 and Fig. S8

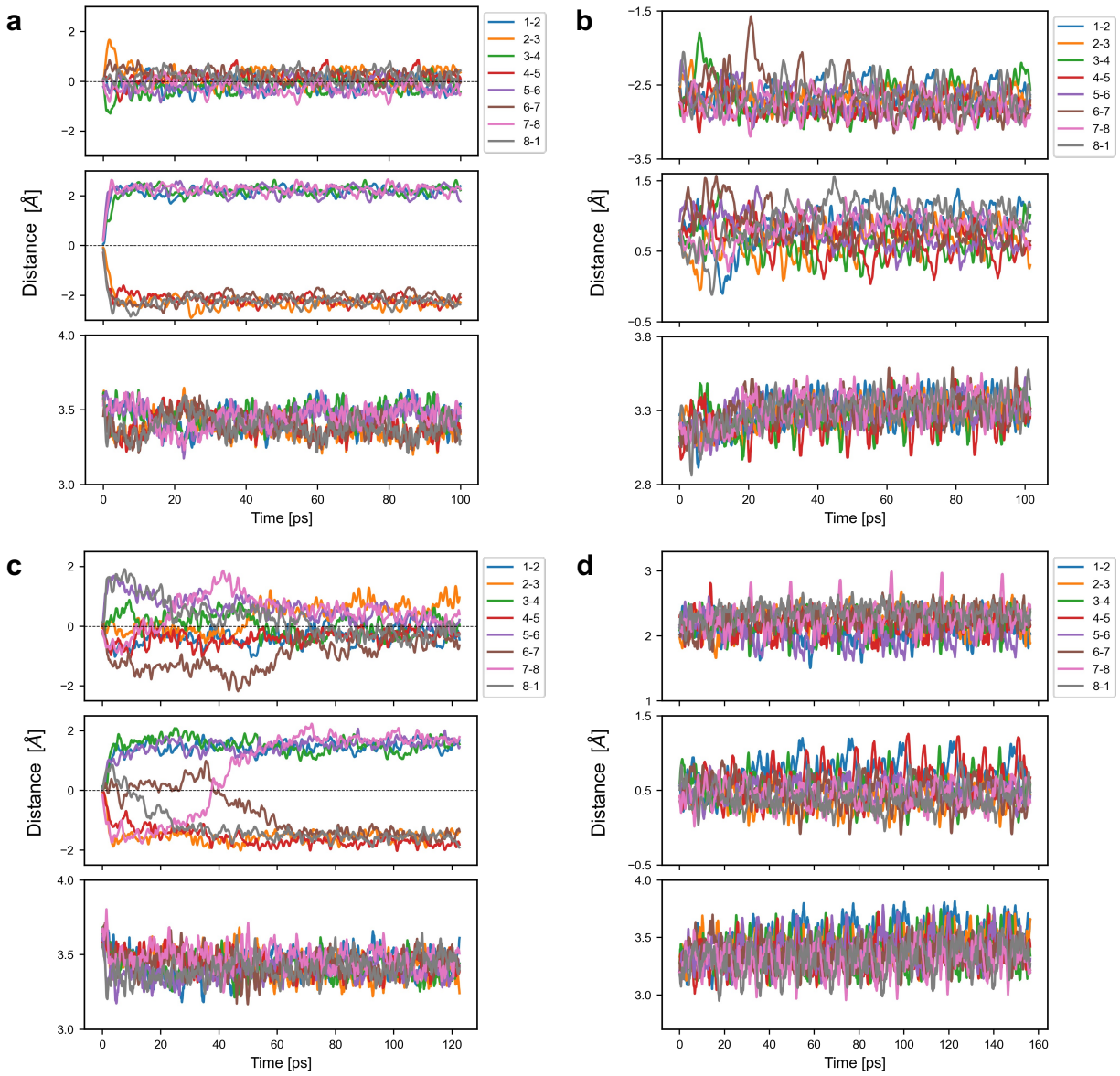

Fig. S10. Distances in x, y and z between adjacent layers in different stacking modes of **a.** zigzag Tp-Azo, **b.** inclined Tp-Azo, **c.** zigzag DAAQ-TFP and **d.** inclined DAAQ-TFP corresponding to Fig. S9. In each subplot, it shows the distances between adjacent layers in the order of x, y and z. The black dashed line corresponds to $y = 0$. The layer numbers correspond to the numbers in Fig. 5 and Fig. S8.

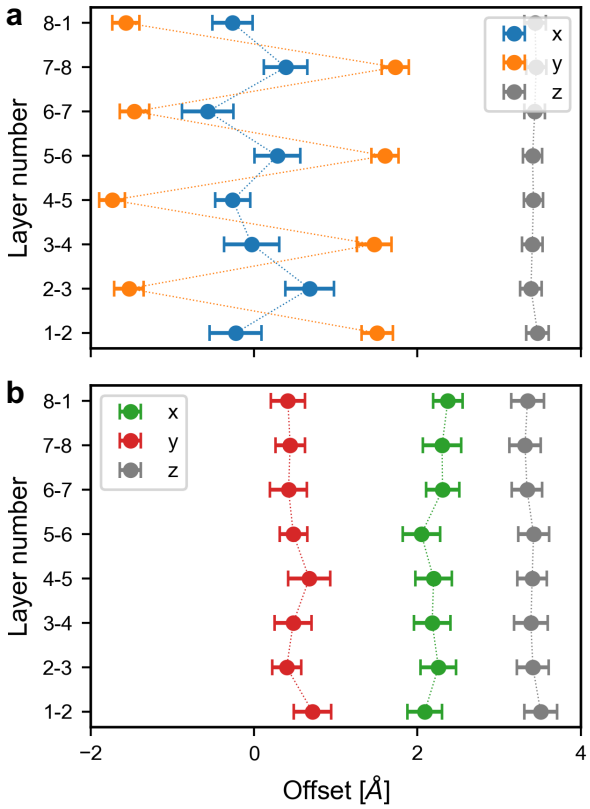

Fig. S11. Offsets in x, y, z directions between adjacent layers in zigzag (**a**) and inclined (**b**) DAAQ-TFP. The layer numbers correspond to the numbers in Fig. S8.

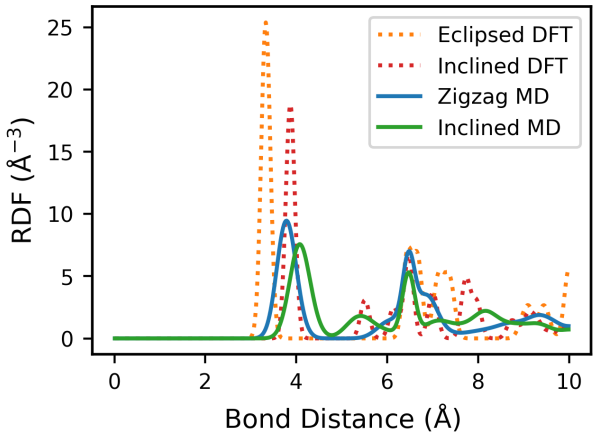

Fig. S12. RDF of N-N bonds in initially eclipsed, inclined and zigzag DAAQ-TFP. The orange and red dotted lines represent the RDF of the initial eclipsed and inclined DAAQ-TFP configurations. The blue and green solid lines are the average RDF in the trajectory after the structures have stabilized.

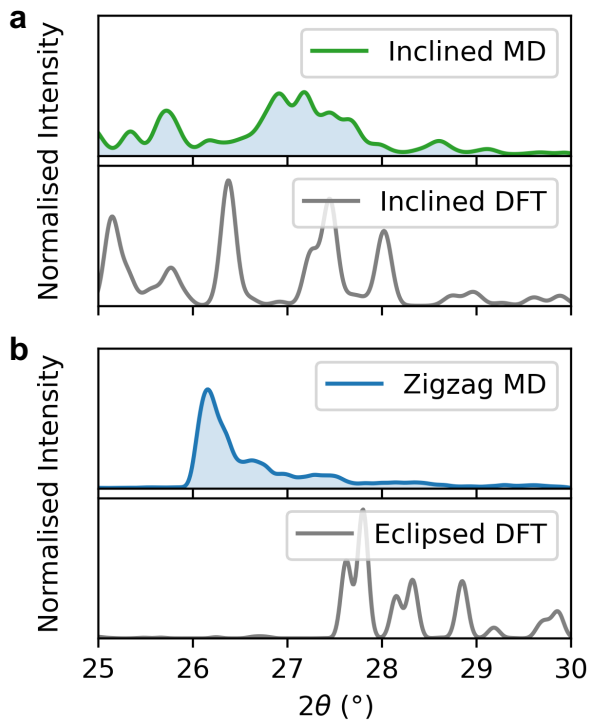

Fig. S13. Simulated average diffraction patterns of **a** initially eclipsed Tp-Azo, **b** inclined Tp-Azo were generated from pymatgen package, by selecting structures every 5 ps in the MD trajectory. The diffraction patterns are zoomed into where the 2θ range of $25 \sim 30°$ from Fig. 8. The bottom diffraction pattern in each plot is from the initial structure of each COF.

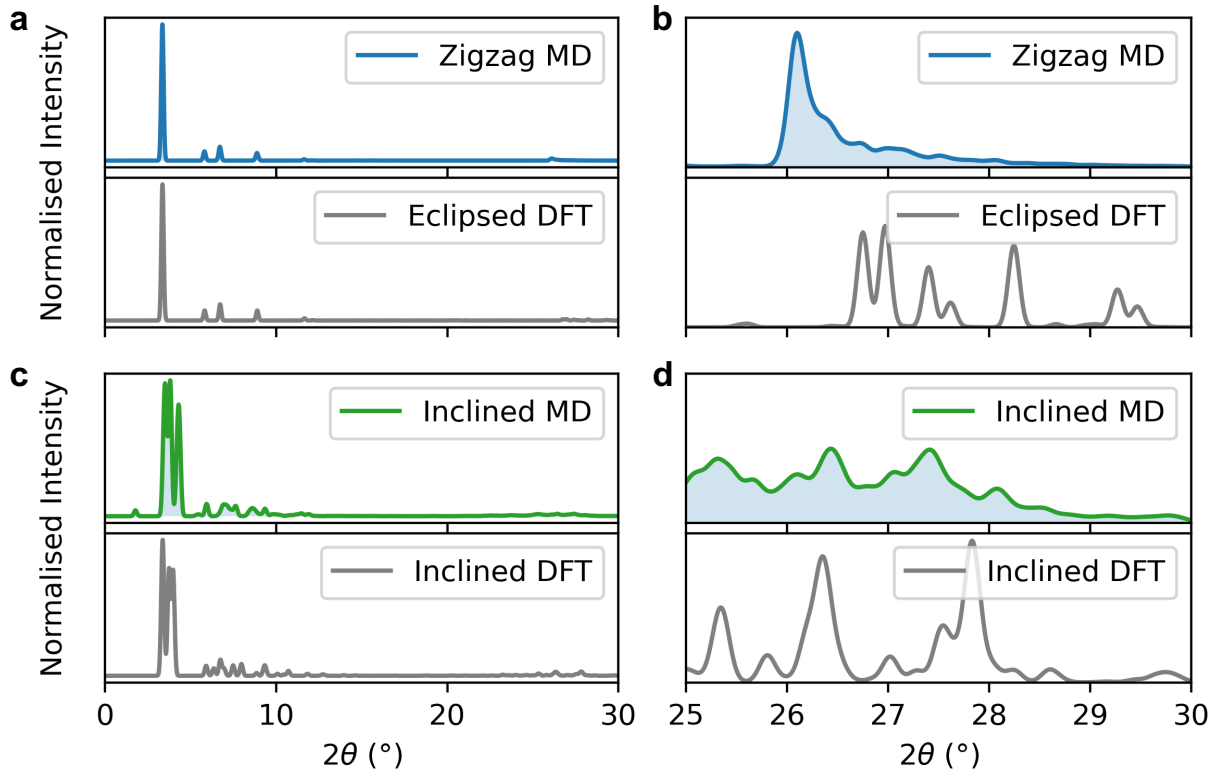

Fig. S14. Simulated average diffraction patterns of **a, b** initially eclipsed Tp-Azo, **c, d** inclined DAAQ-TFP were generated from pymatgen package, by selecting structures every 5 ps in the MD trajectory. **a,c** cover the 2θ range of $0 \sim 30°$. **b,d** are zoomed into where the 2θ range of $25 \sim 30°$. The bottom diffraction pattern in each plot is from the initial structure of each COF.

The peaks of the zigzag Tp-Azo structure are at $2\theta = \sim 3.05°$ (200), $\sim 5.26°$ (220), $\sim 6.10°$ (400), $\sim 10.97°$ (420), and $\sim 26.00°$ (008), which has a good agreement with the positions and shapes of the experimental PXRD.[S1]

In zigzag DAAQ-TFP, 2θ positions are at $\sim 3.35°$, $\sim 5.80°$, $\sim 6.69°$, $\sim 25.98°$, which have a good agreement with experimental PXRD pattern of DAAQ-TFP.[S2]

In inclined Tp-Azo, peaks are at 2θ of $\sim 3.51°$ (020), $\sim 4.27°$ (2-20), $\sim 5.66°$ (220), $\sim 6.95°$ (2-40), $\sim 29.85°$ (008).

In inclined DAAQ-TFP, 2θ are at $\sim 1.77°$ (010), $\sim 3.53°$ (020), $\sim 3.83°$ (200), $\sim 4.34°$ (220), $\sim 5.95°$ (2-20), $\sim 7.07°$ (040), $\sim 8.69°$ (440), $\sim 28.43°$ (008). Only in the diffraction patterns of structure *No.* 3 and 4 of inclined DAAQ-TFP, there is a tiny peak at 2θ of $\sim 1.77°$.

For the simulated zigzag XRD patterns, the (001) reflections (scattering angle $\sim 26°$) aggregate into a broad peak (Fig. S13**a** and Fig. S14**b**). This broadening effect corresponds to the oscillating zigzag offsets between the layers.[S3]

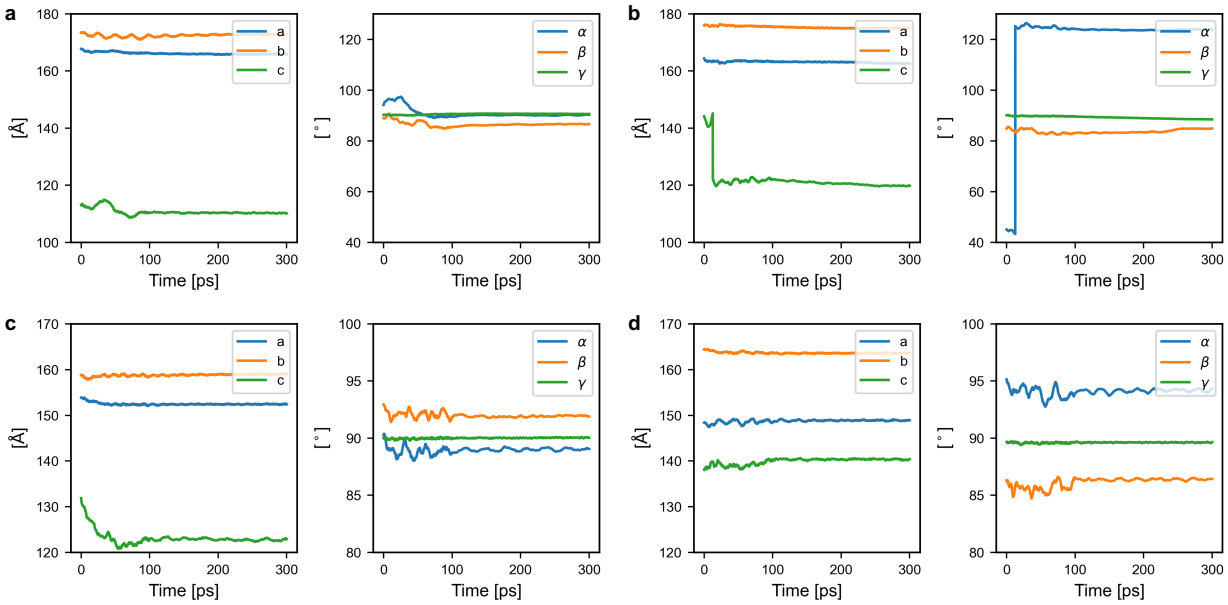

Fig. S15. Evolution of the lattice parameters of a, b, c, α, β, γ in zigzag and inclined Tp-Azo (**a** and **b**) and DAAQ-TFP (**c** and **d**) in the trajectory of the classical force field.

Tab. S2. Equilibrated lattice parameters of supercells of **a.** zigzag Tp-Azo, **b.** inclined Tp-Azo, **c.** zigzag DAAQ-TFP and **d.** inclined DAAQ-TFP structures from classical force field. The initial data is the lattice parameters of supercell from MLFF. The final data is the average values of the trajectory (Fig. S15) from classical force field after equilibrium. There are 48 layers in each supercell. 'Z' and 'I' are represented for the zigzag and inclined initial configurations, respectively. Δ represents the difference between the lattice parameters calculated from classical force field and MLFF.

Lattice parameters							
System		a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
a. Z-Tp-Azo	Initial	157.5	163.7	158.4	90.0	90.0	90.0
	Final	165.9	172.8	110.2	90.3	86.6	90.6
	Δ(%)	5.3	5.6	30.4	0.3	3.8	0.7
b. I-Tp-Azo	Initial	163.6	174.1	203.2	51.0	98.6	89.8
	Final	162.6	175.0	145.9	41.0	84.9	88.4
	Δ(%)	0.6	0.5	28.2	19.7	13.9	1.5
c. Z-DAAQ-TFP	Initial	149.2	155.0	172.8	90.0	90.0	90.0
	Final	152.4	159.0	122.9	89.1	91.9	90.0
	Δ(%)	2.2	2.6	28.9	1.0	2.1	0.0
d. I-DAAQ-TFP	Initial	152.1	157.5	194.0	57.9	79.0	90.2
	Final	148.9	163.6	210.0	42.0	86.4	89.7
	Δ(%)	2.1	3.9	8.2	27.4	9.4	0.6

Tab. S2 exhibits there are small difference between the lattice parameters of a, b, γ of the initial and final supercells in the classical MD. However, the differences related to c axis including c, α, β are much large. These differences come from the exaggerated configurations for both stacking modes of two COFs.

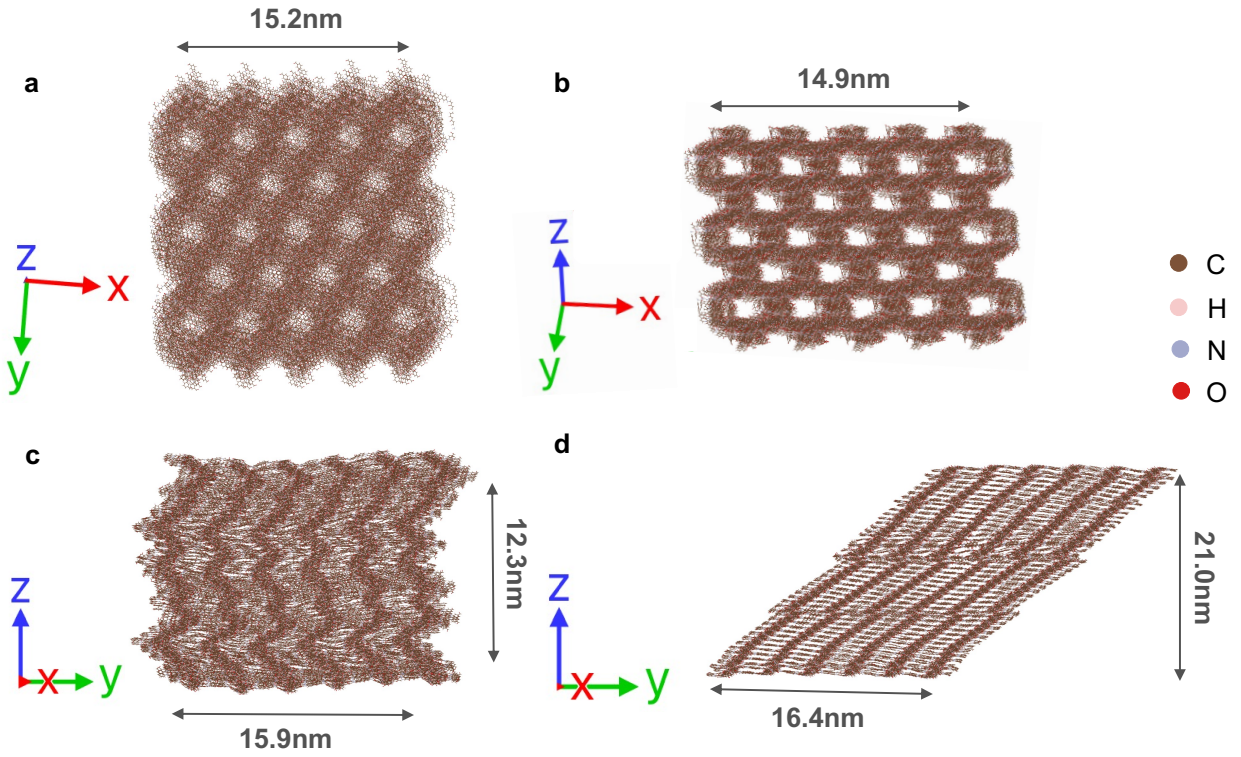

Fig. S16. Snapshots of the supercells of **a,c** zigzag DAAQ-TFP, **b,d** inclined DAAQ-TFP at equilibrium from classical force field.

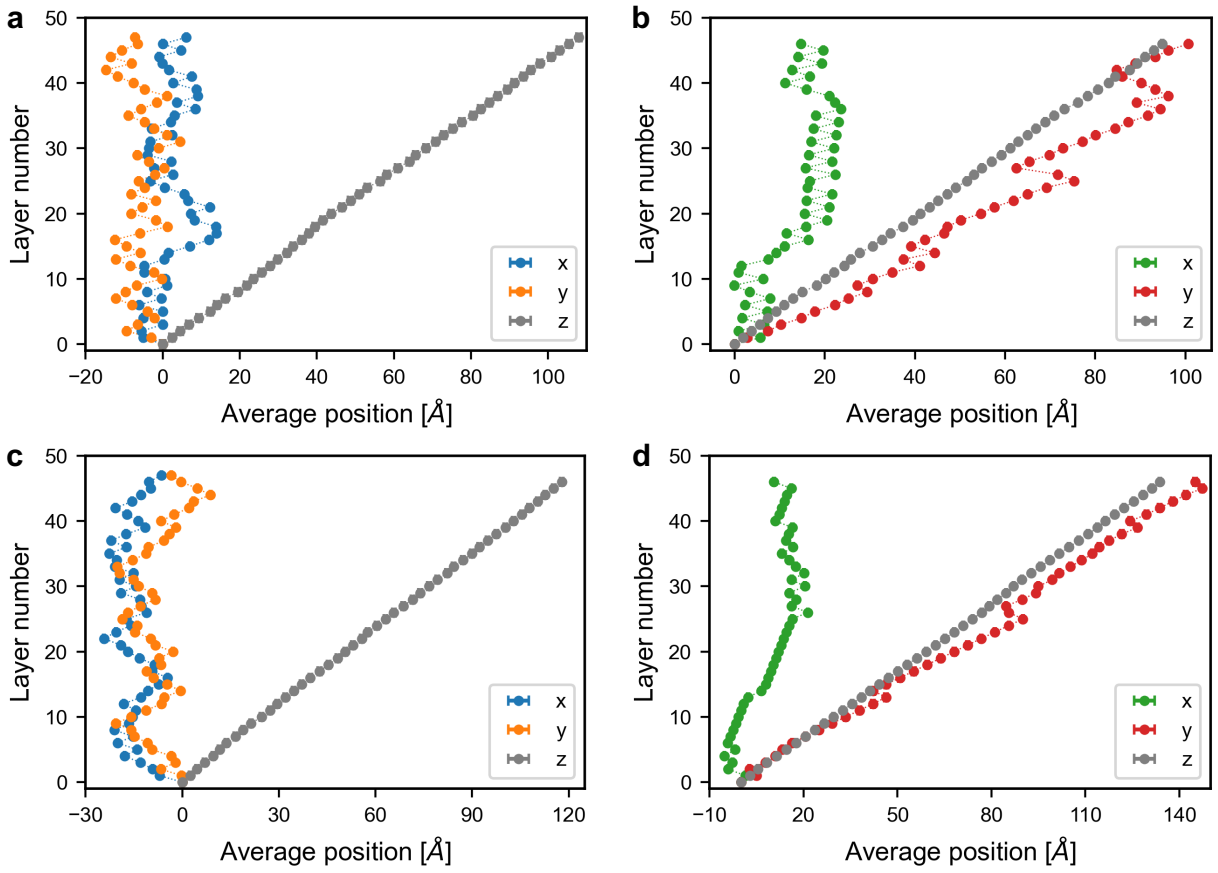

Fig. S17. Averaged positions of layers in x, y and z directions of **a.** zigzag Tp-Azo, **b.** inclined Tp-Azo, **c.** zigzag DAAQ-TFP, **d.** inclined DAAQ-TFP at equilibrium from classical force field.

References

(S1) Chandra, S.; Kundu, T.; Kandambeth, S.; BabaRao, R.; Marathe, Y.; Kunjir, S. M.; Banerjee, R. Phosphoric Acid Loaded Azo (- N N-) Based Covalent Organic Framework for Proton Conduction. *J. Am. Chem. Soc.* **2014**, *136*, 6570–6573.

(S2) DeBlase, C. R.; Silberstein, K. E.; Truong, T.-T.; Abruña, H. D.; Dichtel, W. R. β-Ketoenamine-linked covalent organic frameworks capable of pseudocapacitive energy storage. *J. Am. Chem. Soc.* **2013**, *135*, 16821–16824.

(S3) Natraj, A.; Ji, W.; Xin, J.; Castano, I.; Burke, D. W.; Evans, A. M.; Strauss, M. J.; Ateia, M.; Hamachi, L. S.; Gianneschi, N. C., et al. Single-Crystalline Imine-Linked Two-Dimensional Covalent Organic Frameworks Separate Benzene and Cyclohexane Efficiently. *J. Am. Chem. Soc.* **2022**, *144*, 19813–19824.