**PCCP Manuscript ID: CP-ART-12-2016-008449**

**TITLE: NUCLEATION AND GROWTH OF 2D COVALENT ORGANIC FRAMEWORKS: POLYMERIZATION AND CRYSTALLIZATION OF COF MONOMERS**

**Authors: B.T. Koo, R. Heden, P. Clancy; Cornell University**

**SUPPLEMENTAL INFORMATION**

1. **Supplemental Tables and Figures**
2. **Quantum mechanical, molecular mechanical modeling details**
3. **Validation of the SCC-DFTB model**
4. **Free energy calculations**
5. **Crystallization free energy solvent corrections**
6. **Layer-sliding energetics and step edge effects**
7. **Force Field Parameterization**

## 1. Supplemental Tables and Figures

**Table S1.** Equilibrium path sampling windows and simulation parameters for the first boronate ester condensation reaction with methanol catalyst. We used a window size of 0.25 Å, and 0.05 Å for finer windows near the transition state.

| Lower (Å) | Upper (Å) | Trajectory length (fs) | $\Delta t$ (fs) | Lower (Å) | Upper (Å) | Trajectory length (fs) | $\Delta t$ (fs) |
|---|---|---|---|---|---|---|---|
| -2.025 | -1.775 | 20 | 2 | -0.450 | -0.400 | 4 | 1 |
| -1.825 | -1.575 | 20 | 2 | -0.425 | -0.375 | 4 | 1 |
| -1.625 | -1.375 | 20 | 2 | -0.400 | -0.350 | 4 | 1 |
| -1.425 | -1.175 | 20 | 2 | -0.375 | -0.325 | 4 | 1 |
| -1.225 | -0.975 | 20 | 2 | -0.350 | -0.300 | 4 | 1 |
| -1.025 | -0.775 | 20 | 2 | -0.325 | -0.275 | 4 | 1 |
| -0.825 | -0.575 | 20 | 2 | -0.300 | -0.250 | 4 | 1 |
| -0.625 | -0.375 | 20 | 2 | -0.275 | -0.225 | 4 | 1 |
| Finer Windows | | | | -0.250 | -0.200 | 4 | 1 |
| 0.425 | 0.625 | 20 | 2 | -0.225 | -0.175 | 4 | 1 |
| 0.575 | 0.825 | 20 | 2 | -0.200 | -0.150 | 4 | 1 |
| 0.775 | 1.025 | 20 | 2 | -0.175 | -0.125 | 4 | 1 |
| 0.975 | 1.225 | 20 | 2 | -0.150 | -0.100 | 4 | 1 |

| | | | |
|---|---|---|---|
| -0.125 | -0.075 | 4 | 1 |
| -0.100 | -0.050 | 4 | 1 |
| -0.075 | -0.025 | 4 | 1 |
| -0.050 | 0.000 | 4 | 1 |
| -0.025 | 0.025 | 4 | 1 |
| 0.000 | 0.050 | 4 | 1 |
| 0.025 | 0.075 | 4 | 1 |
| 0.050 | 0.100 | 4 | 1 |
| 0.075 | 0.125 | 4 | 1 |
| 0.100 | 0.150 | 4 | 1 |
| 0.125 | 0.175 | 4 | 1 |
| 0.150 | 0.200 | 4 | 1 |
| 0.175 | 0.225 | 4 | 1 |
| 0.200 | 0.250 | 4 | 1 |
| 0.225 | 0.275 | 4 | 1 |
| 0.250 | 0.300 | 4 | 1 |
| 0.275 | 0.325 | 4 | 1 |
| 0.300 | 0.350 | 4 | 1 |
| 0.325 | 0.375 | 4 | 1 |
| 0.350 | 0.400 | 4 | 1 |
| 0.375 | 0.425 | 4 | 1 |
| 0.400 | 0.450 | 4 | 1 |

**Table S2.** Equilibrium path sampling windows and simulation parameters for the first boronate-ester condensation reaction with water catalyst. We used a window size of 0.25 Å, and 0.05 Å for finer windows near the transition state.

| Lower (Å) | Upper (Å) | Trajectory length (fs) | $\Delta t$ (fs) | Lower (Å) | Upper (Å) | Trajectory length (fs) | $\Delta t$ (fs) |
|---|---|---|---|---|---|---|---|
| -2.225 | -1.975 | 20 | 2 | -0.450 | -0.400 | 4 | 1 |
| -2.025 | -1.775 | 20 | 2 | -0.425 | -0.375 | 4 | 1 |
| -1.825 | -1.575 | 20 | 2 | -0.400 | -0.350 | 4 | 1 |
| -1.625 | -1.375 | 20 | 2 | -0.375 | -0.325 | 4 | 1 |
| -1.425 | -1.175 | 20 | 2 | -0.350 | -0.300 | 4 | 1 |
| -1.225 | -0.975 | 20 | 2 | -0.325 | -0.275 | 4 | 1 |
| -1.025 | -0.775 | 20 | 2 | -0.300 | -0.250 | 4 | 1 |
| -0.825 | -0.575 | 20 | 2 | -0.275 | -0.225 | 4 | 1 |
| -0.625 | -0.375 | 20 | 2 | -0.250 | -0.200 | 4 | 1 |
| **Finer Windows** | | | | -0.225 | -0.175 | 4 | 1 |
| 0.425 | 0.625 | 20 | 2 | -0.200 | -0.150 | 4 | 1 |
| 0.575 | 0.825 | 20 | 2 | -0.175 | -0.125 | 4 | 1 |
| 0.775 | 1.025 | 20 | 2 | -0.150 | -0.100 | 4 | 1 |
| 0.975 | 1.225 | 20 | 2 | -0.125 | -0.075 | 4 | 1 |
| | | | | -0.100 | -0.050 | 4 | 1 |
| | | | | -0.075 | -0.025 | 4 | 1 |
| | | | | -0.050 | 0.000 | 4 | 1 |
| | | | | -0.025 | 0.025 | 4 | 1 |
| | | | | 0.000 | 0.050 | 4 | 1 |
| | | | | 0.025 | 0.075 | 4 | 1 |
| | | | | 0.050 | 0.100 | 4 | 1 |
| | | | | 0.075 | 0.125 | 4 | 1 |
| | | | | 0.100 | 0.150 | 4 | 1 |

| | | | |
|---|---|---|---|
| 0.125 | 0.175 | 4 | 1 |
| 0.150 | 0.200 | 4 | 1 |
| 0.175 | 0.225 | 4 | 1 |
| 0.200 | 0.250 | 4 | 1 |
| 0.225 | 0.275 | 4 | 1 |
| 0.250 | 0.300 | 4 | 1 |
| 0.275 | 0.325 | 4 | 1 |
| 0.300 | 0.350 | 4 | 1 |
| 0.325 | 0.375 | 4 | 1 |
| 0.350 | 0.400 | 4 | 1 |
| 0.375 | 0.425 | 4 | 1 |
| 0.400 | 0.450 | 4 | 1 |

**Table S3.** Equilibrium path sampling windows and simulation parameters for the first boronate-ester condensation reaction with no catalyst. We used a window size of 0.25 Å, and 0.05 Å for finer windows near the transition state.

| Lower (Å) | Upper (Å) | Trajectory length (fs) | $\Delta t$ (fs) | Lower (Å) | Upper (Å) | Trajectory length (fs) | $\Delta t$ (fs) |
|---|---|---|---|---|---|---|---|
| -1.625 | -1.375 | 20 | 2 | -0.650 | -0.600 | 4 | 1 |
| -1.425 | -1.175 | 20 | 2 | -0.625 | -0.575 | 4 | 1 |
| -1.225 | -0.975 | 20 | 2 | -0.600 | -0.550 | 4 | 1 |
| -1.025 | -0.775 | 20 | 2 | -0.575 | -0.525 | 4 | 1 |
| -0.825 | -0.625 | 20 | 2 | -0.550 | -0.500 | 4 | 1 |
| Finer Windows | | | | -0.525 | -0.475 | 4 | 1 |
| 0.825 | 1.025 | 20 | 2 | -0.500 | -0.450 | 4 | 1 |
| 0.975 | 1.225 | 20 | 2 | -0.475 | -0.425 | 4 | 1 |
| 1.175 | 1.425 | 20 | 2 | -0.450 | -0.400 | 4 | 1 |
| 1.375 | 1.625 | 20 | 2 | -0.425 | -0.375 | 4 | 1 |
| 1.575 | 1.825 | 20 | 2 | -0.400 | -0.350 | 4 | 1 |
| 1.775 | 2.025 | 20 | 2 | -0.375 | -0.325 | 4 | 1 |
| | | | | -0.350 | -0.300 | 4 | 1 |
| | | | | -0.325 | -0.275 | 4 | 1 |
| | | | | -0.300 | -0.250 | 4 | 1 |
| | | | | -0.275 | -0.225 | 4 | 1 |
| | | | | -0.250 | -0.200 | 4 | 1 |
| | | | | -0.225 | -0.175 | 4 | 1 |
| | | | | -0.200 | -0.150 | 4 | 1 |
| | | | | -0.175 | -0.125 | 4 | 1 |
| | | | | -0.150 | -0.100 | 4 | 1 |
| | | | | -0.125 | -0.075 | 4 | 1 |
| | | | | -0.100 | -0.050 | 4 | 1 |
| | | | | -0.075 | -0.025 | 4 | 1 |

| | | | |
|---|---|---|---|
| -0.050 | 0.000 | 4 | 1 |
| -0.025 | 0.025 | 4 | 1 |
| 0.000 | 0.050 | 4 | 1 |
| 0.025 | 0.075 | 4 | 1 |
| 0.050 | 0.100 | 4 | 1 |
| 0.075 | 0.125 | 4 | 1 |
| 0.100 | 0.150 | 4 | 1 |
| 0.125 | 0.175 | 4 | 1 |
| 0.150 | 0.200 | 4 | 1 |
| 0.175 | 0.225 | 4 | 1 |
| 0.200 | 0.250 | 4 | 1 |
| 0.225 | 0.275 | 4 | 1 |
| 0.250 | 0.300 | 4 | 1 |
| 0.275 | 0.325 | 4 | 1 |
| 0.300 | 0.350 | 4 | 1 |
| 0.325 | 0.375 | 4 | 1 |
| 0.350 | 0.400 | 4 | 1 |
| 0.375 | 0.425 | 4 | 1 |
| 0.400 | 0.450 | 4 | 1 |
| 0.425 | 0.475 | 4 | 1 |
| 0.450 | 0.500 | 4 | 1 |
| 0.475 | 0.525 | 4 | 1 |
| 0.500 | 0.550 | 4 | 1 |
| 0.525 | 0.575 | 4 | 1 |
| 0.550 | 0.600 | 4 | 1 |
| 0.575 | 0.625 | 4 | 1 |
| 0.600 | 0.650 | 4 | 1 |
| 0.625 | 0.675 | 4 | 1 |

| | | | |
|---|---|---|---|
| 0.650 | 0.700 | 4 | 1 |
| 0.675 | 0.725 | 4 | 1 |
| 0.700 | 0.750 | 4 | 1 |
| 0.725 | 0.775 | 4 | 1 |
| 0.750 | 0.800 | 4 | 1 |
| 0.775 | 0.825 | 4 | 1 |
| 0.800 | 0.850 | 4 | 1 |

**Table S4.** Equilibrium path sampling windows and simulation parameters for the second boronate ester condensation reaction **with methanol or water catalyst**.

| Lower (Å) | Upper (Å) | Trajectory length (fs) | Δt (fs) |
| --- | --- | --- | --- |
| -2.225 | -1.575 | 20 | 2 |
| -1.625 | -1.175 | 20 | 2 |
| -1.225 | -0.975 | 20 | 2 |
| -1.025 | -0.775 | 20 | 2 |
| -0.825 | -0.575 | 20 | 2 |
| -0.625 | -0.425 | 20 | 2 |
| Finer Windows | | | |
| 0.425 | 0.625 | 20 | 2 |
| 0.575 | 1.025 | 20 | 2 |
| 0.975 | 1.225 | 20 | 2 |

| Lower (Å) | Upper (Å) | Trajectory length (fs) | Δt (fs) |
| --- | --- | --- | --- |
| -0.450 | -0.400 | 4 | 1 |
| -0.425 | -0.375 | 4 | 1 |
| -0.400 | -0.350 | 4 | 1 |
| -0.375 | -0.325 | 4 | 1 |
| -0.350 | -0.300 | 4 | 1 |
| -0.325 | -0.275 | 4 | 1 |
| -0.300 | -0.250 | 4 | 1 |
| -0.275 | -0.225 | 4 | 1 |
| -0.250 | -0.200 | 4 | 1 |
| -0.225 | -0.175 | 4 | 1 |
| -0.200 | -0.150 | 4 | 1 |
| -0.175 | -0.125 | 4 | 1 |
| -0.150 | -0.100 | 4 | 1 |
| -0.125 | -0.075 | 4 | 1 |
| -0.100 | -0.050 | 4 | 1 |
| -0.075 | -0.025 | 4 | 1 |
| -0.050 | 0.000 | 4 | 1 |
| -0.025 | 0.025 | 4 | 1 |
| 0.000 | 0.050 | 4 | 1 |
| 0.025 | 0.075 | 4 | 1 |
| 0.050 | 0.100 | 4 | 1 |
| 0.075 | 0.125 | 4 | 1 |
| 0.100 | 0.150 | 4 | 1 |
| 0.125 | 0.175 | 4 | 1 |

| | | | |
|---|---|---|---|
| 0.150 | 0.200 | 4 | 1 |
| 0.175 | 0.225 | 4 | 1 |
| 0.200 | 0.250 | 4 | 1 |
| 0.225 | 0.275 | 4 | 1 |
| 0.250 | 0.300 | 4 | 1 |
| 0.275 | 0.325 | 4 | 1 |
| 0.300 | 0.350 | 4 | 1 |
| 0.325 | 0.375 | 4 | 1 |
| 0.350 | 0.400 | 4 | 1 |
| 0.375 | 0.425 | 4 | 1 |
| 0.400 | 0.450 | 4 | 1 |

**Table S5.** Equilibrium path sampling windows and simulation parameters for the second boronate ester condensation reaction **with no catalyst** present.

| Lower (Å) | Upper (Å) | Trajectory length (fs) | $\Delta t$ (fs) | Lower (Å) | Upper (Å) | Trajectory length (fs) | $\Delta t$ (fs) |
|---|---|---|---|---|---|---|---|
| -1.625 | -1.175 | 20 | 2 | -0.650 | -0.600 | 4 | 1 |
| -1.225 | -0.975 | 20 | 2 | -0.625 | -0.575 | 4 | 1 |
| -1.025 | -0.775 | 20 | 2 | -0.600 | -0.550 | 4 | 1 |
| -0.825 | -0.625 | 20 | 2 | -0.575 | -0.525 | 4 | 1 |
| Finer Windows | | | | -0.550 | -0.500 | 4 | 1 |
| 0.825 | 1.025 | 20 | 2 | -0.525 | -0.475 | 4 | 1 |
| 0.975 | 1.225 | 20 | 2 | -0.500 | -0.450 | 4 | 1 |
| 1.175 | 1.625 | 20 | 2 | -0.475 | -0.425 | 4 | 1 |
| 1.575 | 2.025 | 20 | 2 | -0.450 | -0.400 | 4 | 1 |
| | | | | -0.425 | -0.375 | 4 | 1 |
| | | | | -0.400 | -0.350 | 4 | 1 |
| | | | | -0.375 | -0.325 | 4 | 1 |
| | | | | -0.350 | -0.300 | 4 | 1 |
| | | | | -0.325 | -0.275 | 4 | 1 |
| | | | | -0.300 | -0.250 | 4 | 1 |
| | | | | -0.275 | -0.225 | 4 | 1 |
| | | | | -0.250 | -0.200 | 4 | 1 |
| | | | | -0.225 | -0.175 | 4 | 1 |
| | | | | -0.200 | -0.150 | 4 | 1 |
| | | | | -0.175 | -0.125 | 4 | 1 |
| | | | | -0.150 | -0.100 | 4 | 1 |
| | | | | -0.125 | -0.075 | 4 | 1 |
| | | | | -0.100 | -0.050 | 4 | 1 |
| | | | | -0.075 | -0.025 | 4 | 1 |

| | | | |
|---|---|---|---|
| -0.050 | 0.000 | 4 | 1 |
| -0.025 | 0.025 | 4 | 1 |
| 0.000 | 0.050 | 4 | 1 |
| 0.025 | 0.075 | 4 | 1 |
| 0.050 | 0.100 | 4 | 1 |
| 0.075 | 0.125 | 4 | 1 |
| 0.100 | 0.150 | 4 | 1 |
| 0.125 | 0.175 | 4 | 1 |
| 0.150 | 0.200 | 4 | 1 |
| 0.175 | 0.225 | 4 | 1 |
| 0.200 | 0.250 | 4 | 1 |
| 0.225 | 0.275 | 4 | 1 |
| 0.250 | 0.300 | 4 | 1 |
| 0.275 | 0.325 | 4 | 1 |
| 0.300 | 0.350 | 4 | 1 |
| 0.325 | 0.375 | 4 | 1 |
| 0.350 | 0.400 | 4 | 1 |
| 0.375 | 0.425 | 4 | 1 |
| 0.400 | 0.450 | 4 | 1 |
| 0.425 | 0.475 | 4 | 1 |
| 0.450 | 0.500 | 4 | 1 |
| 0.475 | 0.525 | 4 | 1 |
| 0.500 | 0.550 | 4 | 1 |
| 0.525 | 0.575 | 4 | 1 |
| 0.550 | 0.600 | 4 | 1 |
| 0.575 | 0.625 | 4 | 1 |
| 0.600 | 0.650 | 4 | 1 |
| 0.625 | 0.675 | 4 | 1 |

| | | | |
|---|---|---|---|
| 0.650 | 0.700 | 4 | 1 |
| 0.675 | 0.725 | 4 | 1 |
| 0.700 | 0.750 | 4 | 1 |
| 0.725 | 0.775 | 4 | 1 |
| 0.750 | 0.800 | 4 | 1 |
| 0.775 | 0.825 | 4 | 1 |
| 0.800 | 0.850 | 4 | 1 |

**Scheme S1.** The second boronate ester condensation reaction, linking HHTP and PBBA. In a similar fashion to the first boronate ester condensation reactions, the participation of a protic solvent to facilitate proton transfer transforms the transition state from a strained 4-membered ring to a less strained 6-membered ring. Upon reacting, a water molecule is released. Green, blue, and red colours indicate HHTP, PBBA, and the product respectively. Orange denotes catalytic species.
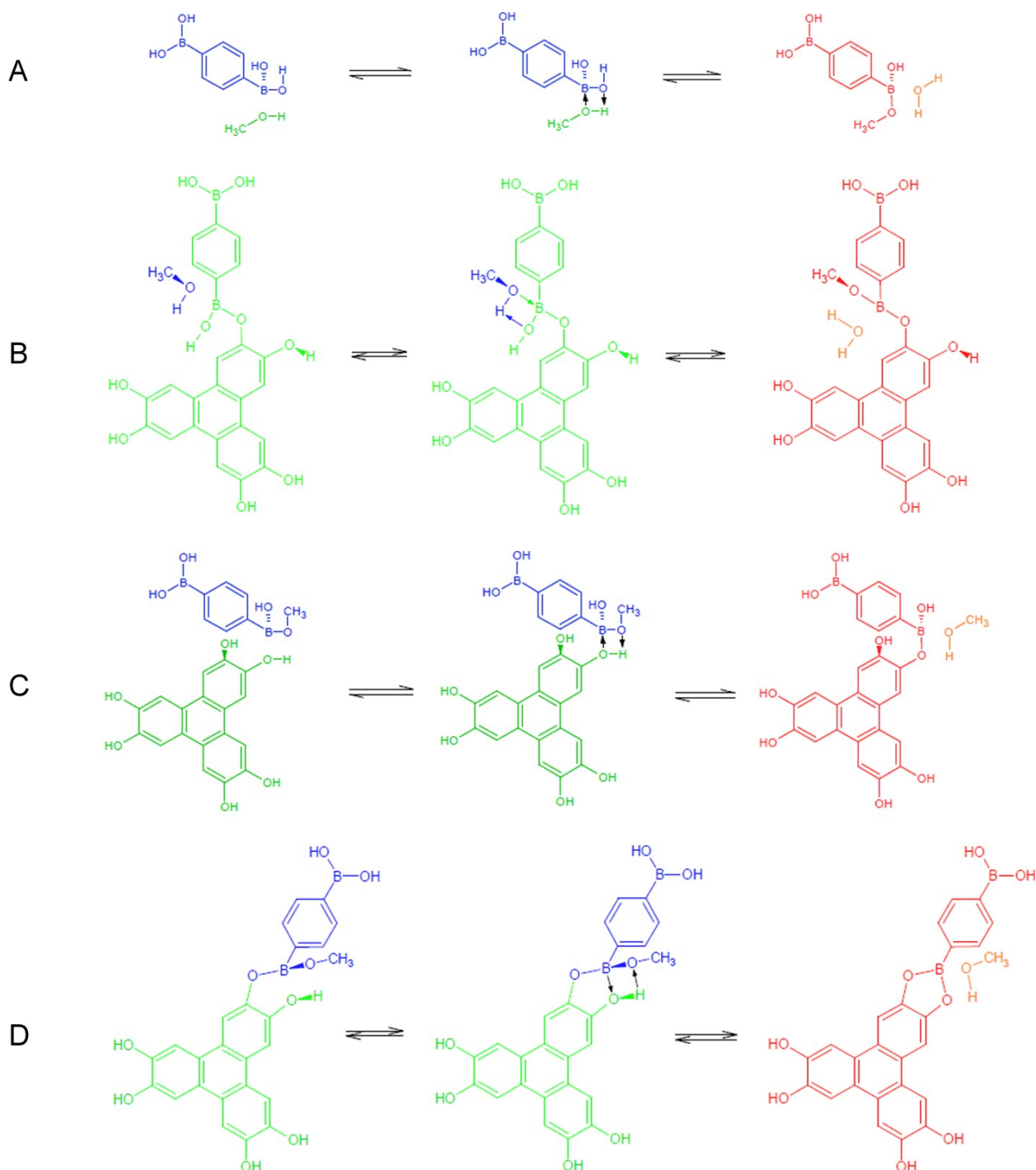
**Scheme S2.** Competing methylation reactions and additional reaction pathways. a) methylation of boronic acid. b) methylation of a boronic acid compound. c) boronate ester condensation of methylated boronic acid and HHTP. d) unimolecular boronate ester condensation of a methylated compound of boronic acid and HHTP. Green, blue, and red indicate HHTP, PBBA, and the product, respectively. Orange denotes catalytic species.
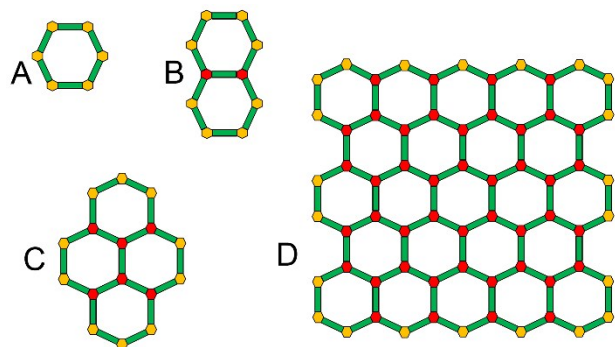
**Figure S1.** Four closed-ring macrocyclic sheets tested for their binding free energy to identical stacked counterparts. Left hand side: (A) A single ring, (B) two fused rings, and (C) four fused rings. Right hand side: (D) a 5x5 sheet of rings. Orange and red hexagons represent HHTP as linked to two (so called "exterior") or three ("interior") PBBA monomers, respectively. Green segments represent the PBBA "linker" monomers.
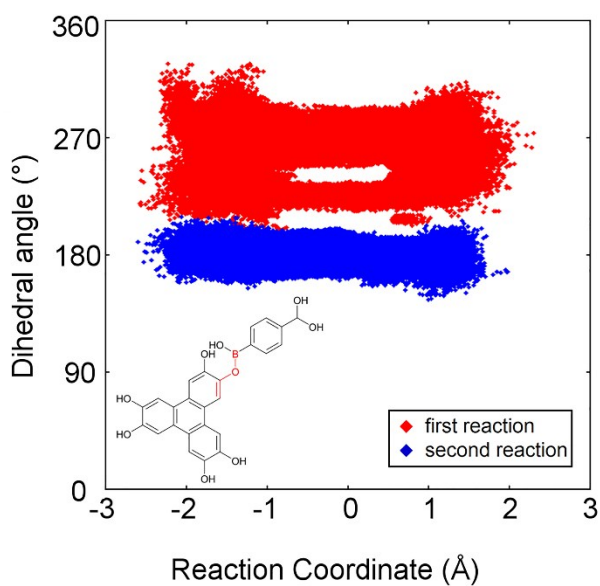


**Figure S2.** Sampling of the key dihedral angle (highlighted in red in the inset molecule) during both the first boronate ester condensation reaction (red) and the second boronate ester condensation reaction (blue). The preferred value of the dihedral (red) for the first reaction has an average value of $260° \pm 22°$, and that for the second reaction has an average dihedral value of $179° \pm 7°$.

## 2.1. Quantum mechanical calculations

We adopted the semi-empirical, dispersion-corrected, self-consistent charge density functional tight-binding method (SCC-DFTB) to model the boronate ester condensation reaction based on its past success in accurately modeling COF formation energies and layer stacking energetics.[1,2] SCC-DFTB is a second-order Taylor expansion of the Kohn-Sham density functional theory total energy[3,4], and is orders of magnitude faster than density functional theory (DFT). This increased speed allows us to explore otherwise computationally intractable problems such as free energies of reaction. We also apply dispersion corrections based on the universal force field (UFF)[4] to accurately model van der Waals interactions. To validate use of this approach,

we compared SCC-DFTB without dispersion corrections to density functional theory at the B3LYP/6-31G(d,p)[5,6] level of theory. We also compared SCC-DFTB to other common semi-empirical quantum mechanical methods, including AM1, PM3, and PM6[7-9].

To locate transition state structures, we used the quadratic synchronous transit algorithm (QST3)[10] as implemented in Gaussian09[11], which converges on a transition state by searching along a parabola connecting reactants and products and performing an eigenvector-following algorithm. We calculate the energetic transition between reactants, transition states, and products with an intrinsic reaction coordinate (IRC) scan in Gaussian09. IRC performs a trace using a steepest descent algorithm[12] to connect a known transition state to either reactants or products.

## 2.2. Molecular Mechanics Calculations

Our all-atom molecular dynamics simulations use a combination of Dreiding models[13] to simulate the solute-solute and solute-solvent interactions and the OPLS force field[14] to simulate the solvent-solvent interactions. We apply all-atom molecular dynamics whenever QM/MM modeling is unnecessary, as in the calculation of solvation free energy and binding free energy.

## 3. Validation of the SCC-DFTB model

To validate our extensive use of SCC-DFTB to model the boronate ester condensation reaction, we compared it to a more accurate Density Functional Theory (DFT) approach at the level of B3LYP/6-311++G**. Since this variant of DFT is not dispersion-corrected, we compared it to the dispersion-uncorrected SCC-DFTB. We performed validation tests on simpler, truncated reactant and product species without a homogeneous catalyst present, but we expect similar energetics to the full system. HHTP was replaced by the structurally similar *o*-dihydroxybenzene, and PBBA was replaced by phenylboronic acid. Given the number of degrees of freedom in such a comparison, we considered two pairs of common collective variables to generate potential energy surfaces: The first pair of variables involves the two reactive boron-oxygen bonds that are part of the 4-membered ring in the transition state structure. The second pair involves the two reactive oxygen-hydrogen bonds that track the proton transfer between diol and boronic acid.

## 3.1. Reactive boron-oxygen bonding in SCC-DFTB and DFT

As a result of these calculations, the first potential energy (boron-oxygen) surface of SCC-DFTB differs from that of DFT in the number of local minima, as shown in Figures S3 and S4. The SCC-DFTB potential energy surface shares two local minima with DFT, namely the reactant and product van der Waals configurations located approximately at (1.4 Å, 2.5-3.0 Å) and (2.5-3.0 Å, 1.4 Å), respectively. Without dispersion corrections, the location of the reactant and product van der Waals states can only be estimated, hence a range of separation distances is given. However, the potential energy surface of SCC-DFTB contains two additional intermediate local minima at (1.45 Å, 1.55 Å) and (1.53 Å, 1.47 Å) that are not found by DFT. The SCC-DFTB potential energy surface indicates that these two extraneous states are only metastable with respect to the reactant and product van der Waals states by approximately 3 kcal/mol, but that most of the activation energy comes from the transition between the extraneous intermediate states. An IRC scan with SCC-DFTB indicates an activation energy of 27 kcal/mol from the reactant side and 25 kcal/mol from the product side. The same IRC scan with DFT indicates a similar activation energy of 26 kcal/mol from the reactant side and 25 kcal/mol from the product side. We correct for the existence of extraneous intermediate states in SCC-DFTB by applying SMD[16,17] (as discussed in SI 4.3) to enlarge the distance between reactive boron and oxygen atoms to arrive at DFT-defined minima.
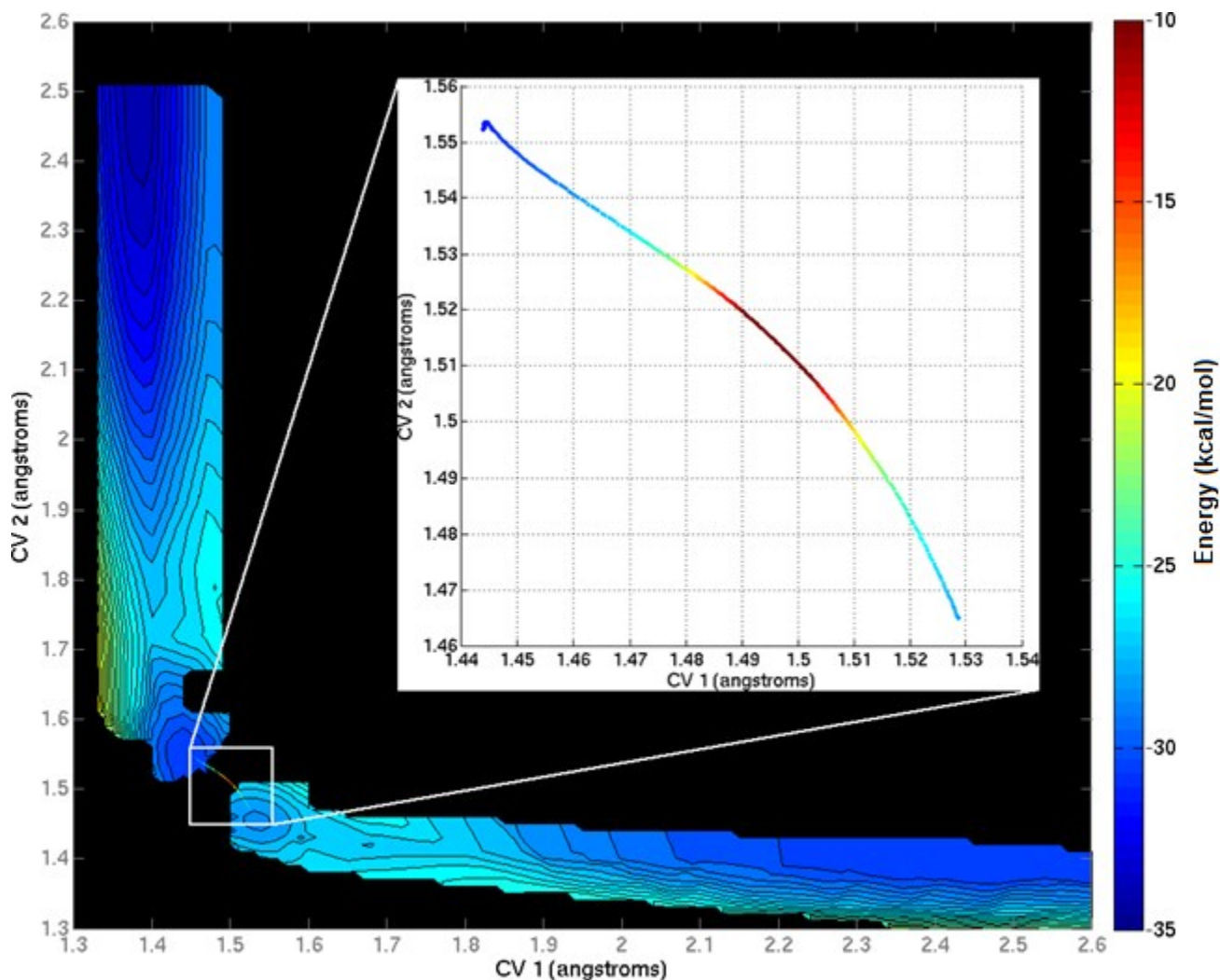
**Figure S3.** A potential energy surface describing the energetics of the first boronate ester condensation reaction between phenylboronic acid (PBBA fragment) and o-dihydroxybenzene (HHTP fragment) in vacuum as modeled by density functional tight binding. Collective variable 1 is the length of one of the B-O bonds in phenylboronic acid, and collective variable 2 is the length of the boron-oxygen bond between the boron in phenylboronic acid and the oxygen in o-dihydroxybenzene. The black and white regions on the potential energy surfaces are not sampled. We superimpose the transition pathway found with an intrinsic reaction coordinate scan onto the contour plot to link the two disjoint contour regions. The transition pathway has end-points at (1.45 Å, 1.55 Å) (the reactant-like state) and (1.53 Å, 1.47 Å) (the product-like state), with a transition state at (1.50 Å, 1.51 Å). The contour region along the left side of the potential energy surface contains the reactant configurations, and the contour region along the bottom of the potential energy surface contains the product configurations.
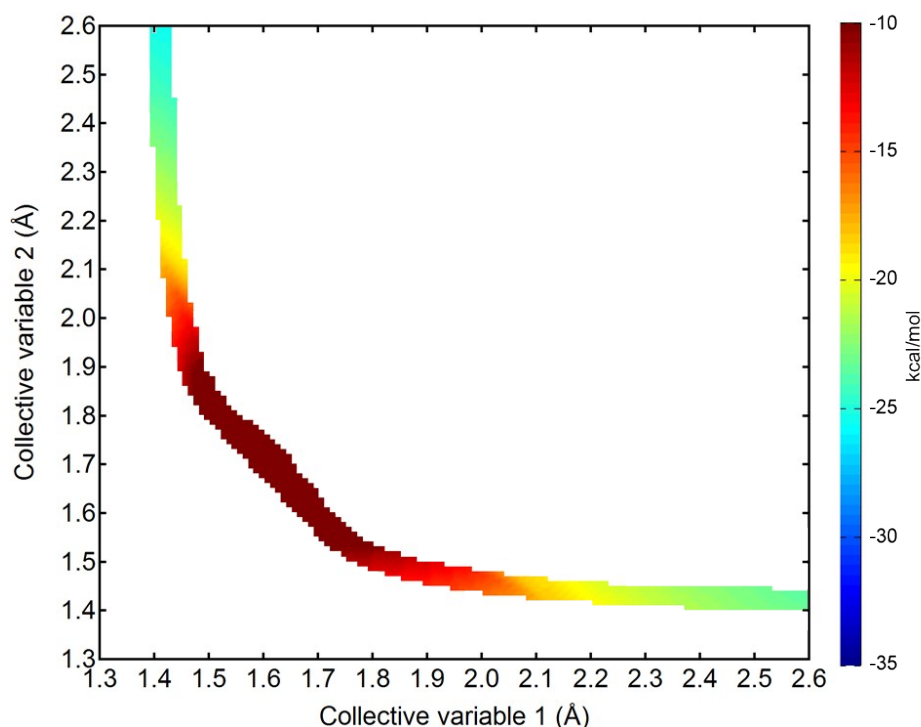
**Figure S4.** A potential energy surface generated for the unassisted first reaction, plotted along the same boron-oxygen coordinates in Figure S3, except modeled with density functional theory (B3LYP/6-311++G**)[15]. There are only two local minima at (1.4 Å, 2.6 Å) (reactants) and (2.6 Å, 1.4 Å) (products) and one saddle point at (1.7 Å, 1.7 Å) (transition state).

## 3.2. Reactive oxygen-hydrogen bonding in SCC-DFTB and DFT

On the reactive oxygen-hydrogen potential energy surface, we found a better match between DFT and SCC-DFTB. The extraneous metastable states found in boron-oxygen phase space were not found in the oxygen-hydrogen phase space, which suggests that SCC-DFTB is suitable for modeling reactive oxygen-hydrogen bonds, but not for modeling reactive boron-oxygen bonds. As shown in Figure S5, the reactant and product van der Waals states are similar in both energy and spatial coordinates. The transition state on the SCC-DFTB potential energy surface differs from the transition state on the DFT potential energy surface by just 0.2 Å in both coordinates, and the transition state energy is similar in both methods.
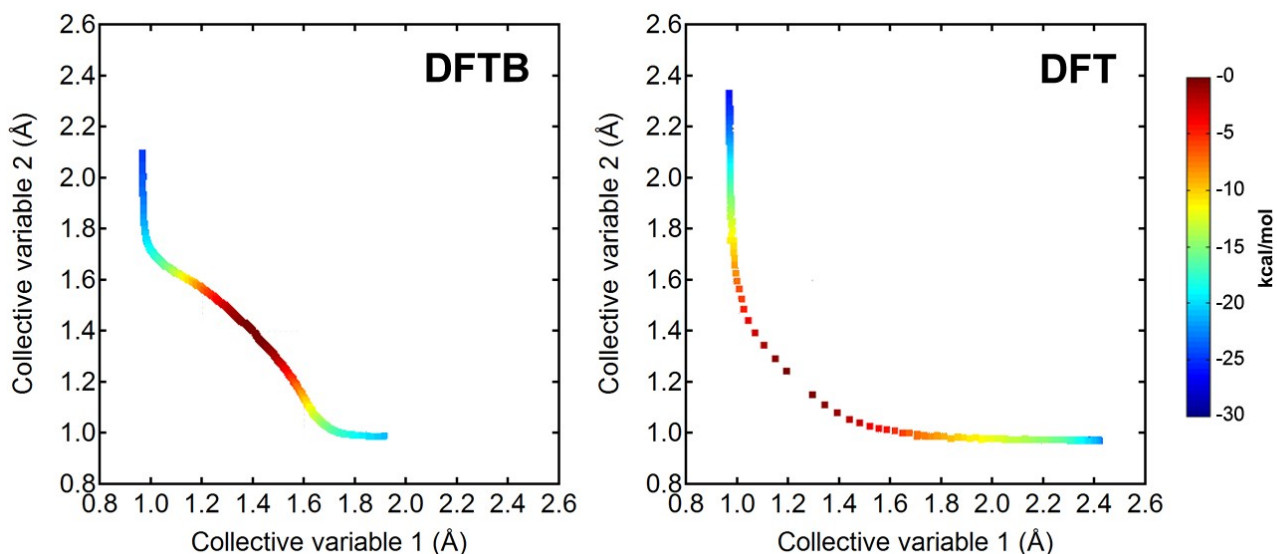
**Figure S5.** Potential energy curves generated *via* an intrinsic reaction coordinate scan for the unassisted first boronate ester condensation, modeled by DFTB (H, B, C, O matscc parameters without dispersion correction) on the left, and DFT (B3LYP/6-311++G**) on the right. The first collective variable (CV 1) is the length of one of the oxygen-hydrogen bonds of dihydroxybenzene, and the second collective variable (CV 2) is the length of the oxygen-hydrogen bond connecting the oxygen from phenylboronic acid and the hydroxyl hydrogen from dihydroxybenzene. The reactants lie on the upper-left end of the curve, and the products lie on the bottom-right end of the curve.

### 3.3. Alternative semi-empirical QM models: PM3, AM1, and PM6

We compared the energetics of other commonly available semi-empirical methods (AM1, PM3, PM6) in boron-oxygen phase space with DFT to identify possible alternative methods for simulating the reaction. Although the number of stationary points (reactants, transition state, and products) in each potential energy surface is equal to that of DFT, their *location* is not. As shown in Figure S6, all three semi-empirical methods underestimate the bond length of CV 1 at the product state (*i.e.,* the length of B-O in phenylboronic acid). The PM6 potential energy surface is the worst predictor of the activation energy, as the transition state energies are 5 and 15 kcal/mol from the reactant and product side, respectively (as opposed to 26 and 25 kcal/mol according to DFT results). Given the unsatisfactory alternatives considered above, we selected the SCC-DFTB approach to model our reaction, and we correct for the extraneous reactant and product states with thermodynamic integration and steered molecular dynamics (SMD)[16,17] to advance the reaction coordinate past the extraneous local minima towards more physically relevant local minima on the potential energy surface.
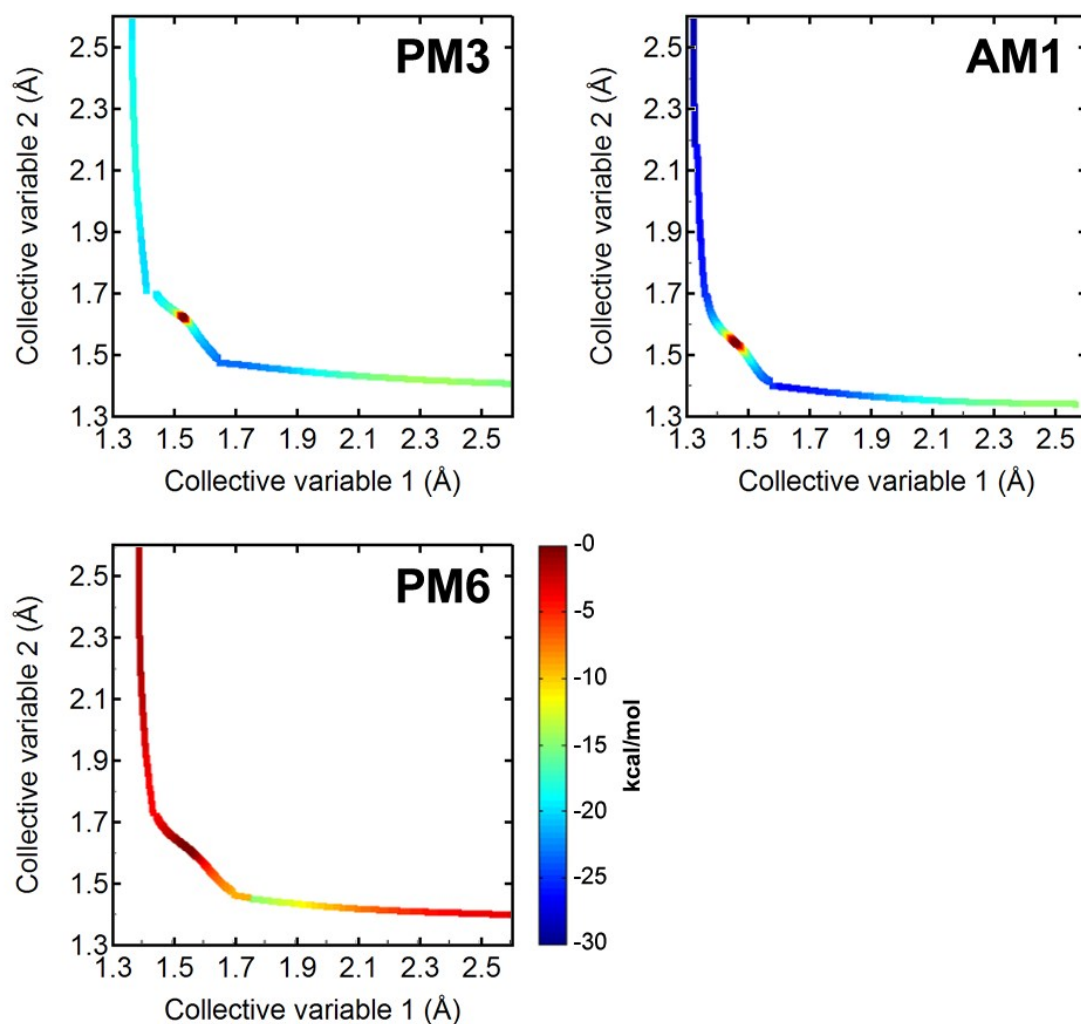
**Figure S6.** Potential energy curves generated via an intrinsic reaction coordinate scan for the unassisted first boronate ester condensation, modeled by PM3 (top left), AM1 (top right), and PM6 (bottom). The collective variables are boron-oxygen bond lengths, in which reactants are located on the top-left of the curve and products are located on the bottom-right of the curve.

## 4. Free energy calculations

### 4.1. Reaction coordinate search for boronate ester formation reaction

The ideal reaction coordinate is a function of the committor probability[18], or the probability a configuration converges to the set of product states when it is assigned velocities randomly sampled from a Boltzmann distribution. Configurations that are transition states by this definition are equally likely to converge to reactants or products and have a committor probability of 0.5. Configurations that belong to the set of reactant or product states have a committor probability of 0.0 and 1.0, respectively, because all such configurations are likely to remain in these states. Reaction coordinates are traditionally defined as a linear combination of collective variables, which are physically intuitive coordinates such as bonds, angles, dihedrals, or trigonometric functions of collective variables. To create a thorough description of the configuration of the reaction site, we include bond differences, angles, and dihedrals involving the atoms labeled in Chart S1, for a total of 29 collective variables (Table S6). Reaction coordinates can be found a priori with aimless shooting[19] and likelihood

maximization[20], with knowledge of just the reactants and products. Aimless shooting is a transition path-sampling algorithm that harvests statistically independent configurations near the transition hypersurface. Likelihood maximization is an algorithm that ranks collective variables based on their orthogonality to this hypersurface.

Aimless shooting harvests collections of configurations near the transition hypersurface by a method known as "shooting and shifting" from an initial seed configuration. A single aimless shooting simulation typically harvests tens of thousands of configurations for post-simulation analysis. A single trajectory is propagated forward and backward from the seed configuration by assigning random velocities according to the Boltzmann distribution at a certain temperature ("shooting"). Each trajectory is categorized as "reactive" if it links reactant and product states, or "non-reactive" if not. Reactive trajectories are sampled further by assigning the next seed configuration a certain small time interval from the initial seed configuration, $\Delta t$, which is set once at the outset of the simulation ("shifting"). Since random velocities are assigned for each seed configuration, the resulting trajectories are independent. Since unreactive trajectories are discarded, sampling is biased towards the transition hypersurface, since configurations far from the transition hypersurface are unlikely to produce reactive trajectories.
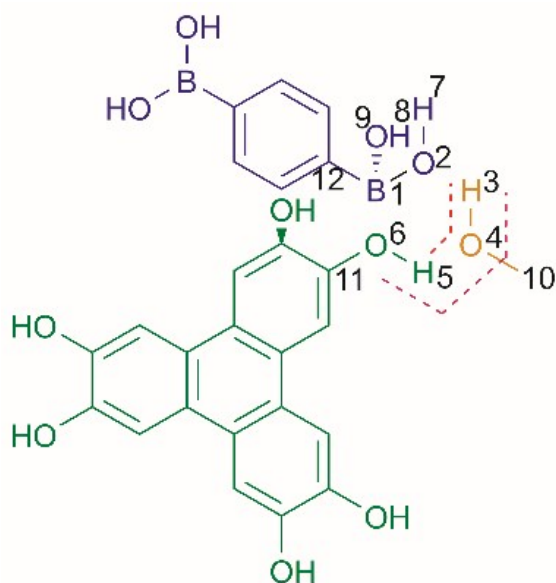


**Chart S1. Collective variable definitions.** Atom labels 1-12 define the collective variables listed in Table S6. The dotted lines indicate the angle (interior dotted red line) and dihedral (exterior dotted purple line) used in the final reaction coordinate definition. Colours of the molecular species as defined in Scheme S1.

**Table S6.** Collective variables for a likelihood maximization approach used to find the best 1-, 2-, and 3-component linear reaction coordinate. The atom indices refer to the atom labels provided in Chart S2.

| CV | Description of CV |
|---|---|
| 1 | Bond 5-6 |
| 2 | Bond 1-6 |
| 3 | Bond 1-2 |
| 4 | Bond 2-3 |
| 5 | Bond 3-4 |
| 6 | Bond 4-5 |
| 7 | Bond 6-11 |
| 8 | Bond 1-12 |
| 9 | Bond 4-10 |
| 10 | Bond 2-7 |
| 11 | Bond 8-9 |
| 12 | Bond difference (Bond 5-6) - (Bond 1-6) |
| 13 | Bond difference (Bond 1-6) - (Bond 1-2) |
| 14 | Bond difference (Bond 1-2) - (Bond 2-3) |
| 15 | Bond difference (Bond 2-3) - (Bond 3-4) |
| 16 | Bond difference (Bond 3-4) - (Bond 4-5) |
| 17 | Bond difference (Bond 4-5) - (Bond 5-6) |
| 18 | Angle 5-6-1 |
| 19 | Angle 6-1-2 |
| 20 | Angle 1-2-3 |
| 21 | Angle 2-3-4 |
| 22 | Angle 3-4-5 |
| 23 | Angle 4-5-6 |
| 24 | Dihedral 5-6-1-2 |

| 25 | Dihedral 6-1-2-3 |
|----|------------------|
| 26 | Dihedral 1-2-3-4 |
| 27 | Dihedral 2-3-4-5 |
| 28 | Dihedral 3-4-5-6 |
| 29 | Dihedral 4-5-6-1 |

$$L(\boldsymbol{\alpha}) = \prod_{k=1}^{B} p_B(r(\boldsymbol{q}^{(k)})) \prod_{k=1}^{\neq B} (1 - p_B(r(\boldsymbol{q}^{(k)}))) \tag{S1}$$

$$p_B = \frac{(1 + tanh(r))}{2}, r(\boldsymbol{q}) = \alpha_0 + \Sigma \alpha_i \cdot q_i \tag{S2}$$

Likelihood maximization (LM) ranks each trial reaction coordinate by performing a non-linear maximization of equation S1. In this equation, B denotes the set of configurations $\boldsymbol{q}$ that converge to products in forward time. Configurations are represented in collective variable coordinates, $\boldsymbol{q}$. The committor probability, $p_B$, is a function of the trial reaction coordinate $\boldsymbol{q}$, typically a linear combination of 1, 2, or 3 collective variables. The functional form of $p_B$ in equation S2 maintains all the qualities of the committor probability and is used here. Each non-linear maximization optimizes the coefficients $\boldsymbol{\alpha}$. Intuitively, the likelihood maximization function combines binary results from all harvested configurations, $\boldsymbol{q}$, to rank the quality of a trial reaction coordinate, $\boldsymbol{q}$. For each $\boldsymbol{q}$, likelihood maximization multiplies either $(1-p_B)$ if $\boldsymbol{q}$ converges to the reactants in forward time, and $p_B$ if $\boldsymbol{q}$ converges to the products in forward time. The expectation is that, for reactants, $(1-p_B)$ is larger than $p_B$, and, for products, $(1-p_B)$ is smaller than $p_B$. So, for good reaction coordinates, more large terms are multiplied together to produce an overall larger LM score. The reaction coordinates determined by LM are given in Table S7.

We use a histogram test (also known as a code density test)[21] to distinguish the "quality" of each reaction coordinate (*i.e.,* how well the reaction coordinate definition maps to committor probability). The histogram test validates each trial reaction coordinate by testing the spread of the committor probability at the transition hypersurface. A suitable reaction coordinate divides reactant and product configurations such that the committor probability at the r = 0 hypersurface is narrowly distributed around 0.5. As shown in Figure S7, all reaction coordinates found by likelihood maximization yield histograms that are, indeed, centered at 0.5 with small 0.2 standard deviations. Therefore, we apply the first reaction coordinate (a bond length difference between two neighboring reactive O-H bonds) to all other boronate ester condensation reactions.

**Table S7**. Definitions of reaction coordinates in terms of collective variables (listed in Table S6)

| | |
|---|---|
| Coordinate 1 | $-0.10 + 3.84\,q_{16}$ |
| Coordinate 2 | $-3.08 + 3.82\,q_{16} + 1.18\,q_{21}$ |
| Coordinate 3 | $-3.47 + 3.78\,q_{16} + 1.31\,q_{21} - 0.37\,q_{28}$ |



RC 1
$\bar{x} = 0.54$
$\sigma = 0.23$

RC 2
$\bar{x} = 0.48$
$\sigma = 0.21$

RC 3
$\bar{x} = 0.49$
$\sigma = 0.21$
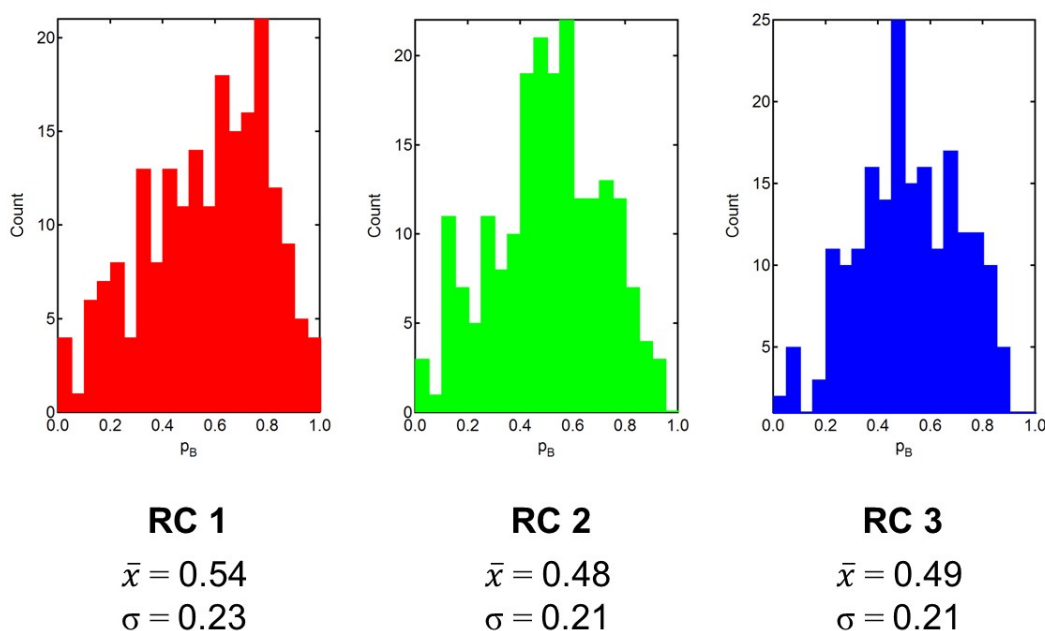
**Figure S7.** Histogram test of the best 1-, 2-, and 3-collective variable reaction coordinates (as displayed in Table S7) found by likelihood maximization, shown from left to right. $\bar{x}$ and $\sigma$ correspond to the mean and standard deviation, respectively.

## 4.2. Transition path sampling

We use equilibrium path sampling (EPS)[22,23] to determine the reaction free energy along a known reaction coordinate. Sampling along the reaction coordinate is typically difficult because of large barriers present between reactant and product states. EPS divides the energetic landscape into small overlapping intervals along the reaction coordinate ("windows") that are easily sampled individually. On steep portions of the energy landscape, windows can be made arbitrarily small to increase the sampling efficiency. EPS follows a similar procedure to aimless shooting in that it requires an initial seed configuration. A trajectory is propagated from the seed and accepted if it intersects with the window. Otherwise, it is rejected. Accepted trajectories spawn new seed configurations through "shifting" – configurations anywhere on the last accepted trajectory are

assigned a new time, t, from which new trajectories are propagated with random Boltzmann velocities. The collection of seed configurations in each window forms individual unweighted **r** (reaction coordinate) histograms, and the collection of histograms are combined using weighted histogram analysis method (WHAM)[24,25] to produce the reaction free energy.

## 4.3. Error correction with vacuum calculations

We found that the parameterization of SCC-DFTB[3,4] tends to overbind boron and oxygen atoms that are not bonded together. The optimized structures obtained by DFT and DFTB show differences in B-O bond lengths greater than 1 Å. However, when we re-optimize boron and oxygen atoms after increasing their initial distance by 1 Å, we obtain the correct van der Waals separation. This indicates a correction factor can be applied by calculating the difference in energy between these overbound states and separated states.

We use in-vacuum calculations[11] to determine the difference in energy between overbound states found by SCC-DFTB and van der Waals states found by DFT. To physically separate overbound boron and oxygen pairs, we applied SMD at a rate of 10 Å/ns. SMD[16,17] is a versatile preliminary technique to calculate the free energy differences between two molecular configurations where the path between them is unknown. SMD applies a spring with spring constant, k, between the centers of mass of two groups of atoms and performs the work of separating these groups of atoms at a constant velocity or force. After separating boron and oxygen with SMD, we performed local structure optimization with UFF dispersion-corrected SCC-DFTB on overbound and van der Waals states from 10 snapshots taken from our transition path sampling calculations.

## 4.4: Step edge effects

For step edge effects, SMD was performed in which the COF fragment in Chart S2 was pulled 20 Å away from a crystal surface in a 4:1 dioxane:mesitylene solvent. TI was then performed using 40 points along the trajectory with 0.5 Å spacing. The results are shown in Figure S8. It is observed that the step edge has little effect on adsorption energies, with the exception of case 3, where the adsorbate is forced out of its normal crystal orientation.
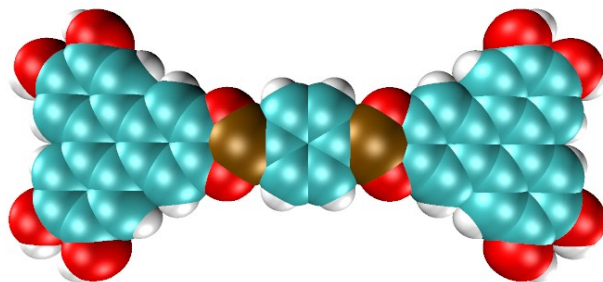


**Chart S2.** COF fragment used for step edge calculations
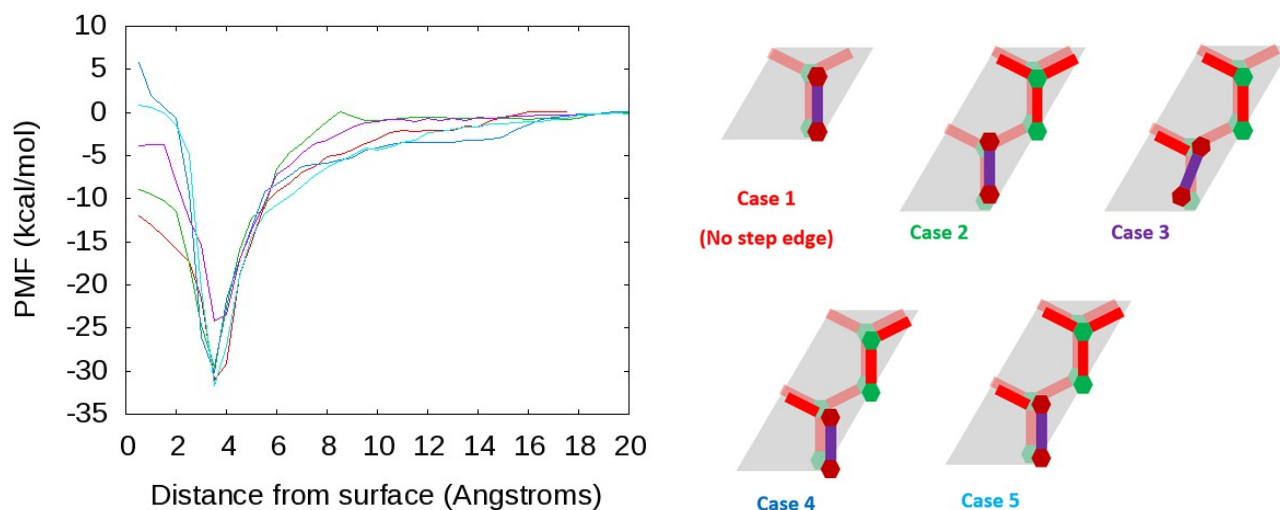
**Figure S8.** Potential of mean force for adsorption onto a flat crystal surface (red), and for adsorption adjacent to the five step edge types shown to the right; the colours of the case numbers correspond to the colour in the PMF plot on the left: Cases 1-5 are shown as red, green purple, blue and cyan, respectively. The gray background denotes the periodic cell, the dark red and green monomers denote the step, and the brown/purple molecule is the adsorbate.

## 5. Crystallization free energy solvent corrections

We used Thermodynamic Integration to compute the solvation free energy of macrocyclic COF-5 rings to understand how solvent stabilizes stacked versus unstacked COF nanosheets. Our implementation of Thermodynamic Integration follows an alchemical pathway between solutes in solution and solutes in vacuum by adjusting a $\lambda$ coupling parameter that indicates the strength of intermolecular interaction between solute and solvent. Due to the instability of conventional intermolecular potentials to $\lambda$ coupling, we use a modified "soft-core" version of Lennard-Jones and Coulomb potentials ($U'$ in equation 3)[26] that approaches a finite energy at zero separation distances. $U_{soft}$ is the summation of a (12-6) Lennard-Jones potential and Coulomb potential with the substitutions dictated in equations S4 and S5. We used the window definitions given in Table S8 for the Thermodynamic Integration. We increased the window density between $\lambda$ values of 0.0 and 0.5 to integrate the free energy peak present in this range with more accuracy.

$$\Delta G = \int \left\langle \frac{\partial U'}{\partial \lambda} \right\rangle_\lambda, \; U' = \lambda^2 U_{so} \qquad (S3)$$

$$r'_{LJ} = \left( \alpha\sigma^6(1-\lambda)^2 + r^6 \right)^{\frac{1}{6}} \qquad (S4)$$

$$r'_{coul} = \left( \alpha(1-\lambda)^2 + r^2 \right)^{\frac{1}{2}} \qquad (S5)$$

S26

We simulate each Thermodynamic Integration window for 1 ns and take snapshots every 1 ps (approximately the correlation time of the system). Since solvation free energy calculations are expensive (given that increasing sizes of solutes require more solvent to surround them), we adopt a geometrical strategy to compute correlations in solvation free energy that can be extrapolated beyond the computational limits of a three- or four- macrocyclic-ringed nanosheet. We assume defect-free, stacked macrocyclic rings can be divided into hybrid components that contribute linearly to the total solvation free energy. To limit statistical noise, we confine our analysis to those of closed macrocyclic rings, examples of which are drawn in Figure S1. We found two hybrid components; one "exterior" and one "interior" (whose meaning is explained in Figure S9) that form the basis for all single-crystalline stacks of equally sized nanosheets. We constructed six simulations to model interior and exterior hybrid components in three stacked arrangements: single (both sides exposed to solvent), surface (one side exposed), and sandwiched (only edges exposed to solvent). Our simulations are "infinite" sheets and "infinite" ribbons in single-, double-, and triple-layer structures. These infinite sheets and ribbons (see Figure S9, b) provided model crystalline structures to eliminate the solvation free energy of each hybrid component. The "infinite sheet" contains four interior hybrid components, and the "infinite ribbon" contains two interior and two exterior hybrid components. The component solvation free energies are listed in Table S9, together with an estimate of the errors associated with Thermodynamic Integration, as calculated from standard errors. We found that surface solvation energies of a single layer are half the solvation free energies of double layers, and that sandwiched solvation energies can be extracted from 3-layer simulations, which are two parts surface and one part sandwiched (see Figure S9).
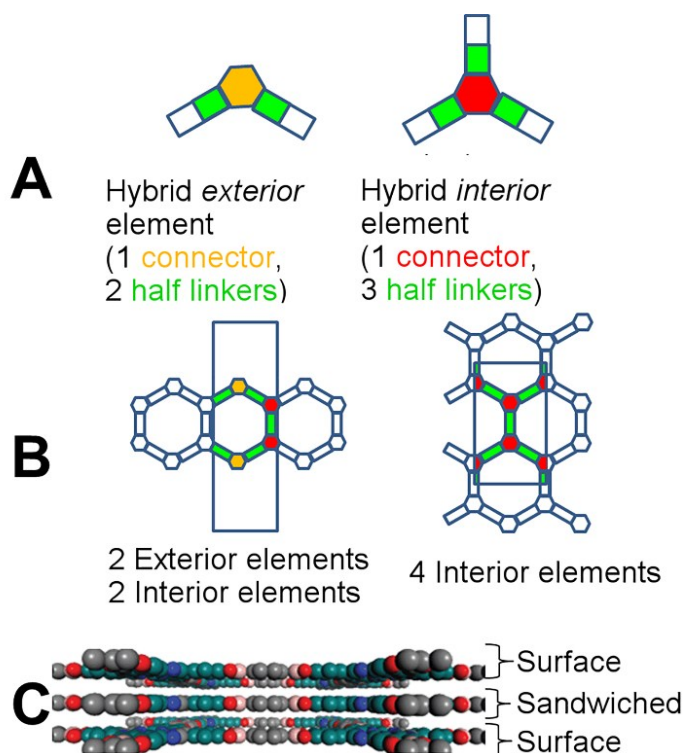


**Figure S9.** Schematics depicting our solvation free energy calculations of closed-ring macrocyclic COF-5 structures in solution. (A) We represent all closed-ring macrocyclic rings as a composite of hybrid exterior and interior elements, under the assumption that the solvation free energy of components is additive, and that the solvation free energy of elements with similar neighbours is the same. For any given COF nanosheet with closed rings, the hybrid exterior element makes up the perimeter of the nanosheet while the hybrid interior element makes up the non-perimeter. (B) Different model systems for calculating the contribution to the solvation free energy of each element in all neighbouring configurations. On the left, the total solvation free energy of two exterior elements and two interior elements is computed. On the right, the total solvation free

energy of four interior elements is computed. (C) Additional model systems for computing single, surface, and sandwiched configurations, which are exposed to different amounts of solvent and should account for different solvation free energies.

**Table S8.** Schedule of λ coupling parameters used in Thermodynamic Integration

| Window | $\lambda_{LJ}$ | $\lambda_{coulomb}$ |
|--------|------|---------|
| 1 | 0.00 | 0.00 |
| 2 | 0.05 | 0.00 |
| 3 | 0.10 | 0.00 |
| 4 | 0.15 | 0.00 |
| 5 | 0.20 | 0.00 |
| 6 | 0.25 | 0.00 |
| 7 | 0.30 | 0.00 |
| 8 | 0.35 | 0.00 |
| 9 | 0.40 | 0.00 |
| 10 | 0.45 | 0.00 |
| 11 | 0.50 | 0.00 |
| 12 | 0.66 | 0.00 |
| 13 | 0.83 | 0.00 |
| 14 | 1.00 | 0.00 |
| 15 | 1.00 | 0.00 |
| 16 | 1.00 | 0.33 |
| 17 | 1.00 | 0.66 |
| 18 | 1.00 | 1.00 |

**Table S9.** Solvation free energies of exterior and interior elements in single, surface, and sandwiched configurations, calculated by integrating the standard errors of intermediate windows from Thermodynamic Integration (TI).

|  | $\Delta G_{solv}$ (kcal/mol) | Error (kcal/mol) |
|---|---|---|
| Single layer exterior | -19.4 | 0.6 |
| Single layer interior | -21.5 | 0.6 |
| Surface exterior | -12.7 | 0.7 |
| Surface interior | -14.7 | 0.7 |
| Sandwiched exterior | -2.4 | 0.8 |
| Sandwiched interior | -8.1 | 0.8 |
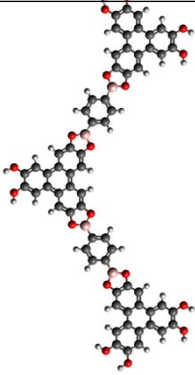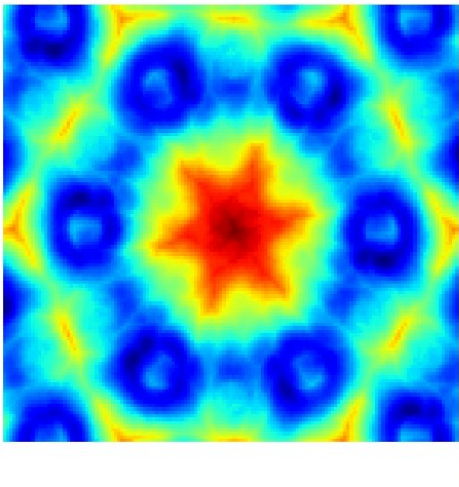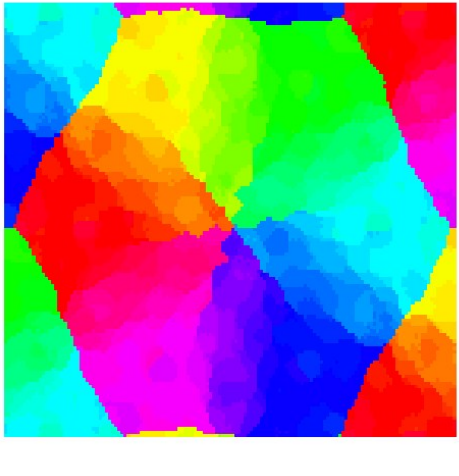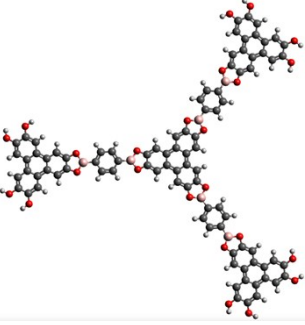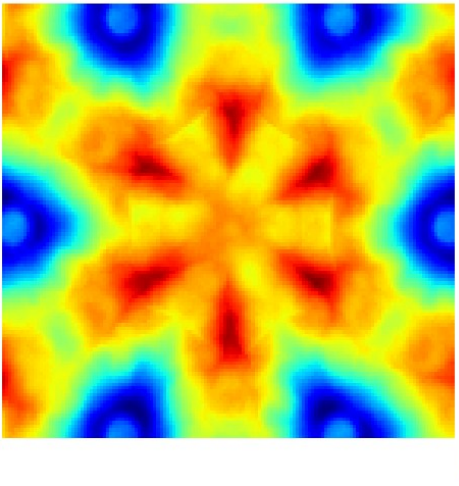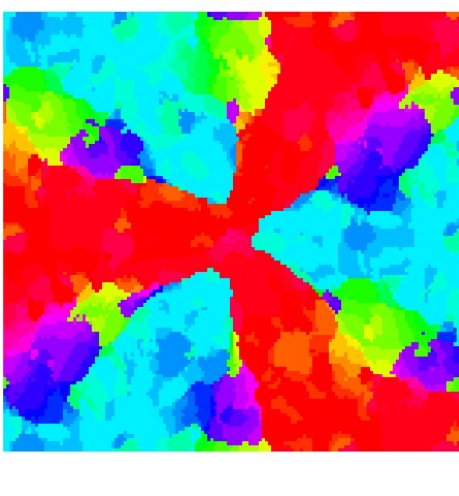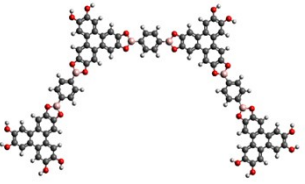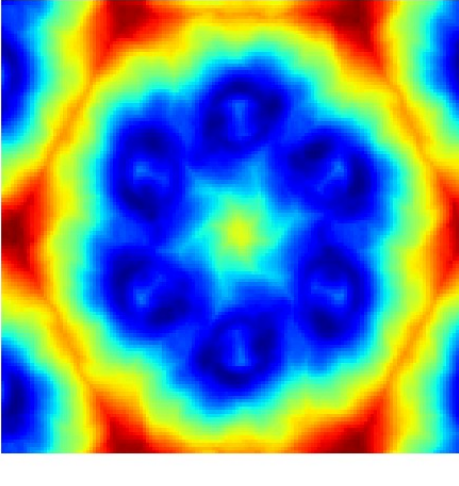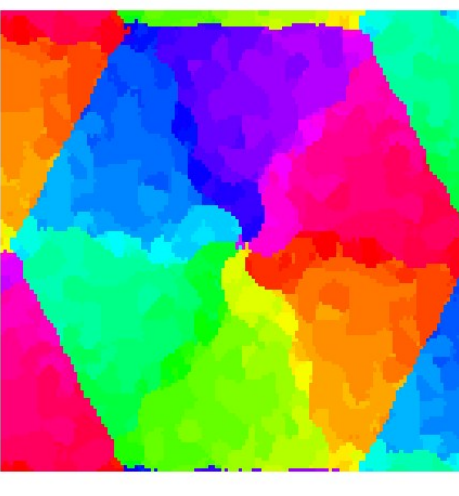
## 6. Layer-sliding energetics

To look at the ease with which stacked entities can slide relative to each other, we performed 4-dimensional potential energy scans in which various oligomers were "rastered" over a crystal surface. In the scans, the centers of mass of the oligomers were varied from an x-offset of -15.5 to 15.5 relative to the center of a pore in 0.25 Å increments, from a y-offset of -17.90 to 17.90 Å relative to the center of the pore in 0.25 Å increments, from 3.0 to 3.9 Å layer spacing from the crystal surface in increments of 0.1 Å, and from a rotation of 0° to 360° in 64 increments (except for molecule 2 in Table S10, which was only rotated from 0° to 120° in 32 increments because of its 3-fold rotational symmetry). The energy at each point was calculated in vacuum using the OPLS force field described in SI 7. For each (x-offset, y-offset) pair, the values of the angle and layer spacing for which the energy was minimized were identified, and the corresponding energies (as well as the angle) are plotted in Table S10. It is apparent that in cases 1, 3, and 4, there is a ring of low energy, where the oligomer can move relatively freely around the perimeter of the pore, but that other monomers will have substantial barriers to motion. This indicates that upon stacking, most layers will be fairly immobile, with the exception of some small fragments, which may diffuse around the perimeters of pores.
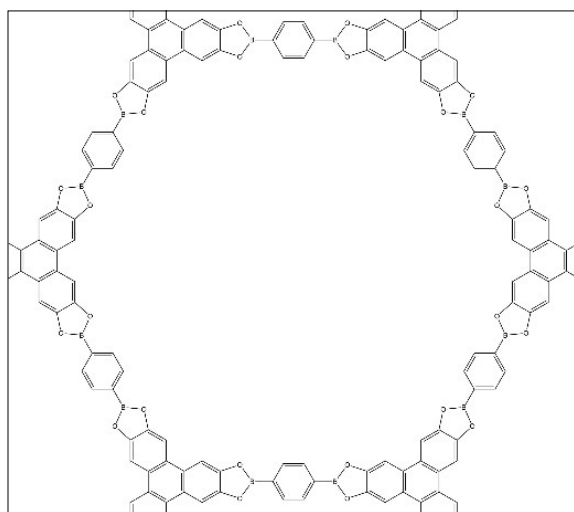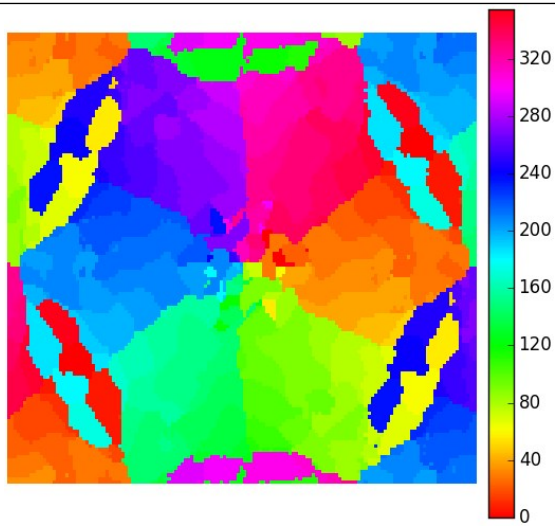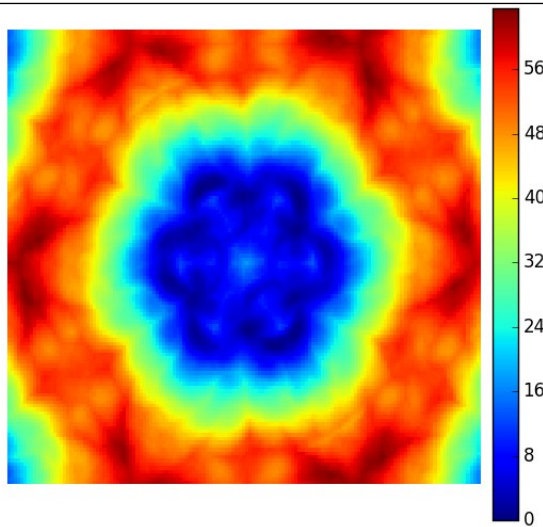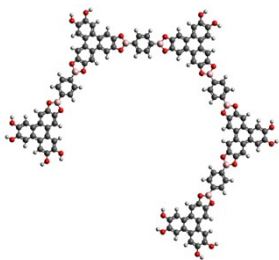


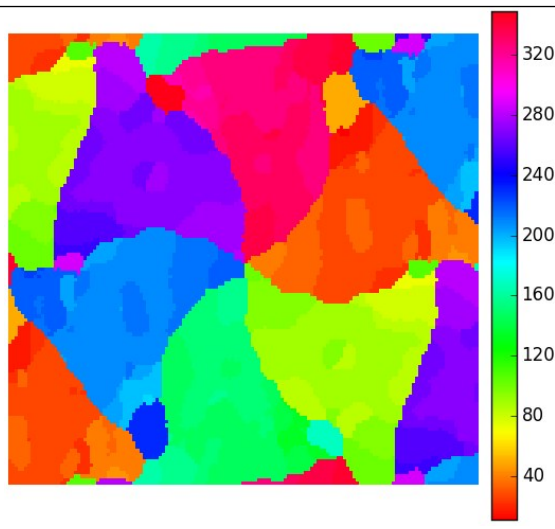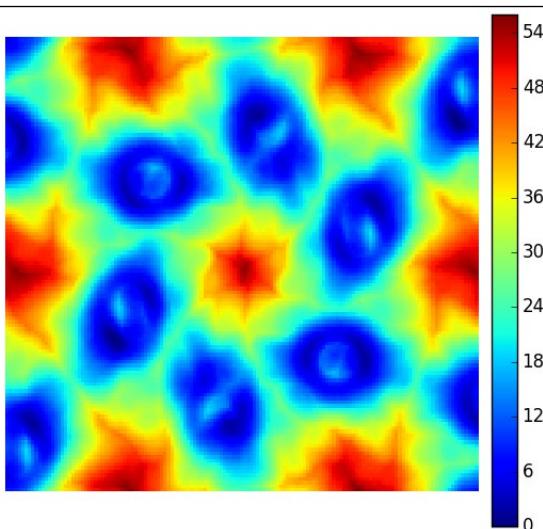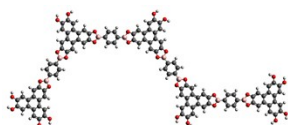**Figure S10.** Approximate bounds of potential energy scans.

**Table S10.** Potential energy surfaces for layer sliding (various oligomers on flat crystal surface)

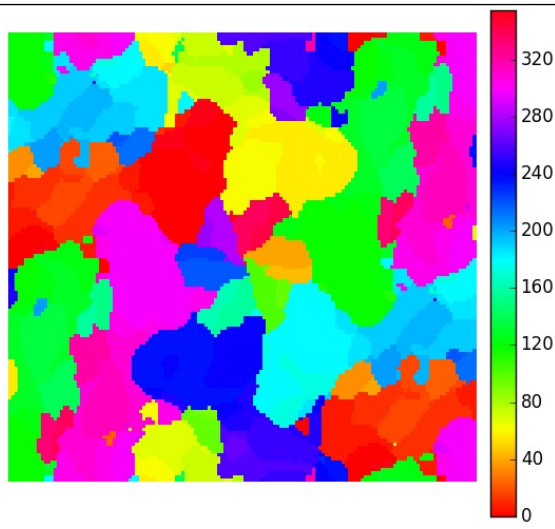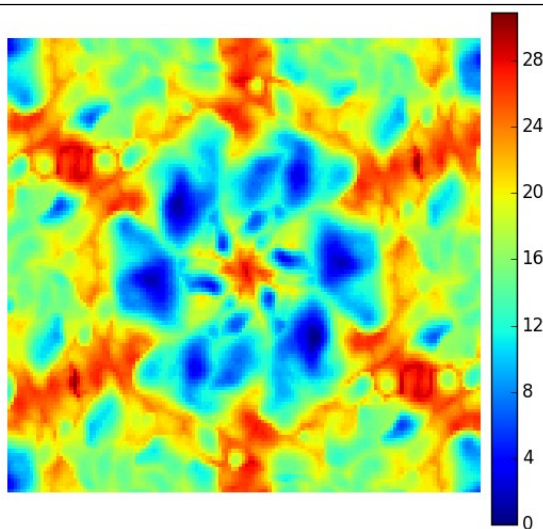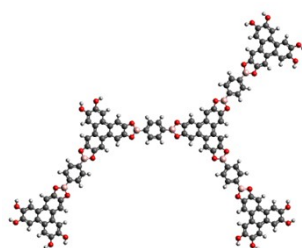| Molecule absorbed on crystal surface | Potential energy surface | |
|---|---|---|
| 1:  |  |  |
| 2:  |  |  |
| 3:  |  |  |

4:

5:

6:

# 7. OPLS Parameterization

In several of our simulations, we use an OPLS parameter set designed to model boronate ester-linked COFs. Interactions between COF atoms were modeled using custom parameters calibrated against B97-D3/def2TZVP calculations[11,27,28], whereas interactions between solvent atoms and interactions between the solvent and COF used preexisting OPLS parameters.[29,30] The atom types were designated as in Figure S11. Atom types 5 and 6 were given the standard OPLS charge values (±0.115), whereas charges for atom types 1-4 were assigned using Merz-Kollman fitting[31] on B97-D3/def2TZVP results. To determine Lennard-Jones parameters, the energies of 180 dimers of the molecule in Figure S11 were calculated using dispersion-corrected DFT. OPLS pair interaction parameters were then calibrated against the DFT energies using Nelder-Mead simplex optimization[32] on 40 different starting guesses. The objective function (a measure of how well the model reproduces the DFT data) was $E = \sum (E_{OPLS} - E_{DFT})^2 * exp(-\frac{E_{DFT}}{Q}), Q = 3 \ kcal/mol$. The parameter set used (Table S11) was selected according to the value of the objective function, and whether the parameters were realistic. Hydrogen atoms were not given Lennard-Jones parameters because when included, they would repel and distort the benzene rings unrealistically.
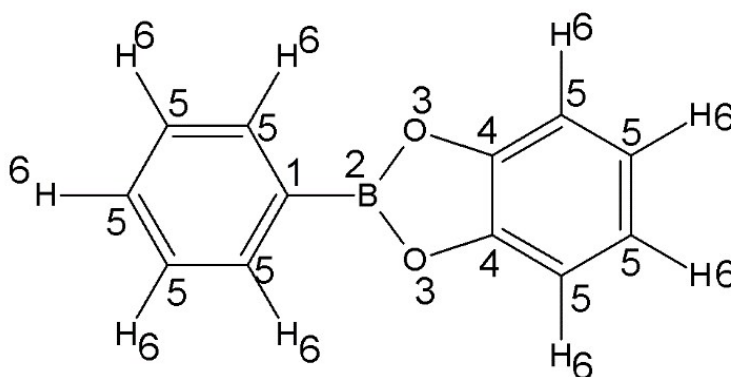


**Figure S11.** Atom types used in OPLS parameterization

**Table S11.** OPLS non-bonding parameters

| Atom Type | Charge | $\sigma$(Å) | $\epsilon$ (kcal/mol) |
|---|---|---|---|
| 1 | -0.0817 | 2.9298 | 0.0858 |
| 2 | 0.4267 | 2.8633 | 0.2497 |
| 3 | -0.4417 | 2.8829 | 0.1975 |
| 4 | 0.2691 | 3.6043 | 0.2586 |
| 5 | -0.115 | 3.2734 | 0.1234 |
| 6 | 0.115 | 0 | 0 |

Bonding parameters were calculated separately, and were calibrated against B3LYP/def2TZVP energies of 26 configurations of the molecule in Figure S11 using the Levenberg–Marquardt optimization algorithm[33] and a similar objective function. The results are tabulated in Table S12, and the forms for the energies in equations S6-8.

**Table S12.** OPLS bonding parameters

| Parameter | Value |
|---|---|
| 1-2 Bond | K=643.6 kcal/mol·Å$^2$, $r_0$=1.54Å |
| 2-3 Bond | K=594.8 kcal/mol·Å$^2$, $r_0$=1.39Å |
| 3-4 Bond† | K=900. kcal/mol·Å$^2$, $r_0$=1.42Å |
| 1-2-3 Angle | K=89.8 kcal/mol·rad$^2$, $\theta_0$=124.74 ° |
| 3-2-3 Angle | K=359.7 kcal/mol·rad$^2$, $\theta_0$=110.50° |
| 5-1-2 Angle† | K=126.0 kcal/mol·rad$^2$, $\theta_0$=120.00° |
| 3-4-4 Angle† | |
| 3-4-5 Angle† | |
| *-1-2-* Torsion | $K_2$=1.364 kcal/mol |
| *-2-3-* Torsion | $K_2$=11.855 kcal/mol |

†From Ref. 32

$$E_{bond} = \frac{K}{2}(r - r_0) \qquad\qquad \text{S6}$$

$$E_{angle} = \frac{K}{2}(\theta - \theta_0), \qquad\qquad \text{S7}$$

$$E_{torsion} = -K_1 cos(\theta) - K_2 \cos(2\theta) - K_3 cos(3\theta) \qquad\qquad \text{S8}$$

We validated the parameterization results by performing a 3D (x-offset, y-offset, and layer separation) potential energy scan. The results suggest that the layers will stack with a small offset of approximately 1.5 Å (see Figure S12), and are consistent with previous DFT scans[34]. Additionally, the minimum energy occurs at a layer separation of between 3.4 and 3.5 Å, which is consistent with X-ray diffraction measurements.[35]
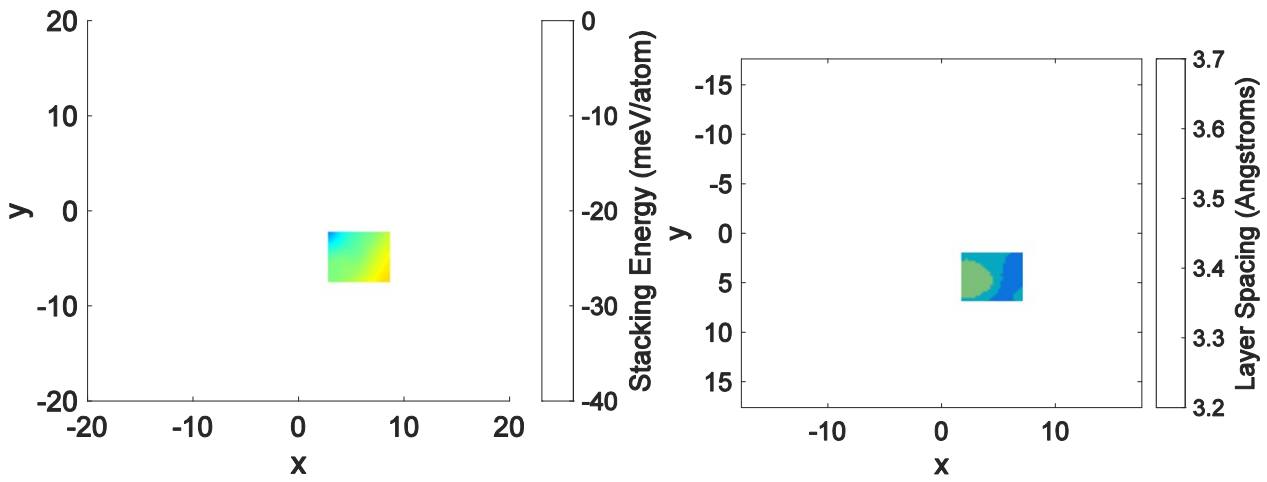
**Figure S12.** Potential Energy (left) and separation of minimal energy (right) as a function of offset between layers.

# REFERENCES

1 B. Lukose, A. Kuc, J. Frenzel and T. Heine, *Beilstein J. Nanotechnol.*, 2010, 1, 60-70.
2 B. Lukose, A. Kuc and T. Heine, *Chem. Eur. J.*, 2011, 17, 2388-2392.
3 M. Elstner and G. Seifert, *Phil. Trans. R. Soc. A*, 2014, 372, 20120483-20120483.
4 L. Zhechkov, T. Heine, S. Patchkovskii, G. Seifert and H. Duarte, *J. Chem. Theory Comput.*, 2005, 1, 841-847.
5 A. Becke, *J. Chem. Phys.*, 1993, 98, 5648.
6 P. Hariharan and J. Pople, *Theor. Chim. Acta*, 1973, 28, 213-222.
7 M. Dewar, E. Zoebisch, E. Healy and J. Stewart, *J. Am. Chem. Soc.*, 1985, 107, 3902-3909. (8)
8 Stewart, J. Optimization Of Parameters For Semiempirical Methods I. Method. *J. Comput. Chem*. 1989, 10, 209-220.
9 J. Stewart, *J. Mol. Model.*, 2007, 13, 1173-1213.
10 C. Peng and H. Bernhard Schlegel, *Isr. J. Chem.*, 1993, 33, 449-454.
11 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, Montgomery, Jr., J. A., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, GAUSSIAN 09, Revision E.01, Gaussian Inc., Pittsburgh, PA, 2009.
12 K. Fukui, *Acc. Chem. Res.*, 1981, 14, 363-368.
13 S. Mayo, B. Olafson and W. Goddard, *J. Phys. Chem.*, 1990, 94, 8897-8909.
14 B. Li, S. Wei and L. Chen, *Mol. Simul.*, 2011, 37, 1131-1142.
15 R. Krishnan, J. Binkley, R. Seeger and J. Pople, *J. Chem. Phys.*, 1980, 72, 650.
16 S. Park and K. Schulten, *J. Chem. Phys*, 2004, 120, 5946.
17 S. Park, F. Khalili-Araghi, E. Tajkhorshid and K. Schulten, *J. Chem. Phys.*, 2003, 119, 3559.
18 A. Ma and A. Dinner, *J. Phys. Chem. B*, 2005, 109, 6769-6779.
19 B. Peters, G. Beckham and B. Trout, *J. Chem. Phys.*, 2007, 127, 034109.
20 B. Peters and B. Trout, *J. Chem. Phys.*, 2006, 125, 054108.
21 B. Knott, M. Haddad Momeni, M. Crowley, L. Mackenzie, A. Götz, M. Sandgren, S. Withers, J. Ståhlberg and G. Beckham, *J. Am. Chem. Soc.*, 2014, 136, 321-329.
22 B. Peters, N. Zimmermann, G. Beckham, J. Tester and B. Trout, *J. Am. Chem. Soc.*, 2008, 130, 17342-17350.
23 R. Radhakrishnan and T. Schlick, *J. Chem. Phys.*, 2004, 121, 2436.
24 S. Kumar, J. Rosenberg, D. Bouzida, R. Swendsen and P. Kollman, *J. Comput. Chem.*, 1992, 13, 1011-1021.
25 J. Hub, B. de Groot and D. van der Spoel, *J. Chem. Theory Comput.*, 2010, 6, 3713-3720.
26 P. Virnau and M. Müller, *J. Chem. Phys.*, 2004, 120, 10925.
27 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, 32, 1456-1465.
28 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, 7, 3297.
29 X. Chen, L. Bartolotti, K. Ishaq and A. Tropsha, *J. Comput. Chem.*, 1994, 15, 333-345.
30 W. Jorgensen, D. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, 118, 11225-11236.
31 B. Besler, K. Merz and P. Kollman, *J. Comput. Chem.*, 1990, 11, 431-439.
32 J. Nelder and R. Mead, *Comput. J.*, 1965, 7, 308-313.
33 D. Marquardt, *J. Soc. Ind. Appl. Math.*, 1963, 11, 431-441.
34 B. Koo, W. Dichtel and P. Clancy, *J. Mater. Chem.*, 2012, 22, 17460.
35 J. Colson, A. Woll, A. Mukherjee, M. Levendorf, E. Spitler, V. Shields, M. Spencer, J. Park and W. Dichtel, *Science*, 2011, 332, 228-231.