

Electronic Supplementary Information

Oligomer-assisted self-assembly of bisurea in organic solvent media

Ching-Hung Wu,^a Ling-Hua Huang^a and Chi-Chung Hua^{*a}

^a*Department of Chemical Engineering, National Chung Cheng University, Chia-Yi 62102, Taiwan*

S1. Description and validation of force fields

The atomic interactions of the bisurea, polyester, and *ortho*-xylene in this study are described by the OPLS (Optimized Potentials for Liquid Simulations) force field¹ and united-atom model, with partial atomic charges adopted from prior work by Cannon *et al.*² for bisurea, OPLS force field³ for polyester, and TraPPE force field^{4,5} for *ortho*-xylene; see chemical structures and charge distributions in Fig. S1. The properties of *ortho*-xylene were thoroughly investigated against known density, self-diffusion coefficient, and viscosity at 1 bar and 298 K. The results, 0.878 g cm⁻³, 20.2×10⁻¹⁰ m² s⁻¹, and 0.59 cp, respectively, are in excellent agreement with the reported values (0.876 g cm⁻³,⁶ 17.5×10⁻¹⁰ m² s⁻¹,⁷ and 0.60 cp⁶). For the polyester and bisurea in this study, since no previous study has focused on the same chemical structures, methyl acetate and 1,3-dimethylurea, respectively, were utilized as the model molecules to help validate the force field. Because hydrogen bonding (HB) interactions play an essential role in dictating the self-assembly behavior of bisureas in solvent media, the net dipoles of methyl acetate and 1,3-dimethylurea in

vacuum were first evaluated, and the results, 1.9 D and 5.2 D, respectively, are in excellent agreement with the reported values (1.7 D⁸ and 5.1 D²). In a further evaluation, the density and viscosity of methyl acetate at 1 bar and 298 K were found to be 0.896 g cm⁻³ and 0.36 cp, respectively, in good agreement with the reported values (0.928 g cm⁻³ and 0.39 cp).⁹ Finally, the crystal structure of 1,3-dimethylurea was investigated under ambient conditions (1 bar and 298 K); see Fig. S2. The lattice parameters $a = 4.3 \text{ \AA}$, $b = 12.0 \text{ \AA}$, and $c = 9.5 \text{ \AA}$ are in excellent agreement with early experimental characterization.¹⁰ While the classical simulations using non-polarizable force fields may occasionally underestimate the strength of HB interactions,^{11,12} the OPLS force field has been found to outperform other classical force fields, i.e., MM2, MM3, and AMBER, in the descriptions of atomic interactions of typical HB molecule pairs, and is comparable to the *ab initio* calculations for the computed stabilization energies and geometries.¹² Moreover, the OPLS force field has previously been used to investigate the stability of self-assembly structures of a bisurea in organic solvent media,¹¹ and good agreement with experimental observations suggests that the effects of HB have been well captured. Overall, the use of carefully verified force field and parameter values in this study should lend credible simulation results reported in the main text.

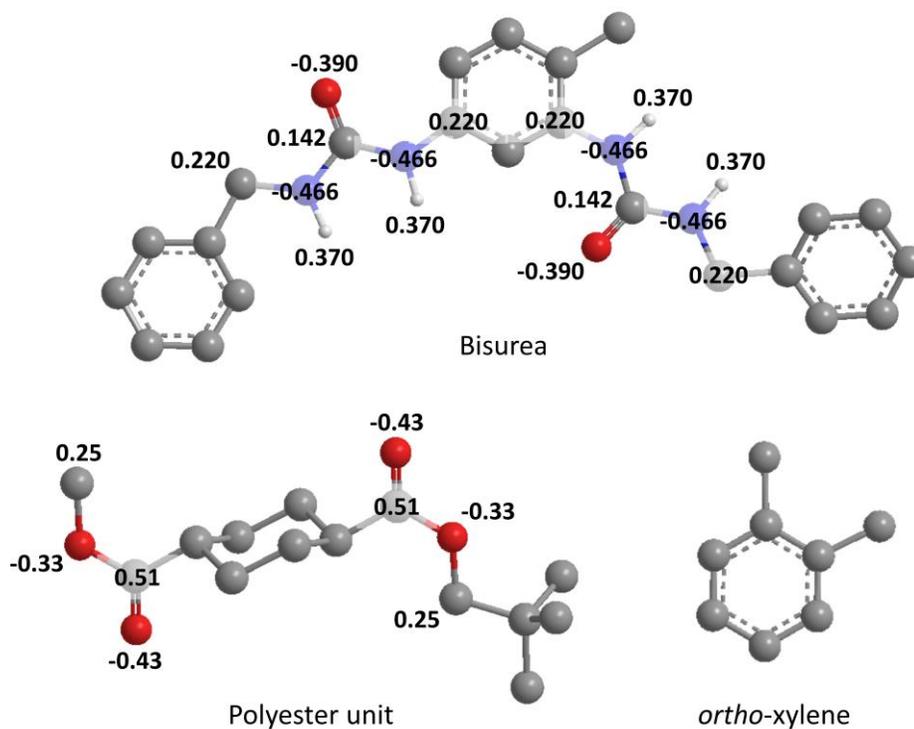


Fig. S1 Chemical structures of the bisurea (2,4-toluene diisocyanate-benzylamine, TDI-BA), polyester unit, and *ortho*-xylene molecules in this study, along with the distributions of partial atomic charge in unit of elementary charge (e). The atoms without labeling have zero partial atomic charge.

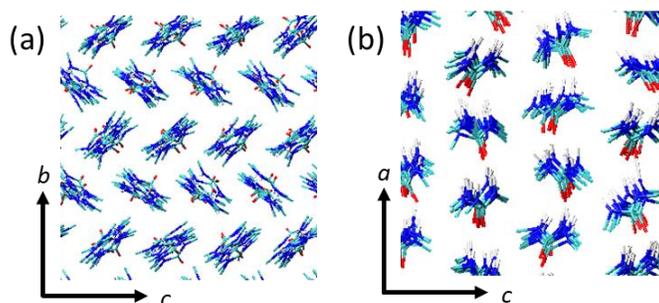


Fig. S2 Snapshots of the crystal structure of 800 1,3-dimethylurea molecules from the AMD simulation performed at 298 K and 1 bar after 5 ns of the simulation: projection of the molecules parallel to the (a) bc and (b) ac planes.

S2. Simulation protocols

To investigate molecular self-assembly in organic solvent media, the bisurea molecules in this study were restrained by weak spring forces during the simulation to keep them from roaming freely in a typically dilute system. The procedure is as follows: Seven bisurea molecules in a row were initially separated by ~ 1 nm for all adjacent pairs, and the full chain of bisurea was fixed at the center of a cubic box consisting of 28 polyester oligomers ($n = 7$) and 455 *ortho*-xylene molecules (1:1 in weight ratio). The system was first equilibrated using the *NPT* ensemble (with a coupling constant of 0.1 ps for both Nosé-Hoover thermostat and MTTK barostat) at 298 K and 400 bar for a duration of 1 ns to accelerate the equilibration process. A thermal annealing process using the *NVT* ensemble was then carried out by gradually increasing the system temperature from 298 K to 800 K at a rate of 100 K ns^{-1} , followed by an equilibration at 800 K for 10 ns and, afterward, a gradual quenching from 800 K back to 298 K at a rate of -100 K ns^{-1} . At this point, the polyester oligomers and *ortho*-xylene molecules were randomly dispersed in the simulation system. The simulation was then carried out under ambient conditions (298 K and 1 bar) for another 300 ns to further relax the molecular structures that might have been affected by the prior thermal and pressure treatment. The final equilibration time adopted here (i.e., 300 ns) was determined by the time convergence of the following dimensionless parameter:¹³

$$\psi = \frac{G_{ii} + G_{jj}}{2G_{ij}} \quad (\text{S1})$$

In the above expression, i and j denote *ortho*-xylene and a monomeric unit of polyester, respectively, and G_{ij} is defined by the following integration of the partial radial distribution function (RDF):

$$G_{ij} = \int_0^{R_0} g_{ij}(r) dr \quad (\text{S2})$$

where g_{ij} is the RDF of the i - j pair, and R_0 is set to 3 nm. Fig. S3 shows that time convergence of ψ was achieved at an equilibration time of ~ 200 ns, and thus 300 ns was selected as the final equilibration time.

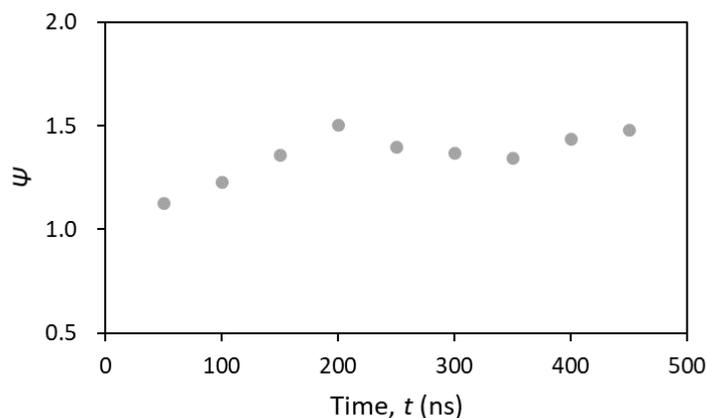


Fig. S3 Time evolution of the dimensionless parameter ψ .

At the end of the equilibration described above, the system size (~ 5.4 nm) is considerably larger than the mean end-to-end distance of a polyester oligomer (~ 3.2 nm) plus the cutoff distance, thereby preventing any self-interaction of the oligomer via periodic boundary conditions. It should also be noted that most N–H units of the bisurea molecules have now been captured by the C=O or C–O–C units of polyesters. Henceforth, the (*previously fixed*) adjacent pair of bisurea molecules was restrained by weak spring force with a force constant of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ (see Fig. S4). Meanwhile, the separation distance was gradually reduced from ~ 1.0 nm to ~ 0.42 nm with a pulling rate of $10^{-5} \text{ nm ps}^{-1}$ during a period of 60 ns to expedite the self-assembly process. The minimum separation distance (i.e., 0.42 nm) was determined by the average distance between two consecutive bisurea molecules in a stable filament, which is in excellent agreement with previous experimental¹⁴ and computational¹⁵ estimates. Afterward, the bisurea molecules were kept at the

same separation distance for a relatively long simulation time of 1000 ns, during which hydrogen bonds were formed between bisurea molecules. The final self-assembly structures for both binary and ternary systems were basically independent of the pulling rate and force constant utilized, as basically identical results were observed with a slower pulling rate of 10^{-6} nm ps $^{-1}$ and a smaller force constant of 100 kJ mol $^{-1}$ nm $^{-2}$.

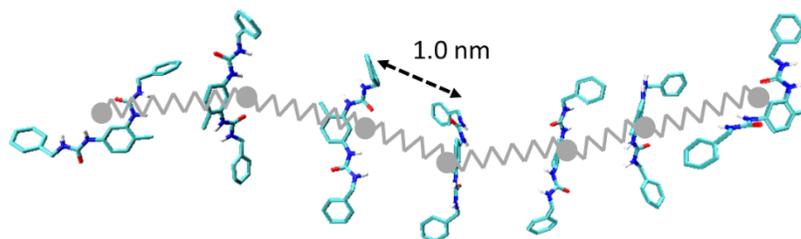


Fig. S4 Schematic illustration of the setup for restraining all adjacent bisurea molecules by weak spring forces, with an initial separation distance of 1.0 nm.

S3. Dynamics of hydrogen-bond formation and dissociation in the ternary system

The requirements of H \cdots O distance < 0.3 nm and N–H \cdots O angle $> 130^\circ$ must be simultaneously met to be regarded as effective hydrogen-bond formation.¹⁶⁻²⁰ The average hydrogen-bond length is found to be 0.219 nm, in good agreement with previous results.^{21,22} Fig. S5 presents the time-dependent counts of HB for the bisurea-polyester and bisurea-bisurea pairs, with $t = 0$ denoting the beginning when the separation distance between adjacent bisurea pairs was restrained at 0.42 nm. From then on, the dissociation of the bisurea-polyester pair and the association between the bisurea-bisurea pair should be evident.

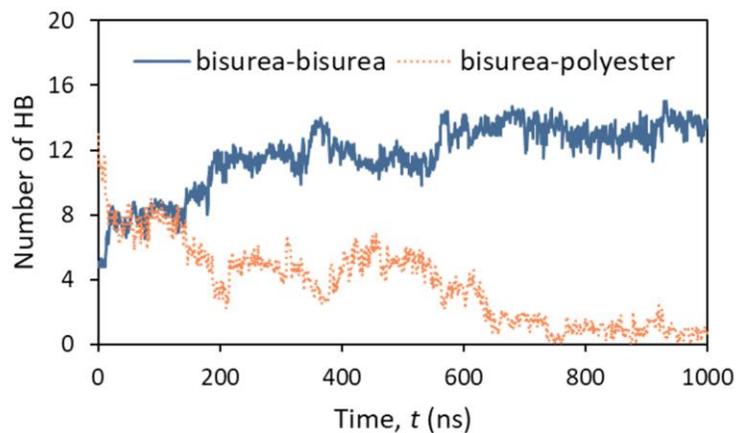


Fig. S5 Time-dependent counts of HB for the bisurea-polyester and bisurea-bisurea pairs. Note that most HB for the bisurea-polyester pair takes place in the later stage of equilibration (i.e., $t < 0$ in this plot).

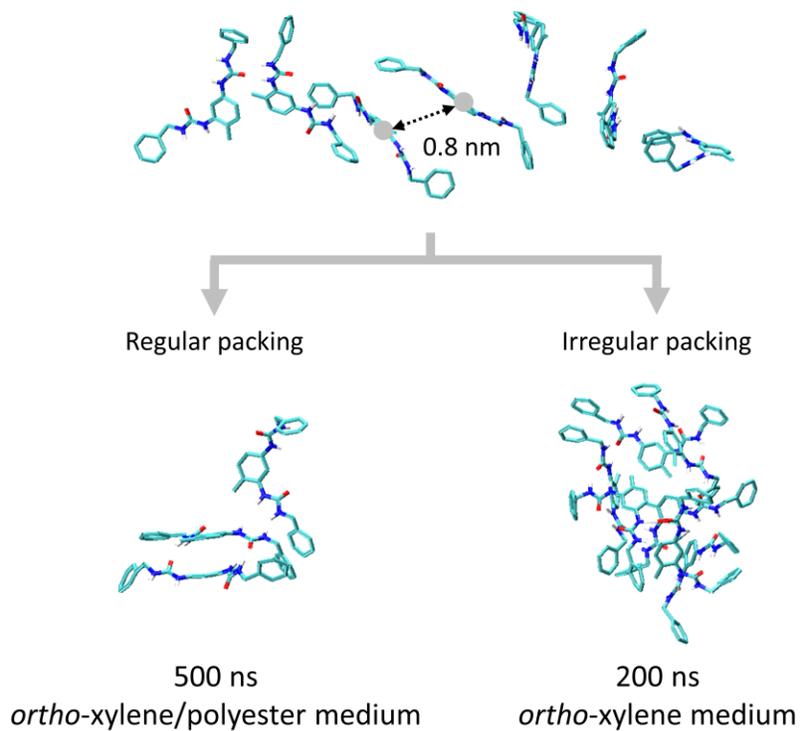


Fig. S6 Comparison of bisurea associations without imposing restraints for the binary (right) and ternary (left) systems.

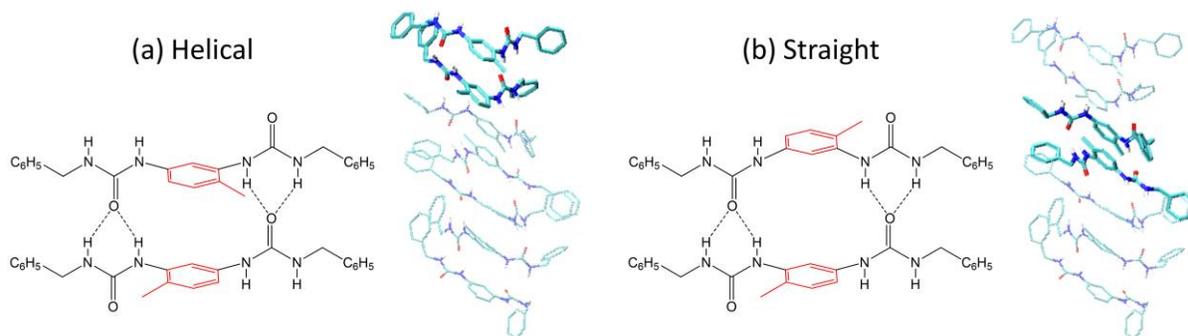


Fig. S7 Illustration of two different self-assembly conformations in the bisurea filament formed in the ternary system. In the case presented here, only 2 out of the 6 pairs belong to the helical conformation.

S4. Evaluation of diffusivities of bisurea and *ortho*-xylene

The mass-center diffusivities (D) of the bisurea and *ortho*-xylene molecules, respectively, in various systems can be computed by the time integration of the (local) velocity autocorrelation function (VAF):

$$D = \frac{1}{3} \int_0^{\infty} \langle v(t) \cdot v(0) \rangle dt \quad (\text{S3})$$

where $v(t)$ and $v(0)$ are the center-of-mass velocities of the probed molecule at times t and 0, respectively. Fig. S8a and b show the results for the *ortho*-xylene and bisurea molecules, respectively, and the diffusivity in each case is estimated when the value of D reaches a plateau at times 4~5 ps.

More often, the diffusivities are evaluated through the time-dependent mean-square displacements (MSDs):

$$\text{MSD} = \langle |r(t) - r(0)|^2 \rangle = 6Dt \quad (\text{S4})$$

where $r(t)$ and $r(0)$ denote the positions of the mass center of an *ortho*-xylene molecule at times t and 0, respectively. Fig. S9 shows the simulation data on MSD vs. t for *ortho*-xylene in two different systems, and the diffusivity in each case is estimated when MSD/t reaches a plateau at times 10~20 ns.

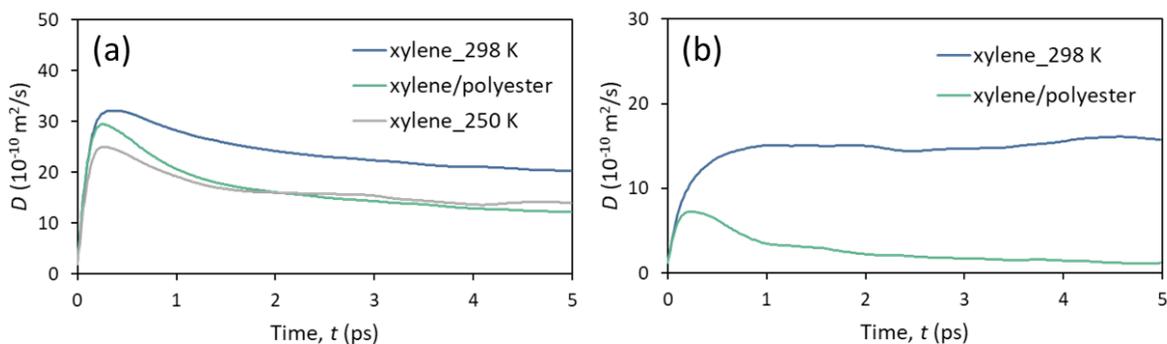


Fig. S8 Computed diffusivities as a function of the integration time for (a) *ortho*-xylene in the bulk system at two different temperatures (298 K (blue), 250 K (grey)) and in the *ortho*-xylene/polyester system at 298 K (green); (b) bisurea in the binary (blue) and ternary (green) system at 298 K, where no restraints are imposed on the bisurea molecule.

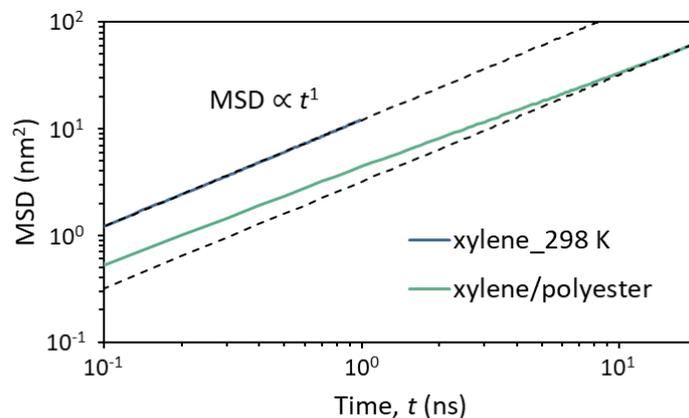


Fig. S9 Time-dependent MSDs of *ortho*-xylene in the bulk (blue) and *ortho*-xylene/polyester (green) system at 298 K, with the dashed lines representing the fits from eqn S4.

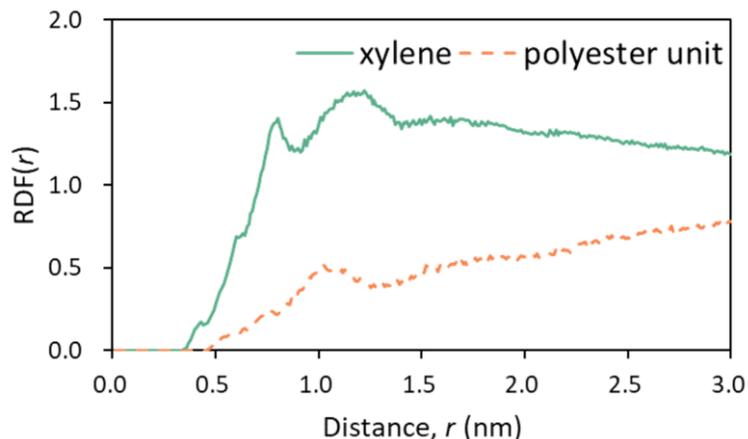


Fig. S10 Radial distribution functions showing the solvent and oligomer distribution near the self-assembly filament.

S5. Calculation of binding free energies

A variety of computational schemes including thermodynamic integration (TI),²³ free energy perturbation (FEP),²⁴ and umbrella sampling (US)²⁵ can be used to calculate binding free energies. Among them, US is commonly used to compute the potential of mean force (PMF) based on rigorous probability calculations along a given coordinate. It requires a well-defined reaction coordinate (ζ) represented by one or a few variables. Intensive conformational sampling is then performed by enforcing external biasing potentials at the conformations along ζ within a series of overlapping windows. Finally, the unbiased PMF can be reconstructed by removing the biasing potentials. The external biasing potential $u_i(\zeta)$ at the window i is a harmonic function:

$$u_i(\zeta) = \frac{K}{2} (\zeta - \zeta_i^{\text{ref}})^2 \quad (\text{S5})$$

where ζ_i^{ref} is the reference position of the respective window i , and K is the spring force constant.

The weighted histogram analysis method (WHAM) was used to remove the biasing potentials and

reconstruct the PMF.²⁶ The methodical details of US have been thoroughly derived and described in Kästner's classic work.²⁷ Recently, we have described a systematic and credible method for determining the protocol parameters for the US simulation.²⁸

The US and WHAM simulations were performed using the GROMACS package (version 2018.1).²⁹ We began with hydrogen-bonded configurations of bisurea-bisurea and bisurea-ester pairs, respectively, solvated in the *ortho*-xylene medium. The hydrogen-bonded configurations were first equilibrated using the *NPT* ensemble (with a coupling constant of 0.1 ps for both Nosé-Hoover thermostat and MTTK barostat) at 298 K and 1 bar for a duration of 1 ns. The reaction coordinate was defined as the separation distance between the mass centers of two bisurea molecules or those of bisurea and ester, extending from $\zeta = 0.1$ nm to $\zeta = 1.6$ nm. The simulation box size, ~ 5.4 nm, was set to be at least twice larger than the maximum separation distance in order to prevent the effect of overlapping due to the periodic boundary condition. Accordingly, the simulation box contains 769 *ortho*-xylene molecules. Different configurations along the reaction pathway were obtained using the GROMACS pull code, with a pulling rate of 0.001 nm ps⁻¹.^{28,30} For each umbrella window, the system was first equilibrated for 5 ns at 298 K and 1 bar. As the PMF reaches time convergence, another 15 ns of simulation was utilized for the sampling. A harmonic force constant for the biasing potential was set to be 1000 kJ mol⁻¹ nm⁻²,²⁸ and the window width ($\Delta\zeta$) is 0.1 nm.^{28,31,32} The reaction pathways were sampled in three independent runs, with different initial atoms' positions and velocities to ensure a comprehensive sampling.³³

In the WHAM simulation, the bin number is set to be 200, the tolerance for iteration is 10^{-6} , and the temperature is 298 K. To exclude a purely entropic contribution to the PMF due to the rotation of two restrained groups, the PMF was corrected by a factor $k_B T \ln(4\pi\zeta^2)$, which removes the entropic decrease in the PMF due to the increase in the number of configurations on a sphere

of radius ζ .^{26,34} Fig. S11 presents the results of PMF for the bisurea-bisurea and bisurea-ester pairs in the *ortho*-xylene medium. The binding free energies were determined by the well depths to be -36 ± 2 and $-8 \pm 1 k_B T$, respectively, where k_B is the Boltzmann constant and $T = 298$ K.

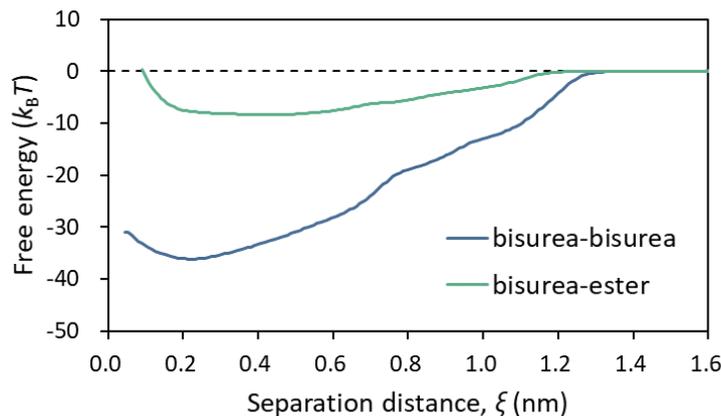


Fig. S11 PMFs for the bisurea-bisurea (blue) and bisurea-ester (green) pairs in the *ortho*-xylene medium.

S6. Persistence length (l_p) analysis

Pre-arranged, non-restrained, and ordered bisurea filaments comprised of 30 bisurea molecules were initially fixed at the center of a rectangular box consisting of 152 polyester oligomers ($n = 7$) and 2466 *ortho*-xylene molecules for the ternary system, and 4176 *ortho*-xylene molecules for the binary system. The system was first equilibrated using the *NPT* ensemble (with a coupling constant of 0.1 ps for both Nosé-Hoover thermostat and MTTK barostat) at 298 K and 400 bar for a duration of 1 ns to accelerate the equilibration process. A thermal annealing process using the *NVT* ensemble was then carried out by gradually increasing the system temperature from 298 K to 800 K at a rate of 100 K ns⁻¹, followed by a further equilibration at 800 K for 10 ns and, afterward, a gradual quenching from 800 K back to 298 K at a rate of -100 K ns⁻¹. The simulation was then

carried out under ambient conditions (298 K and 1 bar) for 300 ns to further relax the molecular structures that might have been affected by the prior thermal and pressure treatment. At the end of the equilibration described above, the system size was about $7.6 \times 7.6 \times 15.1 \text{ nm}^3$. The (*previously fixed*) bisurea filament is then relaxed and undergoes equilibration for a period of 50 ns and 100 ns for the binary and ternary systems, respectively. The same time intervals were subsequently used to collect data for the analysis of persistence length in the individual systems, based on three independently prepared systems for obtaining time and ensemble averages, with configurations that self-interact via periodic boundary conditions discarded.

According to Flory, l_p is the average projection of the end-to-end distance onto the first bond of the chain.³⁵ Thus, it can be calculated as the sum of projection of all bond vectors onto the first bond vector:

$$l_p = \left\langle \sum_{i=1}^{N-1} \vec{l}_i \cdot \vec{l}_1 \right\rangle / \langle l \rangle \quad (\text{S6})$$

where N and l is the total number of repeated units and the bond length, respectively. This method is suitable for all chain models and independent of the state of the chains, i.e., self-avoiding chains or ideal (Gauss) chains.^{36,37}

References

- 1 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225-11236.
- 2 W. R. Cannon, J. D. Madura, R. P. Thummel and J. A. McCammon, *J. Am. Chem. Soc.*, 1993, **115**, 879-884.
- 3 G. Rossi, I. Giannakopoulos, L. Monticelli, N. K. J. Rostedt, S. R. Puisto, C. Lowe, A. C. Taylor, I. Vattulainen and T. Ala-Nissila, *Macromolecules*, 2011, **44**, 6198-6208.
- 4 C. D. Wick, M. G. Martin and J. I. Siepmann, *J. Phys. Chem. B*, 2000, **104**, 8008-8016.
- 5 N. Rai and J. I. Siepmann, *J. Phys. Chem. B*, 2013, **117**, 273-288.
- 6 D. R. Lide, *CRC Handbook of Chemistry and Physics, 90th Edition*, Taylor & Francis, 2009.
- 7 B. Rousseau and J. Petracic, *J. Phys. Chem. B*, 2002, **106**, 13010-13017.
- 8 M. L. P. Price, D. Ostrovsky and W. L. Jorgensen, *J. Comput. Chem.*, 2001, **22**, 1340-1352.
- 9 T. M. Aminabhavi and K. Banerjee, *J. Chem. Eng. Data*, 1998, **43**, 514-518.
- 10 C. Näther, C. Döring, I. Jess, P. G. Jones and C. Taouss, *Acta Crystallogr. B*, 2013, **69**, 70-76.
- 11 B. G. Alvarenga, K. Bernardino, A. F. de Moura and E. Sabadini, *J. Mol. Model.*, 2018, **24**, 154-163.
- 12 R. S. Paton and J. M. Goodman, *J. Chem. Inf. Model.*, 2009, **49**, 944-955.
- 13 C. Li and A. Strachan, *Polymer*, 2018, **149**, 30-38.
- 14 B. Isare, S. Pensec, M. Raynal and L. Bouteiller, *C. R. Chim.*, 2016, **19**, 148-156.
- 15 P. Brocorens, M. Linares, C. Guyard-Duhayon, R. Guillot, B. Andrioletti, D. Suhr, B. Isare, R. Lazzaroni and L. Bouteiller, *J. Phys. Chem. B*, 2013, **117**, 5379-5386.
- 16 H. A. Karimi-Varzaneh, P. Carbone and F. Müller-Plathe, *Macromolecules*, 2008, **41**, 7211-7218.
- 17 H. A. Karimi-Varzaneh, P. Carbone and F. Müller-Plathe, *J. Chem. Phys.*, 2008, **129**, 154904.
- 18 I.-C. Yeh, B. C. Rinderspacher, J. W. Andzelm, L. T. Cureton and J. La Scala, *Polymer*, 2014, **55**, 166-174.
- 19 T. L. Chantawansri, I.-C. Yeh and A. J. Hsieh, *Polymer*, 2015, **81**, 50-61.
- 20 R. J. Gowers and P. Carbone, *J. Chem. Phys.*, 2015, **142**, 224907.
- 21 L. Born and H. Hespe, *Colloid & Polymer Sci.*, 1985, **263**, 335-341.
- 22 E. Yildirim and M. Yurtsever, *Comput. Theor. Chem.*, 2014, **1035**, 28-38.
- 23 A. P. Bhati, S. Wan, D. W. Wright and P. V. Coveney, *J. Chem. Theory Comput.*, 2017, **13**, 210-222.
- 24 M. Aldeghi, A. Heifetz, M. J. Bodkin, S. Knapp and P. C. Biggin, *Chem. Sci.*, 2016, **7**, 207-218.
- 25 I. Buch, S. K. Sadiq and G. De Fabritiis, *J. Chem. Theory Comput.*, 2011, **7**, 1765-1772.
- 26 J. S. Hub, B. L. de Groot and D. van der Spoel, *J. Chem. Theory Comput.*, 2010, **6**, 3713-3720.

- 27 J. Kästner, *WIREs Comput. Mol. Sci.*, 2011, **1**, 932-942.
- 28 C. H. Wu and C. C. Hua, *Macromolecules*, 2022, **55**, 5382-5389.
- 29 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1-2**, 19-25.
- 30 J. A. Lemkul and D. R. Bevan, *J. Phys. Chem. B*, 2010, **114**, 1652-1660.
- 31 R. Menichetti, K. Kremer and T. Berau, *Biochem. Biophys. Res. Commun.*, 2018, **498**, 282-287.
- 32 A. Masunov and T. Lazaridis, *J. Am. Chem. Soc.*, 2003, **125**, 1722-1730.
- 33 W. You, Z. Tang and C.-e. A. Chang, *J. Chem. Theory Comput.*, 2019, **15**, 2433-2443.
- 34 R. M. Neumann, *Am. J. Phys.*, 1980, **48**, 354-357.
- 35 P. J. Flory, *Statistical Mechanics of Chain Molecules*, Interscience, New York, USA, 1969.
- 36 P. Cifra, *Polymer*, 2004, **45**, 5995-6002.
- 37 J.-Z. Zhang, X.-Y. Peng, S. Liu, B.-P. Jiang, S.-C. Ji and X.-C. Shen, *Polymers*, 2019, **11**, 295-306.