Supporting Information

Atomistic Structure Generation of Covalent Triazine-Based Polymers by Molecular Simulation

Ce Song,[ac] Fangyuan Hu,*[b] Zhaoliang Meng,[a] Shengming Li,[d] Wenlong Shao,[c] Tianpeng Zhang,[b] Siyang Liu,[b] and Xigao Jian*[abc]

[a] *School of Mathematical Sciences, Dalian University of Technology, No.2 Linggong Road, Gaoxin District, Dalian, Liaoning, China. E-mail: jian4616@dlut.edu.cn*

[b] *School of Materials Science and Engineering, State Key Laboratory of Fine Chemicals, Dalian University of Technology, No.2 Linggong Road, Gaoxin District, Dalian, Liaoning, China. E-mail: hufangyuan@dlut.edu.cn*

[c] *State Key Laboratory of Fine Chemicals, Liaoning Province Engineering Research Centre of High Performance Resins, Dalian University of Technology, No.2 Linggong Road, Gaoxin District, Dalian, Liaoning, China.*

[b] *School of Innovation and Entrepreneurship, Dalian University of Technology, No.2 Linggong Road, Gaoxin District, Dalian, Liaoning, China*

Table of content

1. Final modification process for the simulated network.

The unlinked building units are removed from the system at first. Consequently, there remains two kinds of unlinked L2 atoms (displayed in Fig. S1 (a)) in the system, which will be treated separately. Fig. S1 demonstrates the modification process diagrammatically. Specifically, when these two kinds of unlinked L2 atoms are detected in the present system, the atoms in the grey regions displayed in Fig. S1 will be removed from the system, and then the remained -CN units are deemed as the unreacted cyano groups. Additionally, H atoms bonded to the unlinked L1 atoms will be replaced by the cyano groups. Moreover, the corresponding force field type needs further adjustment according to the chemical structures and the applied force field. Finally, the monomers separated from the network would be removed. The whole modification process is automatically conducted by an in-house Perl script.

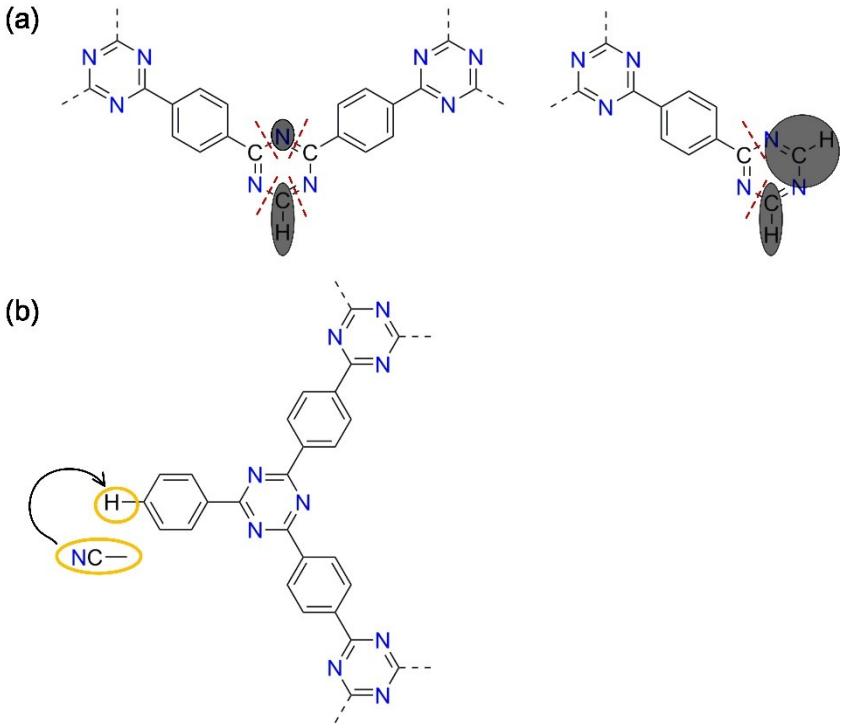

Fig. S1 Diagram of the final modification process for the simulated network.

2. Details of the modeling algorithm and the simulation method.

The initial simulation boxes were generated by the Amorphous Cell module embedded in BIOVIA Materials Studio 2017R2. 300 monomer building units and 200 triazine rings (displayed in Fig. 1) are randomly packed in the simulation cell at a density of 0.4 g/cm^3 for the starting configuration. Therefore, the desired number of bond formations for the linking cycles is set as 600 herein.

Both the geometry optimizations and MD simulations involved in this study are performed using the Forcite module with COMPASS II force field.[1] The geometry optimizations during the linking cycles and after the final modification process consist of 1000 and 2000 steps, respectively, with fine quality settings and the smart algorithm applied for energy minimization calculations. The designed geometry optimization steps have been proved to be capable of optimizing the formed bonds during the linking cycles and the local structures altered by the final modification process to the realistic structures. Specifically, the capability of the geometry optimization steps on optimizing and relaxing the networks is assessed by the length distributions of specific bonds. Fig. S2 and Fig. S3 show the length distributions of the C-C bonds formed during the linking cycles and the C-N bonds in the cyano groups generated during the final modification process (both displayed in the subplots and highlighted in yellow) after the geometry optimization steps, respectively. The narrow distributions around the equilibrium bond lengths manifest the efficiency of the designed geometry optimization steps. The NVT MD simulations during the linking cycles use the Nosé-Hoover thermostat with a time step of 1 fs at 900 K, which is close to the temperature of the actual reactions. Specifically, the NVT MD simulations employed for the searching approaches (i.e., md1 type in Fig. 3) comprise 5000 steps, and the md2 type MD simulations (in Fig. 3) are composed of 10000 steps. The md1 type MD simulation is used to refresh the system for searching for the potential pair satisfying the bonding criteria during the linking cycles. Considering the efficiency of the modelling approach and the computational costs, we set the length of md1 type MD simulation as 5000 steps. Previous work by Colina *et al* has manifested that the NVT MD steps, although short, offer adequate time for a slight relaxation of the network during the linking cycles, which is beneficial to the formation of the realistic lengths of bonds.[2] Thus, the md2 type MD simulation is inserted into the proposed modelling approach. Note that the same lengths of the md1 and md2 type NVT MD simulations have also been employed in the Polymatic algorithm, which has been successfully implemented for a broad class of polymer networks with similar size of the simulated systems as ours in this work.[3-5] The MD simulations during the gradual compression and relaxation process (depicted in Table 1)

utilize the Nosé-Hoover thermostat and Berendsen barostat with a time step of 1 fs. Additionally, atom-based summation methods with the cutoff distance of 5 Å and 15 Å are used for the van der Waals interactions for the geometry optimizations and MD simulations during and at the end of the linking cycles, respectively. And all electrostatic interactions are calculated by the Particle-Particle Particle-Mesh (PPPM) method with fine quality settings.

The proposed modeling algorithm is conducted by an in-house Python script, and the automatic process of searching the systems for the potential linking pairs and adjustment and modification of the systems is carried out by in-house Perl scripts, which can be implemented in the Materials Studio package.

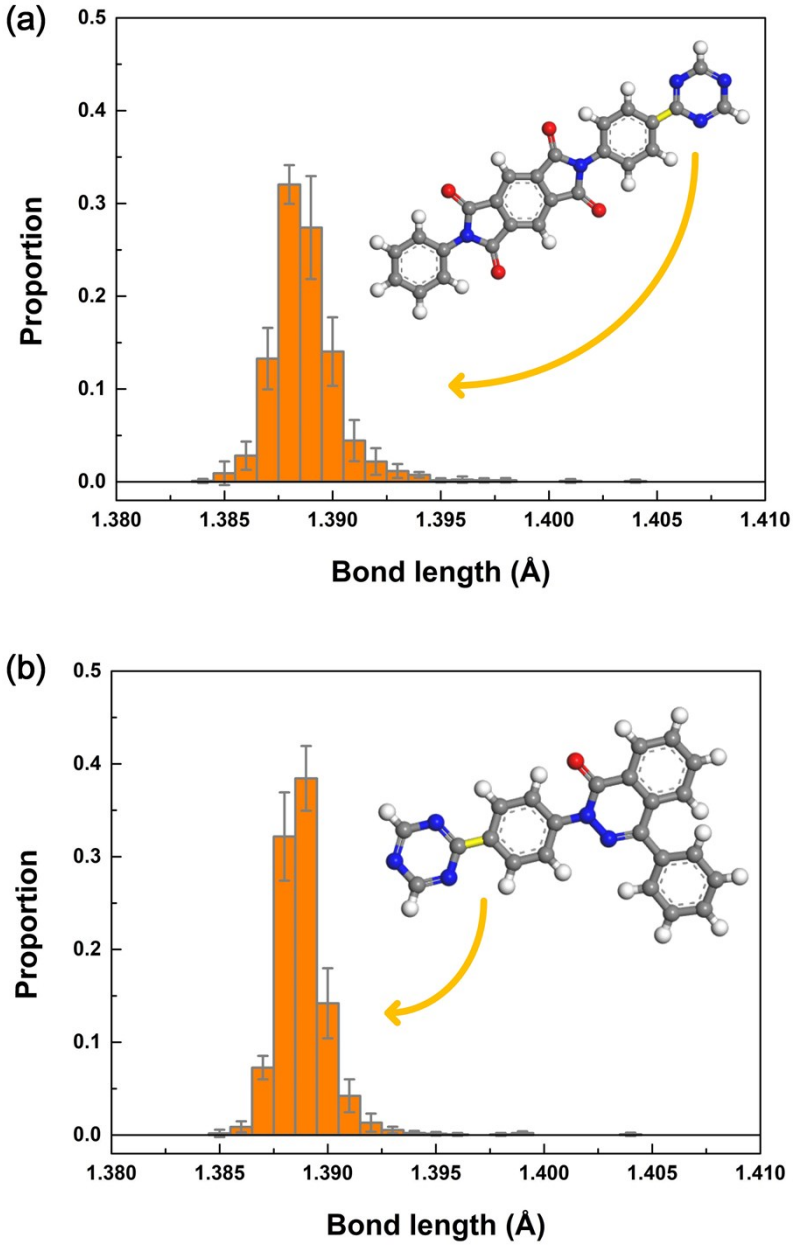

Fig. S2 Histograms of the length distributions of the C-C bonds formed during the linking cycles (displayed in the subplot and highlighted in yellow) after the geometry

optimization steps for (a) MPCF-1 and (b) MPCF-2 simulated models. The results represent the averages values of the five independent models with error bars representing the standard deviations.

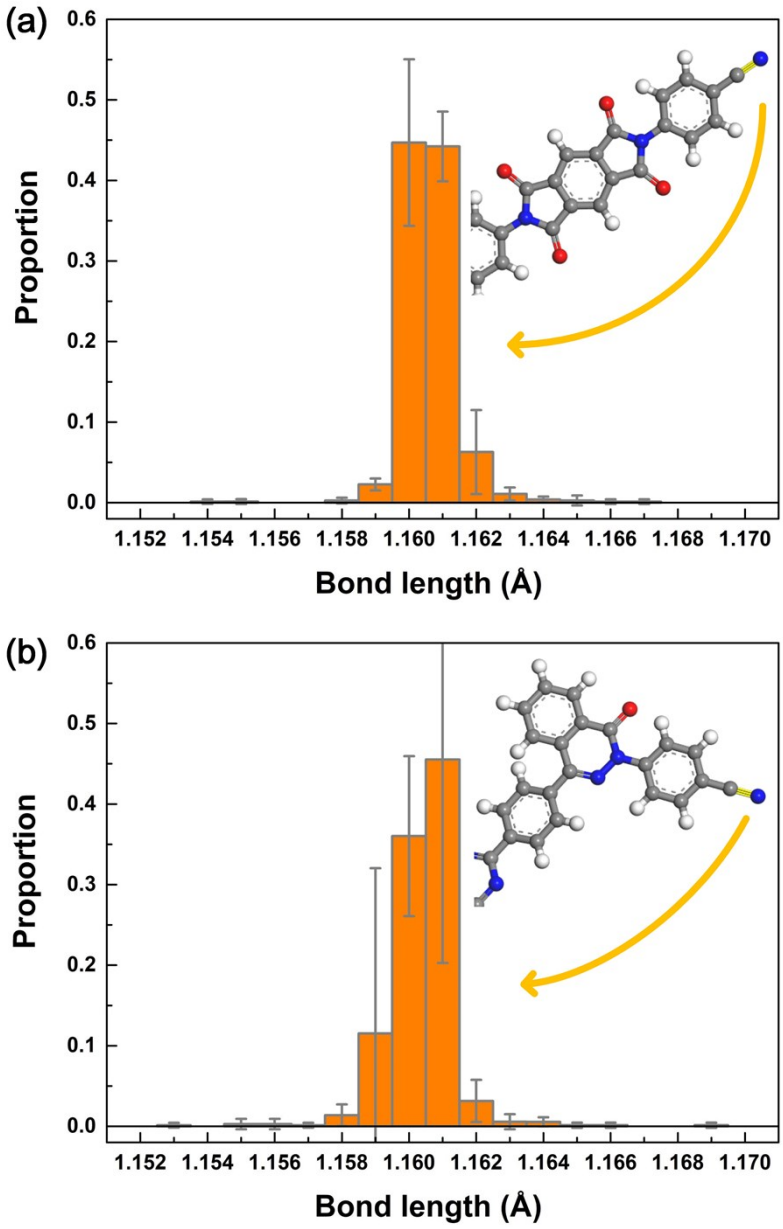

Fig. S3 Histograms of the length distributions of the C-N bonds in cyano groups generated during the final modification process (displayed in the subplot and highlighted in yellow) after the geometry optimization steps for (a) MPCF-1 and (b) MPCF-2 simulated models. The results represent the averages values of the five independent models with error bars representing the standard deviations.

3. Simulated models with accessible surface area.

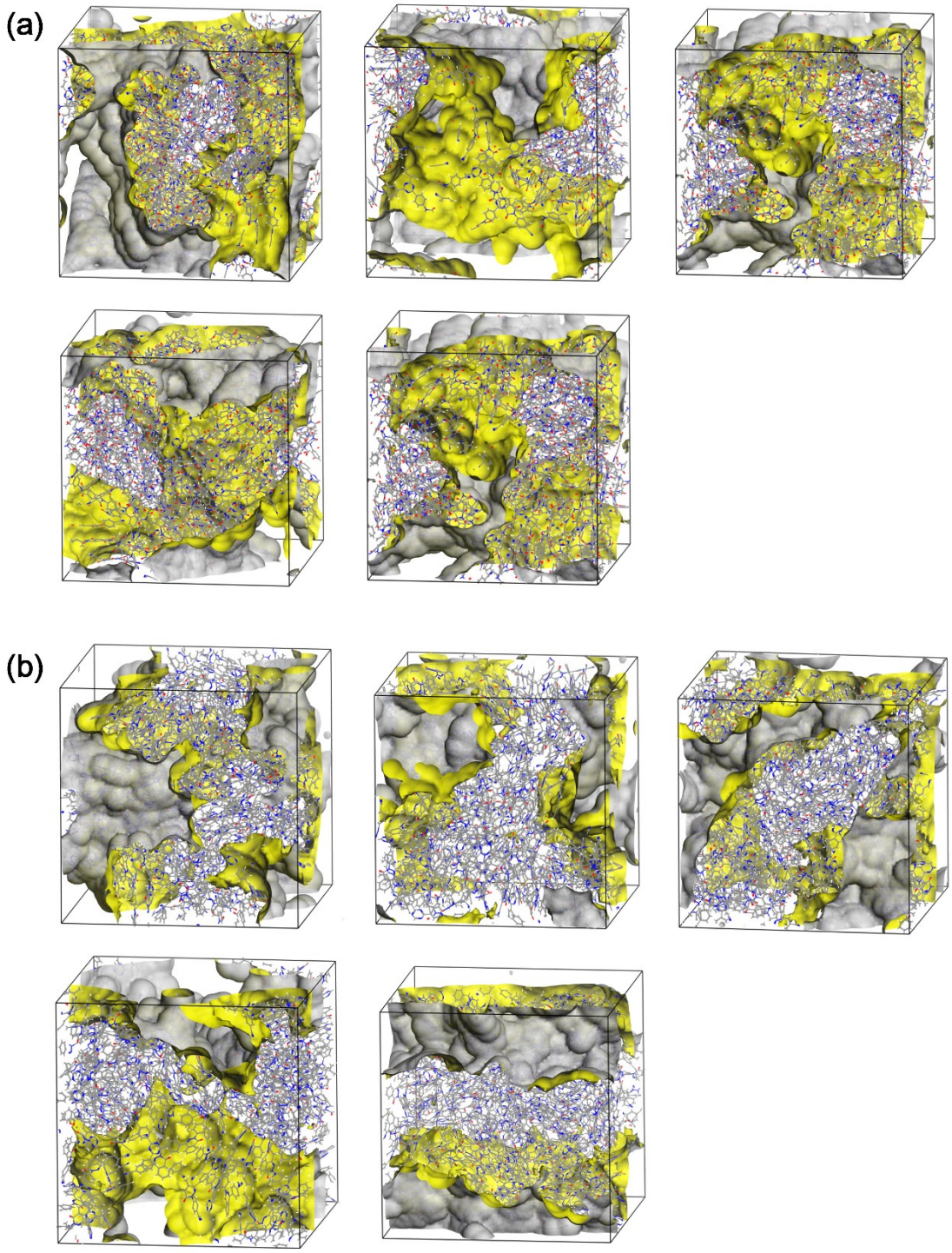

Fig. S4 Simulated accessible surface areas for five independent models of (a) MPCF-1 and (b) MPCF-2.

4. Simulation details for the structure factors.

The structure factors for the simulated models were carried out by the ISSAC software[6] with periodic boundary conditions. The structure factor S(q) is provided by the following equation:

$$S(q) = 1 + 4\pi\rho \int_0^\infty r^2 \frac{sinqr}{qr}(g(r) - 1)dr, \quad (1)$$

where g(r) represents the radial distribution function. And g(r) was calculated by 1000 steps with a smoothing factor of 0.1.

References

1. H. Sun, Z. Jin, C. Yang, R. L. Akkermans, S. H. Robertson, N. A. Spenley, S. Miller and S. M. Todd, *Journal of molecular modeling*, 2016, **22**, 47.
2. L. J. Abbott and C. M. Colina, *Macromolecules*, 2011, **44**, 4511-4519.
3. L. J. Abbott, K. E. Hart and C. M. Colina, *Theoretical Chemistry Accounts*, 2013, **132**.
4. M. E. Fortunato and C. M. Colina, *SoftwareX*, 2017, **6**, 7-12.
5. G. Kupgan, T. P. Liyana-Arachchi and C. M. Colina, *Polymer*, 2016, **99**, 173-184.
6. S. Le Roux and V. Petkov, *Journal of Applied Crystallography*, 2010, **43**, 181-185.