# On the Mechanism of Dynamic Polymerization via Recycled ss-DNA Templated Assembly of Non-natural Bases

*Jianbin Lin, Mathieu Surin, David Beljonne, Xianwen Lou, Joost L. J. van Dongen, Albertus P. H. J. Schenning**

*Contribution from Functional Organic Materials and Devices, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, Laboratory of Macromolecular and Organic Chemistry, Eindhoven University of Technology, P.O. Box 513, 5600,The Netherlands and from Laboratory for Chemistry of Novel Materials, University of Mons – UMONS, 20 Place du Parc 7000 Mons, Belgium*

E-mail: a.p.h.j.schenning@tue.nl

## Materials and Methods

***General methods***

$^{1}$H NMR and $^{13}$C NMR were recorded at room temperature on a Varian 300 or Varian Mercury 400. Chemical shifts are given in ppm (d) relative to tetramethylsilane. Abbreviations used are s = singlet, d = doublet, dd = double doublet, t = triplet and m = multiplet. Infrared (IR) spectra were run on a Perkin Elmer 1600 FT-IR spectrometer. MALDI-TOF MS spectra were measured on a Perspective DE Voyager spectrometer utilizing a-cyano-4-hydroxycinnamic acid matrix. CD and UV-vis were recorded on a JASCO 815 equipped with a Peltier temperature controller, PFD-425S. Gel permeation chromatography was performed on a Shimadzu LC10 system equipped with a photodiode array detector, using a Shodex KW402.5-4Fcolumn (300 × 6.4 mm i.d.) with 1M NaAC as the eluent at a flow rate of 0.85mL/min (T = 278K). The HPLC-UV-MS measurements were carried out on an LCQ Fleet instrument (Thermo Finnigan) using a GraceSmart RP 18 column (50 mm × 2.1 mm, 3 µm particle) with water/acetonitrile as the eluent.

***Sample preparation***

All guest molecules-containing samples were prepared by adding **Tn** in buffer to solid guest molecules and heating to 70 °C for at least 5 minutes.

***Materials***

Trityl tetraethylene glycol *p*-tosyl diether, **2** was synthesized according to literature procedures.[1] The ssDNA was supplied, HPLC purified and freeze-dried by MWG Biotech AG. All solvents, purchased from Acros Chimica or Sigma-Aldrich-Fluka, were of p.a. quality. Dry DMF and THF were obtained by distillation. Deuterated solvents were from Cambridge Isotope Laboratories. All other chemicals were commercially available and were used without purification.

***Molecular modeling methodology: molecular dynamics and quantum-chemical calculations***

The CHARMM force field for nucleic acids[2] was used to simulate **Tn-G1**, **Tn-G2**, and **Tn**-polymer structures, following the methodology reported in ref.5 of the manuscript. The oligothymine template was built using starting from the canonical B-helix structure of dsDNA. Starting from this geometry, the whole $T_n$-G complexes were constructed step-by-step with consecutive energy minimizations by placing the diaminotriazine or dihydrazine-triazine moieties forming three hydrogen-bonds with each thymine. For the hydrazone polymer, the modeled structures were started from B-helix conformation of **T40** surrounded by right-handed helical hydrazone 10-mer in relatively extended conformation, because of the matching in the length (a fully-extended hydrazone polymer 10-mer being around 14 nm-long while a dT40 template being around 13.6 nm-long) and in view of the GPC results (see Fig. S6). Molecular

dynamics simulations were performed with a time step of 1 fs in the canonical ensemble *N,V,T* (Berendsen thermostat) in the Generalized Born implicit water model with a simple smoothing function.[3] A heating and pre-equilibration period was used prior to attain a MD temperature of 300 K and the simulation was run at this temperature for 10 ns. For each system, more than 10 conformations were extracted from MD snapshots (from 2 ns to 10 ns) and each was subject to excited-state quantum-chemical calculations. Although the supramolecular models obtained are not definitive structures (in particular because the approximations on the effect of the medium), this approach gives a reasonable conformational sampling with respect to the calculation cost. The calculation of the excitonic CD spectra proceeded in two steps. First, the lowest 30 excited states of the 40 naphtalene derivatives involved in **$G_1$-T40** were computed at the INDO/SCI level (using an active space of 20 occupied times 20 empty molecular orbitals). Then, an excitonic Hamiltonian encompassing a total of 40x30 basis functions (30 localized excitations per molecule) was built on the basis of INDO/SCI[4] excitation energies and exciton couplings. The latter were calculated as Coulomb interactions between transition densities, thus going beyond the usual point dipole model.[5] Diagonalization of this Hamiltonian yielded a set of 1200 exciton states α with energies $\hbar\omega_\alpha$ and wavefunctions $|\psi_\alpha>$, for which the rotational strength $R_\alpha$ is computed as:[6]

$$R_\alpha = \frac{\hbar\omega_\alpha}{c}\sum_{i,n}\sum_{j,n'}\frac{\langle\psi_\alpha|\hat{\underline{\mu}}_{i,n}|G\rangle\times\langle G|\hat{\underline{\mu}}_{j,n'}|\psi_\alpha\rangle\cdot(\underline{r}_n-\underline{r}_{n'})}{\left|\underline{\mu}_{i,n}\right|\left|\underline{\mu}_{j,n'}\right|} \quad ,$$

where c is the speed of light, $\mu_{i,n}$ the transition dipole moment from the ground state $|g>$ to the excited state $|i>$ of molecule *n* along the stack, $\hat{\underline{\mu}}_{i,n} = \underline{\mu}_{i,n}(|i,n\rangle\langle g| + h.c.)$ the corresponding dipole operator, $|G>$ the ground state of the helical stack (product state of all $|g>$), and $|\psi_\alpha\rangle = \sum_{i,n} c_{i,n}^{\alpha}|i,n\rangle$ the exciton state wavefunctions expanded in terms of the $c_{i,n}^{\alpha}$ eigenvectors. The CD response at input frequency $\omega$ is calculated on the basis of the rotational strengths as:

$$CD(\omega) = <\sum_{\alpha} R_\alpha\, G(\omega-\omega_\alpha)> \ ,$$

where $G(\omega-\omega_\alpha)$ is a Gaussian function centered around $\omega_\alpha$ with variance $\sigma$=0.1 eV. The brackets denote a configurational average over the positional and energetic disorder as explored during the MD simulations. Here, a total of 20 supramolecular helical structures were used; this approach was found to yield CD spectra that are stable with respect to configurational averaging.

This approach, however, only accounts for through-space excitonic couplings between chromophores and cannot therefore be applied to multi-chromophoric systems involving through-bond interactions as the hydrazone polymer. In that case, we therefore turned to supermolecular calculations. These were performed at the TD-DFT (long-range corrected LC-wPBE functional[7] and 6-31g(d) basis set) level, as implemented in Gaussian09,[8] on the basis of the last snapshot (i.e. at 10 ns) of a MD simulationon the 10-mer.

***Synthesis***

Scheme S1. Chemical synthesis of the naphthalene guest molecule **$G_2$**.

a) 1.1 eq. nBuLi, THF. -78 $^{O}$C. b) 1.0 eq. cyanuric chloride -78 $^{O}$C – r.t. c) 2.5 eq. $NH_2NH_2$. DIEA, reflux

**3**

Under argon atmosphere at 0$^{o}$C tetraethylene glycol methyl *p*-tosyl ether (**2**, 3.62g, 10mmol) in 5mL dry DMF was added dropwise to a solution of 6-bromo-2-naphthol (**1**, 1.16g, 5mmol) and KOH (0.84g, 15mmol) in dry DMF, after 15 minutes the reaction mixture was heated to 80$^{o}$C for 6 hrs. The reaction mixture was added to 40mL 1 N HCl and extracted with chloroform. The collected organic phase was washed with water and dried with $MgSO_4$. After column

chromatography (EtOAc), tetraethylene glycol methyl-6-bromo-2-naphthyl ether (1.95g, 91% yield) was obtained as a white solid. $^{1}$H NMR $\delta$ (400 MHz, $CDCl_3$): 3.31 (s, 3H, OC$H_3$), 3.35(t, 2H, OC$H_2$), 3.58-3.69(m, 8H, OC$H_2$), 3.73(t, 2H, OC$H_2$), 3.92(t, 2H, OC$H_2$), 4.23(t, 2H, OC$H_2$), 7.10(d, 1H, Na*H*, $J$ = 2.5Hz), 7.19(dd, 1H, Na*H*, $J_1$ = 9.0Hz and $J_2$ = 2.5Hz), 7.49(dd, 1H, Na*H*, $J_1$ = 8.7Hz and $J_2$ = 2.0Hz), 7.58(d, 1H, Na*H*, $J$ = 8.7Hz), 7.10(d, 1H, Na*H*, $J$ = 9.0Hz), 7.10(d, 1H, Na*H*, $J$ = 2.0Hz). $^{13}$C NMR $\delta$ (75 MHz, $CDCl_3$): 59.0, 67.5, 69.6, 70.5, 70.6(3), 70.8, 71.9, 106.7, 117.0, 120.0, 128.4(2), 129.5, 129.6, 130.0, 132.9, 157.03. MALDI-TOF MS (M = 412.09) M/z = 413.95 $[M+H]^+$.

**$G_2$**

To a solution of tetraethylene glycol methyl-6-2-naphthyl ether (**3**, 1.85g, 4.5mmol) in anhydrous THF (40mL) was added n-BuLi (3.2mL, 4.95mmol, 1.57 M solution in hexane) at -78$^{o}$C under argon and the reaction mixture allowed to warm to room temperature and stirred for 2 hrs. The resulting solution was slowly added to a solution of cyanuric chloride (0.83g, 4.5mmol) in THF at -78$^{0}$C. The mixture was then warmed to room temperature and stirred for another 9 hrs. A THF solution of Hydrazine (10mL, 10mmol, 1.0 M solution in THF) were added to the mixture and refluxed for 6 hrs. After cooling to room temperature, NaOH (0.4g, 10mmol) in water (10mL) was added. After separation and aqueous layer was extracted with AcOEt. The combined organic phases were dried over $MgSO_4$. After removal of the solvent, purified by preparative reversed phase HPLC (C18 column, water/acetonitrile), and precipitation from a concentrated chloroform solution added to ether yielded **$G_2$** (570 mg, 28% yield). $^{1}$H NMR $\delta$ (400 MHz, $CDCl_3$): 3.36 (s, 3H, OC$H_3$), 3.53(t, 2H, OC$H_2$), 3.62-3.67(m, 8H, OC$H_2$), 3.72(t, 2H, OC$H_2$), 3.84(t, 2H, OC$H_2$), 4.13(t, 2H, OC$H_2$), 7.07(s, 1H, Na*H*), 7.10(d, 1H, Na*H*, $J$ = 8.0Hz), 7.62(s, 1H, Na*H*), 7.73(d, 1H, Na*H*, $J$ = 7.2Hz), 8.26(s, 1H, Na*H*), 8.69(s, 1H, Na*H*). $^{13}$C NMR $\delta$ (75 MHz, $CDCl_3$): 59.0, 67.3, 69.6, 70.5, 70.6, 70.8, 71.9, 106.4, 119.2, 125.4, 126.5, 128.2, 128.8, 130.8, 131.4, 136.5, 158.0, 167.7, 170.5. IR (KBr); $\nu$/cm$^{-1}$ = 840, 885, 913, 1105, 1189, 1223, 1259, 1352, 1387, 1421, 1484, 1585, 2870, 3291. MALDI-TOF MS (M = 473.24) M/z = 474.13 $[M+H]^+$.

Scheme S2. Chemical synthesis of the naphthalene guest molecule **$G_3$**.

**G3**

To a solution of **$G_2$** (94.6 mg, 2.0mmol) in Acetone (10mL) was added $HCO_2H$ (0.1mL) and refluxed for 1 hr. After cooling to room temperature, removed the solvent and precipitated from concentrated chloroform solution added to ether yielded **$G_3$** (110 mg, 100% yield). $^{1}$H NMR $\delta$ (400 MHz, $CDCl_3$): 2.01 (s, 6H, N=C(C$H_3$)$_2$), 2.16 (s, 6H, N=C(C$H_3$)$_2$), 3.33 (s, 3H, OC$H_3$), 3.50(t, 2H, OC$H_2$), 3.60-3.72(m, 8H, OC$H_2$), 3.75(t, 2H, OC$H_2$), 3.90(t, 2H, OC$H_2$), 4.20(t, 2H, OC$H_2$), 7.15(d, 1H, Na*H*, $J$ = 2.4Hz), 7.19(dd, 1H, Na*H*, $J_1$ = 8.9Hz and $J_2$ = 2.4Hz), 7.76(d, 1H, Na*H*, $J$ = 8.9Hz), 7.87(d, 1H, Na*H*, $J$ = 9.0Hz), 8.18(s, 2H, N*H*), 8.45(d, 1H, Na*H*, $J$ = 8.1Hz), 8.94(s, 1H, Na*H*). $^{13}$C NMR $\delta$ (75 MHz, $CDCl_3$): 16.4, 25.8, 59.0, 67.5, 69.7, 70.5, 70.6(3), 70.9, 71.9, 106.6, 119.4, 125.7, 126.6, 128.4, 129.6, 131.0, 131.3, 136.8, 151.4, 158.2, 164.6, 170.6. MALDI-TOF MS (M = 553.30) M/z = 554.23 $[M+H]^+$.

***Figures***

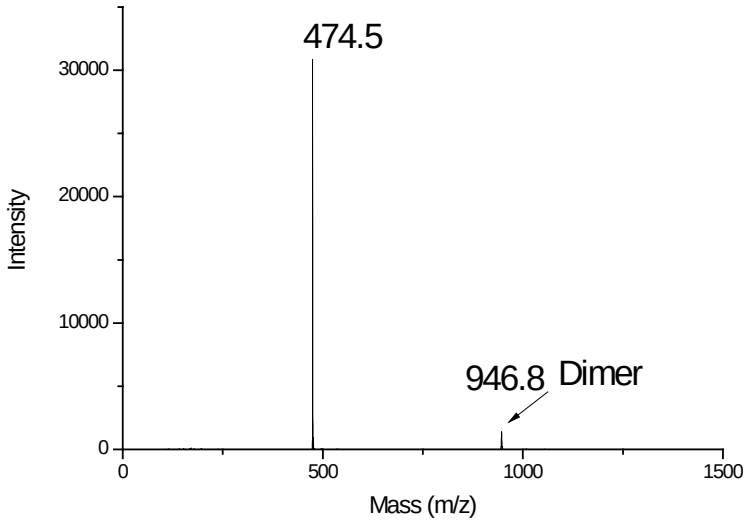


**Fig. S1** ESI Mass spectrum of $\mathbf{G_2}$

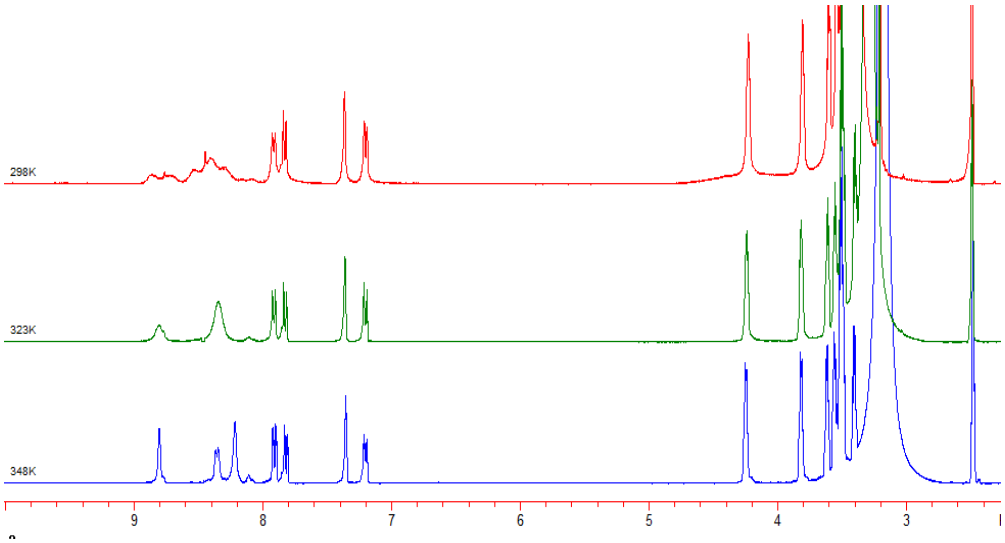


**Fig. S2** Temperature-dependent $^1$H-NMR spectra of $\mathbf{G_2}$ in DMSO-d6 (0.5mM).

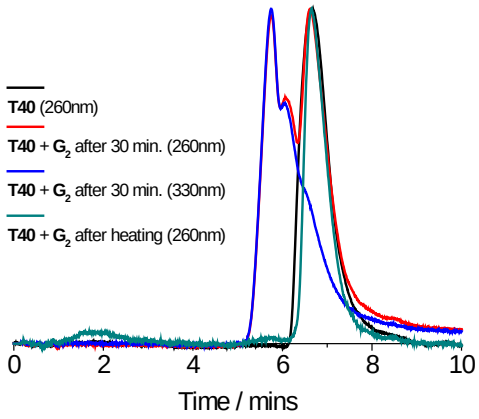


**Fig. S3** GPC chromatograms of the **T40**, $\mathbf{G_2}$ + **T40** + glyoxal at different stages. ([$\mathbf{G_2}$] = 0.5mM, $[\mathbf{T}]_{\mathbf{T40}}$ = 0.34mM, [$MgCl_2$] = 1mM, [glyoxal] = 0.5mM, [aniline] = 1mM, pH = 6.2).

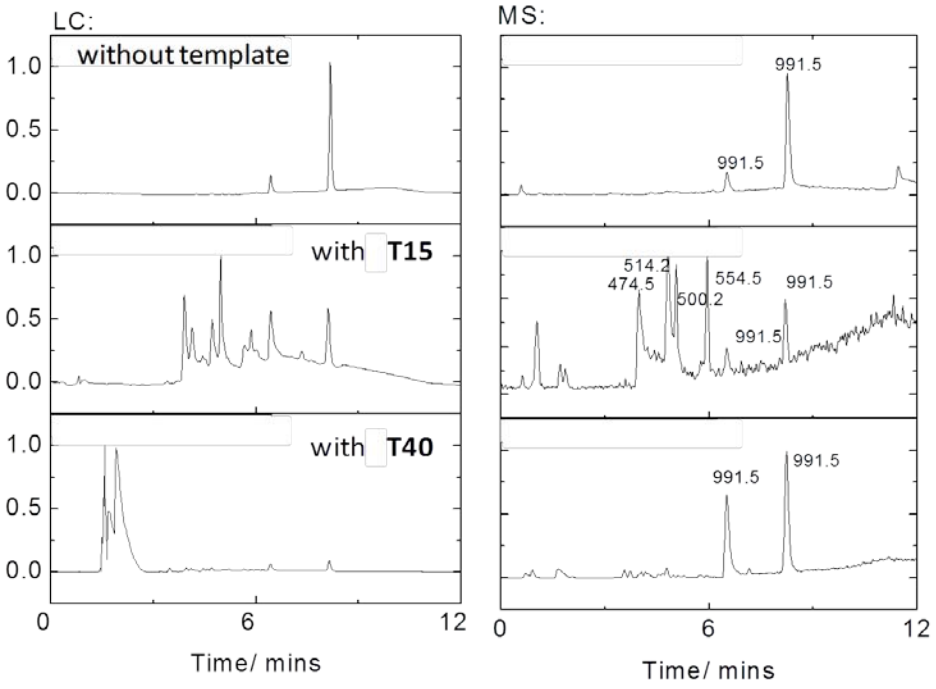


**Fig. S4** LC - MS chromatograms of $\mathbf{G_2}$ + glyoxal without template, and with the **T15** and **T40** template. ([$\mathbf{G_2}$] = 0.5mM, $[\mathbf{T}]_{\mathbf{Tn}}$ = 0.34mM, [$MgCl_2$] = 1mM, [glyoxal] = 0.5mM, [aniline] = 1mM, pH = 6.2. HPLC-traces were monitored at 330 nm corresponding to the naphthalene chromophore.

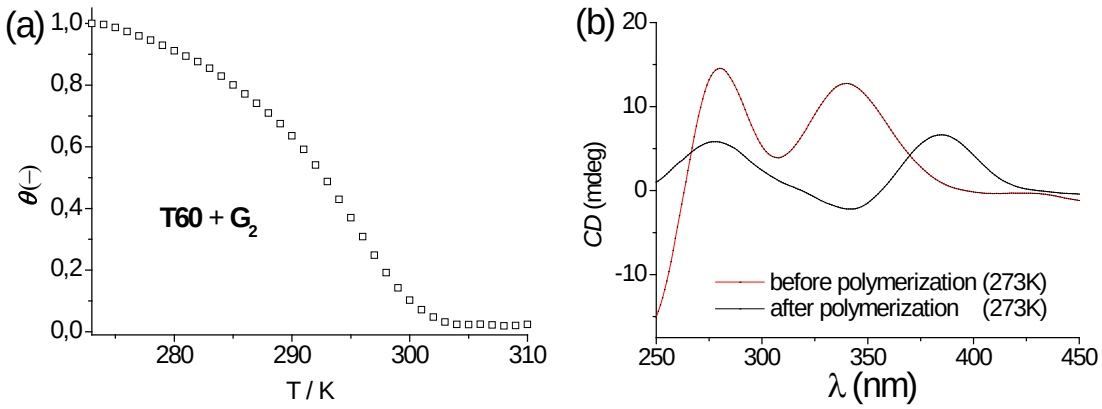


**Fig. S5** a) Normalized CD cooling curve monitored at λ = 330 nm for **T60** with $\mathbf{G_2}$ ([$\mathbf{G_2}$] = 0.5mM, $[\mathbf{T}]_{\mathbf{T60}}$ = 0.34mM, [$MgCl_2$] = 1mM, [aniline] = 1mM at pH = 6.2). (1K/Min) b) CD spectra before and after templated polymerization of $\mathbf{G_2}$ with **T60** as template ([$\mathbf{G_2}$] = 0.5mM, $[\mathbf{T}]_{\mathbf{T60}}$ = 0.34mM, [$MgCl_2$] = 1mM, [Glyoxal] = 0.5mM, [aniline] = 1mM, pH = 6.2).

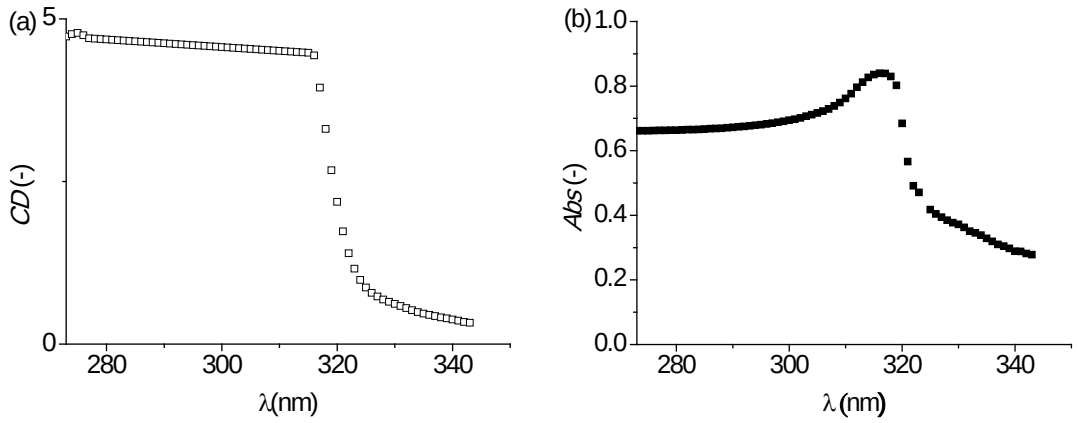


**Fig. S6** (a) CD and (b) UV-Vis heating curves monitored at λ = 375 nm of the hydrazone polymer with **T40**.

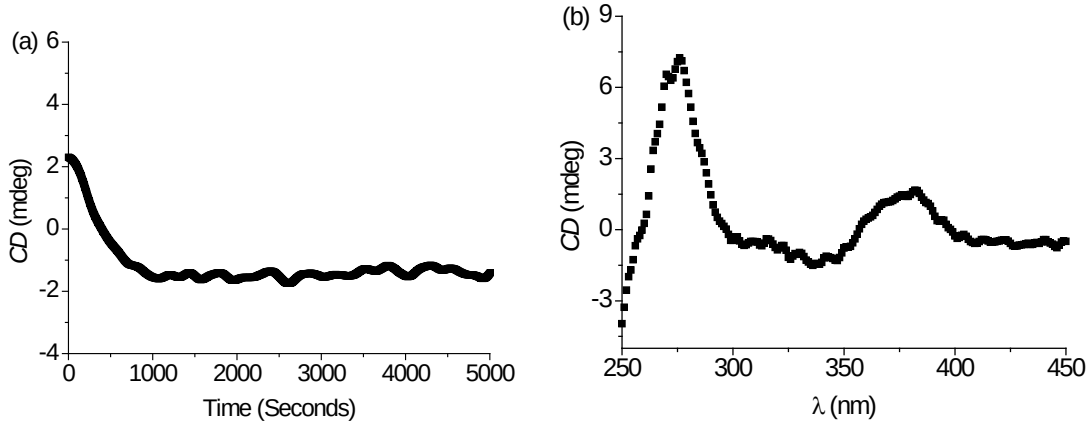


**Fig. S7** (a) An additional template polymerization reaction of the $\mathbf{G_2}$ with glyoxal in the presence of recycled **T40** versus time monitored at 330 nm. (b) corresponding CD spectrum measured at 273K.

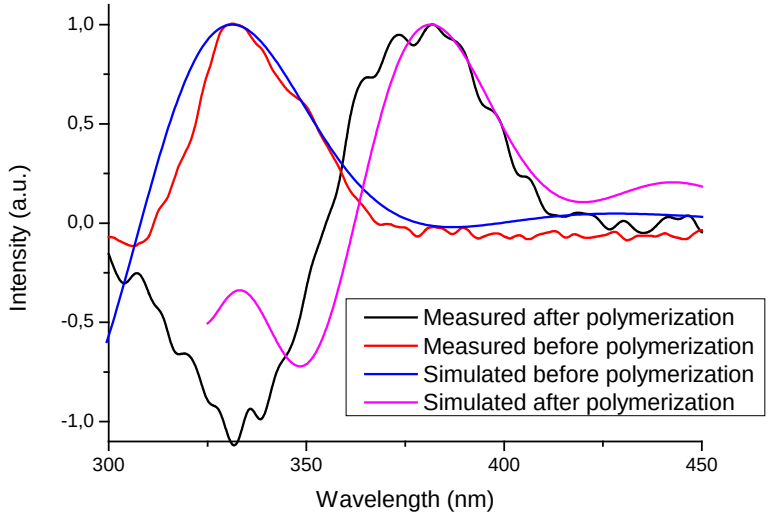


**Fig. S8** Experimental and simulated CD spectra before and after polymerization of the **G T40** systems. In this figure, the figure 2d and figure 3c of main text are combined.

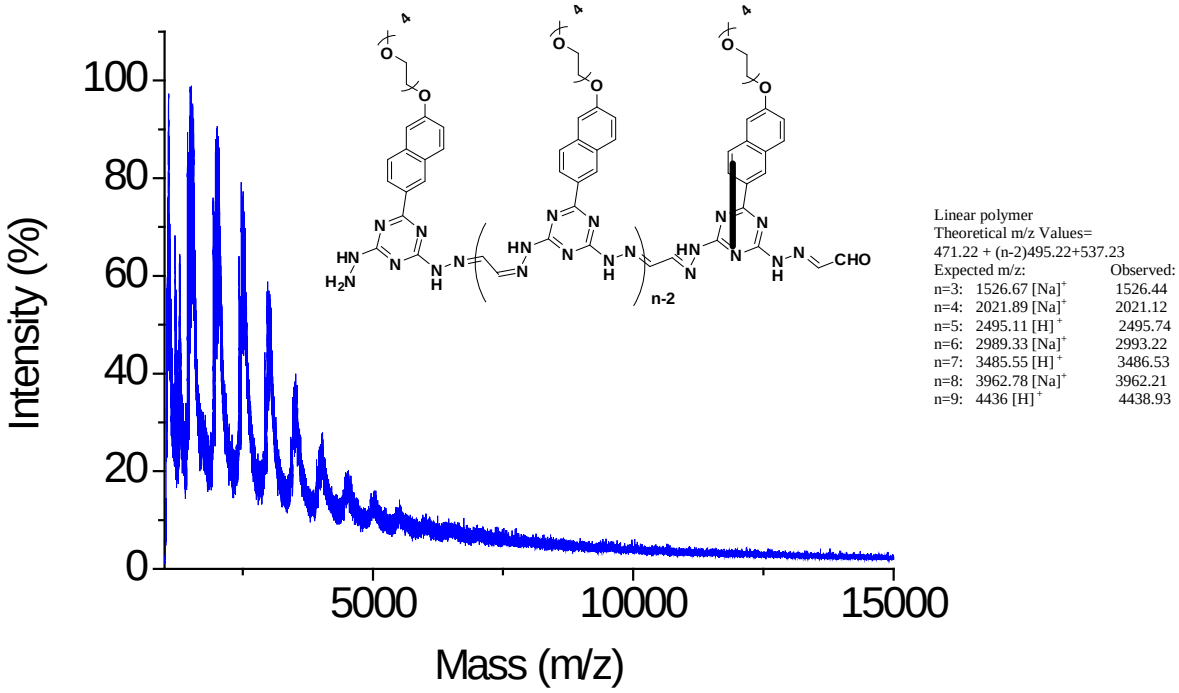


**Fig. S9** MALDI-TOF spectrum of the polymer precipitate dissolved in DMF with TFA (1%).

***References***

1 R.A. Schultz, B.D. White, D. M. Dishong, K.A. Arnold and G.W. Gokel, *J. Am. Chem. Soc*. 1985, **107**, 6659.

2 a) A.D. Mackerell, J. Wiorkiewicz-Kuczera and M. Karplus, *J. Am. Chem. Soc.* 1995, **117**, 11946. b) A.D. MacKerell, N. Banavali and N. Foloppe, *Biopolymers* 2001, **56**, 257, and references therein.

3 a) W. Im, M.S. Lee and C.L. Brooks III, *J. Comput. Chem.* 2003, **24**,1691. b) W. Im, M. Feig, C.L and Brooks III *Biophys. J.* 2003, **85**, 2900.

4 J. Ridley and M. Zerner, *Theoretica Chimica Acta* 1973, **32**, 111-134.

5 D. Beljonne, J. Cornil, R. Silbey, P. Millie and J.L. Brédas, *J. Chem. Phys.* 2000, **112**, 4749-4758.

6 F. C. Spano, S. C. J. Meskers, E. Hennebicq and D. Beljonne, *J. Am. Chem. Soc.* 2007, **129**, 7044–7054.

7 Y. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai and K. Hirao, *J. Chem. Phys.*, 2004, **120**, 8425.

8 Gaussian 09, Revision A.02, 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, J.A. Montgomery, Jr., 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, Inc., Wallingford CT, 2009.