

Electronic Supplementary Information (ESI) for

**Self-complementary hydrogen-bonded duplexes and helices
based on bis(pyrrolyl) carbonylhydrazide derivatives**

**Dehui Hu, Zhipei Yang, Guoqi Zhang, ‡* Min Liu, Junfeng Xiang, Tongling Liang,
Jinshi Ma and Guoqiang Yang***

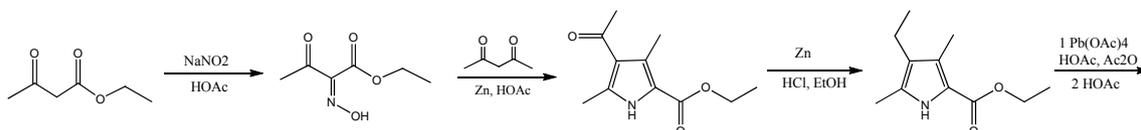
*Beijing National Laboratory for Molecular Sciences, CAS Key Laboratory of Photochemistry,
Institute of Chemistry, Chinese Academy of Sciences, Beijing, 100190, China. Tel: 0086-10-82617263;
Email: gqyang@iccas.ac.cn*

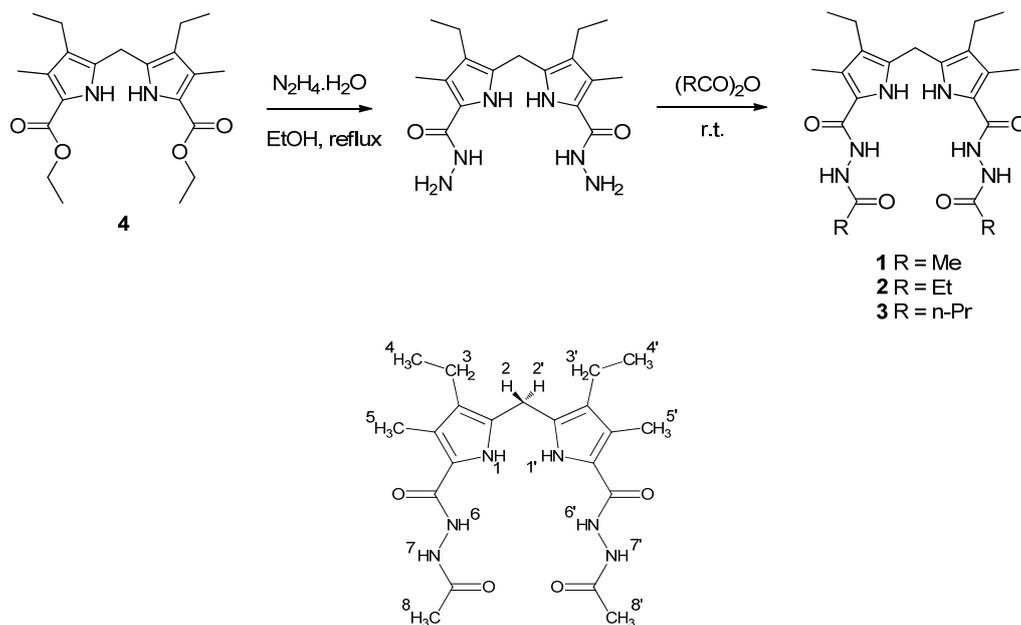
Experimental section

General

All reagents and solvents for syntheses and analyses are of analytical grade and used as received. Diethyl-5,5'-methylenebis(4-ethyl-3-methyl-1H-pyrrole-2-carboxylate) (**4**) was prepared by a literature procedure (Scheme S1).^{1,2} Melting points were determined on a Yanaco MP-500 micro-melting point apparatus. Electron spray mass spectra (ESI-MS) were recorded on Waters Micromass Q-TOF mass spectrometer, MALDI-TOF MS measurements were carried out on a Bruker BIFLEX III spectrometer. FT-IR spectra were recorded on a DIGILAB FTS3500 IR spectrometer. Samples for C, H, N and S analyses were dried under vacuum and the analysis were performed with a Carlo Erba-1106 elemental analyzer. ¹H and ¹³C NMR spectra were recorded on Bruker AVANCE 400 MHZ, 600 MHZ NMR spectrometers, chemical shifts are referenced to residual solvent peaks with respect to TMS = δ 0 ppm.

X-ray crystallography. Accurate unit cell parameters were determined by a least-square fit of 2θ values, and intensity data sets were measured on Rigaku Raxis Rapid IP diffractometer with Mo K α radiation ($\lambda = 0.71073 \text{ \AA}$). The intensities were corrected for Lorentz and polarization effects, but no corrections for extinction were made. All structures were solved by direct methods. The non-hydrogen atoms were located in successive difference Fourier synthesis. The final refinement was performed by full matrix least-squares methods with anisotropic thermal parameters for non-hydrogen atoms on F^2 . The hydrogen atoms were added theoretically and riding on the concerned atoms.





Scheme S1. Syntheses of 1-3 and numbering scheme for NMR spectroscopic assignments of 1.

4, 5,5'-methylenebis(4-ethyl-3-methyl-1H-pyrrole-2-carbohydrazide)

Diethyl 5,5'-methylenebis(4-ethyl-3-methyl-1H-pyrrole-2-carboxylate) (3.5 g, 0.1 mol) was dissolved in heated ethanol (25 ml) and an excess of hydrazine hydrate (85%, 15 ml) was added. The solution was refluxed for 36 h with plenty of white precipitate emerging. The solid was collected and washed with ethanol to afford the product (2.5 g, 74%). mp = 246-248 °C. ¹H-NMR (400 MHz, DMSO- *d*₆) δ 10.54 (s, 2H), 8.45 (s, 2H), 4.26 (s, 4H), 3.73 (s, 2H), 2.29 (q, *J* = 7.3 Hz, 4H), 2.14 (s, 6H), 0.86 (t, *J* = 7.3 Hz, 6H).

1, 5,5'-methylenebis(N'-acetyl-4-ethyl-3-methyl-1H-pyrrole-2-carbohydrazide)³

5,5'-methylenebis(4-ethyl-3-methyl-1H-pyrrole-2-carbohydrazide) (500 mg, 1.44 mmol) was suspended in acetic anhydride (10 ml) and then stirred at room temperature for 4h. The insoluble solid was filtered and washed repeatedly by ether to give a white powder (562 mg, 95%), which was further purified by recrystallizing from ethanol to give

colorless crystals (460 mg, 74.0%). mp = 274-275 °C. ¹H-NMR (400 MHz, DMSO-*d*₆) δ10.63 (s, 2H), 9.74 (s, 2H), 9.12 (s, 2H), 3.80 (s, 2H), 2.31 (q, *J*=7.2 Hz, 4H), 2.18 (s, 6H), 1.88 (s, 6H), 0.89 (t, *J*=7.2 Hz, 6H). ¹H-NMR (600 MHz, CDCl₃) δ10.88 (s, 1H, H^{6/6'}), 10.62 (s, 1H, H^{7/7'}), 10.44 (s, 1H, H^{6/6'}), 9.99 (s, 1H, H^{1/1'}), 9.57 (s, 1H, H^{1/1'}), 7.95 (s, 1H, H^{7/7'}), 3.97 (d, *J*=17.3 Hz, 1H, H^{2/2'}), 3.78 (d, *J*=17.3 Hz, 1H, H^{2/2'}), 2.45 (q, *J*=6.4 Hz, 2H, H^{3/3'}), 2.36 (overlapping, 5H, H^{3/3'+5/5'}), 2.04 (s, 3H, H^{8/8'}), 1.99 (s, 3H, H^{5/5'}), 1.66 (s, 3H, H^{8/8'}), 1.08 (t, *J*=7.0 Hz, 3H, H^{4/4'}), 1.00 (t, *J*=7.1 Hz, 3H, H^{4/4'}). ¹³C-NMR (101 MHz, DMSO-*d*₆) δ168.39, 160.63, 128.45, 122.89, 122.12, 118.62, 22.30, 20.55, 16.62, 15.47, 10.40. MALDI-TOF *m/z*: 453.6 [M + Na⁺] (calc. 453.2), 469.4 [M + K⁺] (calc. 469.2). ESI-MS *m/z*: 883.7 [2M + Na⁺] (the highest peak, calc. 883.5), 861.6 [2M + H⁺] (calc. 861.5), 453.3 [M + Na⁺] (base peak, calc. 453.2), 431.3 [M + H⁺] (calc. 431.2). Found C 58.35, H 7.12, N 19.50; C₂₁H₃₀N₆O₄ requires C 58.59, H 7.02, N 19.52%,

2, 5,5'-methylenebis(N'-propionyl-4-ethyl-3-methyl-1H-pyrrole-2-carbohydrazide)

5,5'-methylenebis(4-ethyl-3-methyl-1H-pyrrole-2-carbohydrazide) (500 mg, 1.44 mmol) was suspended in propionic anhydride (10 ml) and then stirred at room temperature for 4h. The insoluble solid was filtered and washed repeatedly by ether to give a white powder (582 mg, 88.2%), which was further purified by recrystallizing from ethanol to give colorless crystals (367 mg, 55.6%). mp = 275-276 °C. ¹H-NMR (400 MHz, DMSO-*d*₆) δ10.64(s, 2H), 9.71(s, 2H), 9.14(s, 2H), 3.82(s, 2H), 2.33(q, *J*=7.2 Hz, 4H), 2.20(s, 6H), 2.15(q, *J*=7.6 Hz, 4H), 1.06(t, *J*=7.6 Hz, 6H), 0.92(t, *J*=7.2 Hz, 6H). ¹H-NMR (400 MHz, CDCl₃) δ10.92 (s, 1H), 10.70 (s, 1H), 10.47 (s, 1H), 10.16 (s, 1H), 9.66 (s, 1H), 7.86 (s, 1H), 4.01 (d, *J*=17.5 Hz, 1H), 3.72 (d, *J*=17.5 Hz, 1H), 2.38 (m, 8H), 2.00

(s, 3H), 1.69 (s, 3H), 1.17 (t, $J=7.2$ Hz, 3H), 1.09 (t, $J=7.1$ Hz, 3H), 1.00 (t, $J=6.9$ Hz, 3H), 0.80 (t, $J=7.1$ Hz, 3H). ^{13}C -NMR (101 MHz, DMSO- d_6) δ 172.26, 160.85, 128.40, 122.93, 122.13, 118.65, 26.47, 22.32, 16.62, 15.47, 10.39, 9.73. MALDI-TOF m/z : 481.2 [M + Na $^+$] (calc. 481.3), 497.2 [M + K $^+$] (calc. 497.2). Found C 60.26, H 7.58, N 18.42; C $_{23}$ H $_{34}$ N $_6$ O $_4$ requires C 60.24, H 7.47, N 18.33%.

3, 5,5'-methylenebis(N'-butyryl-4-ethyl-3-methyl-1H-pyrrole-2-carbohydrazide)

5,5'-methylenebis(4-ethyl-3-methyl-1H-pyrrole-2-carbohydrazide) (500 mg, 1.44 mmol) was suspended in butyric anhydride (10 ml) and then stirred at room temperature for 4h. The insoluble solid was filtered and washed repeatedly by ether to give a white powder (682 mg, 97.4%), which was further purified by recrystallizing from ethanol to give colorless crystals (485 mg, 69.3%). mp =260-262 °C. ^1H -NMR (400 MHz, DMSO- d_6) δ 10.64(s, 2H), 9.70(s, 2H), 9.15(s, 2H), 3.81(s, 2H), 2.33(q, $J=7.2$ Hz, 4H), 2.20(s, 6H), 2.15(q, $J=7.6$ Hz, 4H), 1.57(m, 4H), 0.91(m, 12H). ^1H -NMR (400 MHz, CDCl $_3$) δ 10.91 (s, 1H), 10.75 (s, 1H), 10.50 (s, 1H), 10.13 (s, 1H), 9.59 (s, 1H), 7.86 (s, 1H), 4.00 (d, $J=17.8$ Hz, 1H), 3.73 (d, $J=17.8$ Hz, 1H), 2.38 (m, 8H), 2.33 (q, $J=7.1$ Hz, 3H), 2.00 (s, 3H), 1.87 (s, 3H), 1.67 (q, $J=7.2$ Hz, 3H), 1.09 (t, $J=6.8$ Hz, 3H), 1.00 (t, $J=7.1$ Hz, 3H), 0.95 (t, $J=7.2$ Hz, 3H), 0.78 (t, $J=6.8$ Hz, 3H). ^{13}C -NMR (101 MHz, DMSO- d_6) δ 171.36, 160.66, 128.39, 122.91, 122.12, 118.66, 35.15, 22.31, 18.54, 16.62, 15.48, 13.56, 10.40. MALDI-TOF m/z : 509.3 [M + Na $^+$] (calc. 509.3), 525.3 [M + K $^+$] (calc. 525.3). Found C 60.65, H 7.70, N 17.01; C $_{25}$ H $_{38}$ N $_6$ O $_4$ ·0.5H $_2$ O requires C 60.58, H 7.93, N 16.96%.

References:

- 1 L. Cheng, J. Ma, *Synth. Commun.*, 1994, **24**, 2771-2775.
- 2 D. L. Timothy, *Tetrahedron*, 1998, **54**, 359-374.
- 3 H. Fischer, O. Sus, F. G. Weilguny, *J. Liebig. Ann. Chem.* 1930, **481**, 159-192.

Table S1 The hydrogen bond parameters [\AA , $^\circ$] for the double helix of **1**·2THF.

D-H \cdots A	d(D-H)	d(H \cdots A)	d(D \cdots A)	\angle (DHA)	Symmetry code
N(1)-H(1) \cdots O(8)	0.86	1.93	2.775(6)	166.8	x,y+1,z
N(2)-H(2) \cdots O(5)	0.86	1.98	2.827(5)	167.8	x,y+1,z
N(3)-H(3) \cdots O(5)	0.86	2.12	2.867(5)	145.4	x,y+1,z
N(4)-H(4) \cdots O(6)	0.86	2.13	2.918(5)	152.9	x,y+1,z
N(5)-H(5) \cdots O(6)	0.86	2.04	2.806(6)	148.4	x,y+1,z
N(6)-H(6) \cdots O(3)	0.86	2.04	2.848(6)	155.6	-x+1,-y+2,-z
N(7)-H(7) \cdots O(4)	0.86	1.91	2.758(6)	167.3	x,y-1,z
N(8)-H(8) \cdots O(1)	0.86	1.97	2.832(6)	174.9	x,y-1,z
N(9)-H(9) \cdots O(1)	0.86	2.08	2.895(5)	157.8	x,y-1,z
N(11)-H(11) \cdots O(2)	0.86	2.03	2.877(6)	166.6	x,y-1,z
N(10)-H(10) \cdots O(2)	0.86	2.23	3.026(4)	153.9	x,y-1,z
N(12)-H(12) \cdots O(7)	0.86	2.04	2.836(6)	154.3	-x+1,-y,-z+1

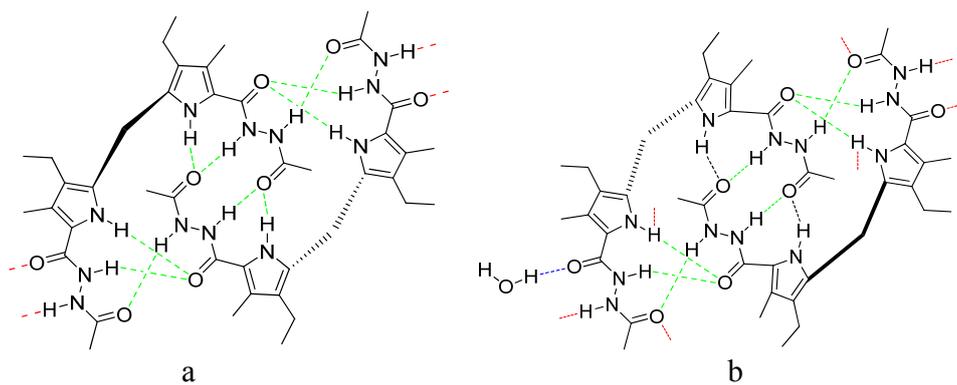


Fig. S1. The schematic representation of intermolecular and inter-helical N-H \cdots O hydrogen bonds in **1**·2THF (a) and **1**·0.5H₂O (b).

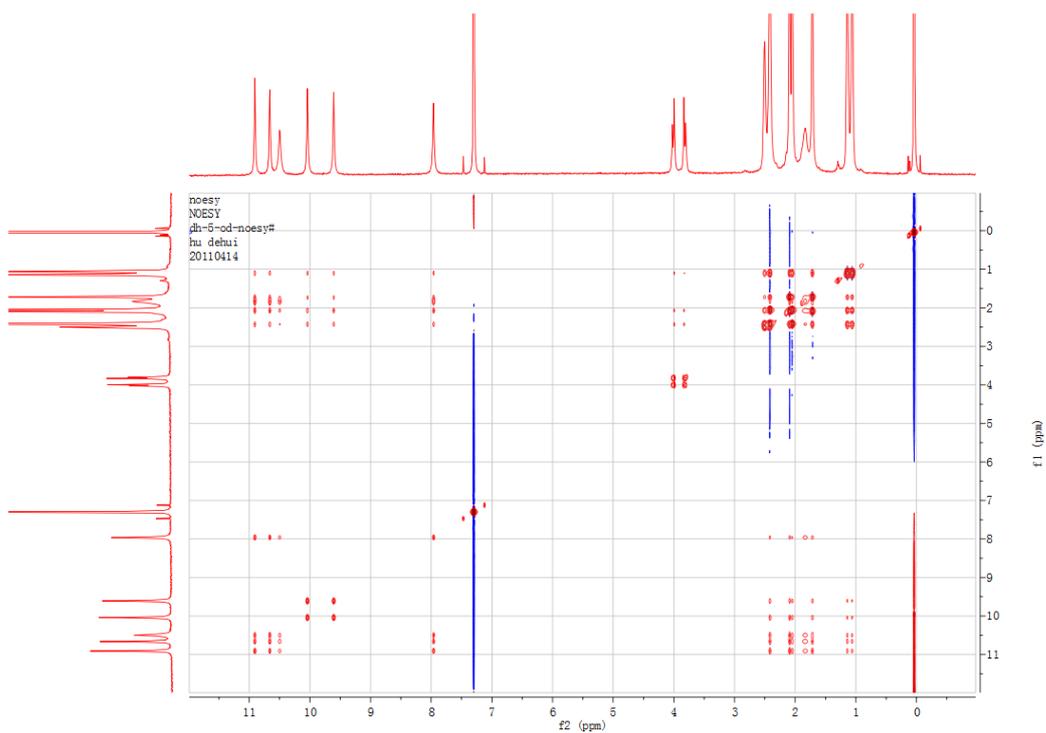


Fig. S2. 2D NOESY spectrum (600 MHz) of **1** in CDCl_3 . (mixing time= 0.80 sec).

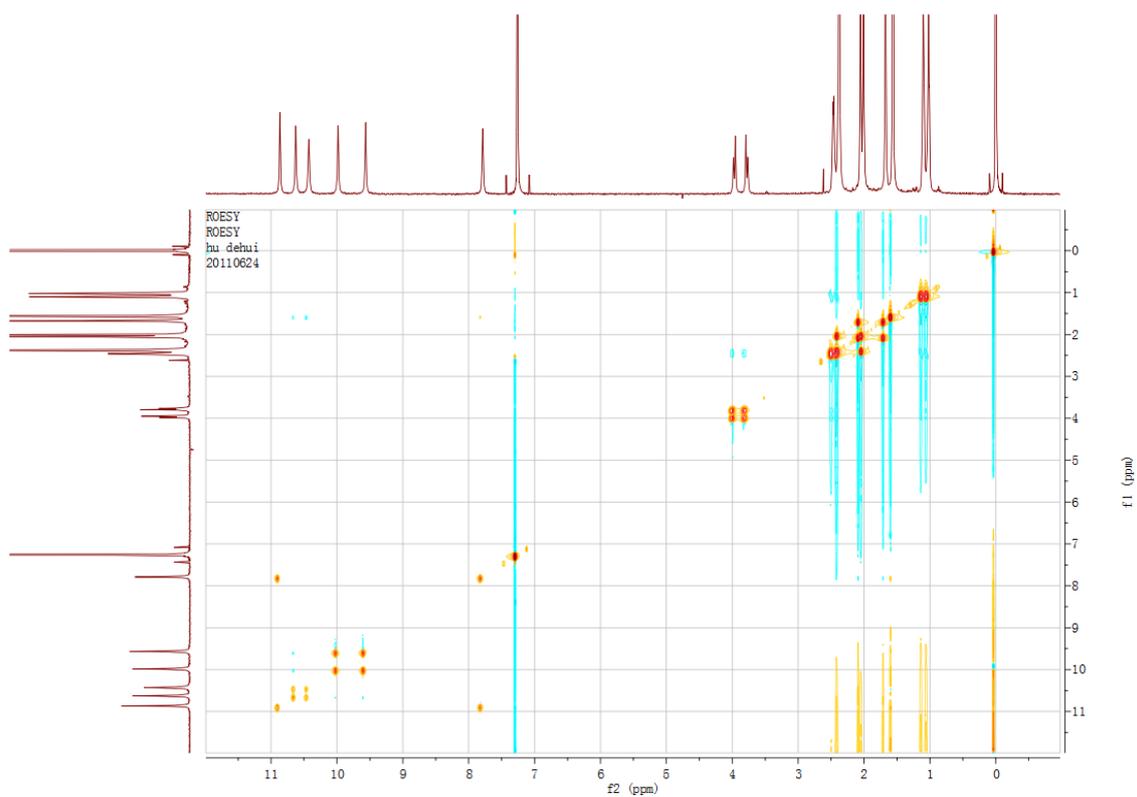


Fig. S3. 2D ROESY spectrum (600 MHz) of **1** in CDCl₃ (mixing time=200 000.00 usec).

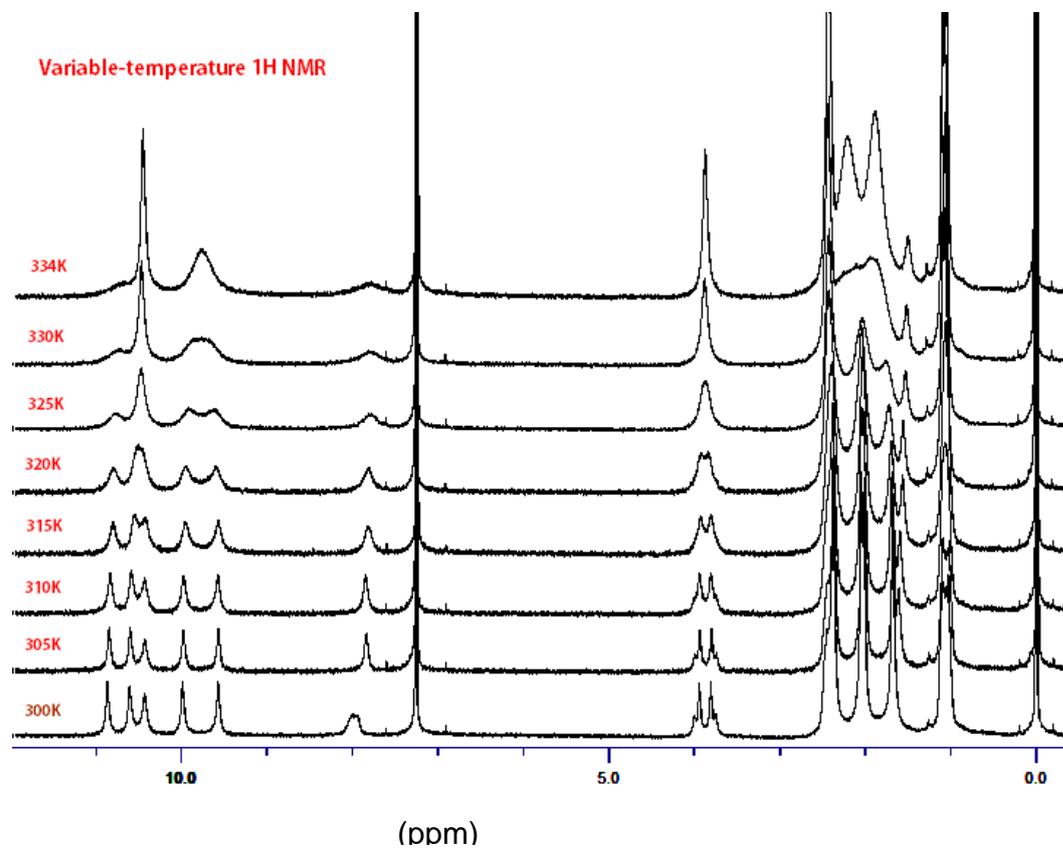


Fig. S4. Variable-temperature ^1H NMR (300 MHz) spectra of **1** in CDCl_3 .

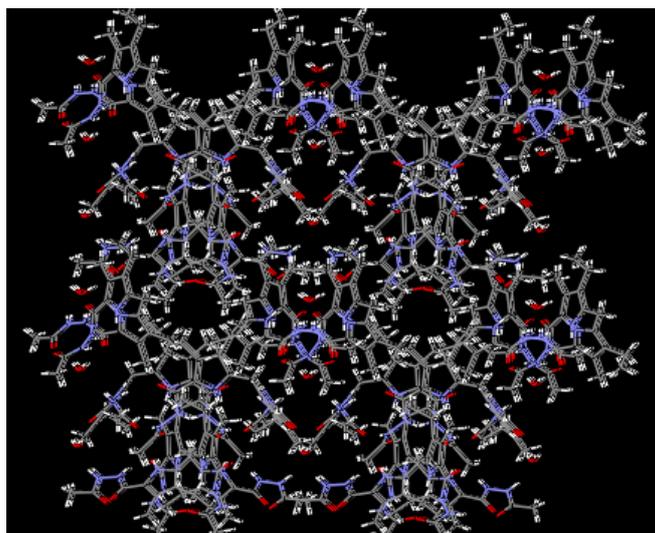


Fig. S5. The cell packing of $1 \cdot 0.5\text{H}_2\text{O}$ along crystallographic a-axis.

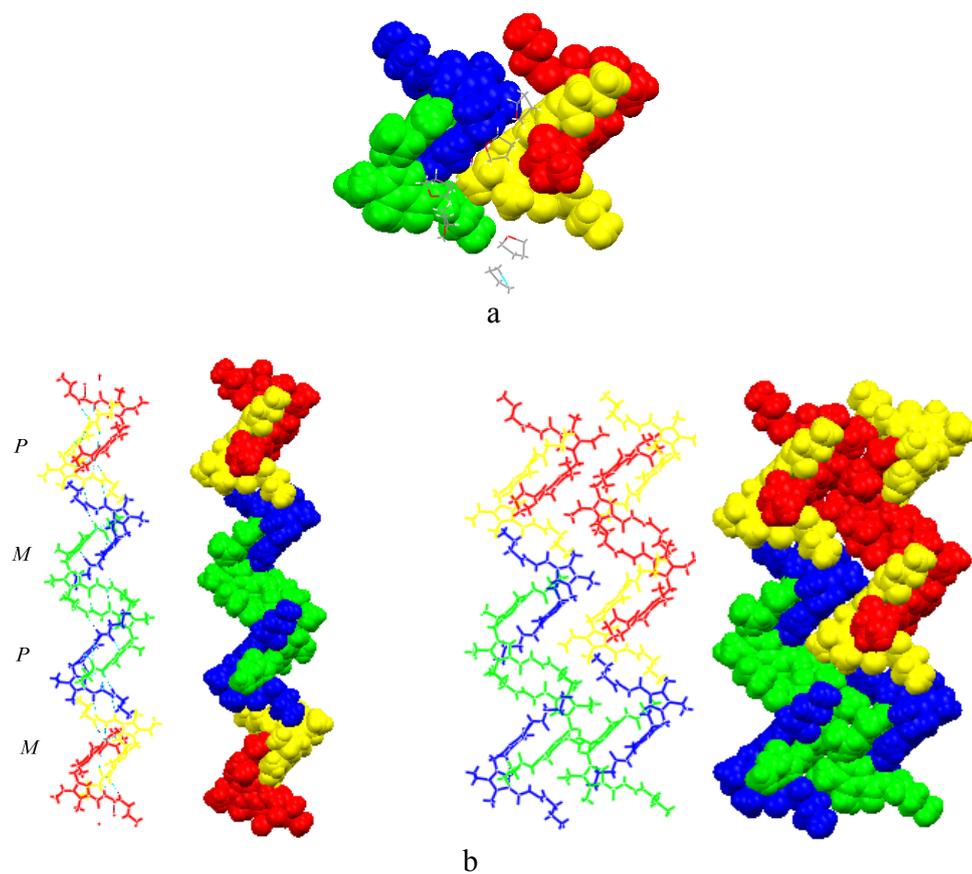


Fig. S6. (a) Unit cell of $3 \cdot 2\text{THF}$. (b) Stick model and space-filling representation of *meso*-helices in $3 \cdot 2\text{THF}$.

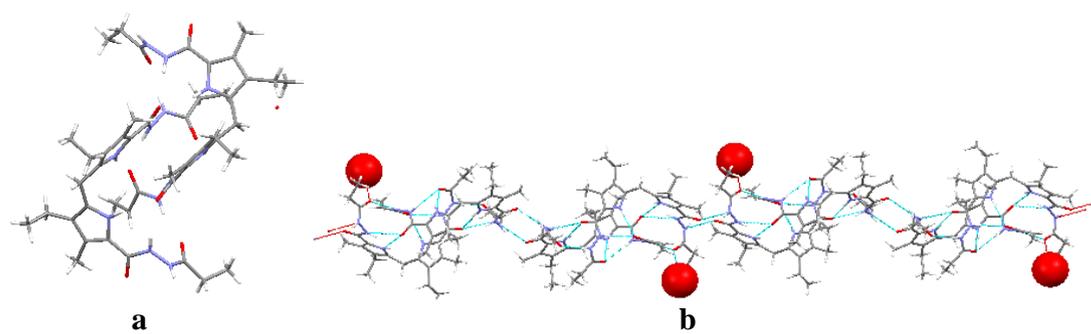


Fig. S7. (a) The Double helical structure formed by two independent molecules in $2 \cdot 0.25\text{H}_2\text{O}$; (b) The intermolecular and inter-helical N-H...O hydrogen bonds in $2 \cdot 0.25\text{H}_2\text{O}$, H₂O oxygen atoms were highlighted as red circles.