

Electronic Supplementary Information

for

Selectivity in pairing of model nucleobases implicates the solid state for the plausibly prebiotic supramolecular chemistry

by

Tomislav Stolar,^a Stipe Lukin,^a Martin Etter,^b Maša Rajić Linarić,^c Krunoslav Užarević^a Ernest Meštrović,^{*d} and Ivan Halasz^{*a}

^a Ruđer Bošković Institute, Bijenička c. 54, 10000 Zagreb, Croatia

^b Deutsches Elektronen-Synchrotron (DESY), 22607 Hamburg, Germany

^c PLIVA Croatia Ltd., TAPI R&D, 10000 Zagreb, Croatia

^d Xellia Pharmaceuticals, Slavonska avenija 24/6, 10000 Zagreb, Croatia

E-mail: ernest.mestrovic@xellia.com, ivan.halasz@irb.hr

List of Figures

S1	9-mG:1-mC Watson-Crick nucleobase pair forms chains through N–H···O and N–H···N hydrogen-bonded interactions.	5
S2	Chains further form hydrogen-bonded tapes in the crystal structure of 9-mG:1-mC Watson-Crick nucleobase pair.	6
S3	Plot of the Rietveld refinement of 9-methylguanine. The right part is enhanced by a factor of 3 for a better view.	8
S4	View of the crystal structure of 9-mG along the crystallographic a-axis.	9
S5	View of the crystal structure of 9-mG along the crystallographic c-axis.	10
S6	PXRD patterns of milling binary mixture of methylated adenine and thymine. Milling for 60 minutes results in nucleobase pairing while milling for 5 minutes results in physical mixture of reactants.	10
S7	PXRD patterns of methylated complementary binary mixtures obtained by milling for 60 minutes, and respective reactants.	11
S8	PXRD patterns of methylated non-complementary binary mixtures obtained by milling for 60 minutes, and respective reactants.	11
S9	PXRD patterns of canonical complementary binary mixtures obtained by milling for 60 minutes, and respective reactants.	12
S10	PXRD patterns of canonical non-complementary binary mixtures obtained by milling for 60 minutes, and respective reactants.	12
S11	Dry heating physical mixture of A and T monitored in situ by synchrotron X-ray diffraction.	13
S12	Dry heating physical mixture of A and U monitored in situ by synchrotron X-ray diffraction.	13
S13	Dry heating physical mixture of G and C monitored in situ by synchrotron X-ray diffraction.	14
S14	Dry heating physical mixture of A and C monitored in situ by synchrotron X-ray diffraction.	14
S15	Dry heating physical mixture of T and C monitored in situ by synchrotron X-ray diffraction.	15
S16	Dry heating physical mixture of T and G monitored in situ by synchrotron X-ray diffraction.	15
S17	Dry heating physical mixture of T and U monitored in situ by synchrotron X-ray diffraction.	16
S18	Dry heating physical mixture of U and C monitored in situ by synchrotron X-ray diffraction.	16
S19	Dry heating physical mixture of A and G monitored in situ by synchrotron X-ray diffraction.	17
S20	Dry heating physical mixture of U and G was monitored in situ by synchrotron X-ray diffraction.	17
S21	Rietveld refinement of PXRD pattern of 9-mA and 1-mT physical mixture heated at 100 °C. New crystalline phase corresponds to the known 9-mA:1-mT nucleobase pair (REFCODE: MTHMAD13).	18
S22	Dry heating physical mixture of 9-mA and 1-mU monitored in situ by synchrotron X-ray diffraction.	18
S23	DSC thermograms of 9-mA and 1-mU physical mixture, and respective reactants. Endotherms corresponding to melting points are labeled.	19
S24	Dry heating physical mixture of 1-mU and 1-mC monitored in situ by synchrotron X-ray diffraction.	19
S25	Dry heating physical mixture of 9-mA and 1-mC monitored in situ by synchrotron X-ray diffraction.	20
S26	Dry heating physical mixture of 9-mA and 9-mG monitored in situ by synchrotron X-ray diffraction.	20
S27	Dry heating physical mixture of 9-mG and 1-mU monitored in situ by synchrotron X-ray diffraction.	21

S28	Dry heating physical mixture of 1-mT and 1-mC monitored in situ by laboratory X-ray diffraction.	21
S29	Dry heating physical mixture of 1-mT and 1-mU monitored in situ by laboratory X-ray diffraction.	22
S30	Dry heating physical mixture of 9-mG and 1-mT monitored in situ by laboratory X-ray diffraction.	22
S31	PXRD patterns of methylated ternary mixture obtained by milling, and respective reactants.	23
S32	PXRD patterns of methylated quaternary mixture obtained by milling, and respective reactants.	23
S33	Dry heating ternary mixture of 9-mA, 1-mT, and 1-mU monitored in situ by laboratory X-ray diffraction.	24
S34	DSC thermogram of G and C physical mixture. Endotherm corresponding to melting point is labeled.	24
S35	DSC thermogram of A and T physical mixture. Endotherm corresponding to melting point is labeled.	25
S36	DSC thermogram of A and U physical mixture. Endotherms corresponding to melting points are labeled.	25
S37	FTIR-ATR spectra of 9-mG:1-mC nucleobase pair, 9-mG, and 1-mC. Spectra were recorded on PerkinElmer SpectrumTwo spectrometer equipped with a diamond cell in a 4000-450 cm^{-1} range.	26
S38	Photograph showing the experimental setup at P02.1 beamline, PETRA III, DESY. Hot-air blower system was placed directly below the measured capillary.	27

S1 Materials

9-methyladenine (9-mA), 1-methylthymine (1-mT) and 9-methylguanine (9-mG) were purchased from Carbosynth. 1-methyluracil (1-mU) and 1-methylcytosine (1-mC) were purchased from Sigma-Aldrich. Adenine (A) was purchased from Alfa Aesar. Thymine (T), uracil (U), guanine (G) and cytosine (C) were purchased from Tokyo Chemicals Industry. All chemicals were used as received.

S2 Mechanochemical milling experiments

To achieve nucleobase pairing by mechanochemistry, all binary combinations of solid nucleobases were milled for 60 minutes in 1:1 molar ratio (there are 10 binary combinations of canonical as well as 10 binary combinations of methylated nucleobases) and were further used for DSC and variable-temperature powder X-ray diffraction experiments, except in case of 9-mA and 1-mT. Milling of 9-mA and 1-mT in 1:1 molar ratio for 60 minutes results in their nucleobase pair, whereas milling for 5 minutes results in their physical mixture which was subsequently used for DSC and dry heating. 9-mA, 1-mT, 1-mU ternary mixture was prepared by milling solids in 1:1:1 molar ratio for 5 minutes as well as quaternary mixture of 9-mA, 1-mT, 9-mG, 1-mC in 1:1:1:1 molar ratio. Milling was performed in a mixer mill (In Solido Technologies) using 14 mL PMMA jars, two 7 mm stainless steel balls and at 30 Hz frequency.

S3 Differential scanning calorimetry

DSC thermograms were collected on a Discovery DSC (TA Instruments) and Perkin Elmer instruments. Around 3 mg of each sample was put in an aluminum crucible. Experiments were performed in an inert N₂ atmosphere using a 5 °C/min ramp.

S4 In situ monitoring of dry heating reactions by X-ray diffraction

Synchrotron X-ray diffraction monitoring of heating in the solid state was performed at P02.1 beamline at PETRA III, DESY. Capillaries of 0.5 mm width were filled with solid binary mixtures and were heated using a hot-air blower system. Calibration of the hot-air blower system was done using a thermocouple. Hot-air blower system was placed directly below the capillaries. Distance between the sample and the detector was 2010 mm. X-rays of $\lambda = 0.20720 \text{ \AA}$ were used and X-ray diffraction pattern was collected on a spinning capillary. Each sample was calibrated at 30 °C and was heated using a 5 °C/min ramp. After reaching the desired temperature set-point, sample was thermally equilibrated after which X-ray diffraction pattern was collected. Exposure time for each collected diffraction pattern was 300 s. Laboratory X-ray diffraction monitoring of heating reactions in the solid state was performed using a TTK 400 Anton Paar thermal camera, Panalytical Aeris diffractometer ($\lambda = 1.54175 \text{ \AA}$) and zero background sample holder.

S5 Crystal structure determination of 9-mG:1-mC cocrystal base pair

For crystal structure determination we used PXRD data collected at room temperature with 0.002 diffraction step in a 4-60 2Theta degrees range on a Bruker diffractometer ($\lambda = 1.54175 \text{ \AA}$). After indexing the powder pattern, we have found a candidate monoclinic P21/c unit cell with a volume of 1308 Å³ and subsequently performed Pawley fitting. Also, after observing some peaks corresponding to pure 9-mG with unknown crystal structure in the final diffraction pattern, it was described as collection of peaks taken from a collected PXRD of a pure 9-mG. Crystal structure was solved by simulated annealing in direct space. Molecules were treated as rigid bodies and their positions and orientations in a unit cell were varied independently. After recognizing a meaningful hydrogen bonding network between 9-mG and 1-mC molecules, Rietveld refinement was finally performed. All calculations were performed using the program TOPAS. The crystallographic information file (CIF), containing also measured and calculated diffraction patterns, has been deposited with the Cambridge Crystallographic Data Centre (CCDC) under

the deposition number 1919850.

Table S1 Crystallographic data for 9-mG:1-mC cocrystal base pair.

molecular formula	C ₁₁ H ₁₄ N ₈ O ₂
formula weight (g mol ⁻¹)	290.30
space group	<i>P</i> 2 ₁ / <i>c</i>
<i>a</i> (Å)	8.8815(5)
<i>b</i> (Å)	14.3878(5)
<i>c</i> (Å)	10.3075(5)
α (°)	90
β (°)	96.707(2)
γ (°)	90
<i>V</i> (Å ³)	1308.1(1)
<i>Z</i>	4
temperature (K)	298
radiation type	Cu <i>K</i> α
wavelength (Å)	1.54175
<i>R</i> _p	0.021
<i>R</i> _{wp}	0.049
GOF	5.11
<i>R</i> (<i>F</i> ²)	0.021

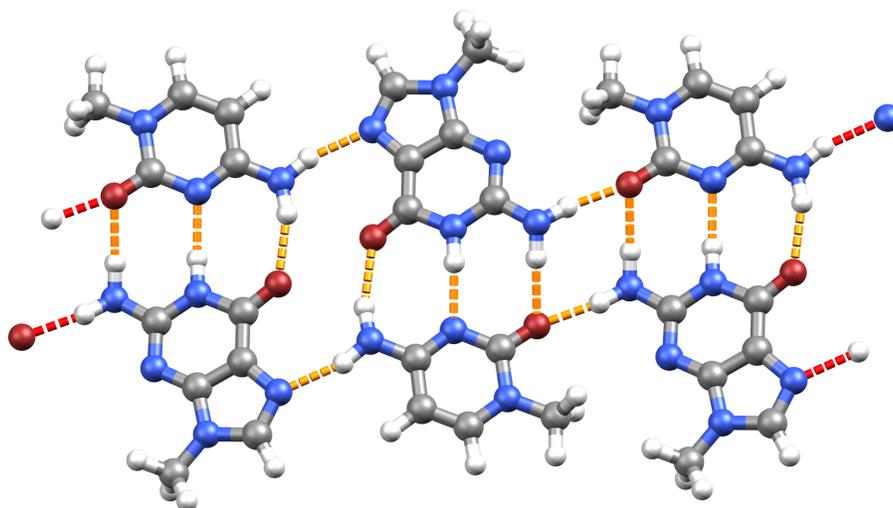


Figure S1 9-mG:1-mC Watson-Crick nucleobase pair forms chains through N–H...O and N–H...N hydrogen-bonded interactions.

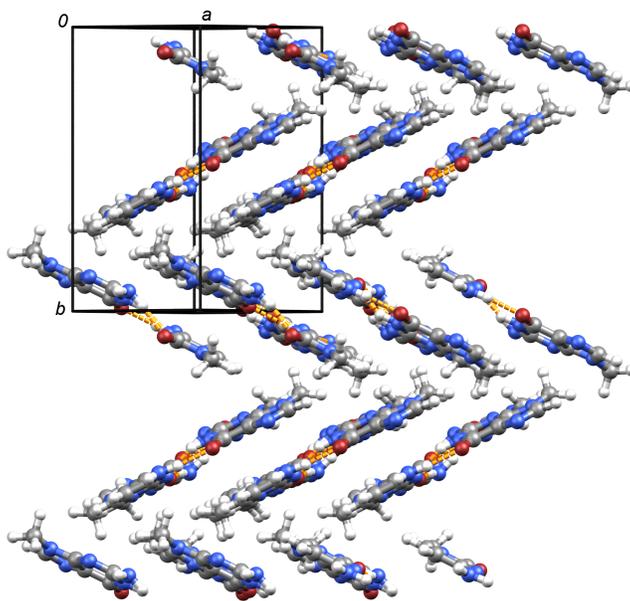


Figure S2 Chains further form hydrogen-bonded tapes in the crystal structure of 9-mG:1-mC Watson-Crick nucleobase pair.

S6 Measurement and solution of the crystal structure of 9-mG

A polycrystalline sample of 9-mG was filled into a glass capillary of 0.8 mm diameter (manufacturer: WJM Glas GmbH, Berlin, Germany), densified and sealed at both ends. The capillary was then measured at the Powder Diffraction and Total Scattering beamline P02.1 of the PETRA III synchrotron at the Deutsches Elektronen-Synchrotron facility in Hamburg, Germany with a Perkin Elmer XRD 1621 area detector at a sample detector distance of approximately 2100 mm to achieve a powder diffraction pattern with highest resolution. In order to increase the statistics for the measurement, the capillary was spun during the entire integration time of 180 seconds. The wavelength for the measurement was determined to be $\lambda = 0.20713$ Å. The obtained 2-dimensional diffraction pattern was integrated into a 1-dimensional powder diffraction pattern using the DAWN software¹ by making use of a calibration measurement of a LaB6 standard (NIST 660 b). Subsequent indexing, crystal structure solution attempts and the final Rietveld refinement were performed with the TOPAS 6 software.² For indexing, the LSI indexing method implemented in TOPAS was used.³ First indexing attempts gave quite rapidly a bunch of monoclinic unit cells with volumes in the range of 2735 Å³ to 2749 Å³. A Pawley refinement⁴ of the best indexed unit cell with space group $P2_1/c$ and a volume of approximately 2743 Å³ gave a perfect fit with an almost flat difference curve and a weighted profile residual (R_{wp}) of 0.759 %. The background in this refinement was modelled by 4 Lorentzian peaks. Since this primitive monoclinic unit cell in space group $P2_1/c$ covered all observed reflections, it was used for a first simulated annealing attempt. In this simulated annealing attempt four rigid body molecules of 9-methylguanine were placed, all with 3 degrees of translation, 3 degrees of rotation and 1 additional degree of freedom for the rotation of the methyl group. The simulated annealing dropped down to a R_{wp} of 7.29 % within a few hundred thousand iterations. However, from a visual inspection of the crystal structure it was clear, that some rigid bodies had to be flipped by 180 degree, as the bonding network between individual molecules looked a bit odd. Therefore, individual rigid bodies were fixed and the simulated annealing was continued in an iterative manner. The final R_{wp} after the improved simulated annealing was 4.439 % and the bonding scheme between the molecules

looked much better. In the following Rietveld refinement,⁵ the internal bond distances and bond angles within the molecules were optimized by changing the rigid body model to an individual atomic site refinement. In order to preserve the geometry of the molecules, harsh distance and angle restraints were applied. Hydrogen atoms were refined using a riding model. Using the in TOPAS integrated fundamental parameter approach for modelling the peak shapes and using 4 Lorentzians for the modelling of the background led the Rietveld refinement converge to a final R_{wp} of 1.758 %. A plot of the final Rietveld refinement can be found in Figure S3. Details of the crystal structure and final residual values can be found in Table S2. Further details, e.g. atomic positions, can be found in the CIF which was deposited with the CCDC under deposition number 1995351.

Table S2 Crystal structure parameters of 9-mG and final residual fit values as defined in TOPAS.

molecular formula	C ₆ H ₇ N ₅ O
space group	$P2_1/c$
a (Å)	7.0933(2)
b (Å)	22.5623(10)
c (Å)	17.2902(10)
α (°)	90
β (°)	97.667(4)
γ (°)	90
V (Å ³)	2742.43(22)
Z	4
R_{exp}	0.0133
R_{wp}	0.01758
R_p	0.01203
R_{Bragg}	0.00671
GOF	1.322

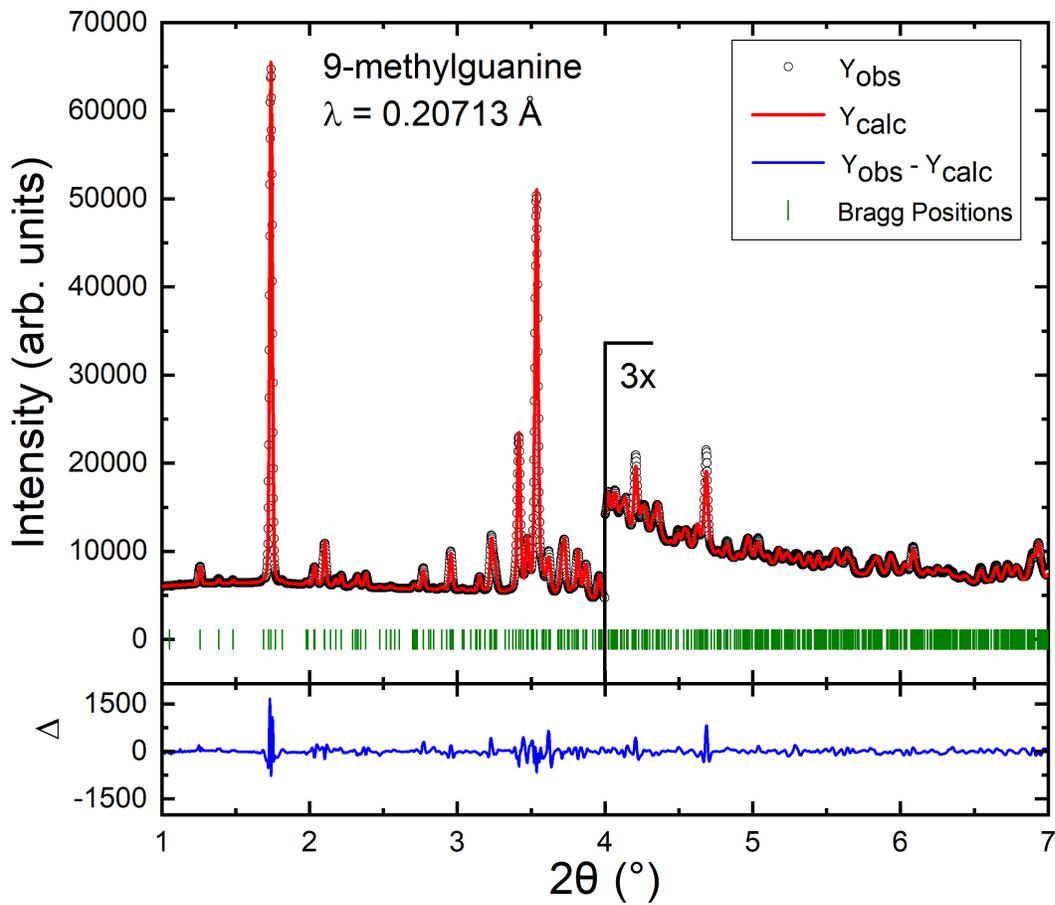


Figure S3 Plot of the Rietveld refinement of 9-methylguanine. The right part is enhanced by a factor of 3 for a better view.

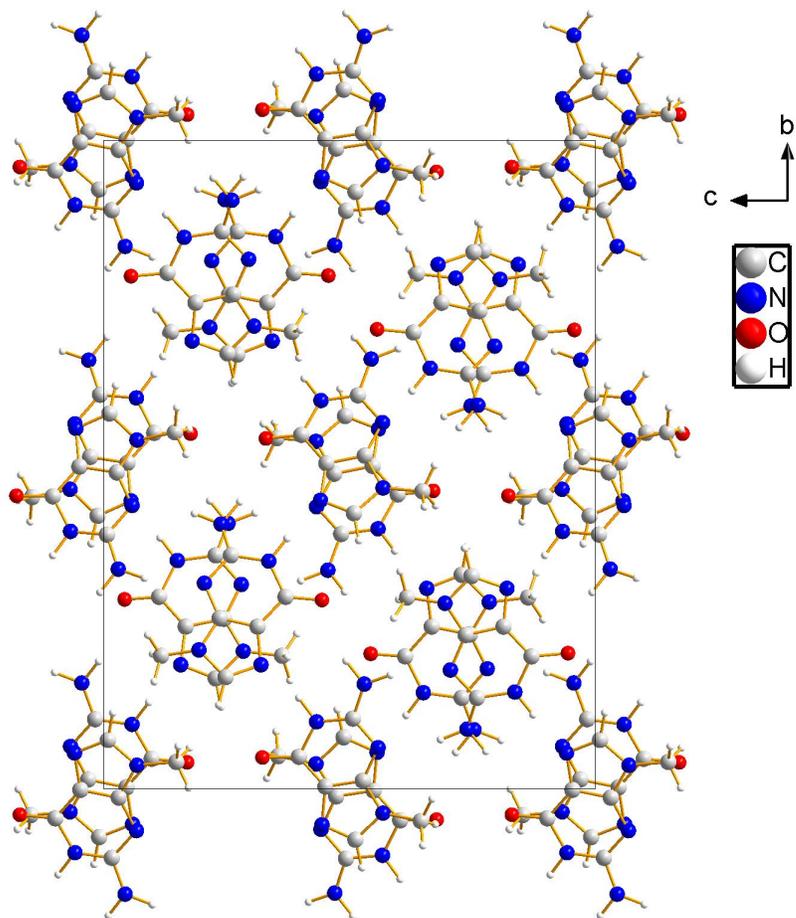


Figure S4 View of the crystal structure of 9-mG along the crystallographic a-axis.

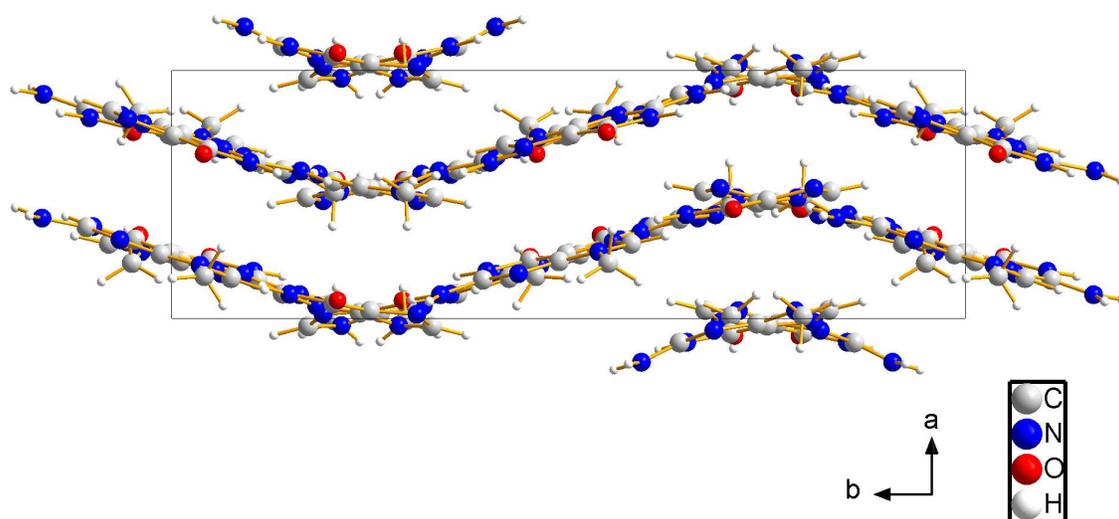


Figure S5 View of the crystal structure of 9-mG along the crystallographic c-axis.

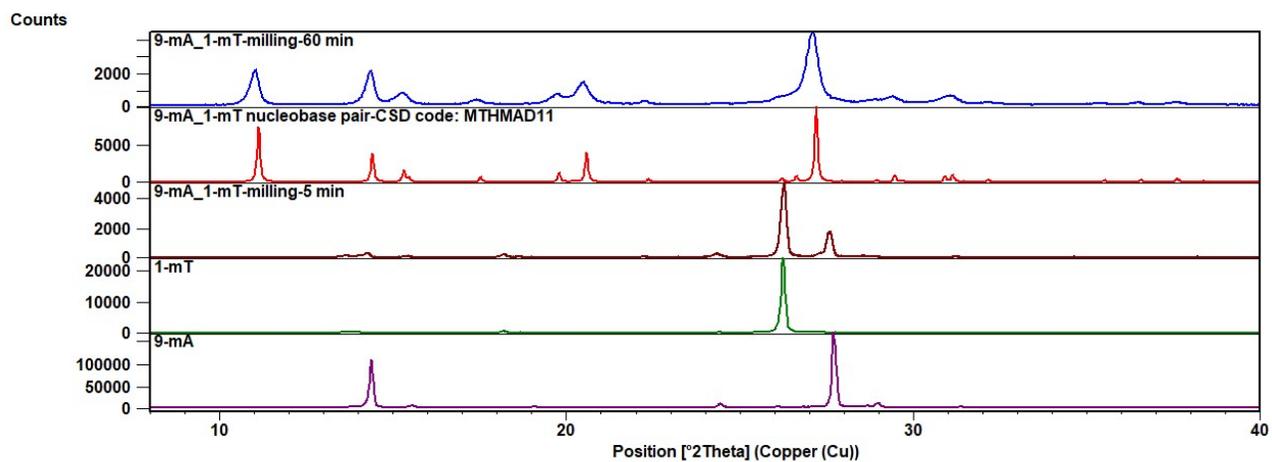


Figure S6 PXRD patterns of milling binary mixture of methylated adenine and thymine. Milling for 60 minutes results in nucleobase pairing while milling for 5 minutes results in physical mixture of reactants.

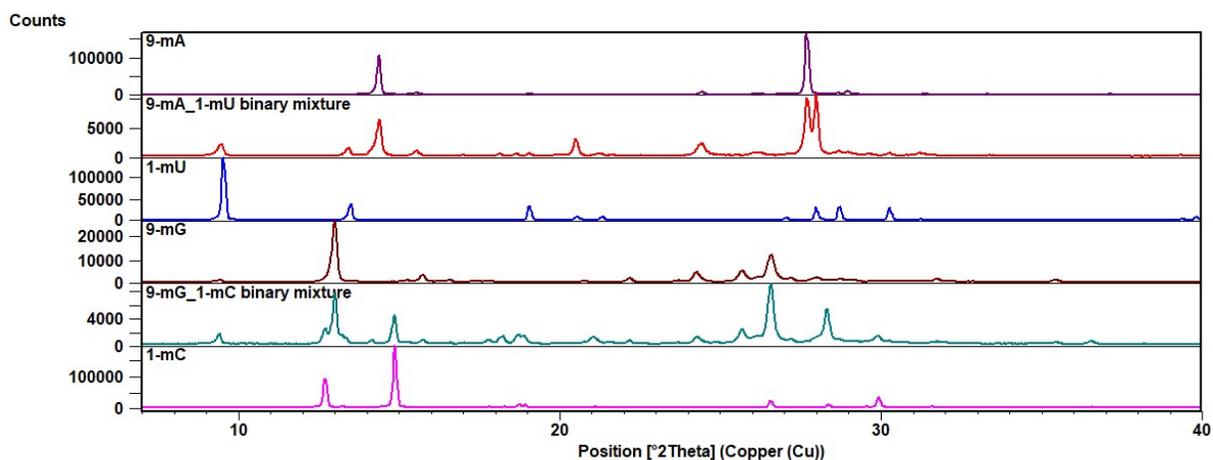


Figure S7 PXR D patterns of methylated complementary binary mixtures obtained by milling for 60 minutes, and respective reactants.

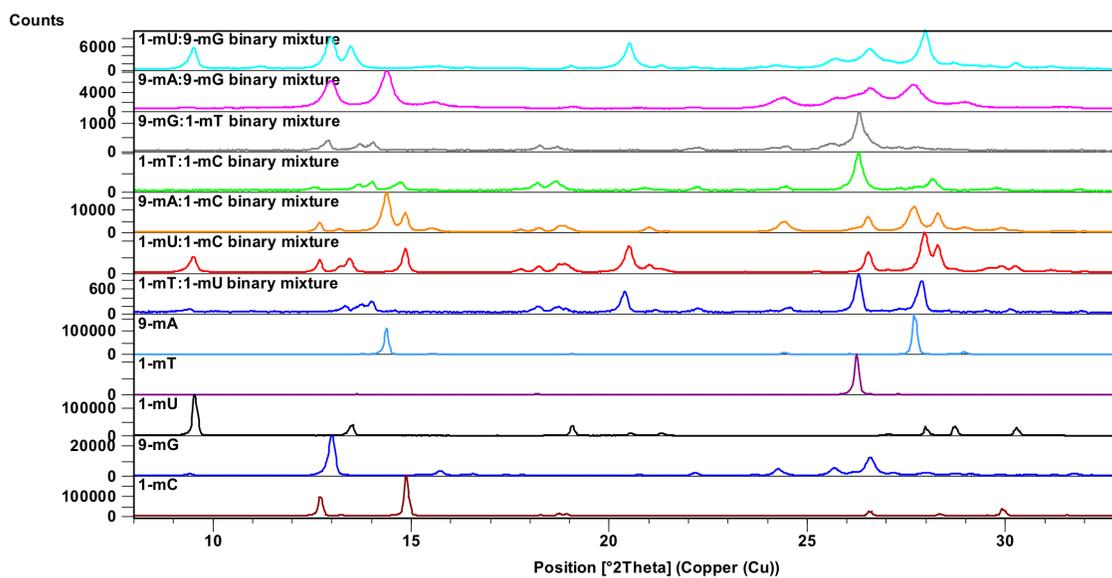


Figure S8 PXR D patterns of methylated non-complementary binary mixtures obtained by milling for 60 minutes, and respective reactants.

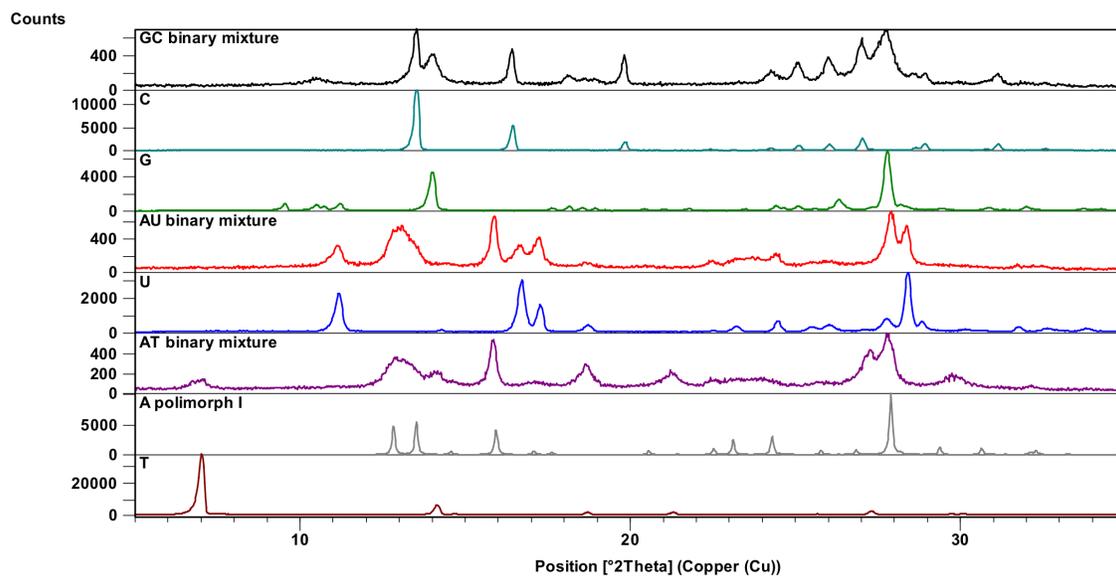


Figure S9 PXR D patterns of canonical complementary binary mixtures obtained by milling for 60 minutes, and respective reactants.

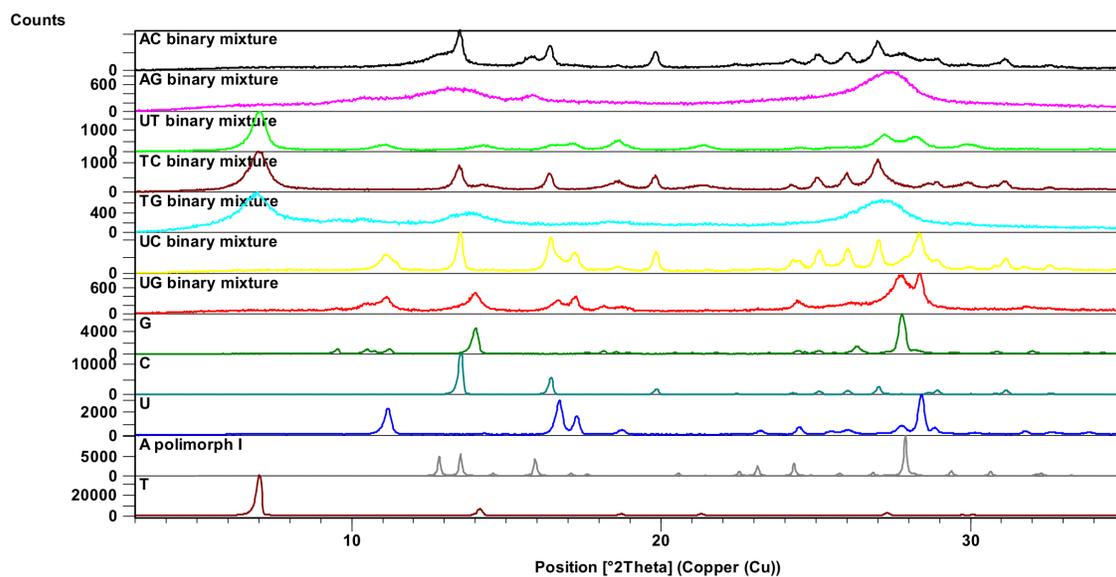


Figure S10 PXR D patterns of canonical non-complementary binary mixtures obtained by milling for 60 minutes, and respective reactants.

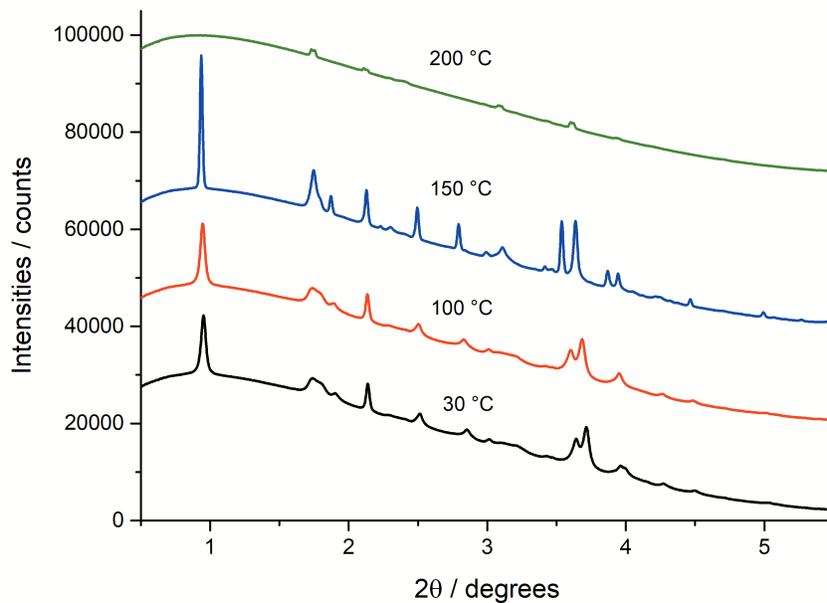


Figure S11 Dry heating physical mixture of A and T monitored in situ by synchrotron X-ray diffraction.

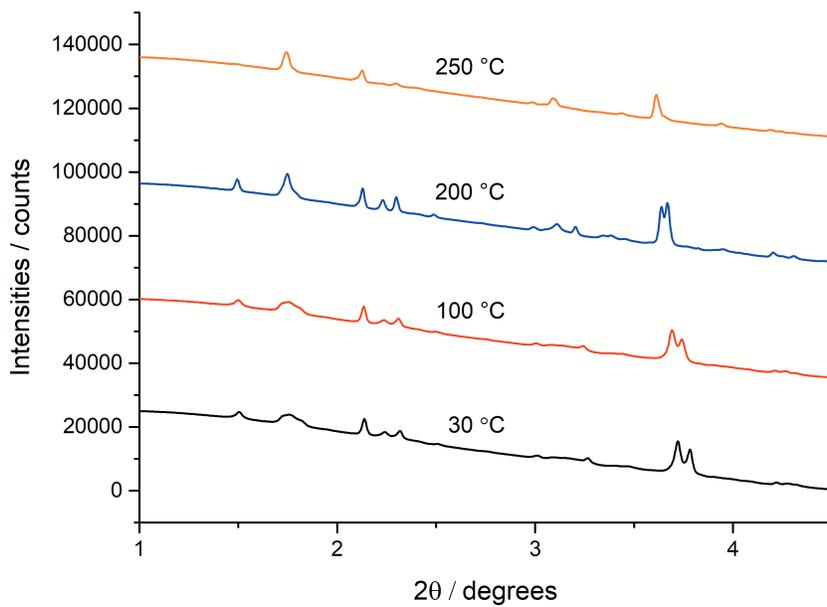


Figure S12 Dry heating physical mixture of A and U monitored in situ by synchrotron X-ray diffraction.

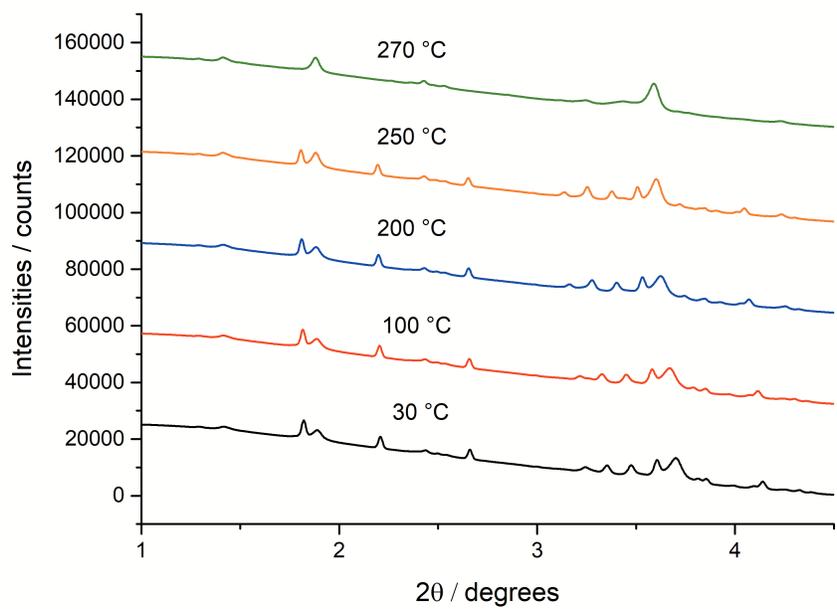


Figure S13 Dry heating physical mixture of G and C monitored in situ by synchrotron X-ray diffraction.

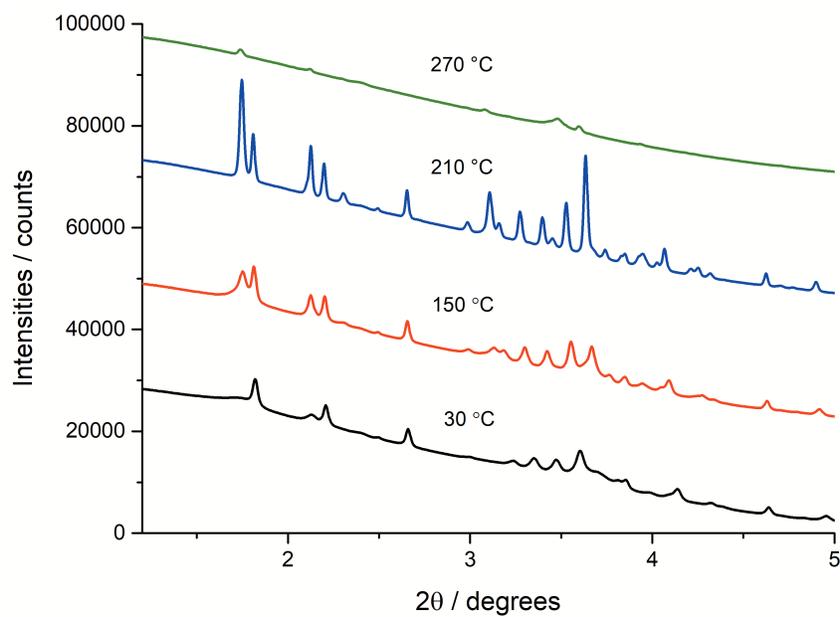


Figure S14 Dry heating physical mixture of A and C monitored in situ by synchrotron X-ray diffraction.

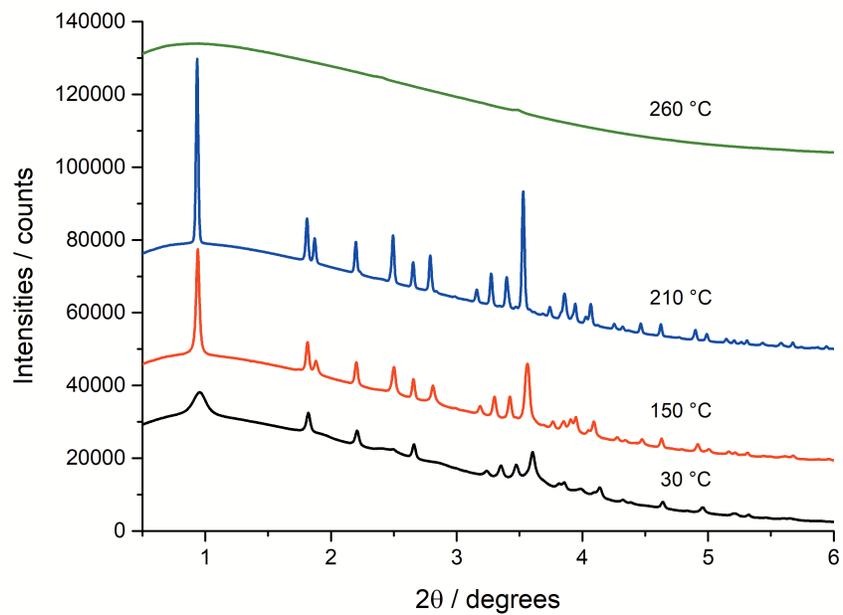


Figure S15 Dry heating physical mixture of T and C monitored in situ by synchrotron X-ray diffraction.

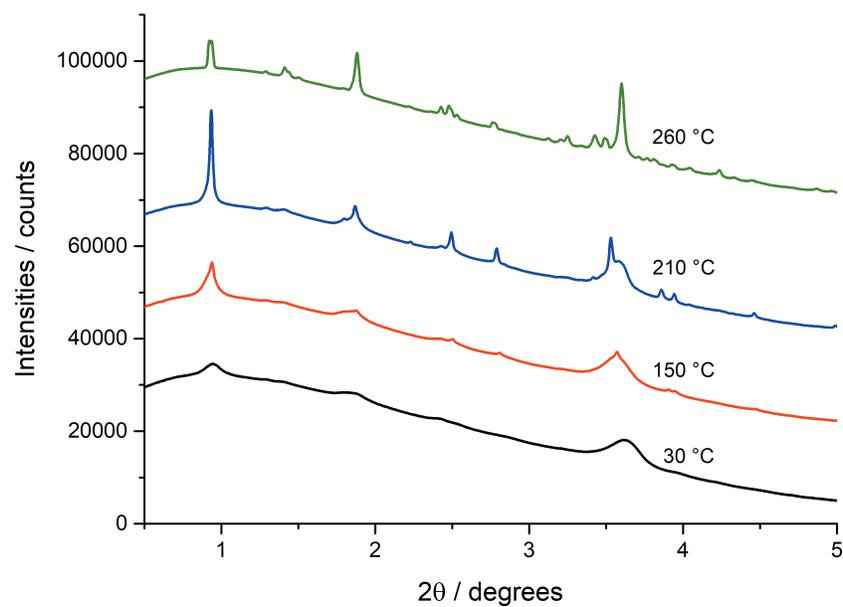


Figure S16 Dry heating physical mixture of T and G monitored in situ by synchrotron X-ray diffraction.

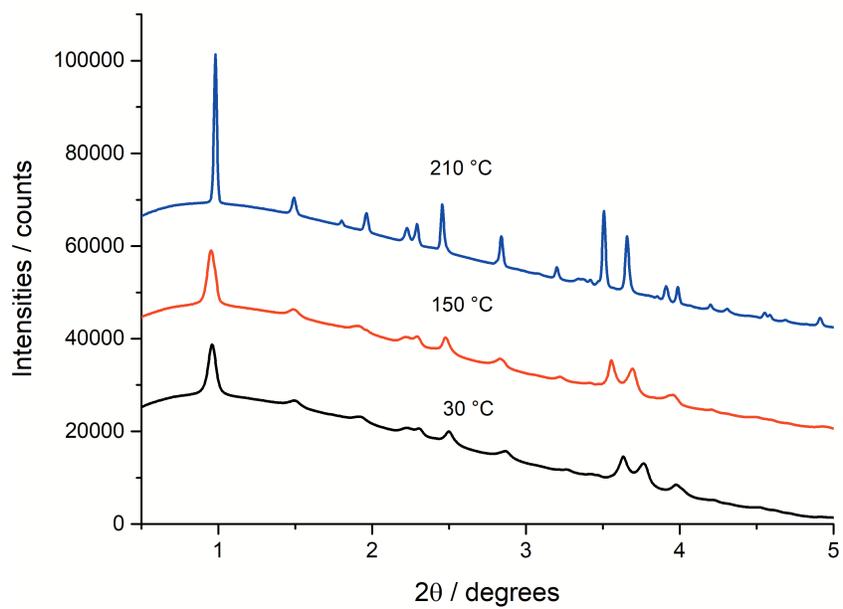


Figure S17 Dry heating physical mixture of T and U monitored in situ by synchrotron X-ray diffraction.

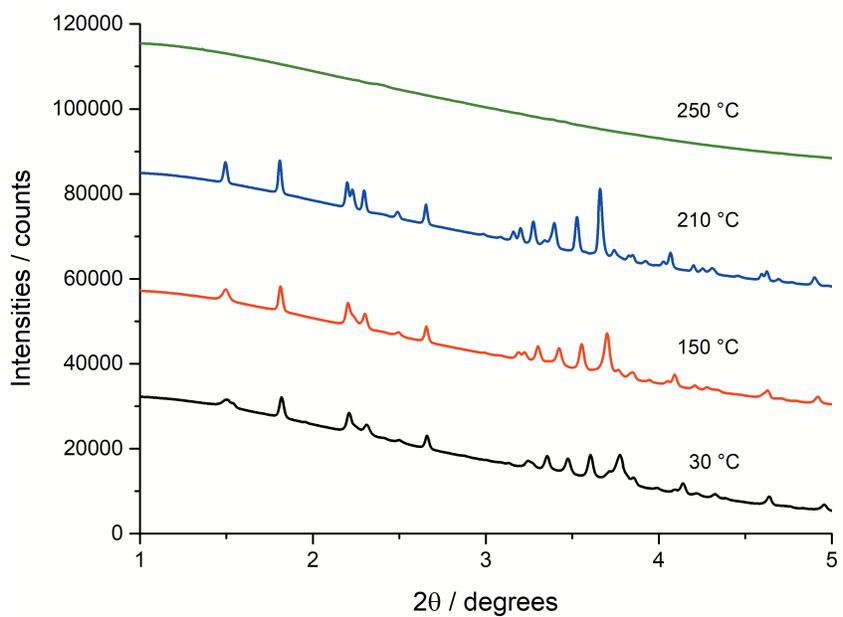


Figure S18 Dry heating physical mixture of U and C monitored in situ by synchrotron X-ray diffraction.

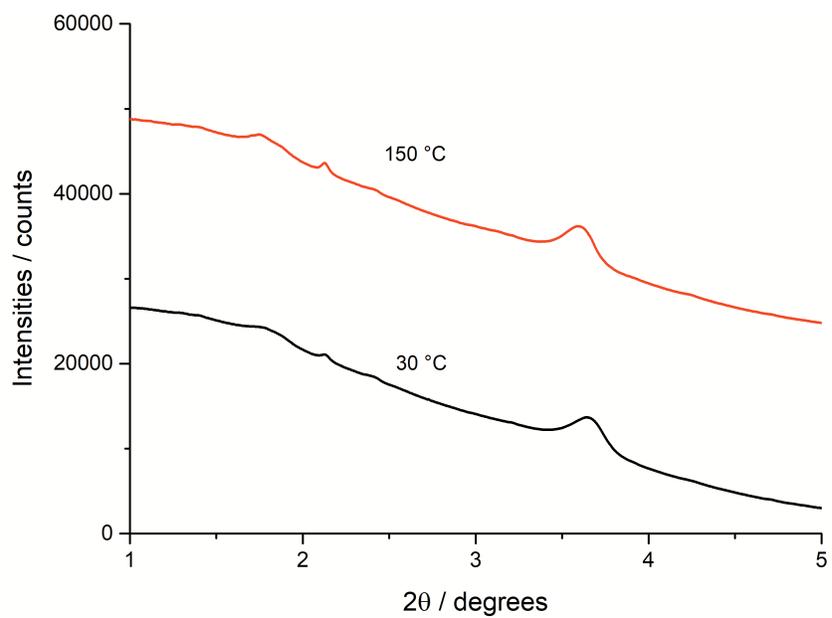


Figure S19 Dry heating physical mixture of A and G monitored in situ by synchrotron X-ray diffraction.

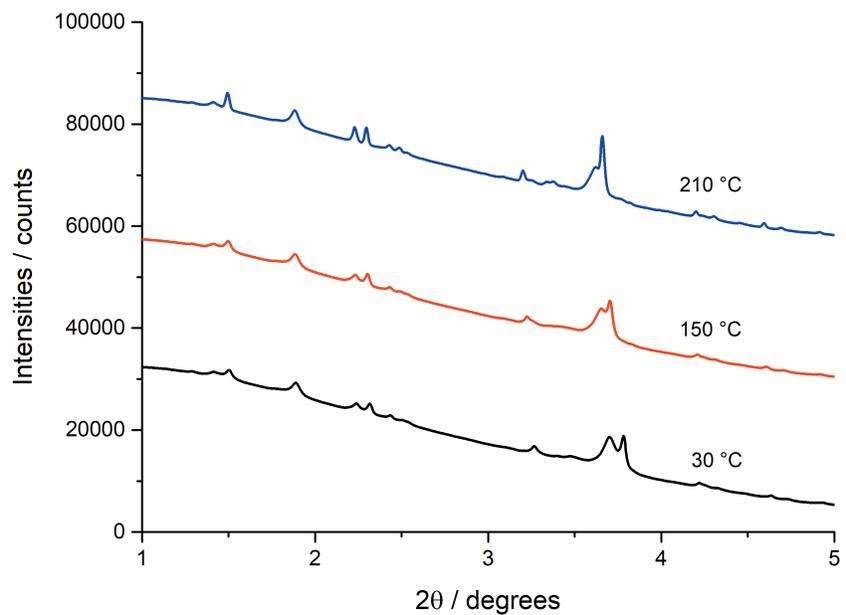


Figure S20 Dry heating physical mixture of U and G was monitored in situ by synchrotron X-ray diffraction.

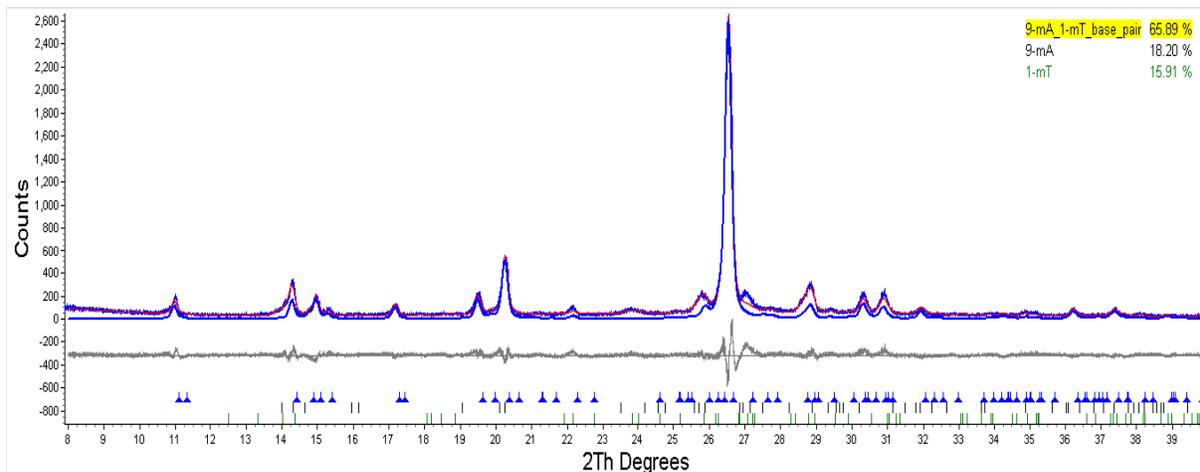


Figure S21 Rietveld refinement of PXR D pattern of 9-mA and 1-mT physical mixture heated at 100 °C. New crystalline phase corresponds to the known 9-mA:1-mT nucleobase pair (REFCODE: MTHMAD13).

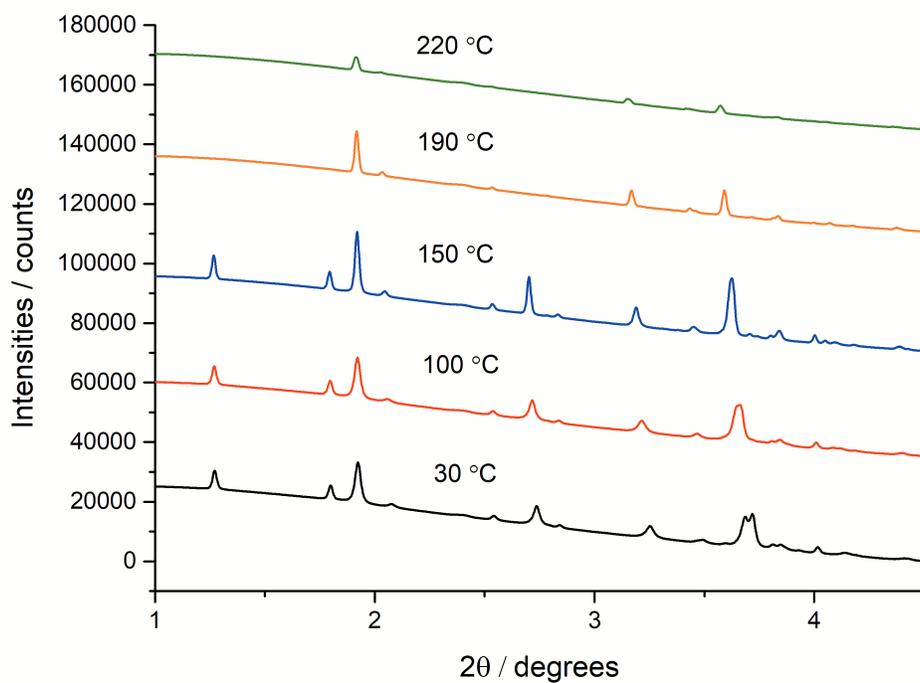


Figure S22 Dry heating physical mixture of 9-mA and 1-mU monitored in situ by synchrotron X-ray diffraction.

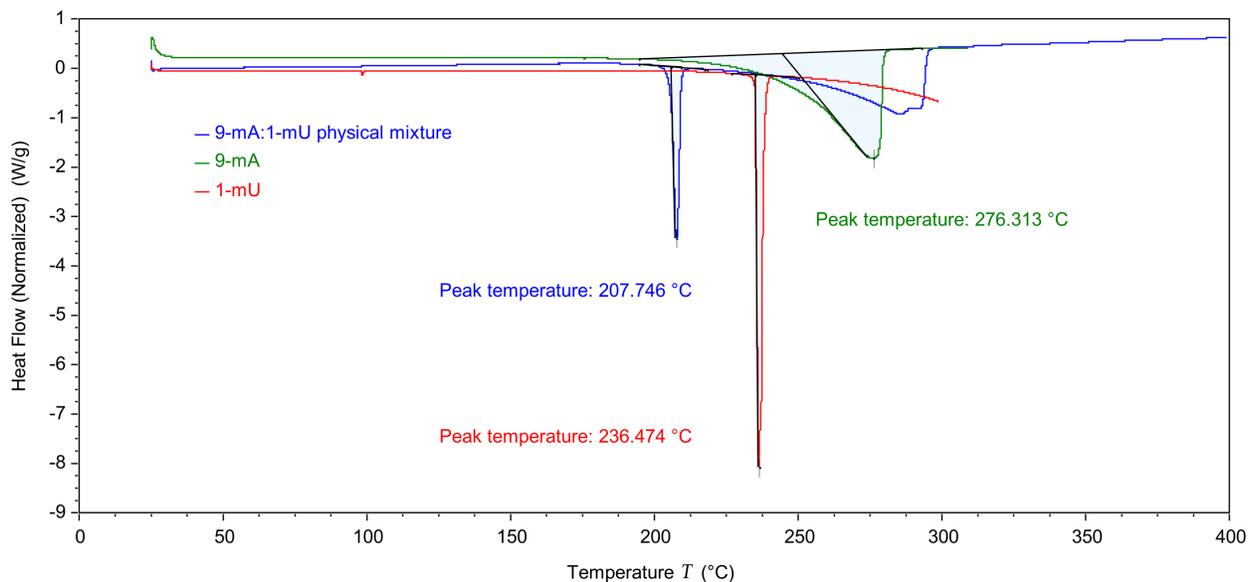


Figure S23 DSC thermograms of 9-mA and 1-mU physical mixture, and respective reactants. Endotherms corresponding to melting points are labeled.

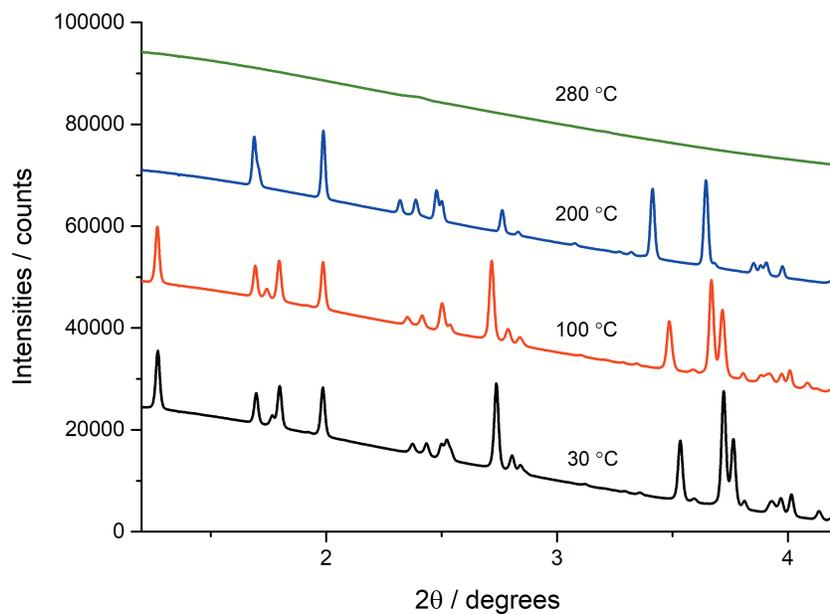


Figure S24 Dry heating physical mixture of 1-mU and 1-mC monitored in situ by synchrotron X-ray diffraction.

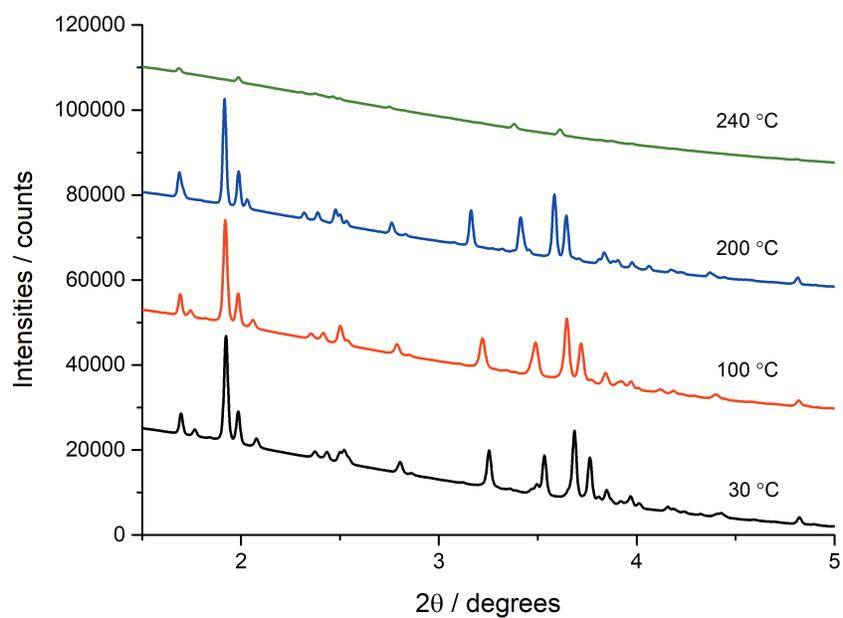


Figure S25 Dry heating physical mixture of 9-mA and 1-mC monitored in situ by synchrotron X-ray diffraction.

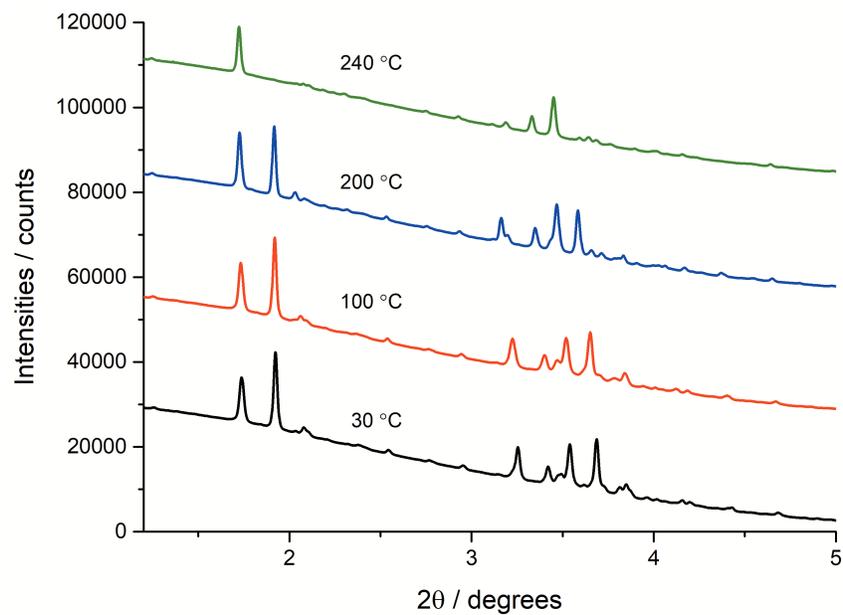


Figure S26 Dry heating physical mixture of 9-mA and 9-mG monitored in situ by synchrotron X-ray diffraction.

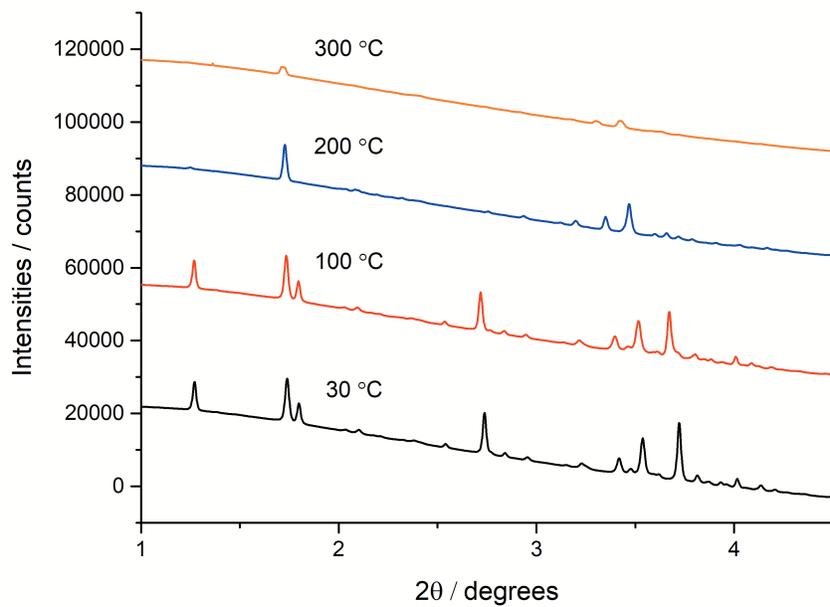


Figure S27 Dry heating physical mixture of 9-mG and 1-mU monitored in situ by synchrotron X-ray diffraction.

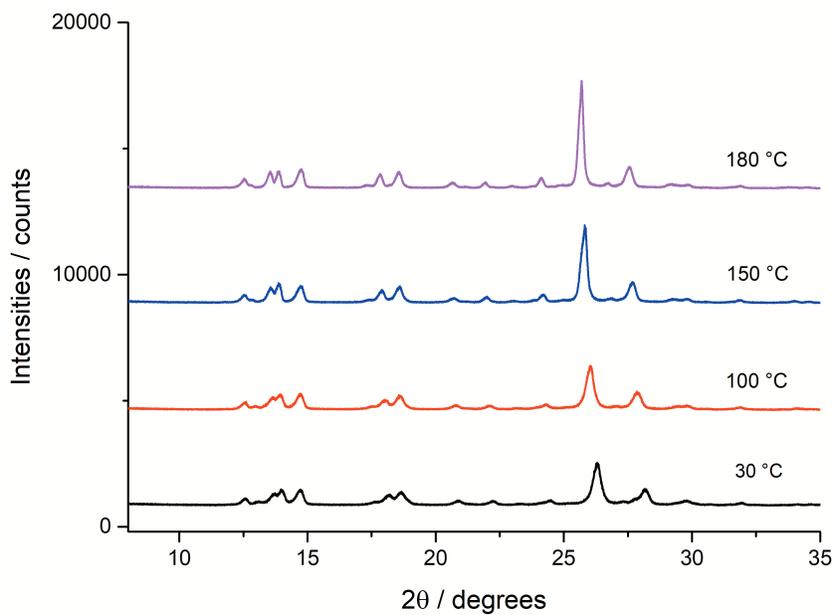


Figure S28 Dry heating physical mixture of 1-mT and 1-mC monitored in situ by laboratory X-ray diffraction.

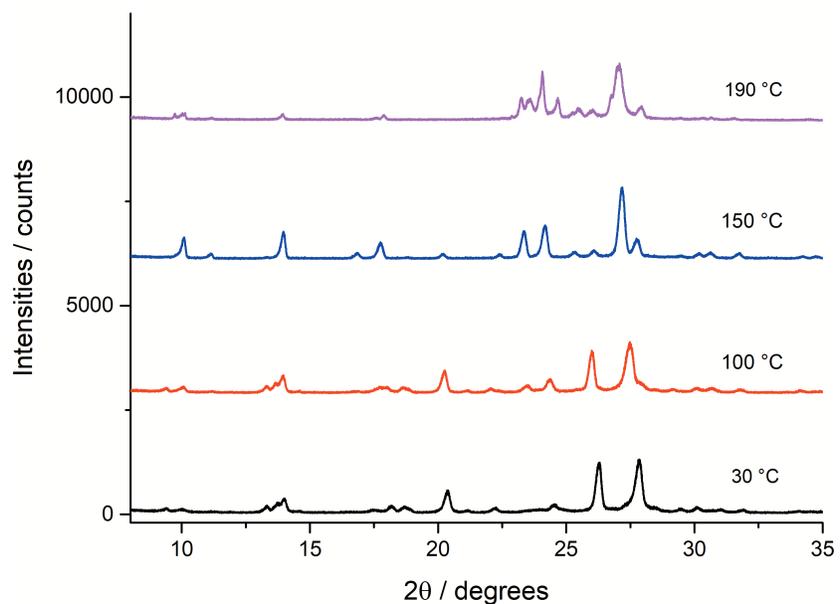


Figure S29 Dry heating physical mixture of 1-mT and 1-mU monitored in situ by laboratory X-ray diffraction.

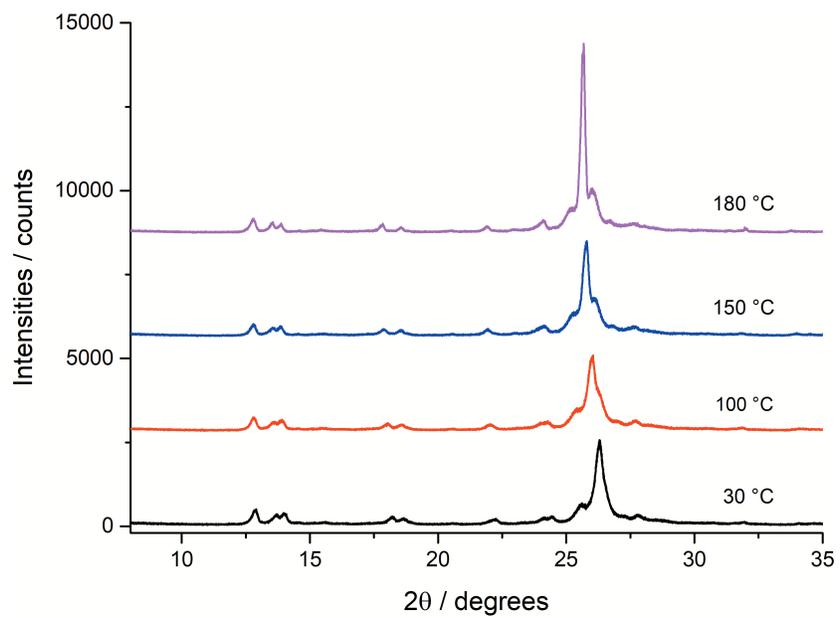


Figure S30 Dry heating physical mixture of 9-mG and 1-mT monitored in situ by laboratory X-ray diffraction.

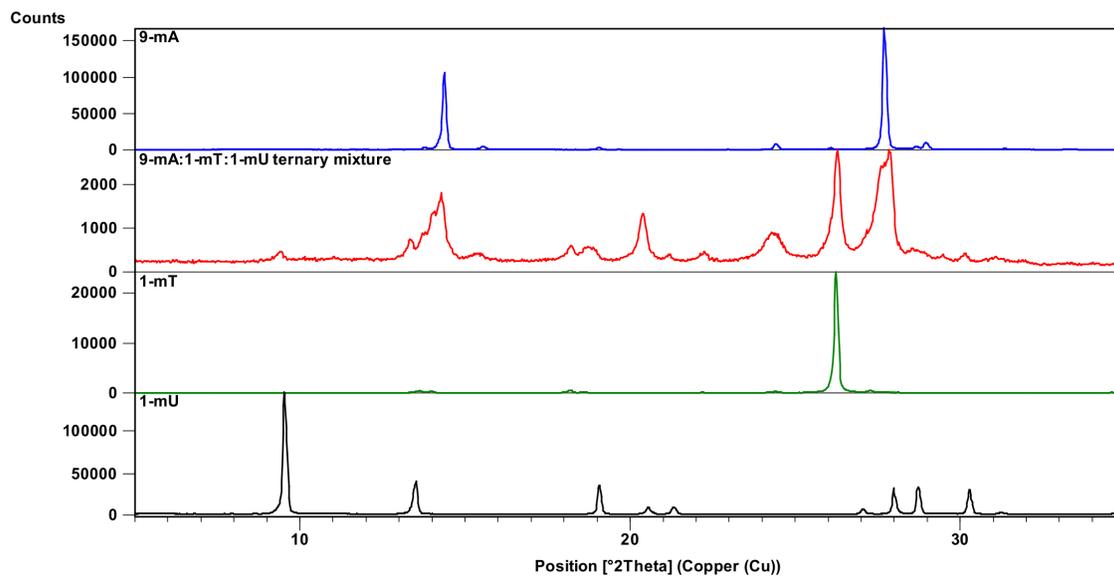


Figure S31 PXR D patterns of methylated ternary mixture obtained by milling, and respective reactants.

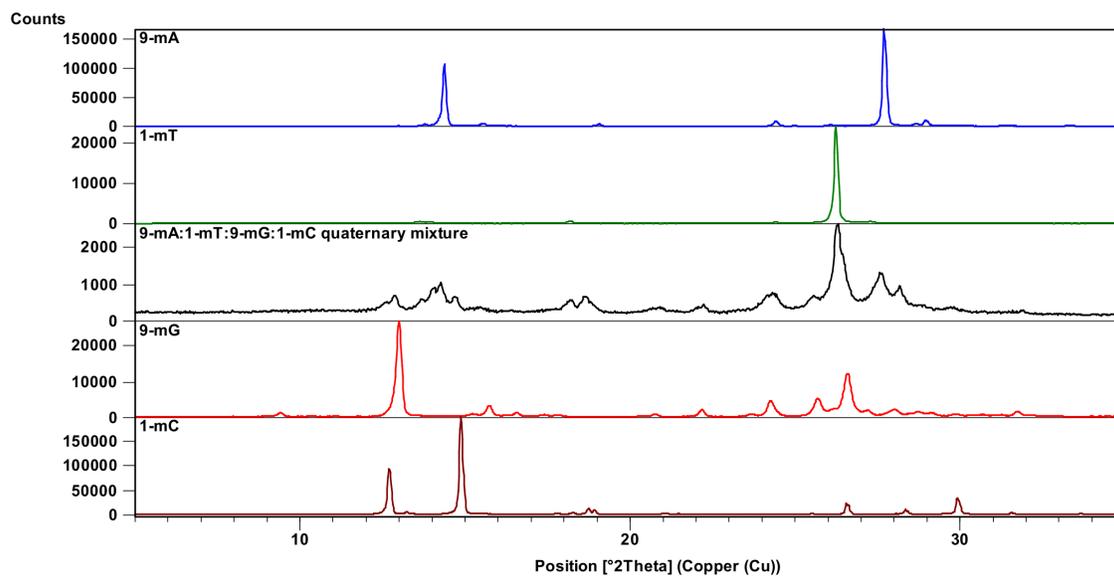


Figure S32 PXR D patterns of methylated quaternary mixture obtained by milling, and respective reactants.

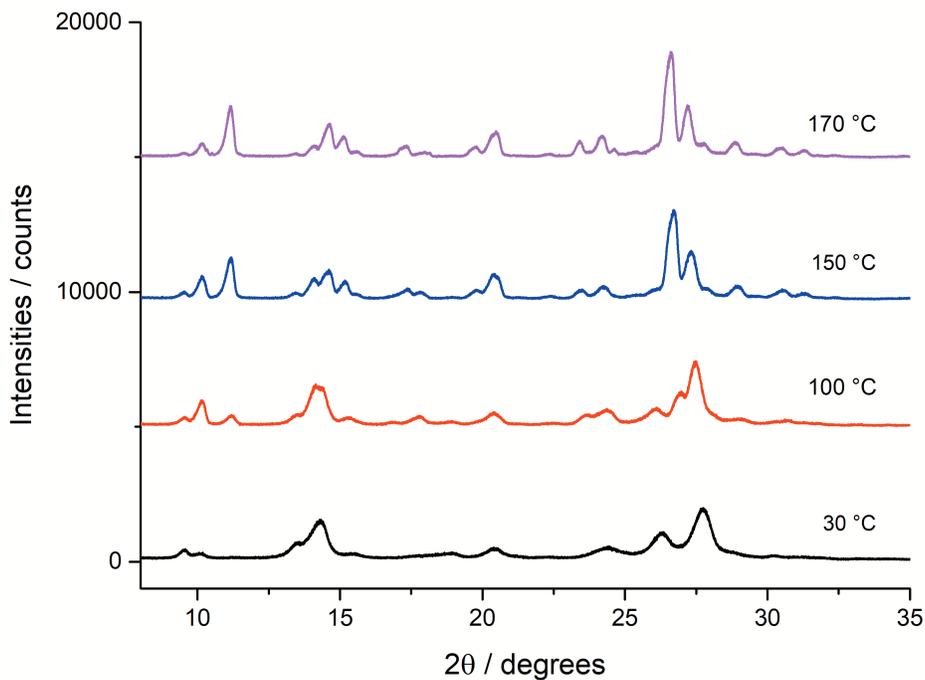


Figure S33 Dry heating ternary mixture of 9-mA, 1-mT, and 1-mU monitored in situ by laboratory X-ray diffraction.

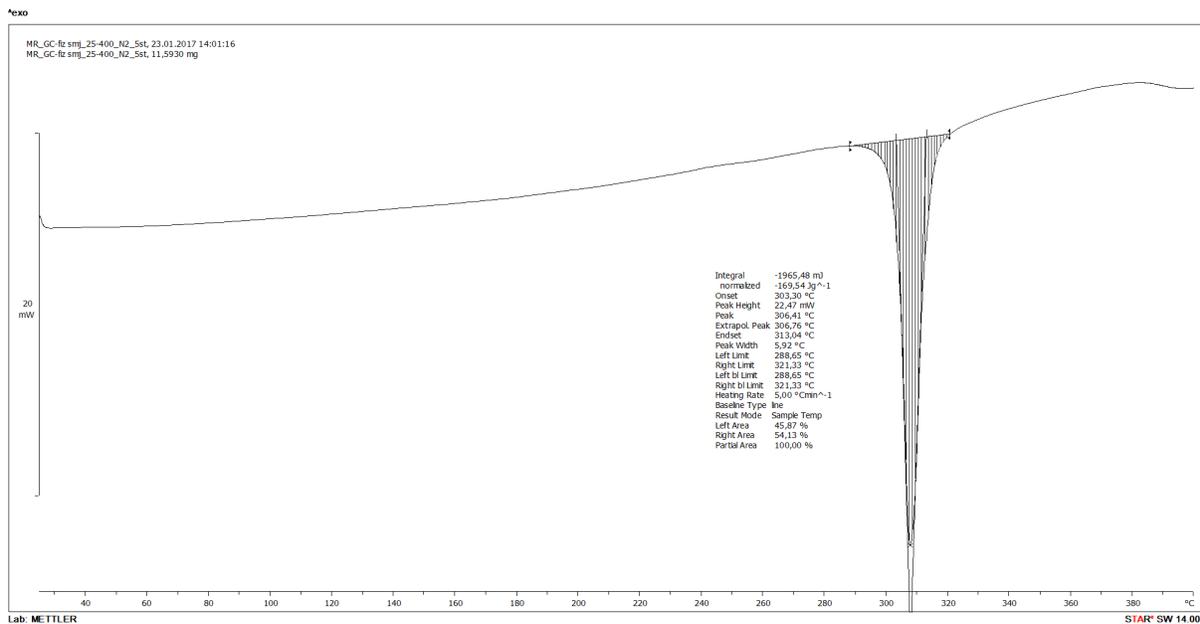


Figure S34 DSC thermogram of G and C physical mixture. Endotherm corresponding to melting point is labeled.

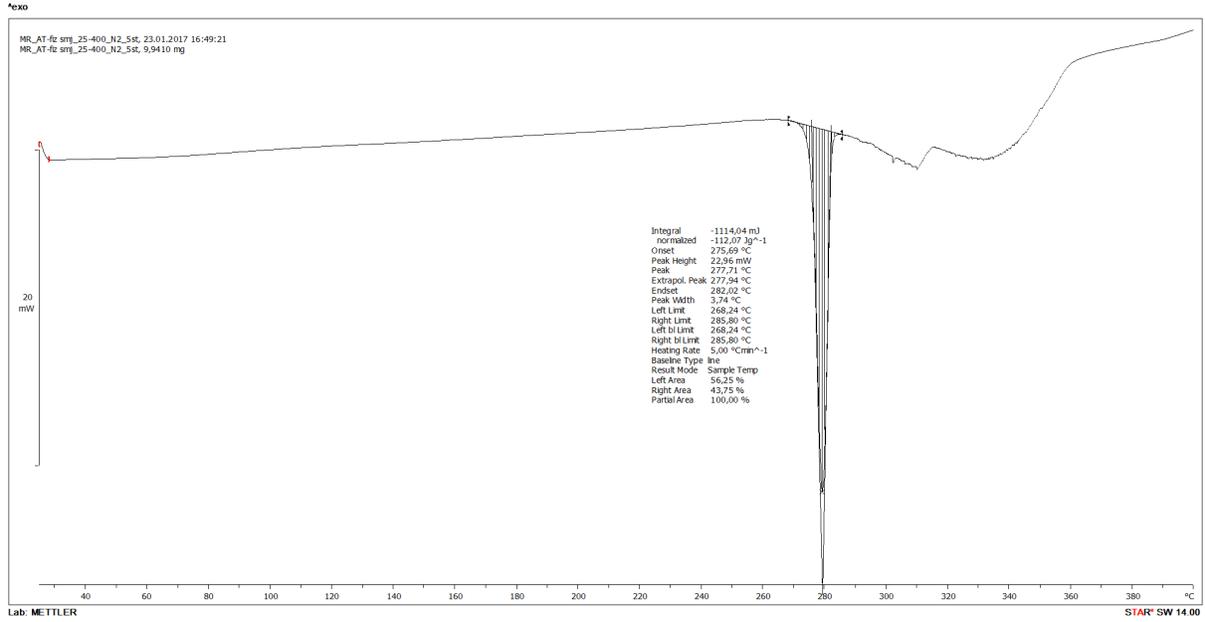


Figure S35 DSC thermogram of A and T physical mixture. Endotherm corresponding to melting point is labeled.

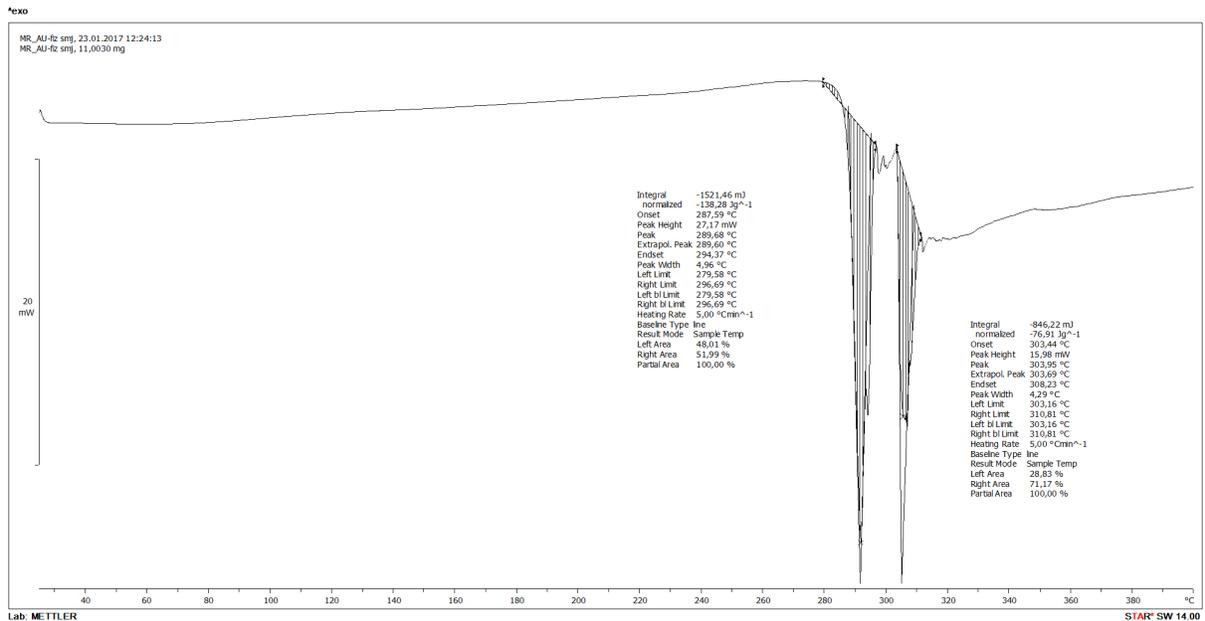


Figure S36 DSC thermogram of A and U physical mixture. Endotherms corresponding to melting points are labeled.

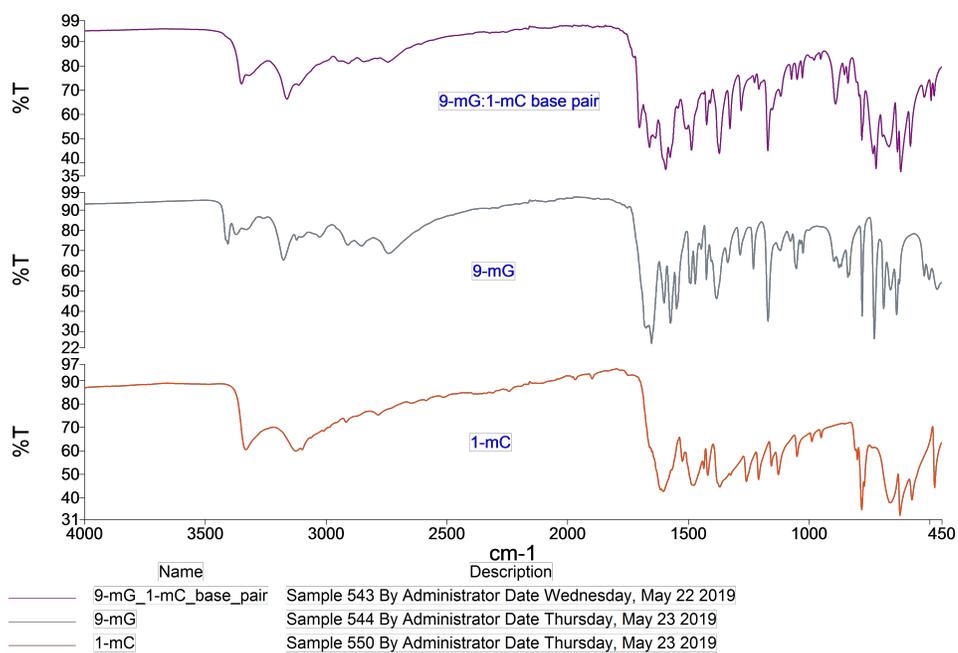


Figure S37 FTIR-ATR spectra of 9-mG:1-mC nucleobase pair, 9-mG, and 1-mC. Spectra were recorded on PerkinElmer SpectrumTwo spectrometer equipped with a diamond cell in a 4000-450 cm⁻¹ range.

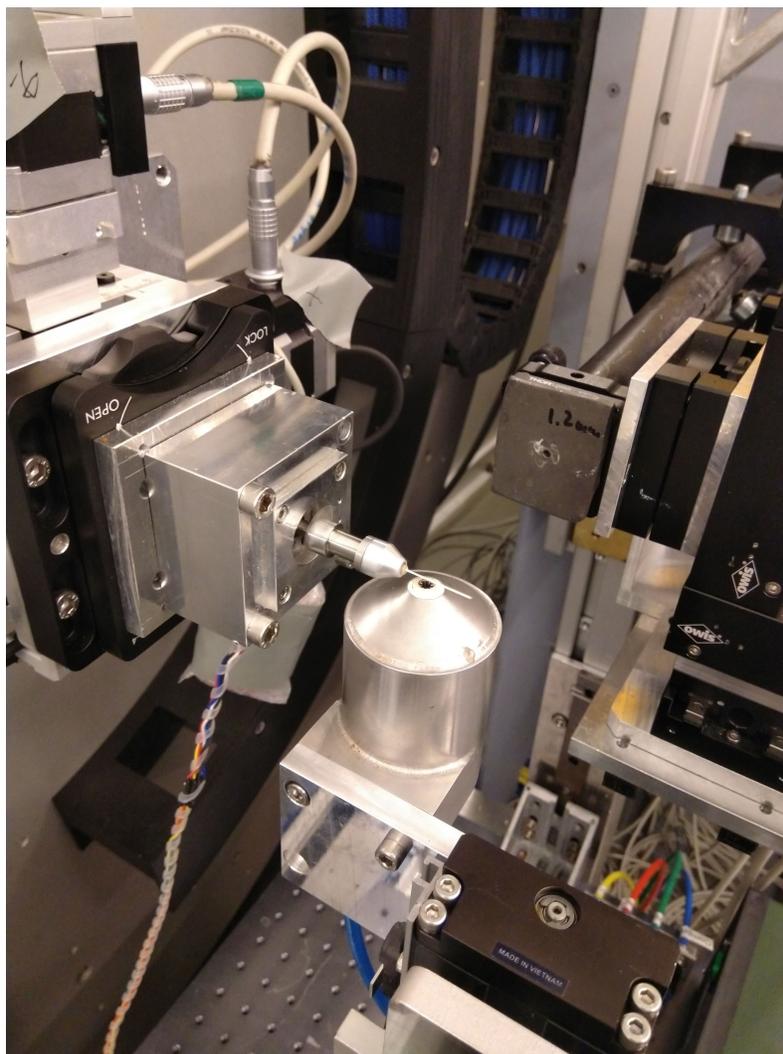


Figure S38 Photograph showing the experimental setup at P02.1 beamline, PETRA III, DESY. Hot-air blower system was placed directly below the measured capillary.

References

- [1] J. Filik, A. W. Ashton, P. C. Y. Chang, P. A. Chater, S. J. Day, M. Drakopoulos, M. W. Gerring, M. L. Hart, O. V. Magdysyuk, S. Michalik, A. Smith, C. C. Tang, N. J. Terrill, M. T. Wharmby and H. Wilhelm, *Journal of Applied Crystallography*, 2017, **50**, 959–966.
- [2] TOPAS 6 software from Bruker AXS 2016.
- [3] A. A. Coelho, *Journal of Applied Crystallography*, 2003, **36**, 86–95.
- [4] G. S. Pawley, *Journal of Applied Crystallography*, 1981, **14**, 357–361.
- [5] H. M. Rietveld, *Journal of Applied Crystallography*, 1969, **2**, 65–71.