Supporting information

A one-pot, water compatible synthesis of pyrimidine nucleobases under plausible prebiotic conditions

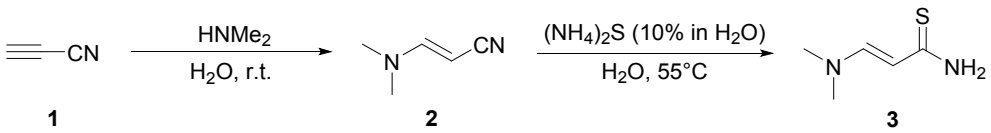
Hidenori Okamura, Sidney Becker, Niklas Tiede,
Stefan Wiedemann, Jonas Feldmann and Thomas Carell*

Center for Integrated Protein Science (CiPS^{M}) at the Department of Chemistry
Ludwig-Maximilians-Universität München, Butenandtstr. 5-13, 81377 München, Germany
E-mail: thomas.carell@lmu.de

1. General

Chemicals were purchased from Sigma-Aldrich, TCI, Fluka, ABCR, Carbosynth or Acros organics and used further purification. The solvents were of reagent grade or purified by distillation. Reactions and chromatography fractions were monitored by qualitative thin-layer chromatography (TLC) on silica gel F_{254} TLC plates from Merck KGaA. Flash column chromatography was performed on Geduran® Si60 (40-63 μm) silica gel from Merck KGaA. NMR spectra were recorded on Bruker AVIIIHD 400 spectrometers (400 MHz). ^{1}H NMR shifts were calibrated to the residual solvent resonances: $CDCl_3$ (7.26 ppm), DMSO-d_6 (2.50 ppm), CD_3OD (4.87 ppm), D_2O (4.79 ppm). ^{13}C NMR shifts were calibrated to the residual solvent: $CDCl_3$ (77.16 ppm), DMSO-d_6 (39.52 ppm), CD_3OD (49.00 ppm). All NMR spectra were analyzed using the program MestRE NOVA 10.0.1 from Mestrelab Research S. L. Normal resolved mass spectra were measured on a LTQ FT-ICR by Thermo Finnigan GmbH. High resolution mass spectra were measured by the analytical section of the Department of Chemistry of the Ludwigs-Maximilians-Universität München on the following spectrometers (ionization mode in brackets): MAT 95 (EI) and MAT 90 (ESI) from Thermo Finnigan GmbH. IR spectra were recorded on a PerkinElmer Spectrum BX II FT-IR system. All substances were directly applied as solids or on the ATR unit. Analytical RP-HPLC was performed on an analytical HPLC Waters Alliance (2695 Separation Module, 2996 Photodiode Array Detector) equipped with the column Nucleosil 120-2 C18 from Macherey Nagel using a flow of 0.5 ml/min, a gradient of 0-70% of buffer B in 30 min was applied for the measurements. Preparative RP-HPLC was performed on a HPLC Waters Breeze (2487 Dual λ Array Detector, 1525 Binary HPLC Pump) equipped with the column VP 250/32 C18 from Macherey Nagel. Using a flow of 5 ml/min, a gradient of 0-70% of buffer B in 30 min was applied for the purifications. Buffer A: 50 mM ammonium formate in H_2O, Buffer B: 50 mM ammonium formate in 80% (v/v) acetonitrile. Detection wavelength: 260 nm.

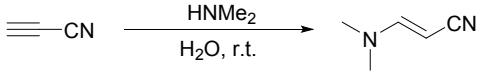
2. A stepwise synthesis of thioamide **3** under plausible prebiotic conditions



1 **2** **3**

Scheme S1. A stepwise synthesis of thioamide **3** from cyanoacetylene **1**

3-(Dimethylamino)acrylonitrile (2)

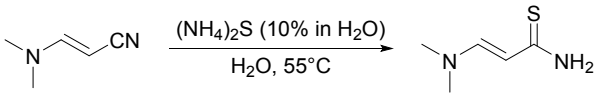


1 **2**

To a solution of dimethylamine (405 µl, 40% in H_2O, 3.19 mmol) in H_2O (16 ml) was added cyanoacetylene **1** (100 µl, 1.60 mmol) at room temperature. After 20 h, the colorless solution was extracted with CH_2Cl_2. The combined organic phase was washed with brine and dried over $MgSO_4$. The solvent was removed under reduced pressure to yield the compound **2** (140 mg, 1.50 mmol, 94%) as a colorless oil. The spectra data matched with the compound **2** purchased from TCI Germany (product number: D2094).

IR (neat, cm^{-1}) 3077 (w), 2915 (w), 2818 (w), 2189 (s), 1623 (s), 1485 (w), 1437 (m), 1420 (w), 1391 (s), 1346 (m), 1287 (m), 1214 (w), 1114 (m), 1066 (m), 1023 (w), 957 (m), 862 (w), 709 (m), ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.21 (d, J = 13.4 Hz, 1H), 3.82 (d, J = 13.5 Hz, 1H), 2.89 (br, 3H), 2.70 (br, 3H), ^{13}C NMR (100 MHz, DMSO-d_6) δ 155.1, 122.7, 58.3, HRMS (EI) calcd. for $C_5H_8N_2^+$ $[M]^+$: 96.0687, found: 96.0683.

3-(Dimethylamino)prop-2-enethioamide (3)

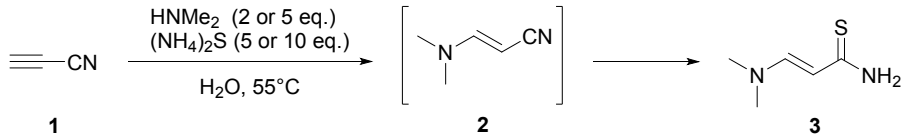


2 **3**

To a solution of $(NH_4)_2S$ (20 ml, 20% in H_2O, mmol) in H_2O (20 ml) was added compound **2** (2 ml, 19.76 mmol), and the reaction mixture was stirred at 55°C. After 12 h, the reaction mixture was cooled to room temperature. The compound **3** crystallized out from the reaction mixture as a yellow solid (1.63 g, 12.51 mmol, 63%).

IR (neat, cm^{-1}) 3337 (w), 3268 (w), 3158 (m), 1600 (s), 1495 (w), 1428 (m), 1372 (m), 1342 (m), 1310 (m), 1278 (s), 1114 (s), 1038 (s), 980 (s), 861 (s), 810 (s), 670 (m), ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.01 (br, 1H), 7.86 (br, 1H), 7.70 (d, J = 12.2 Hz, 1H), 5.20 (d, J = 12.2 Hz, 1H), 2.88 (s, 6H), ^{13}C NMR (100 MHz, DMSO-d_6) δ 194.6, 155.3, 96.5, HRMS (EI) calcd. for $C_5H_{10}N_2S^+$ $[M]^+$: 130.0565, found: 130.0560.

3. One-pot synthesis of thioamide **3** under plausible prebiotic conditions



Cyanoacetylene **1** (20 µl, 0.32 mmol) was added to a solution of dimethylamine (40% in H_2O, 2 or 5 eq.) and ammonium sulfide (20% in H_2O, 5 or 10 eq.) in H_2O (2 ml) at room temperature, and the reaction mixture was shaken at 55°C for 24 h in a thermomixer. The reaction was monitored by HPLC (Fig. S1). The product peak (compound **3**; Retention time: 12.8 min) was isolated and analyzed by NMR measurement by which the structural integrity was confirmed. The yield of the compound **3** from each reaction was determined according to the calibration curve prepared from the synthetically prepared compound **3** (Fig. S5a).(1)

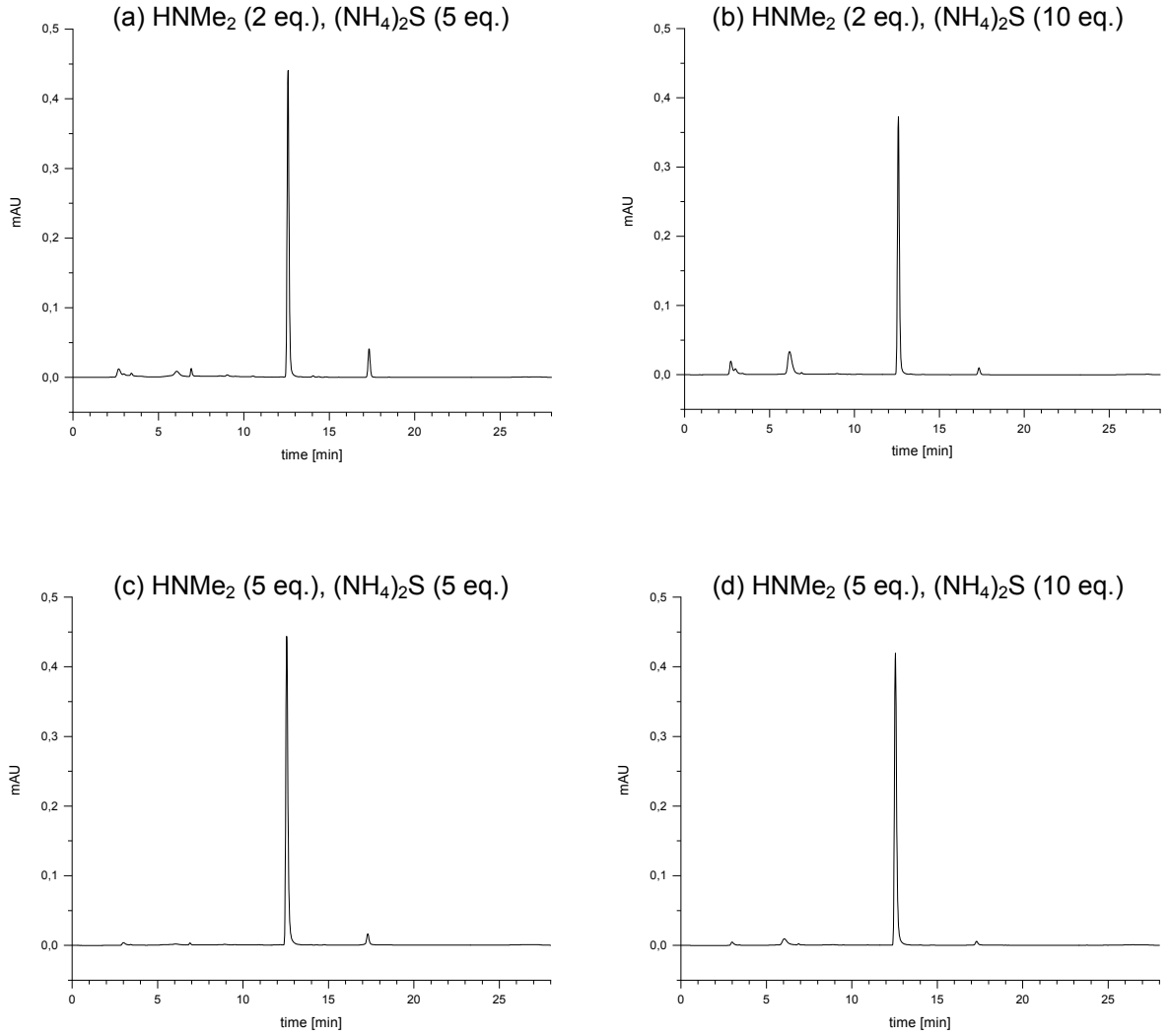
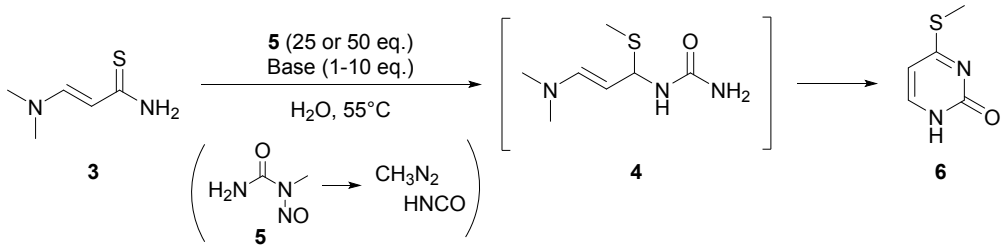


Figure S1. HPLC analysis of the one-pot thioamide **3** syntheses in different equivalents of the reagents.

4. Synthesis of SMePy **6** under plausible prebiotic conditions



To a solution of compound **3** (2 mg, mmol) in H_2O (1 ml) was added base (1~10 eq.) and *N*-methyl-*N*-nitrosourea **5**[2] (25 or 50 eq.), and the reaction mixture was shaken at 55°C for 24 h in a thermomixer. The reaction was monitored by HPLC (Fig. S2). The product peak (compound **6**; Retention time: 13.2 min) was isolated and analyzed by NMR measurement by which the structural integrity was confirmed. The yield of the compound **6** from each reaction was determined according to the calibration curve prepared from the synthetically prepared compound **6** (Fig. S5b).[3]

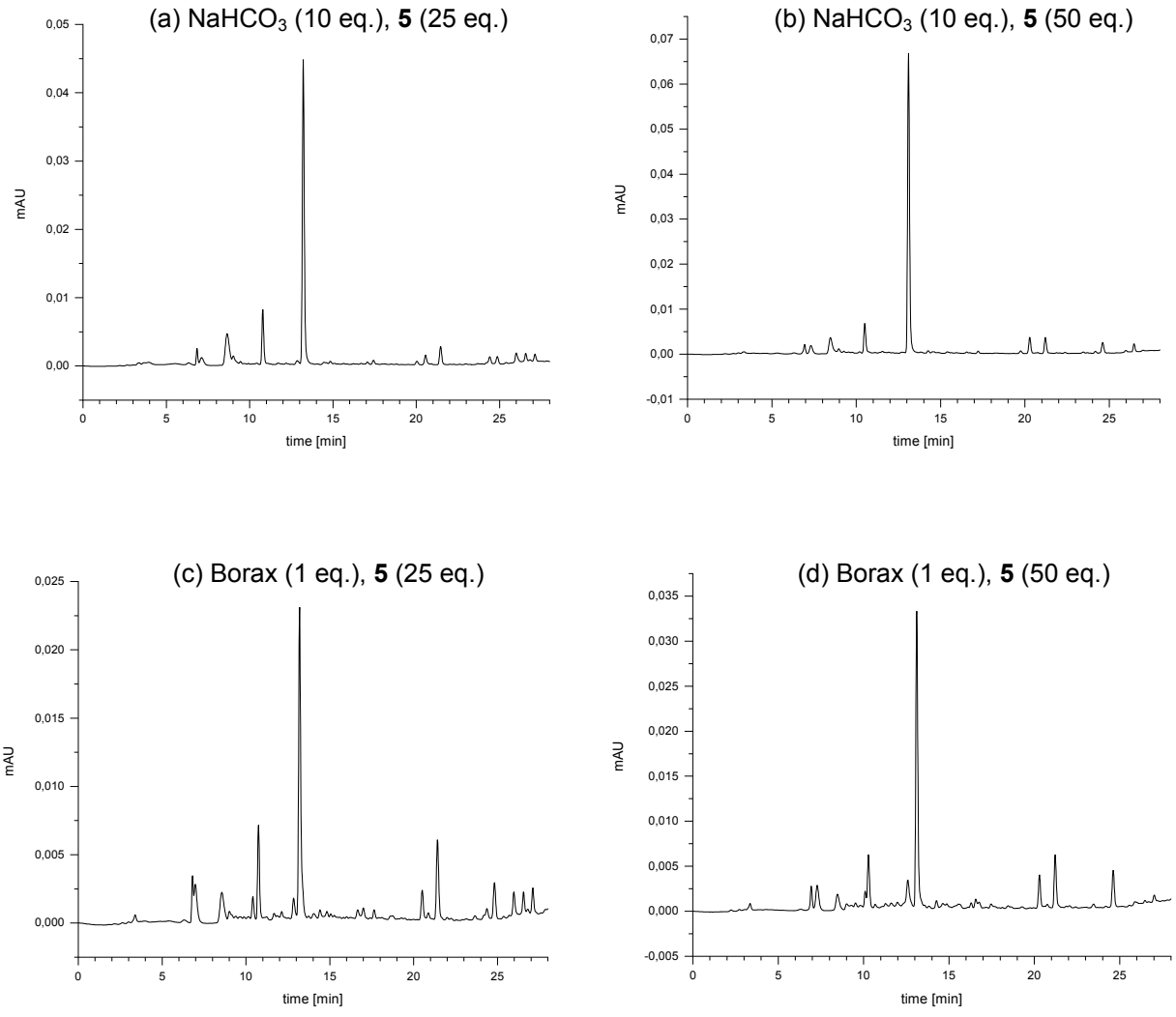


Figure S2. HPLC analysis of the SMePy **6** syntheses in the presence of different kinds of bases.

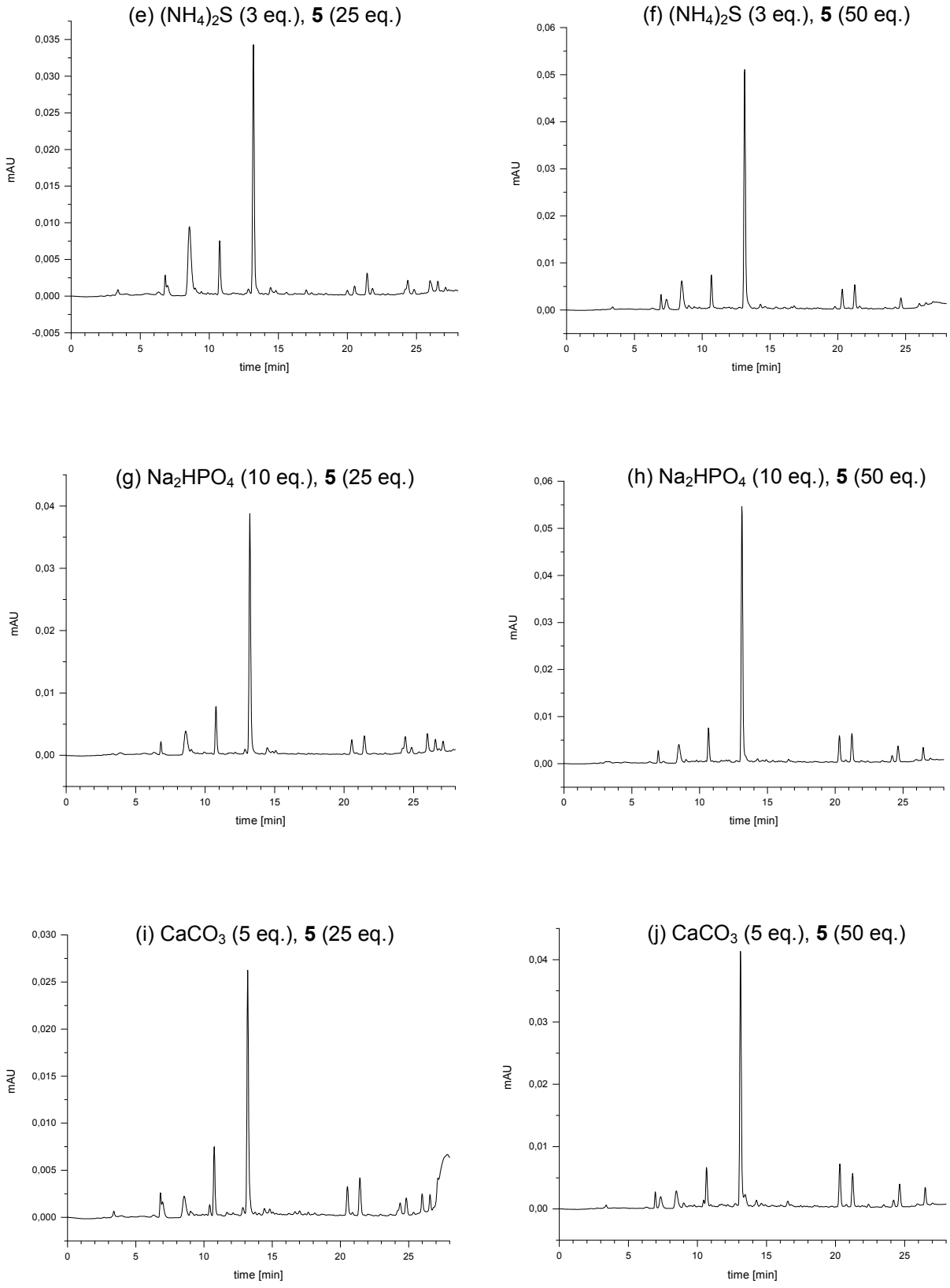
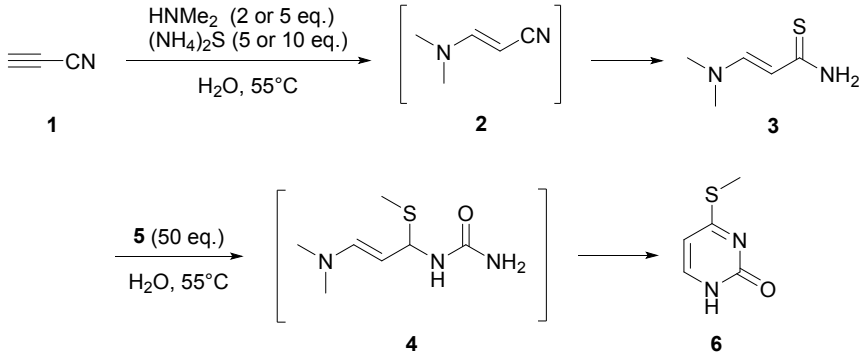


Figure S2 (continued). HPLC analysis of the SMePy **6** syntheses in the presence of different kinds of bases.

5. One-pot synthesis of SMePy **6** under plausible prebiotic conditions



Cyanoacetylene (20 µl, 0.32 mmol) was added to a solution of dimethylamine (40% in H_2O, 2 or 5 eq.) and ammonium sulfide (20% in H_2O, 5 or 10 eq.) in H_2O (2 ml) and samples were shaken at 55°C for 24 h. The solutions were made up to 16 ml with H_2O to dissolve precipitate completely, and 1 ml of the solution was transferred into a new falcon tube. *N*-methyl-*N*-nitrosourea **5** (50 eq.) was added and the suspension was shaken at 55°C in a thermomixer for 24 h. The reaction was monitored by HPLC (Fig. S3). The yield of the compound **6** (Retention time: 13.2 min) from each reaction was determined according to the calibration curve (Fig. S5b).

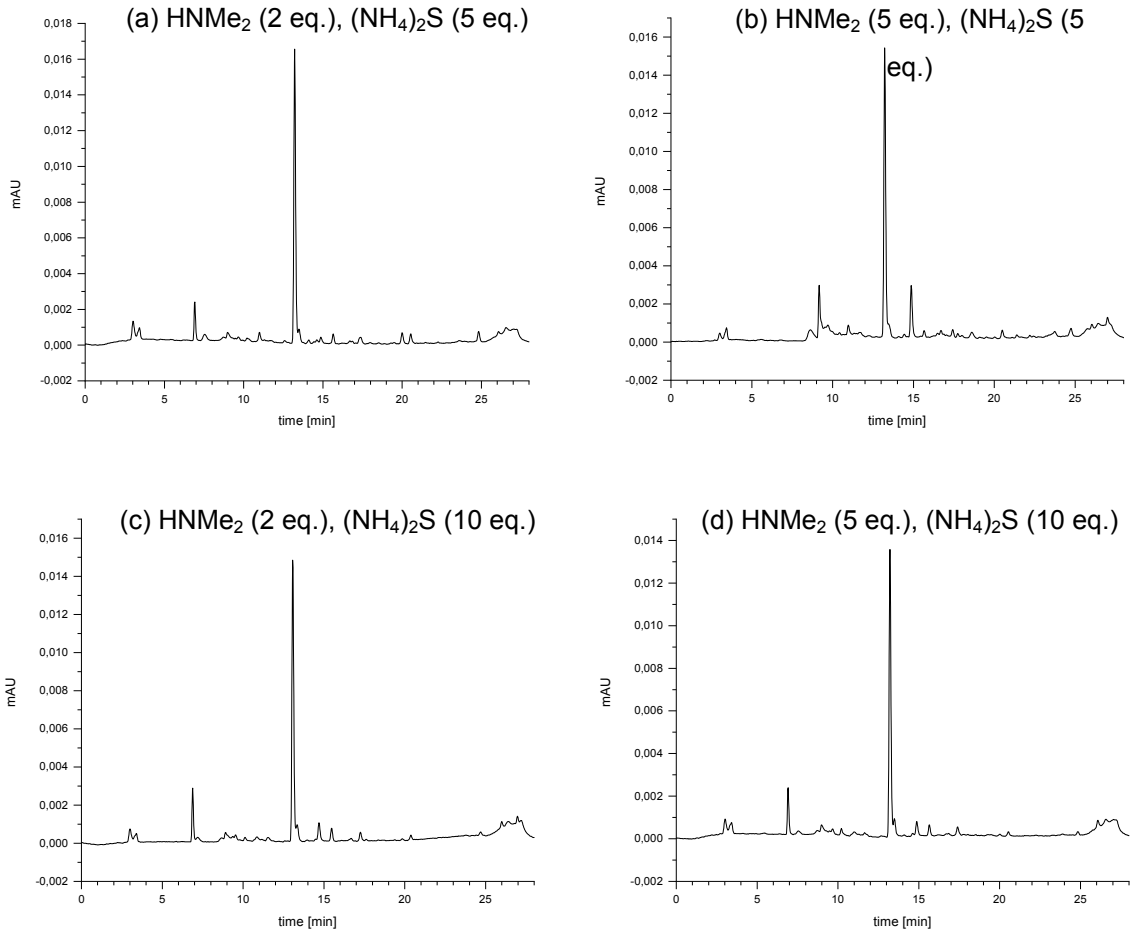
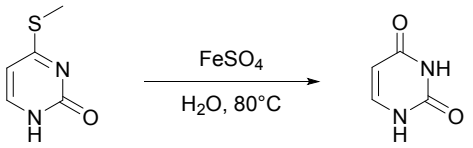


Figure S3. HPLC analysis of SMePy **6** syntheses in one-pot in different equivalents of the reagents.

6. Conversion of SMePy **6** into 4-substituted pyrimidine nucleobases

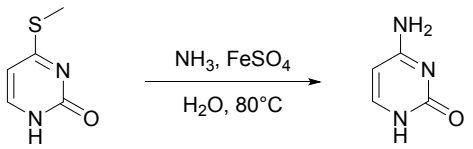
Uracil (7a)



6 **7a**

Compound **6** (10 mg, 70 μmol) was added to a solution of $FeSO_4$-$7H_2O$ (19.6 mg, 70 μmol) in H_2O (5 ml) and the reaction mixture was stirred at 80°C. After 3 days, the reaction mixture was lyophilized. Purification of the crude product by flash column chromatography (CH_2Cl_2:MeOH = 90:10 to 80:20) yielded compound **7a** (7.8 mg, 69 μmol, quant.) as a colorless solid. The analytical data matched those reported previously.[4]

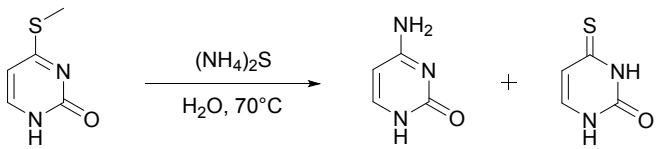
Cytosine (7b)



6 **7b**

Compound **6** (10 mg, 70 μmol) was added to a solution of $FeSO_4$-$7H_2O$ (19.6 mg, 70 μmol) in H_2O (2.6 ml) and ammonia solution (28% in H_2O, 1.7 ml) and the reaction mixture was stirred at 80°C. After 3 days, the reaction mixture was freeze dried. Purification of the crude product by flash column chromatography (CH_2Cl_2:MeOH = 90:10 to 50:50) yielded compound **7b** (6.0 mg, 54 μmol, 77%) as a white foam. The analytical data matched those reported previously.[5]

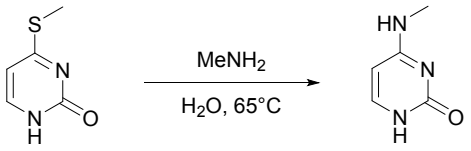
Cytosine (7b) and 4-thiouracil (7c)



6 **7b** **7c**

A solution of the compound **6** (0.8 μmol) in $(NH_4)_2S$ (1 ml, 1, 5 or 10%) was stirred at 70°C for 16 h. The reaction was monitored by HPLC (Fig. S4a-c). The product peaks (compound **7b**; Retention time: 7.0 min, compound **7c**; Retention time: 10.7 min) were isolated and analyzed by NMR measurement by which their structural integrity was confirmed. The analytical data matched those reported previously.[5],[6] The yields of the compound **7b** and **7c** from each reaction were determined according to the calibration curves (Fig. S5c, d).

4-Methylcytosine (7d)

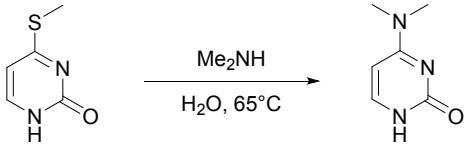


6 **7d**

A solution of compound **6** (2 mg, 14.1 μmol) and methylamine (40% in H_2O, 20 eq.) in H_2O (1 ml) was stirred at 70°C for 36 h. The reaction was monitored by HPLC (Fig. S4d). The product peak (compound **7d**; Retention time: 9.6 min) was isolated and analyzed by NMR measurement by which the structural integrity was confirmed. The yield of the compound **7d** was determined according to the calibration curves (Fig. S5e).

IR (neat, cm^{-1}) 2815 (w), 1620 (s), 1536 (m), 1516 (m), 1416 (m), 1396 (m), 1380 (m), 1312 (m), 1237 (w), 1175 (w), 1175 (w), 1125 (w), 1064 (w), 1016 (w), 970 (w), 924 (m), 788 (m), 775 (m), 729 (m), ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.21 (s, 1H), 7.55 - 7.49 (m, 1H), 7.24 (d, J = 7.0 Hz, 1H), 5.56 (d, J = 7.1 Hz, 1H), 2.73 (d, J = 4.7 Hz, 3H), ^{13}C NMR (100 MHz, DMSO-d_6) δ 165.4, 157.2, 141.6, 93.7, 27.2, HRMS (ESI) calcd. for $C_5H_8N_3O^+$ $[M+H]^+$: 126.0662, found: 126.0662.

4,4-Dimethylcytosine (7e)

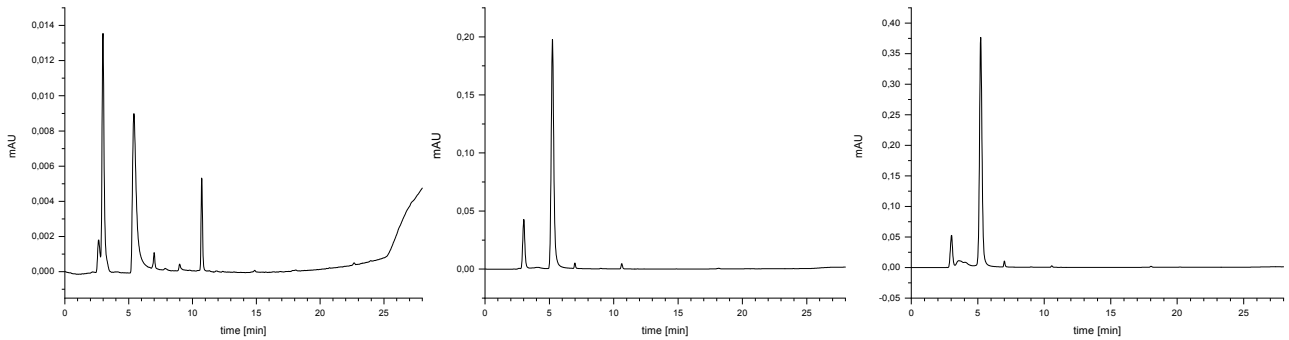


6 **7e**

A solution of compound **6** (2 mg, 14.1 μmol) and dimethylamine (40% in H_2O, 20 eq.) in H_2O (1 ml) was stirred at 70°C for 36 h. The reaction was monitored by HPLC (Fig. S4e). The product peak (compound **7e**; Retention time: 11.6 min) was isolated and analyzed by NMR measurement by which the structural integrity was confirmed. The yield of the compound **7e** was determined according to the calibration curves (Fig. S5f).

IR (neat, cm^{-1}) 3269 (w), 3126 (w), 3046 (m), 1656 (s), 1629 (s), 1604 (s), 1569 (s), 1536 (m), 1471 (m), 1426 (s), 1397 (s), 1340 (s), 1293 (m), 1226 (s), 1153 (m), 1083 (m), 1003 (w), 956 (w), 878 (m), 816 (s), 786 (s), 772 (m), 750 (m), 712 (m), ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.43 (br, 1H), 7.40 (d, J = 7.4 Hz, 1H), 5.87 (d, J = 7.3 Hz, 1H), 3.01 (br, 6H), ^{13}C NMR (100 MHz, DMSO-d_6) δ 164.8, 156.4, 143.2, 90.3, HRMS (ESI) calcd. for $C_6H_{10}N_3O^+$ $[M+H]^+$: 140.0818, found: 140.0818.

(a) 1% $(NH_4)_2S$ (b) 5% $(NH_4)_2S$ (c) 10% $(NH_4)_2S$



(d) 20 eq. H_2NMe (e) 20 eq. $HNMe_2$

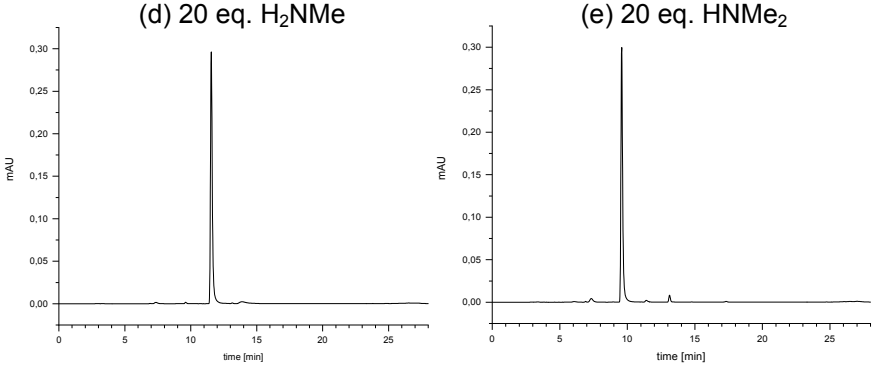
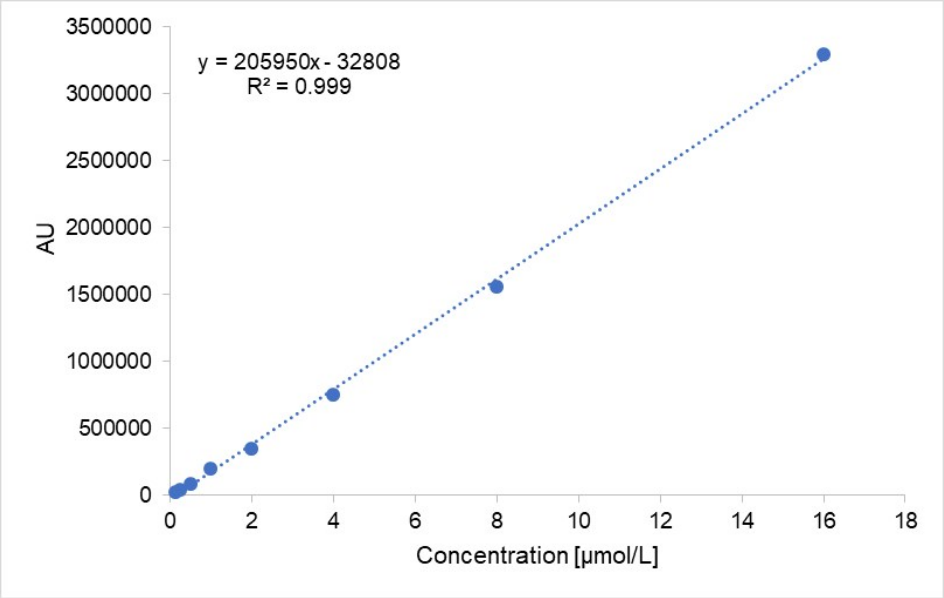


Figure S4. HPLC analysis of 4-substituted pyrimidine syntheses.

7. Calibration curves of the compounds synthesized in this study

To determine the concentration of a compound by HPLC a solution of the respective compound with a known concentration was made three times. The integrals and the arithmetic middle were calculated, and the average area under the curve from three independent experiment was plotted against the concentration. A regression line was generated and its equation was extracted.

(a) Thioamide **3**



(b) SMePy **6**

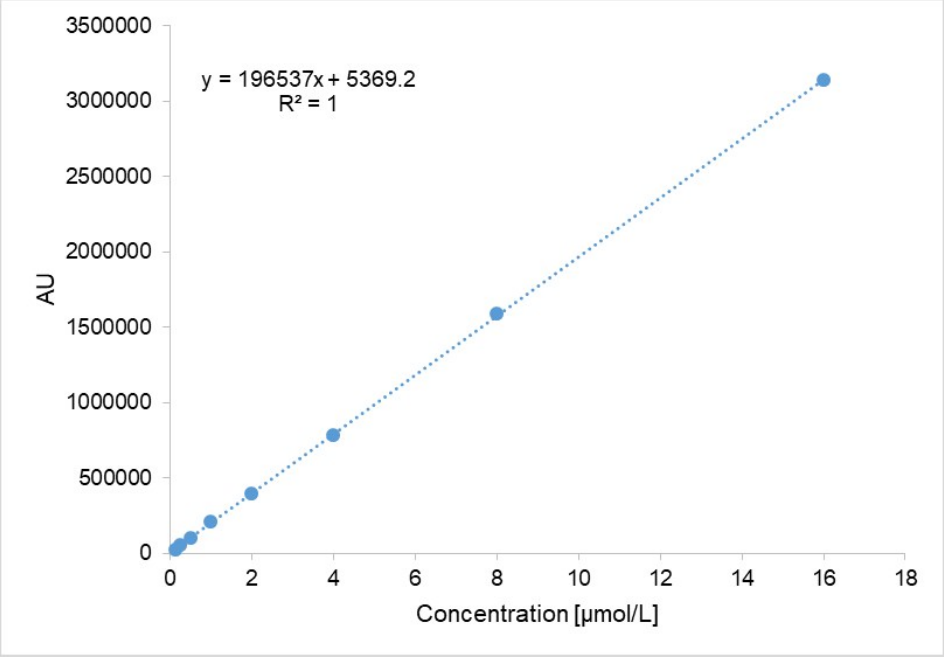
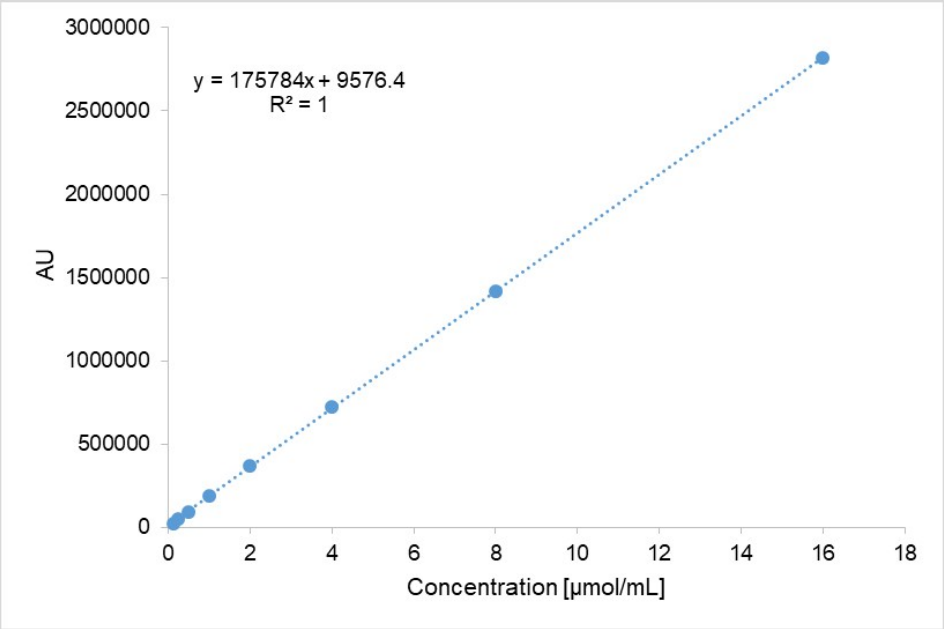


Figure S5. Calibration curves of the compounds synthesized in this study.

(c) Cytosine **7b**



(d) 4-Thiouracil **7c**

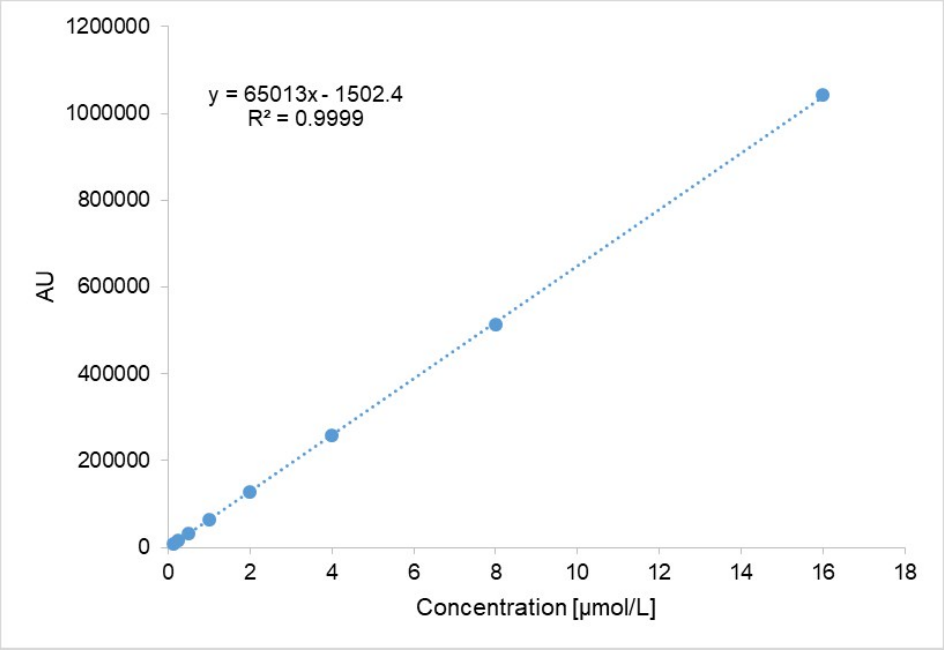
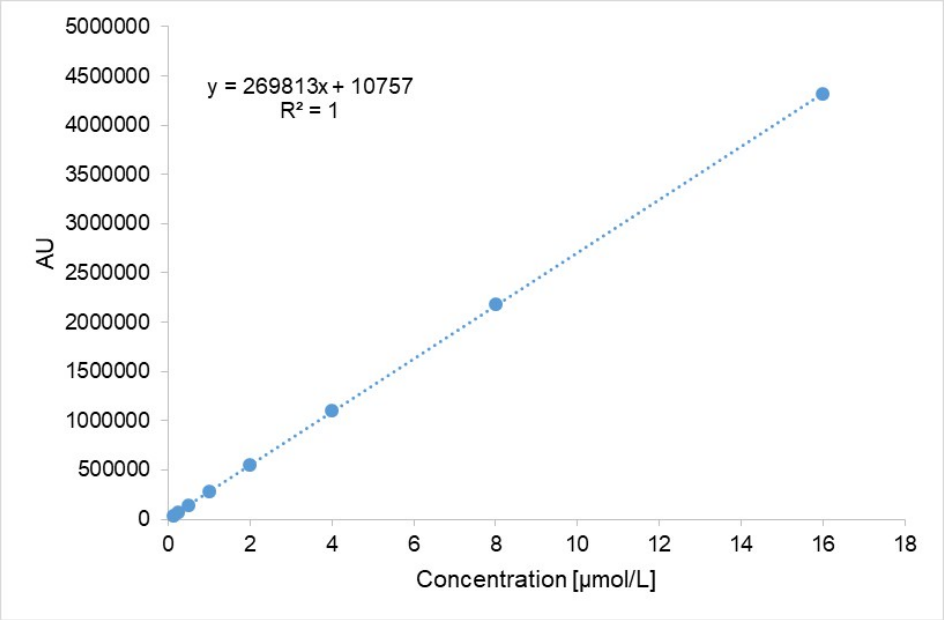


Figure S5 (continued). Calibration curves of the compounds synthesized in this study.

(e) 4-Methylcytosine



(f) 4,4-Dimethylcytosine

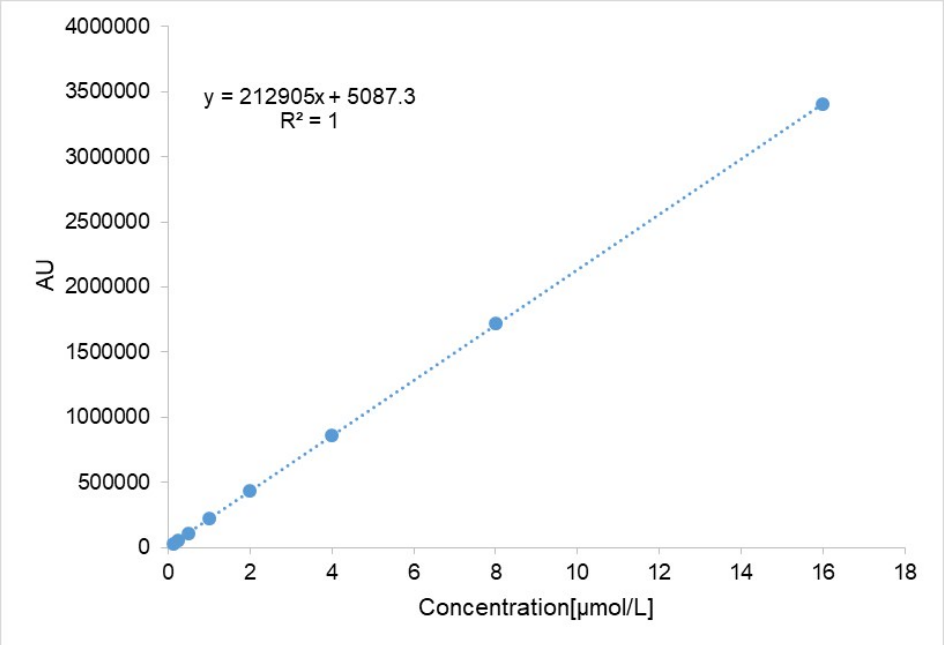
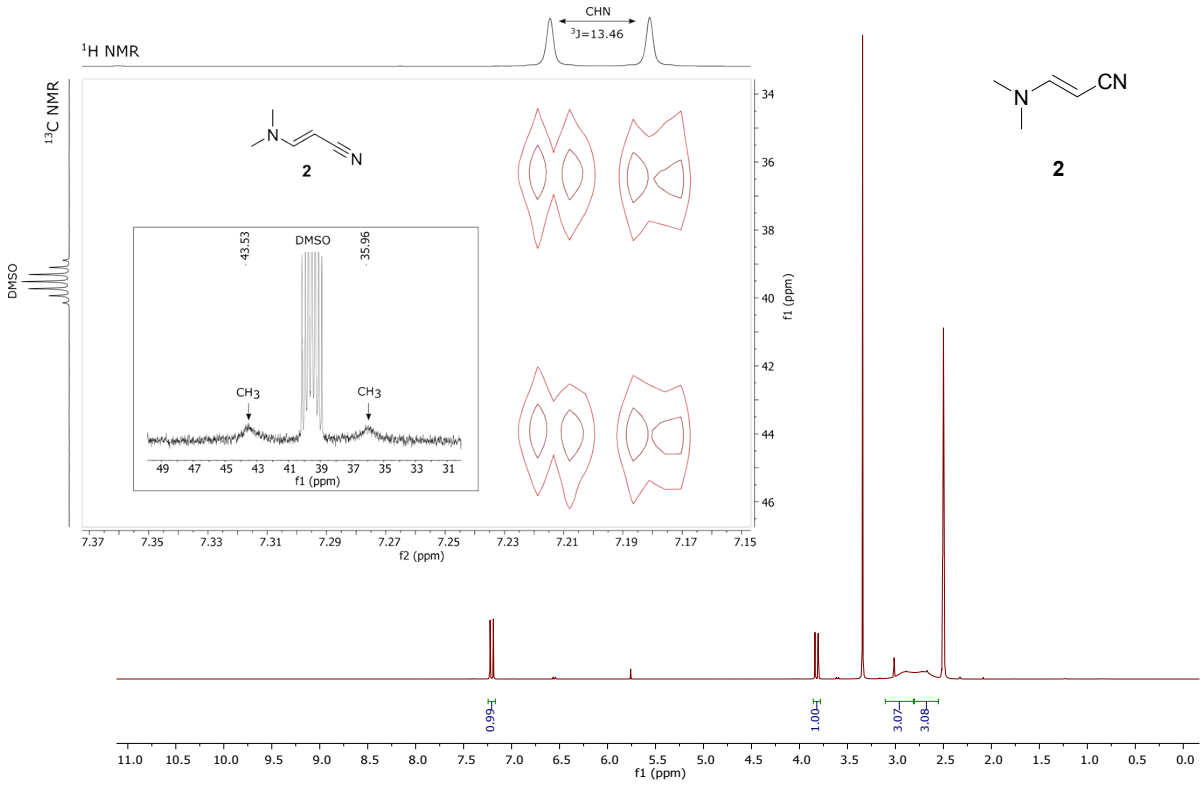
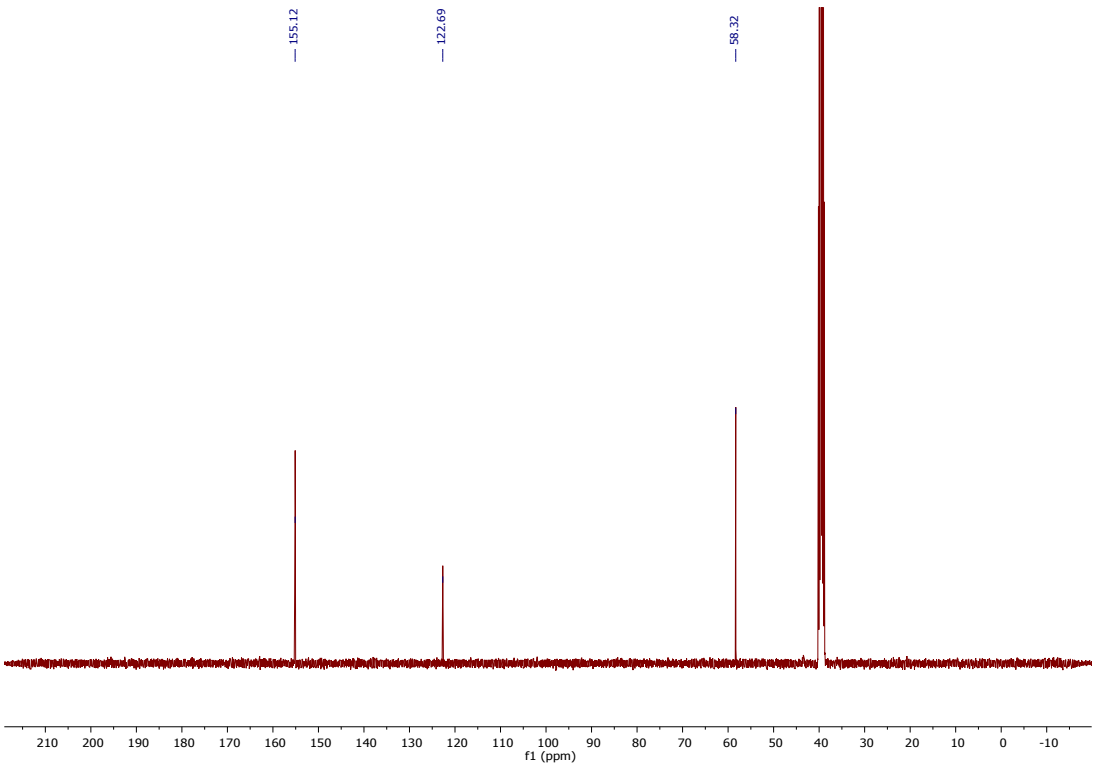


Figure S5 (continued). Calibration curves of the compounds synthesized in this study.

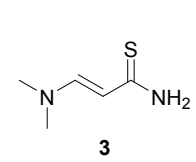
8. ^{1}H and ^{13}C NMR spectra of the compounds

^{1}H NMR, ^{13}C NMR and HMBC of the compound **2**

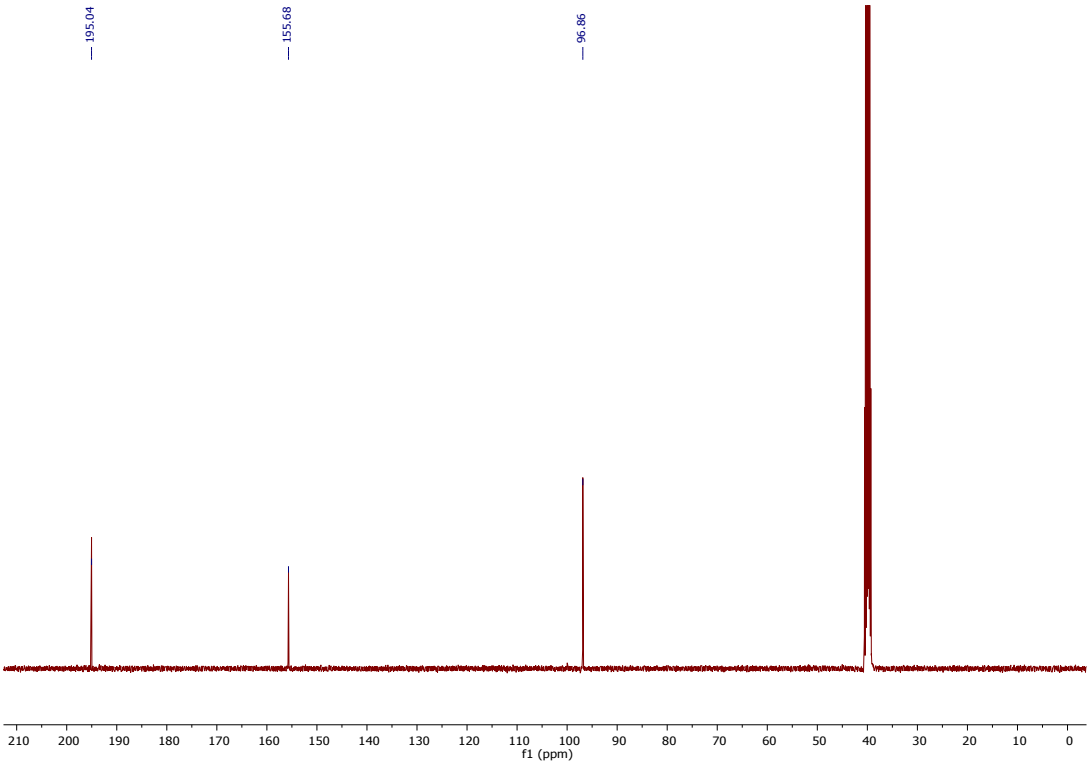




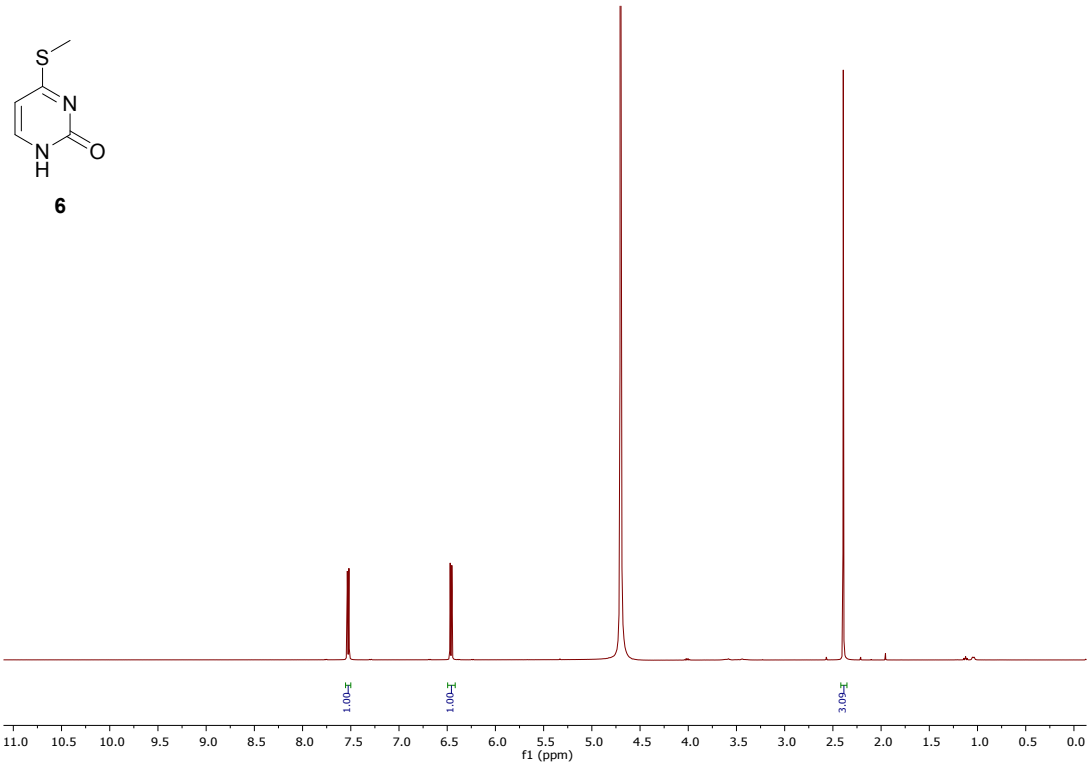
^{1}H NMR and ^{13}C NMR of the compound **3**



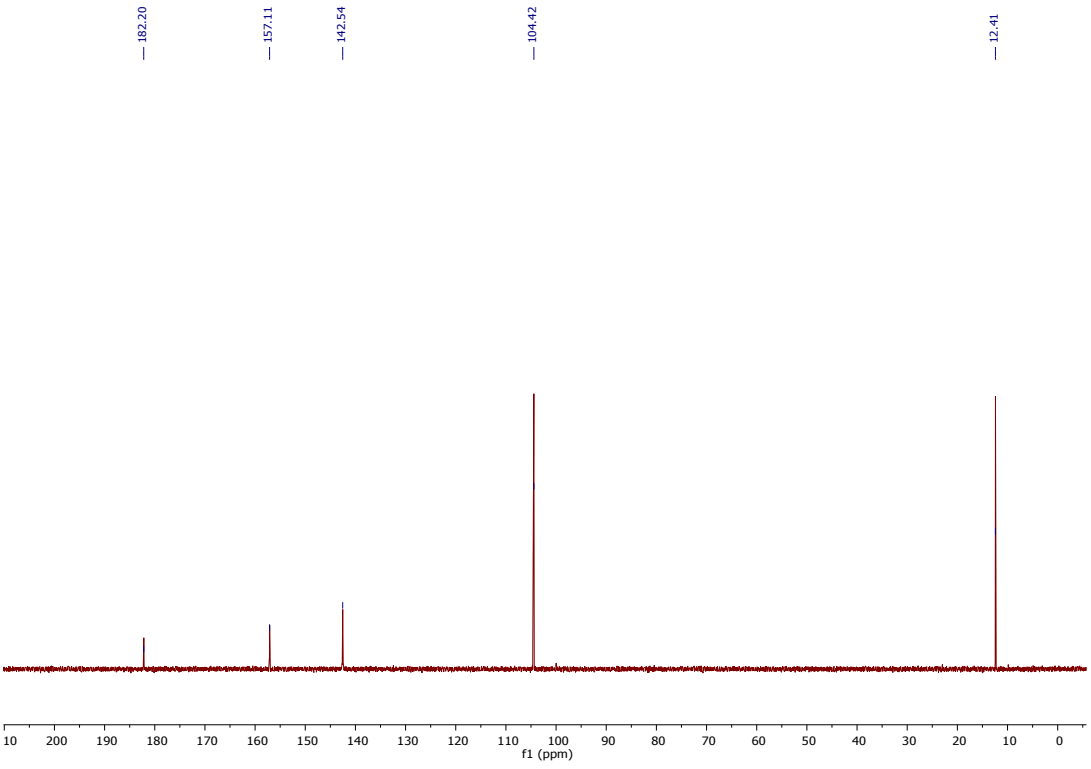
3



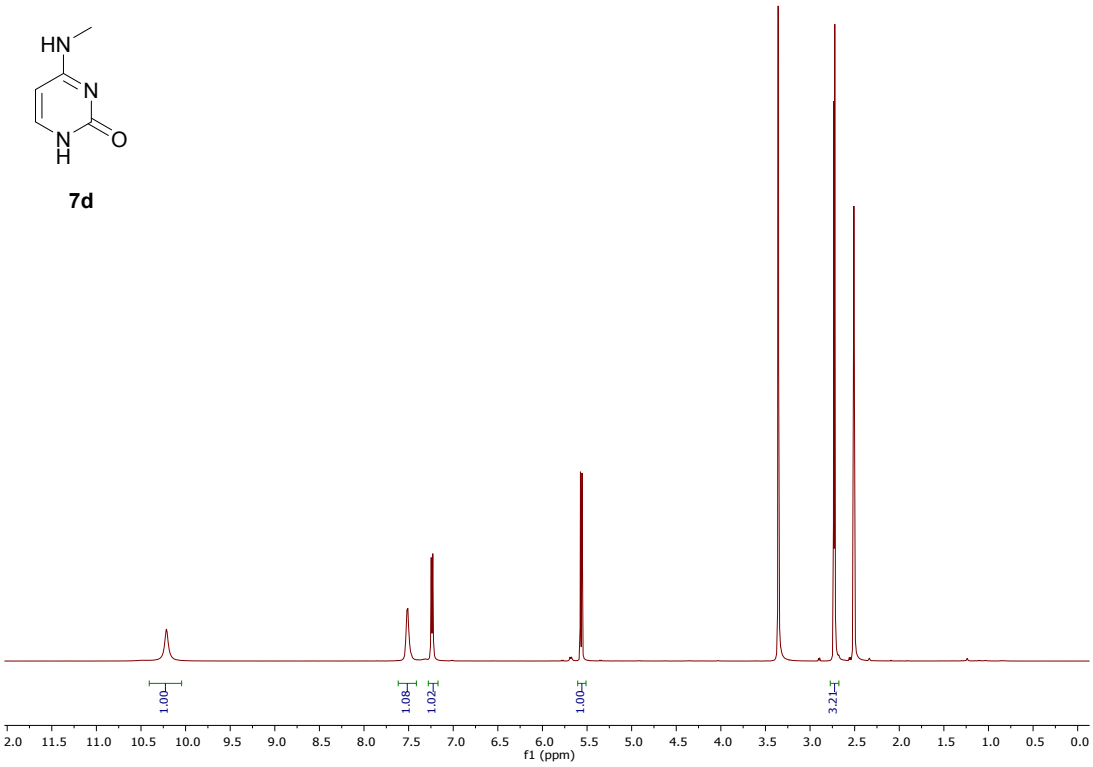
^{1}H NMR and ^{13}C NMR of the compound **6**



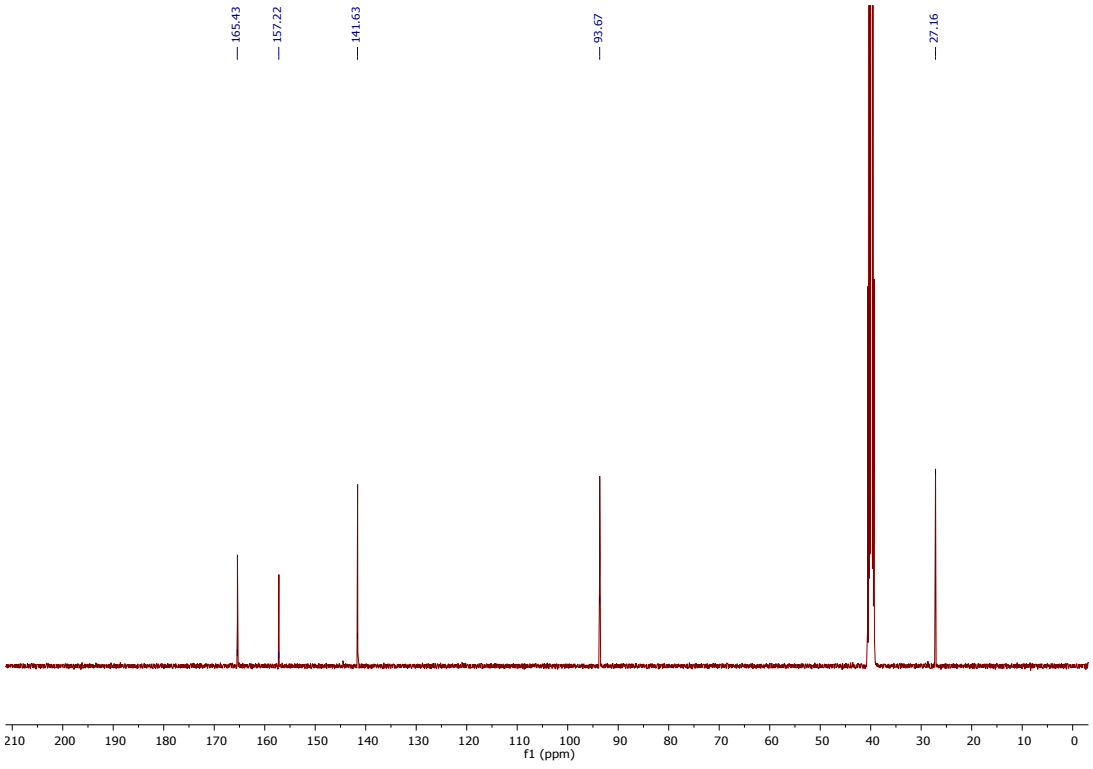
6



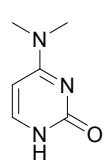
^{1}H NMR and ^{13}C NMR of the compound **7d**



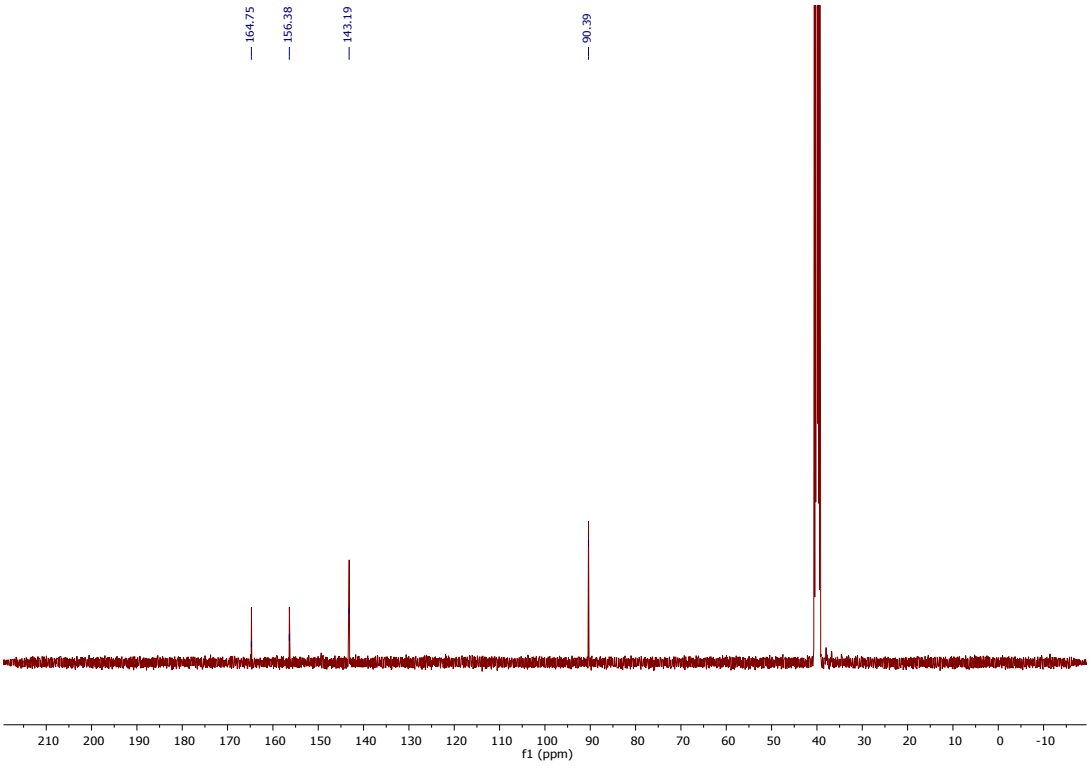
7d



^{1}H NMR and ^{13}C NMR of the compound **7e**

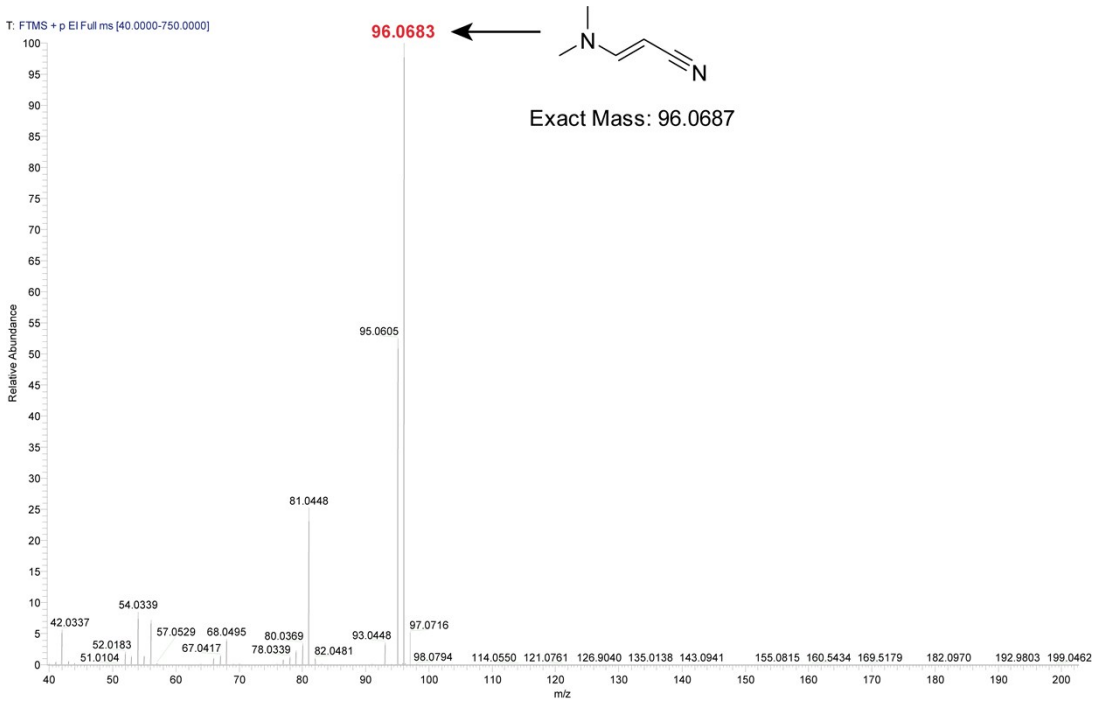


7e

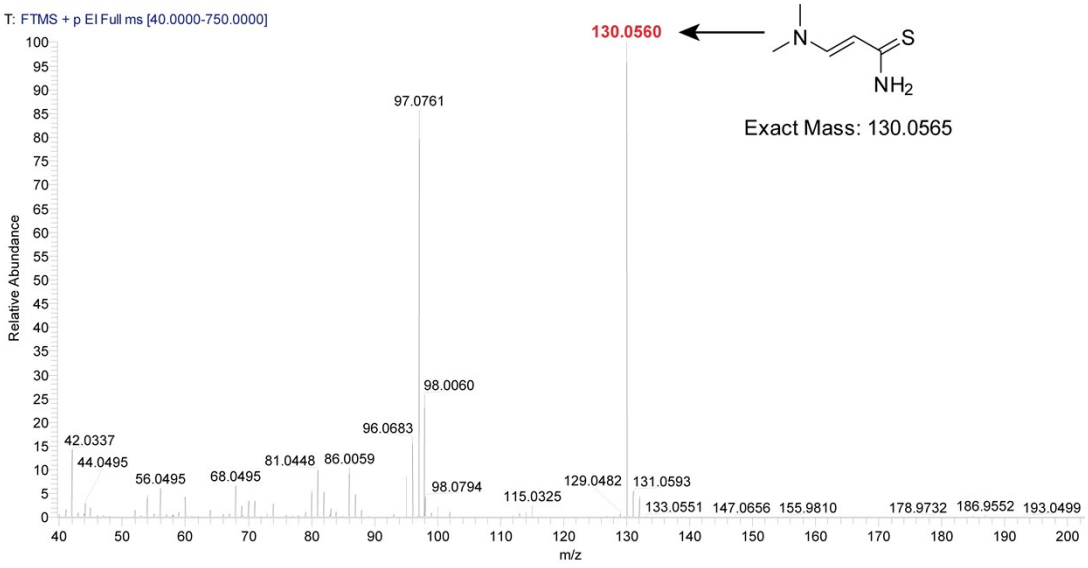


9. HRMS spectra of compound **2** and **3**

HRMS (EI+) spectrum of compound **2**



HRMS (EI+) spectrum of compound **3**



10. References

(1) M. S. M. Pearson, A. Robin, N. Bourgougnon, J. C. Meslin, D. Deniaud, *J. Org. Chem.*, **2003**, *68*, 8583-8587.

(2) B. T. Golding, C. Bleasdale, J. McGinnis, S. Müller, H. T. Rees, N. H. Rees, P. B. Farmer, W. P. Watson, *Tetrahedron*, **1997**, *53*, 4063-4082.

(3) T. J. Delia, M. J. Olsen, G. B. Brown, *J. Org. Chem.*, **1965**, *30*, 2766-2768.

(4) N. P. Dolman, H. M. Troop, J. C. A. More, A. Alt, J. L. Knauss, R. Nistico, S. Jack, R. M. Morley, Z. A. Bortolotto, P. J. Roberts, D. Bleakman, G. L. Collingridge, D. E. Jane, *J. Med. Chem.*, **2005**, *48*, 7867-7881.

(5) A. Shaw, M. D. Shetlar, *J. Am. Chem. Soc.*, **1990**, *112*, 7736-7742.

(6) M. Masaki, K. Kazusaki, K. Shigeo, S. Katsuyoshi, M. Hiroshige, *Bull. Chem. Soc. Jpn*., **1989**, *62*, 2939-2941.