

## Supplementary Information

### A Safety-Catch Protecting Group Strategy compatible with Boc-Chemistry for the Synthesis of Peptide Nucleic Acids (PNAs)

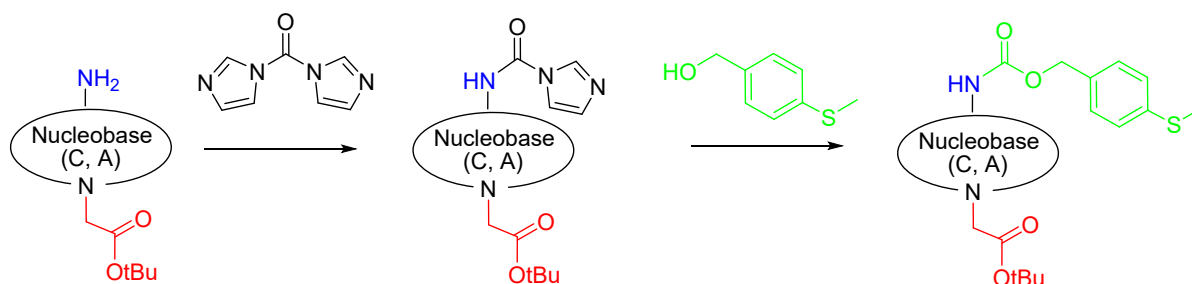
K. P. Nandhini<sup>1,2</sup>, Sikabwe Noki<sup>1,2</sup>, Edikarlos Brasil<sup>2</sup>, Fernando Albericio<sup>2,4,\*</sup> and Beatriz G. de la Torre<sup>1,\*</sup>

<sup>1</sup> KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), School of Laboratory Medicine and Medical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban 4041, South Africa.

<sup>2</sup> Peptide Science Laboratory, School of Chemistry and Physics, University of KwaZulu-Natal, Westville, Durban 4000, South Africa.

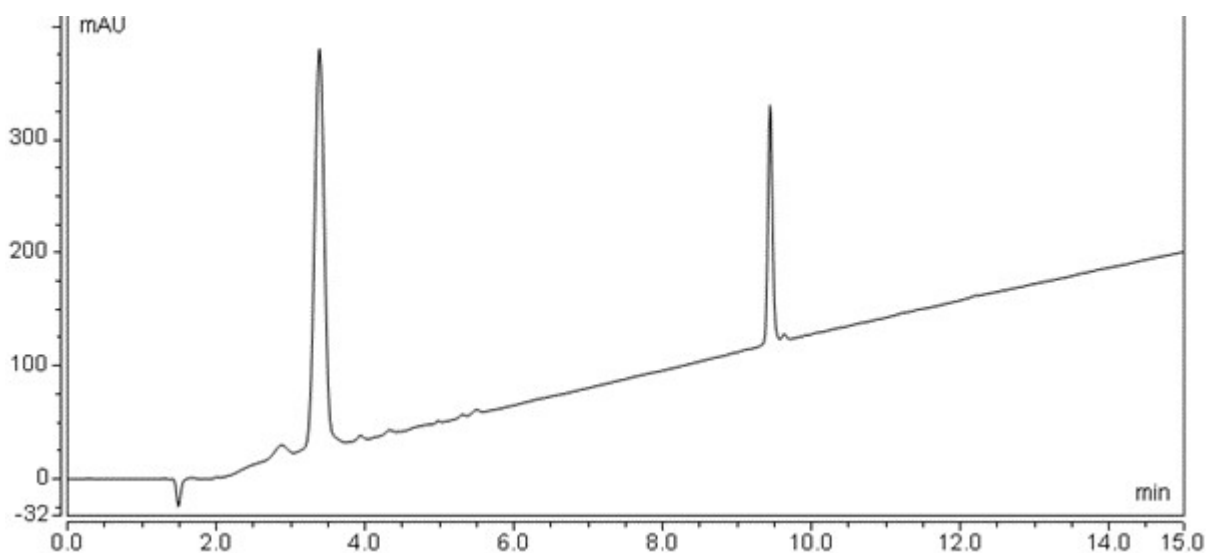
<sup>3</sup> CIBER-BBN, Networking Centre on Bioengineering, Biomaterials and Nanomedicine, and Department of Organic Chemistry, University of Barcelona, Martí i Franqués 1-11, 08028 Barcelona, Spain.

\*Correspondence: albericio@ukzn.ac.za; Tel.: (+27 614009144) (F.A); garciadelatorreb@ukzn.ac.za; Tel.: +27614 475528



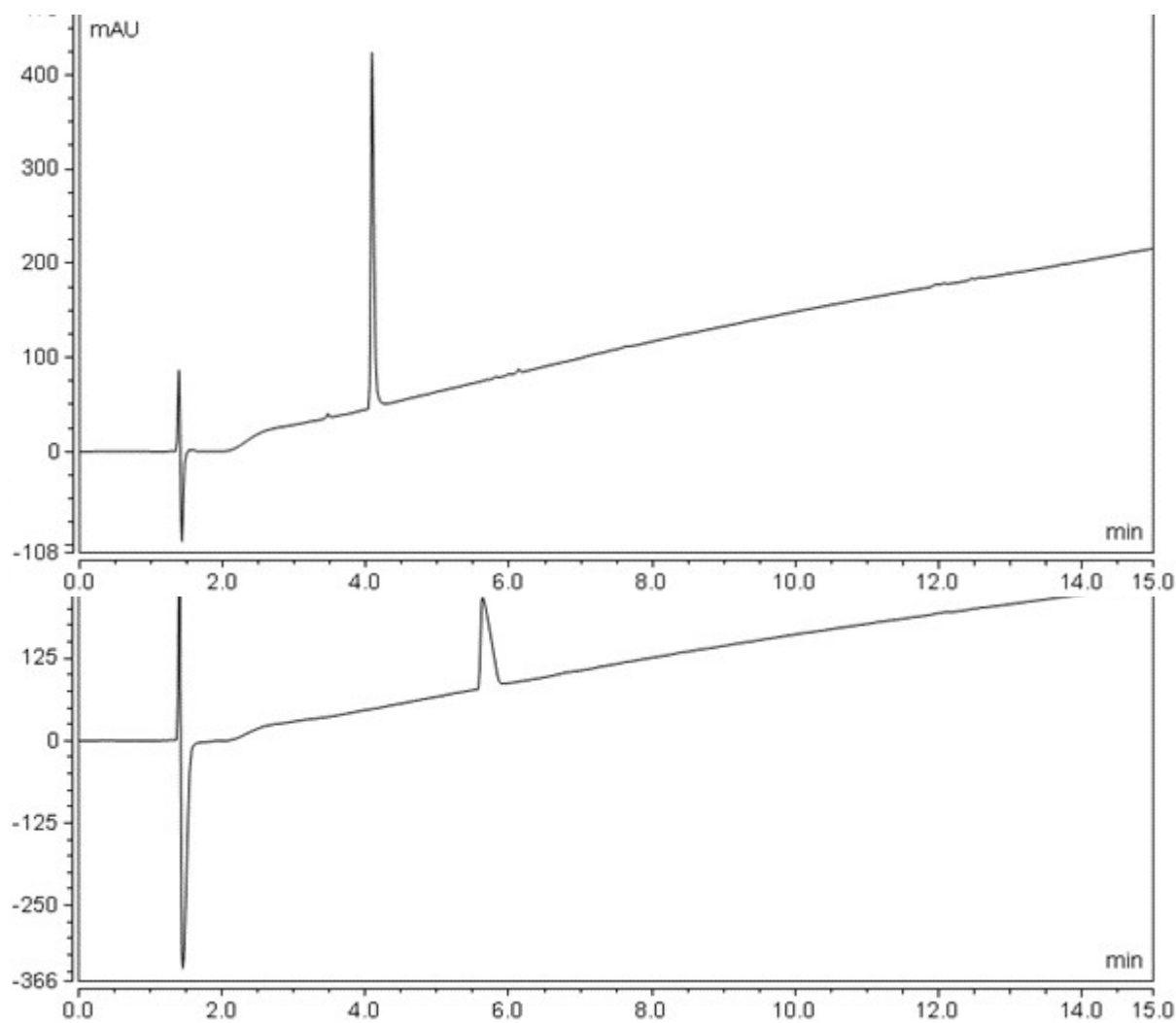
**Scheme S1.** Protection of the exocyclic nitrogen of the nucleobase *via* carbonyldiimidazole (CDI) reaction

#### 1. Chromatograms from LC-MS

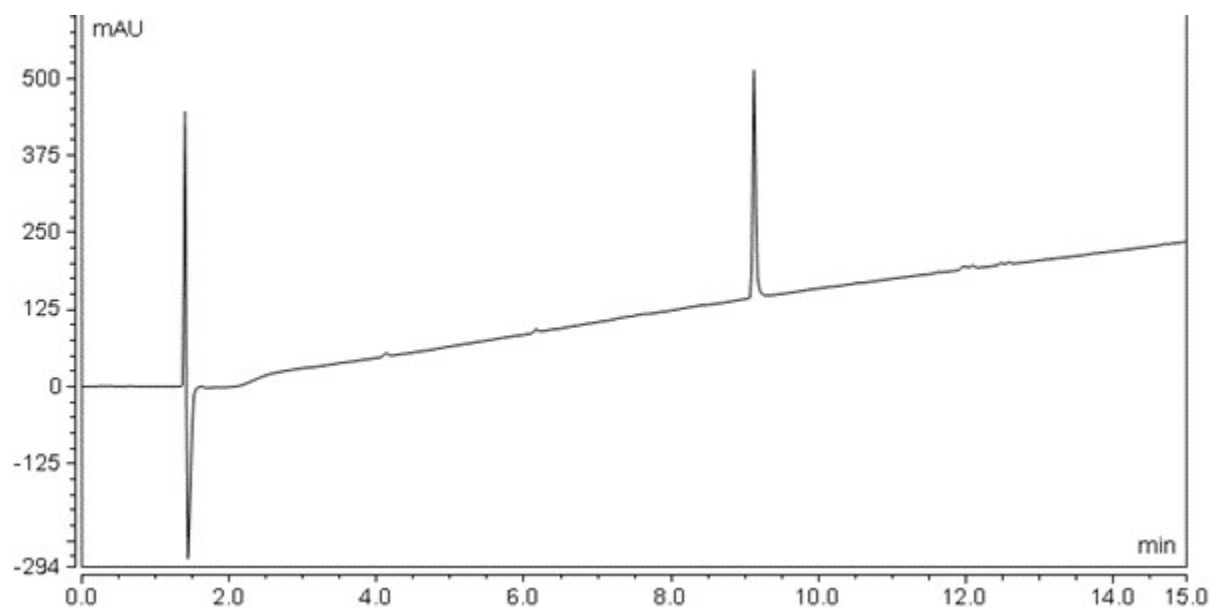


**Figure S1:** Chromatogram of H-PNA(T<sub>4</sub>)-βAla-OH synthesized using Fmoc-PNA(T)-OH on Wang resin; treated twice with 20% piperidine (10 min + 10 min).

**Figure S2:** Chromatogram of Boc-Aeg-OBzl (4).



**Figure S3:** Chromatogram of *tert*-butyl adenine-9-ylacetate (5).



**Figure S4:** Chromatogram of *tert*-butyl (*N*<sup>6</sup>-Mtz-adenine-9-yl)acetate (6).

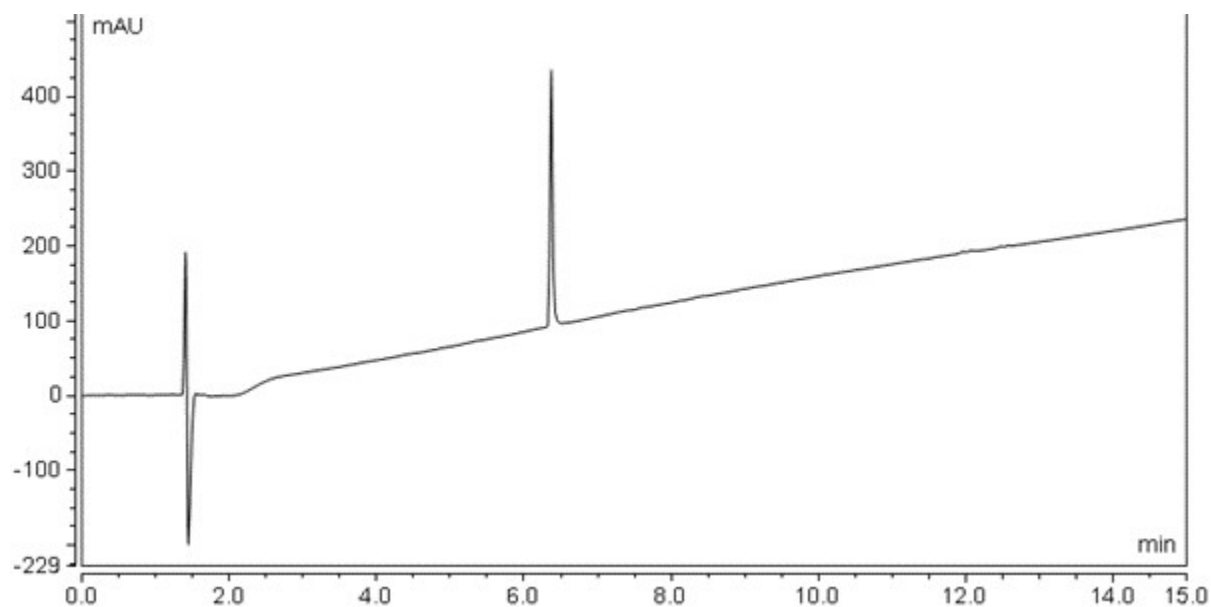


Figure S5: Chromatogram of *tert*-butyl (*N*<sup>6</sup>-Msz-adenine-9-yl)acetate (7).

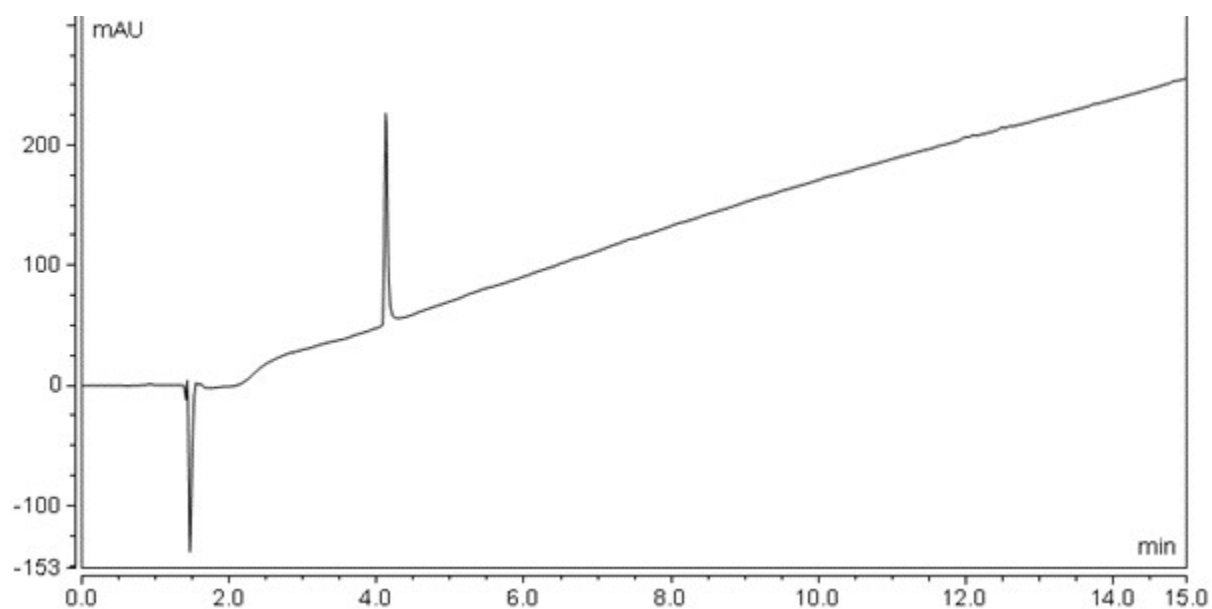


Figure S6: Chromatogram of (*N*<sup>6</sup>-Msz-adenine-9-yl)acetic acid (8).

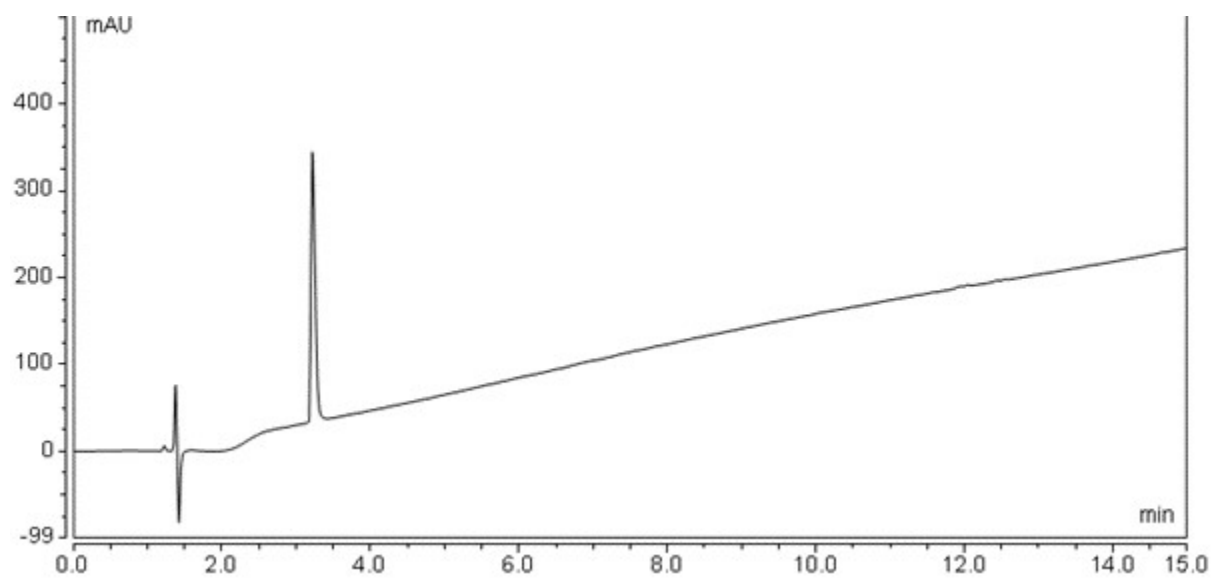
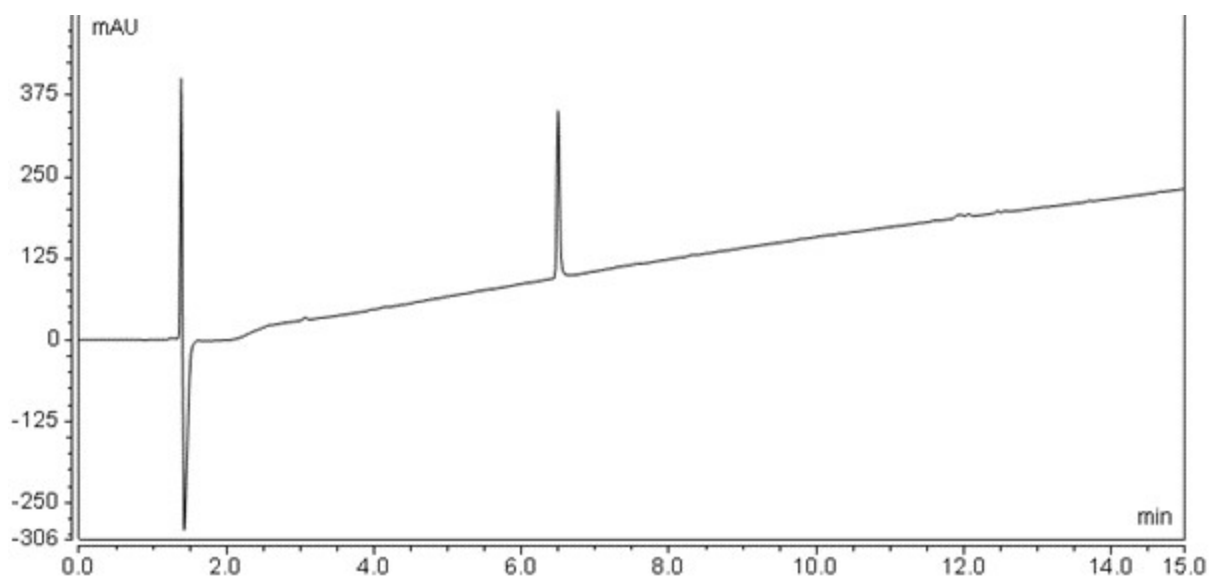
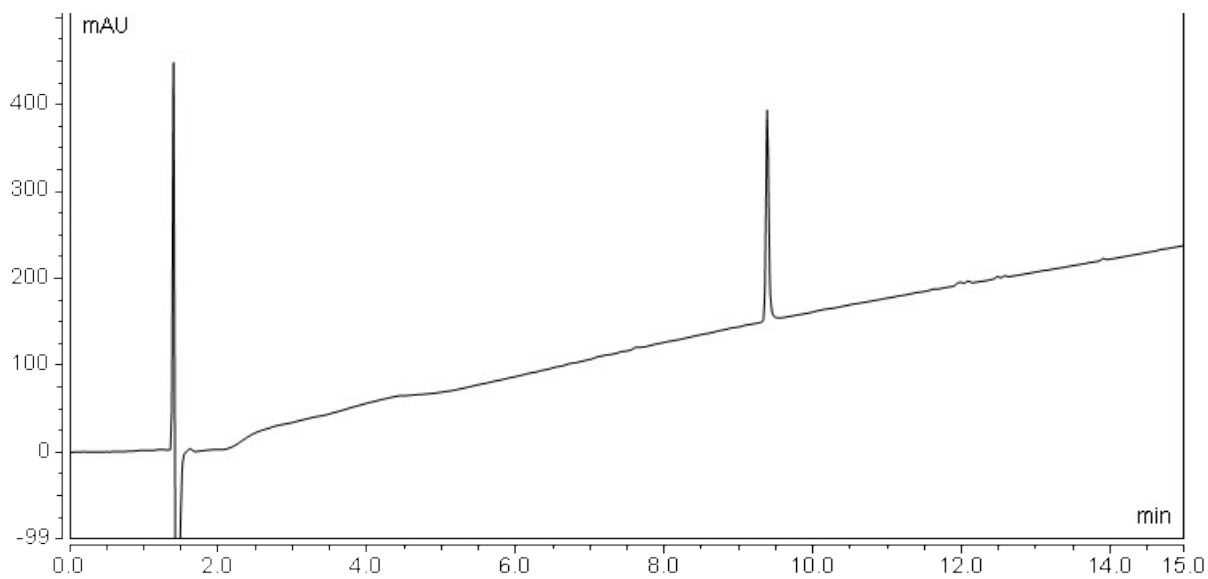


Figure S7: Chromatogram of *tert*-butyl cytosine-1-ylacetate (9).



**Figure S8:** Chromatogram of *tert*-butyl (*N*<sup>4</sup>-Mtz-cytosine-1-yl)acetate (10).

**Figure S9:** Chromatogram of *tert*-butyl (*N*<sup>4</sup>-Msz-cytosine-1-yl)acetate (11).

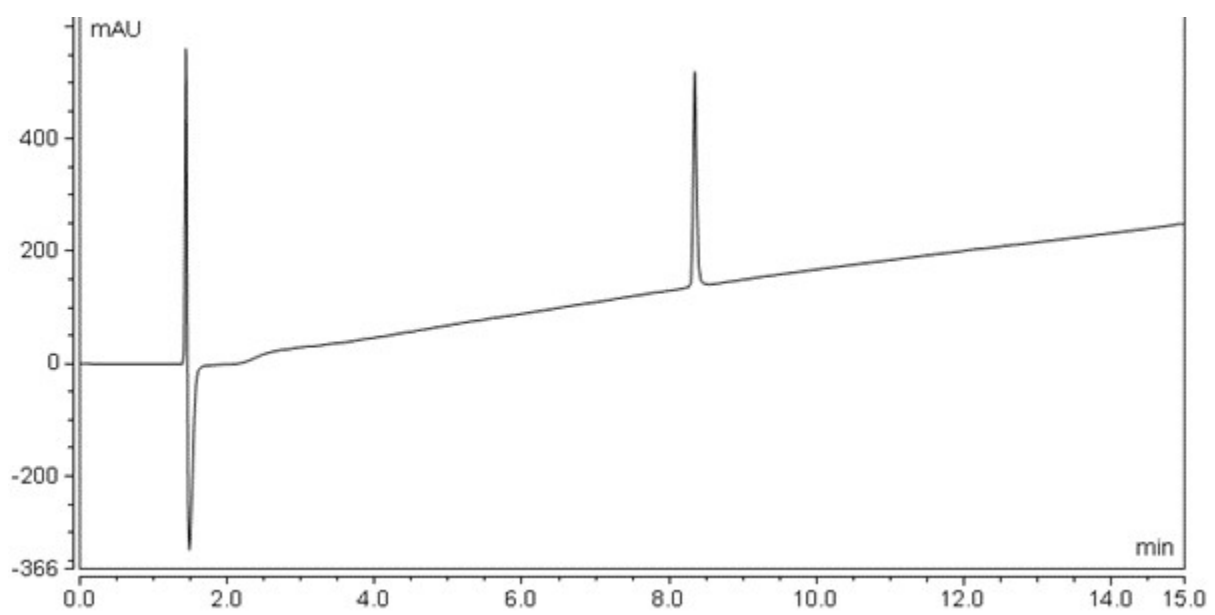
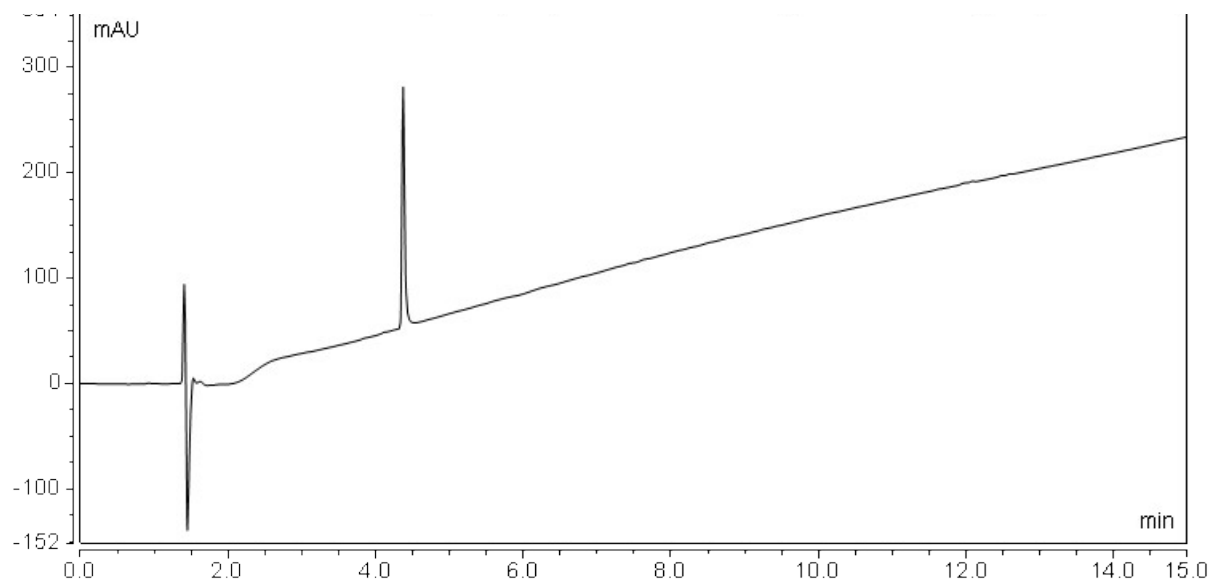


Figure S10: Chromatogram of ( $N^4$ -Msz-cytosine-1-yl)acetic acid (12).

Figure S11: Chromatogram of Boc-PNA-A(Msz)-OBzl (13).

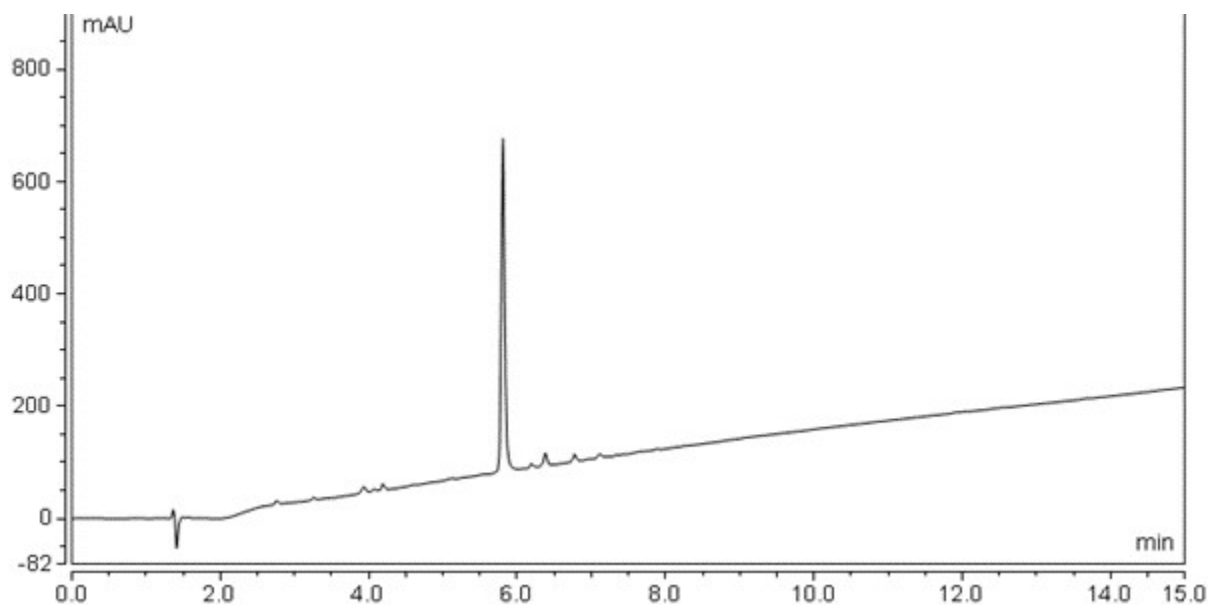


Figure S12: Chromatogram of Boc-PNA-A(Msz)-OH (14).

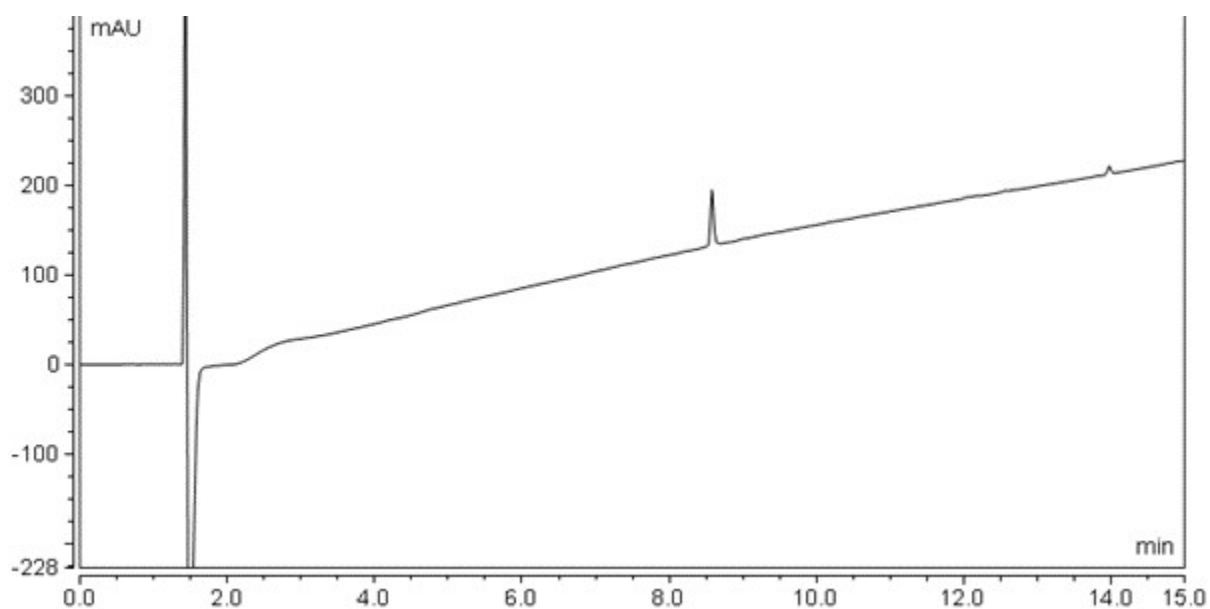


Figure S13: Chromatogram of Boc-PNA-C(Msz)-OBzl (15).

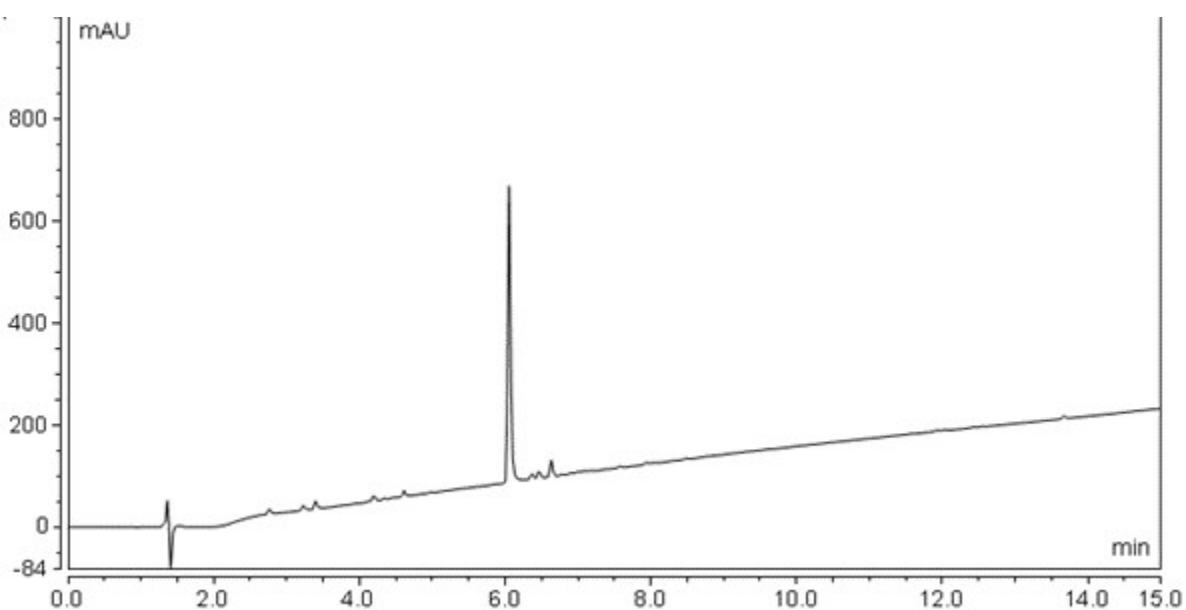
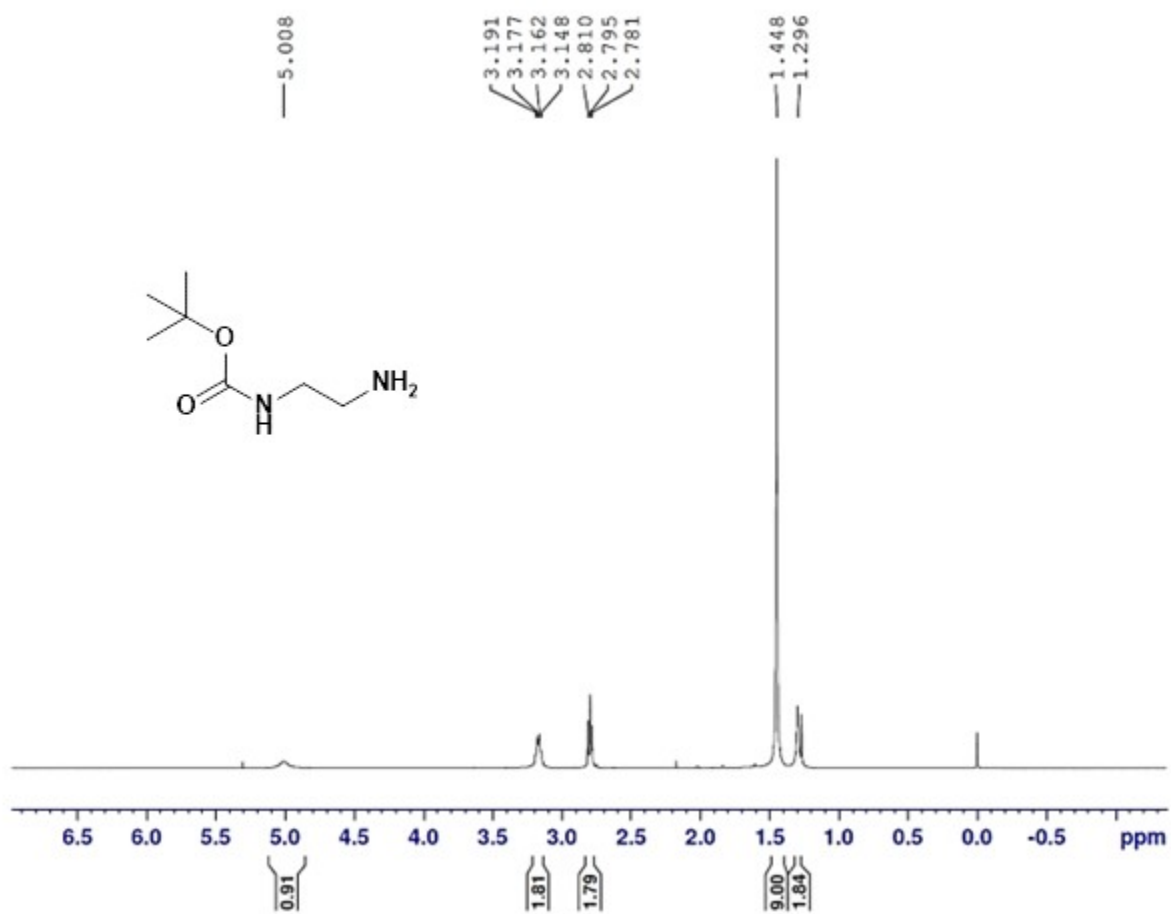


Figure S14: Chromatogram of Boc-PNA-C(Msz)-OH (16).

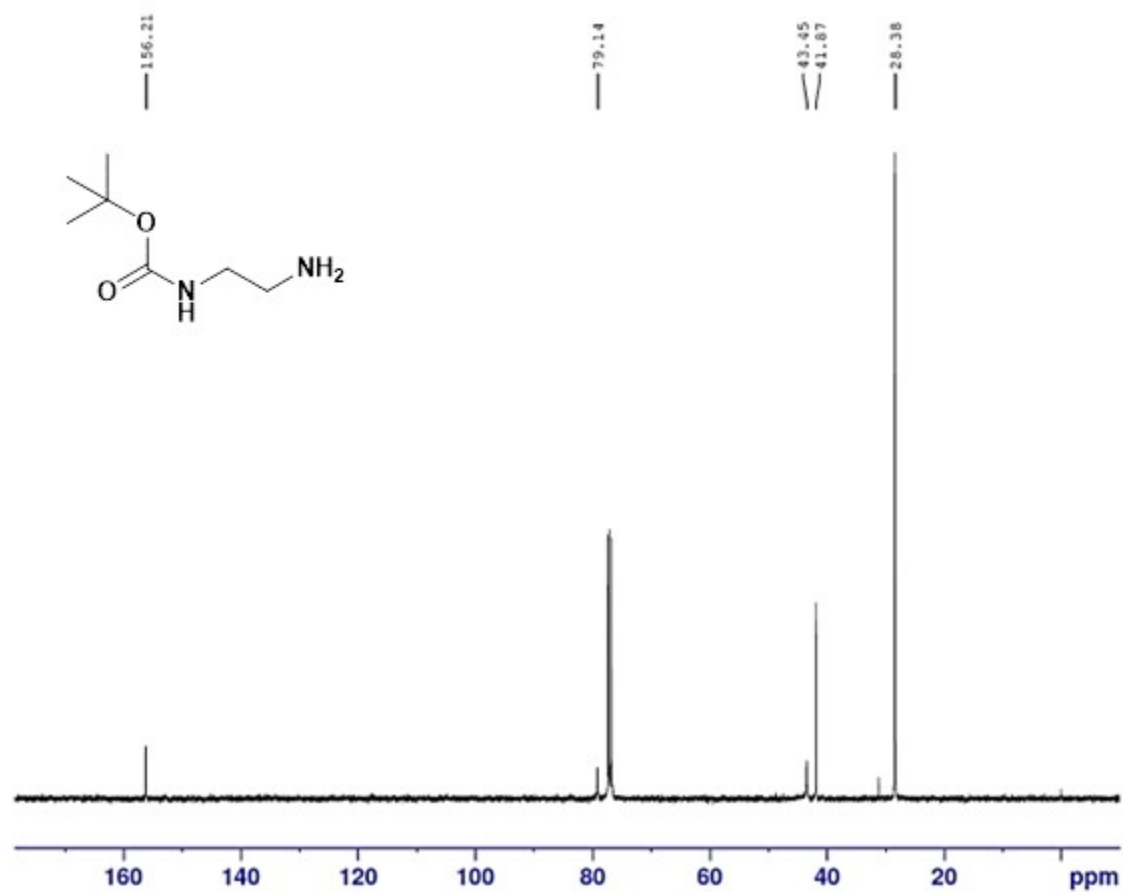
## 2. HRMS and NMR characterization

### Boc-ethylenediamine (3)

$^1\text{H}$  NMR

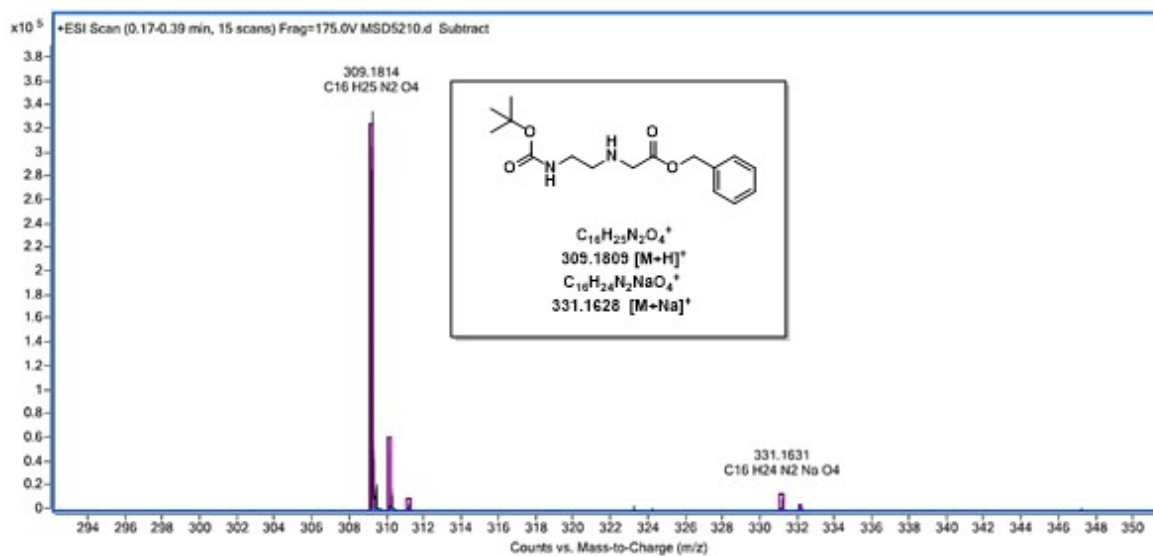


$^{13}\text{C}$  NMR



$^1\text{H}$  NMR ([d6] DMSO, 400 MHz, 25 °C):  $\delta$  = 5.01 ppm (s, 1 H, NH), 2.10 (t,  $J$  = 5.92 Hz, 2 H), 3.17 (q,  $J$  = 5.65 Hz, 2 H), 1.45 (s, 9 H), 1.29 (s, 2 H,  $\text{NH}_2$ );  $^{13}\text{C}$  NMR ([d6] DMSO, 100 MHz, 25 °C):  $\delta$  = 156.2, 79.1, 43.5, 41.9, 28.4.

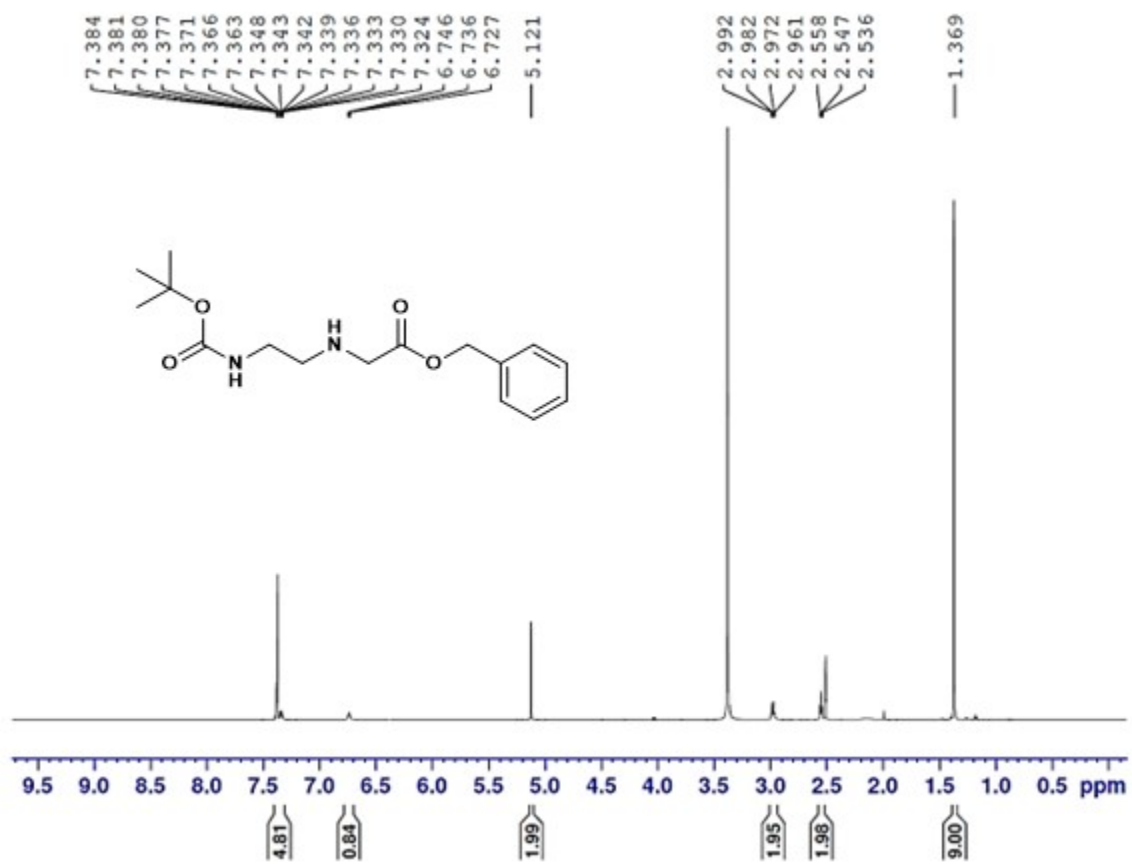
#### Boc-Aeg-OBzl (4)



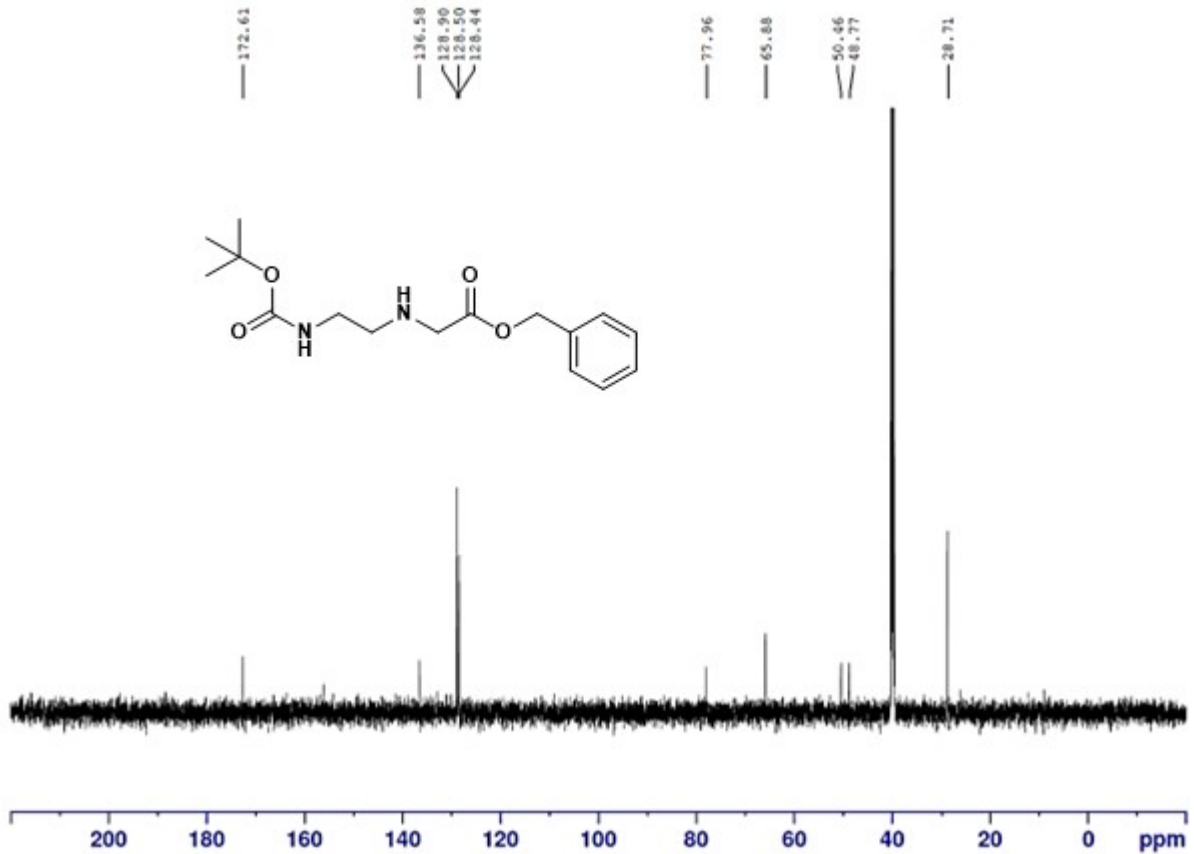
HRMS



# <sup>1</sup>H NMR

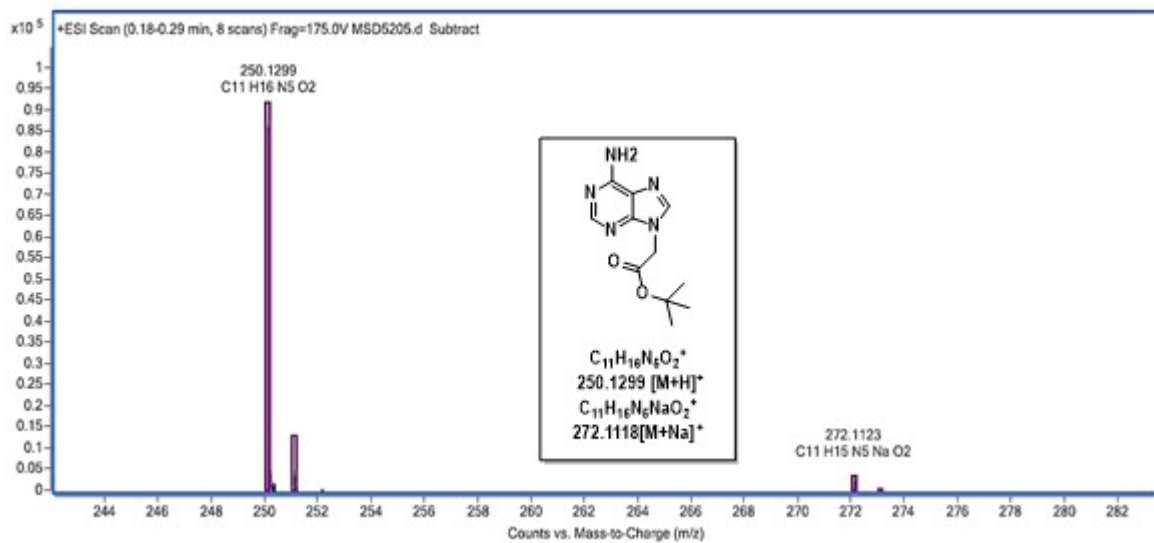


# <sup>13</sup>C NMR



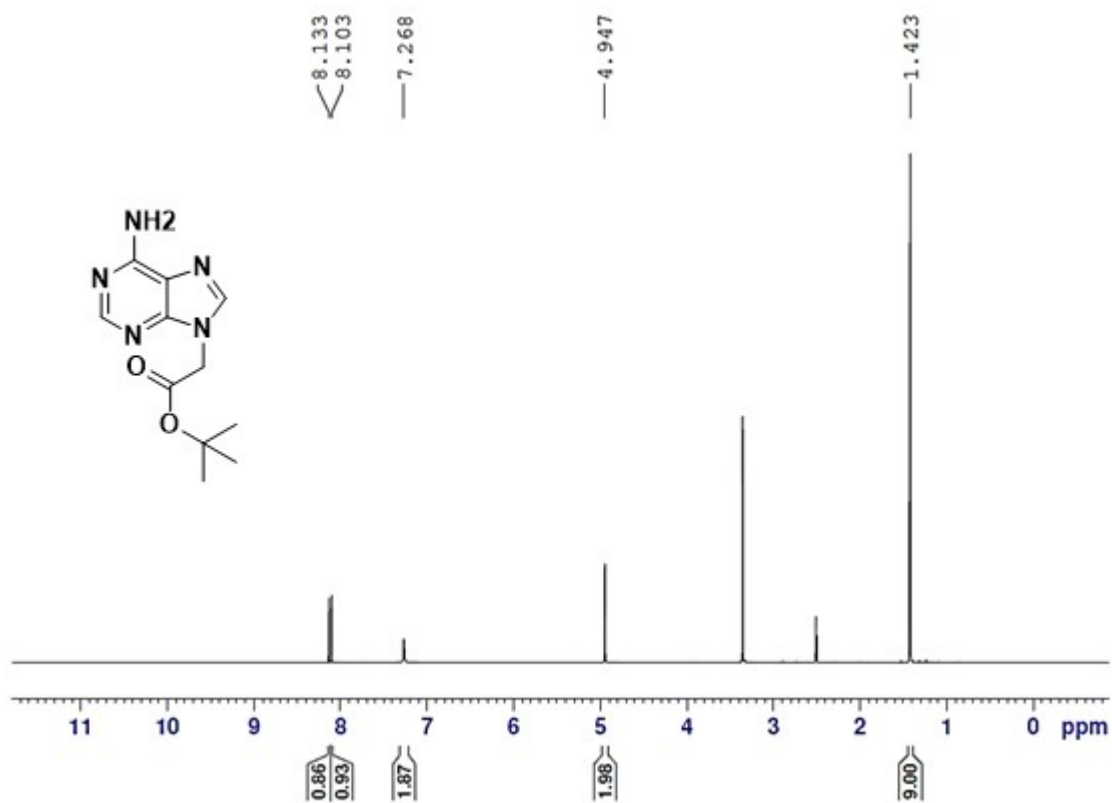
<sup>1</sup>H NMR ([d<sub>6</sub>] DMSO, 400 MHz, 25 °C): δ = 7.38 ppm (m, 5 H, ArH), 6.73 (t, *J* = 5.49 Hz, 1 H, NH), 5.12 (s, 2 H), 2.97 (q, *J* = 6.22 Hz, 2 H), 2.55 (t, *J* = 6.48 Hz, 2 H), 1.37 (s, 9 H); <sup>13</sup>C NMR ([d<sub>6</sub>] DMSO, 100 MHz, 25 °C): δ = 172.6, 136.5, 128.9, 128.5, 128.4, 77.9, 65.8, 50.5, 48.7, 28.7.

### tert-butyl adenine-9-ylacetate (5)

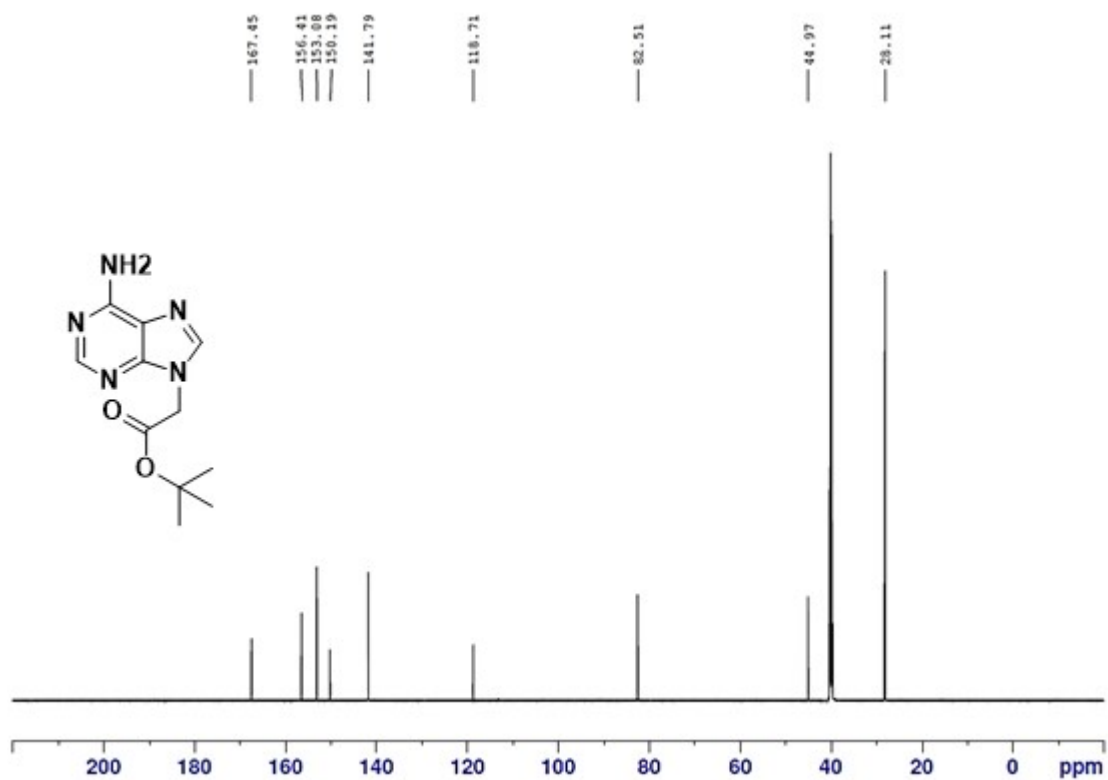


HRMS

<sup>1</sup>H NMR

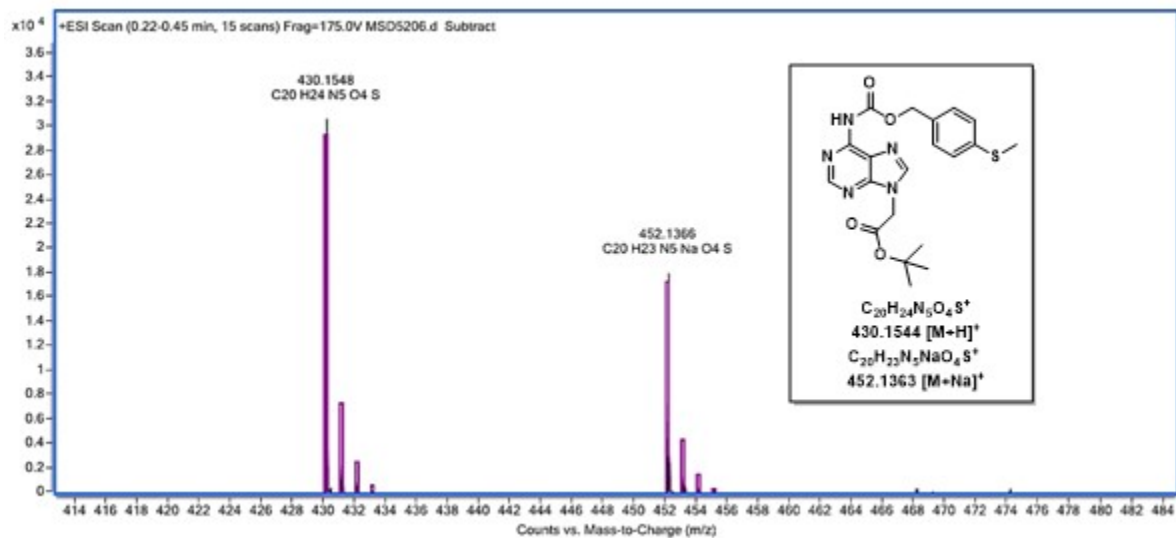


**<sup>13</sup>C NMR**



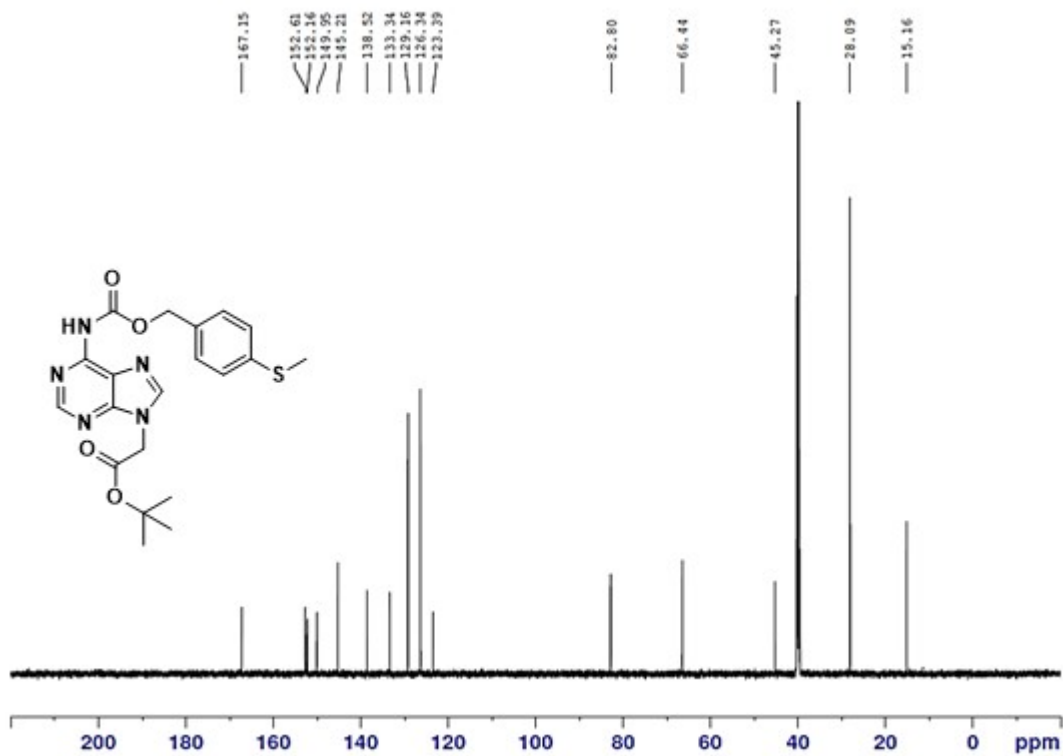
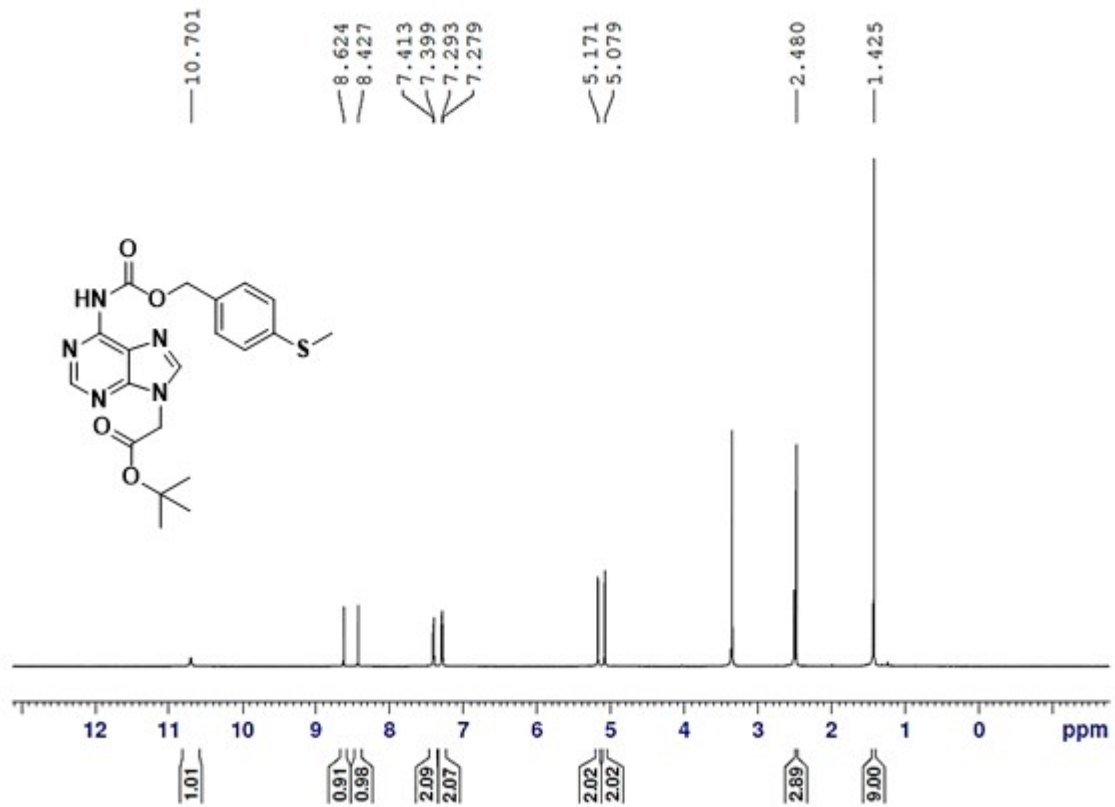
$^1\text{H}$  NMR ([d6] DMSO, 600 MHz, 25 °C):  $\delta$  = 8.13 (s, 1 H), 8.10 (s, 1 H), 7.26 (s, 2 H), 4.94 (s, 2 H), 1.42 (s, 9 H);  $^{13}\text{C}$  NMR ([d6] DMSO, 100 MHz, 25 °C):  $\delta$  = 167.4, 156.4, 153.1, 150.2, 141.8, 118.7, 82.5, 44.9, 28.1.

### *tert*-butyl (*N*<sup>6</sup>-Mtz-adenine-9-yl)acetate (6)



HRMS

$^1\text{H}$  NMR

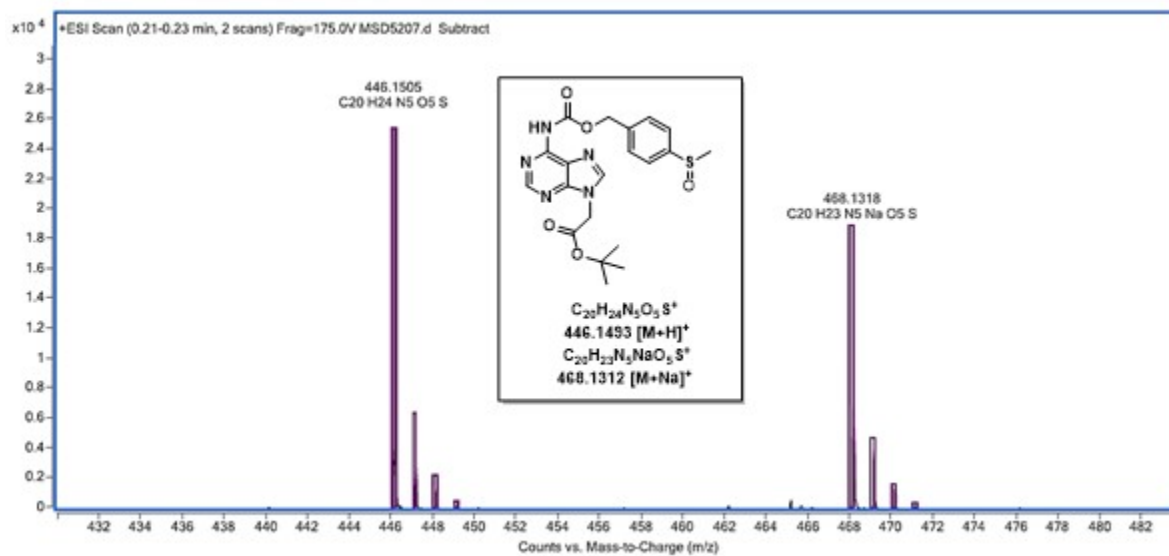


### <sup>13</sup>C NMR

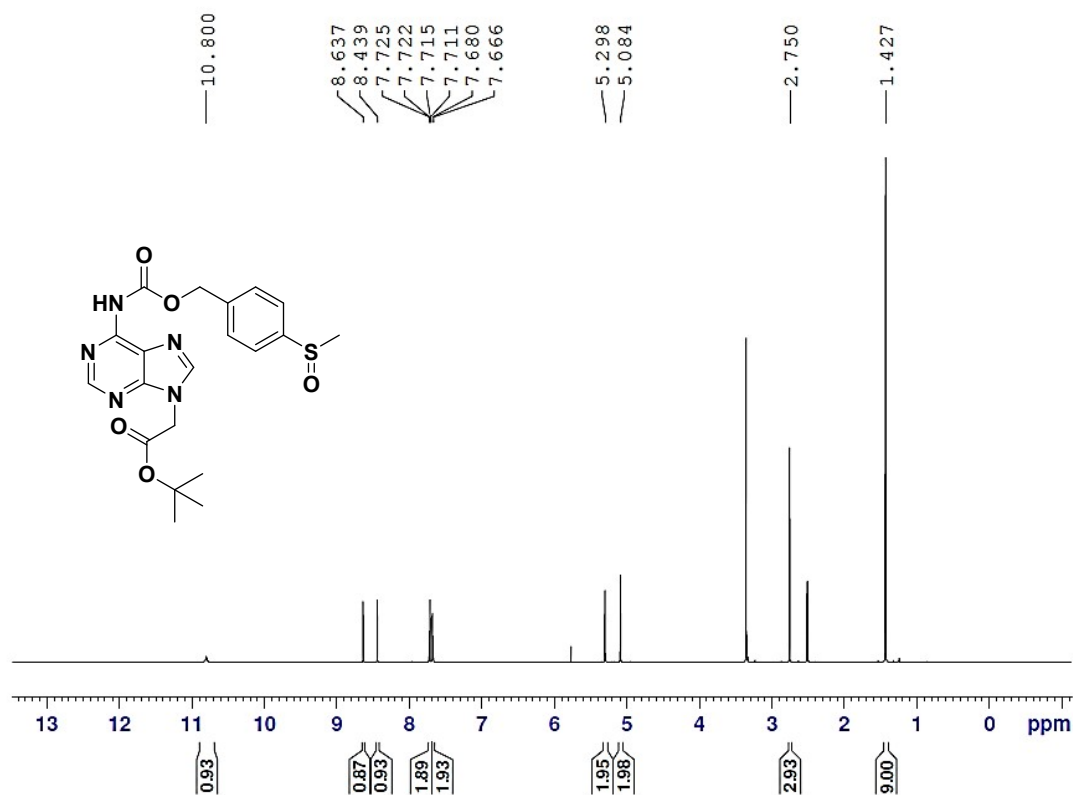
<sup>1</sup>H NMR ([d<sub>6</sub>] DMSO, 600 MHz, 25 °C): δ = 10.7(s, NH, 1 H), 8.62 (s, 1 H), 8.42 (s, 1 H), 7.41 (d, *J* = 8.3 Hz, 2 H), 7.29 (d, *J* = 8.3 Hz, 2 H), 5.17 (s, 2 H), 5.08 (s, 2 H), 2.48(s, 3 H) 1.42 (s, 9 H); <sup>13</sup>C NMR ([d<sub>6</sub>] DMSO, 100 MHz, 25 °C): δ = 167.2, 152.6, 152.2, 149.9, 145.2, 138.5, 133.3, 129.2, 126.3, 123.4, 82.8, 66.4, 45.3, 28.1, 15.2.

# *tert*-butyl (*N*<sup>6</sup>-Msz-adenine-9-yl)acetate (7)

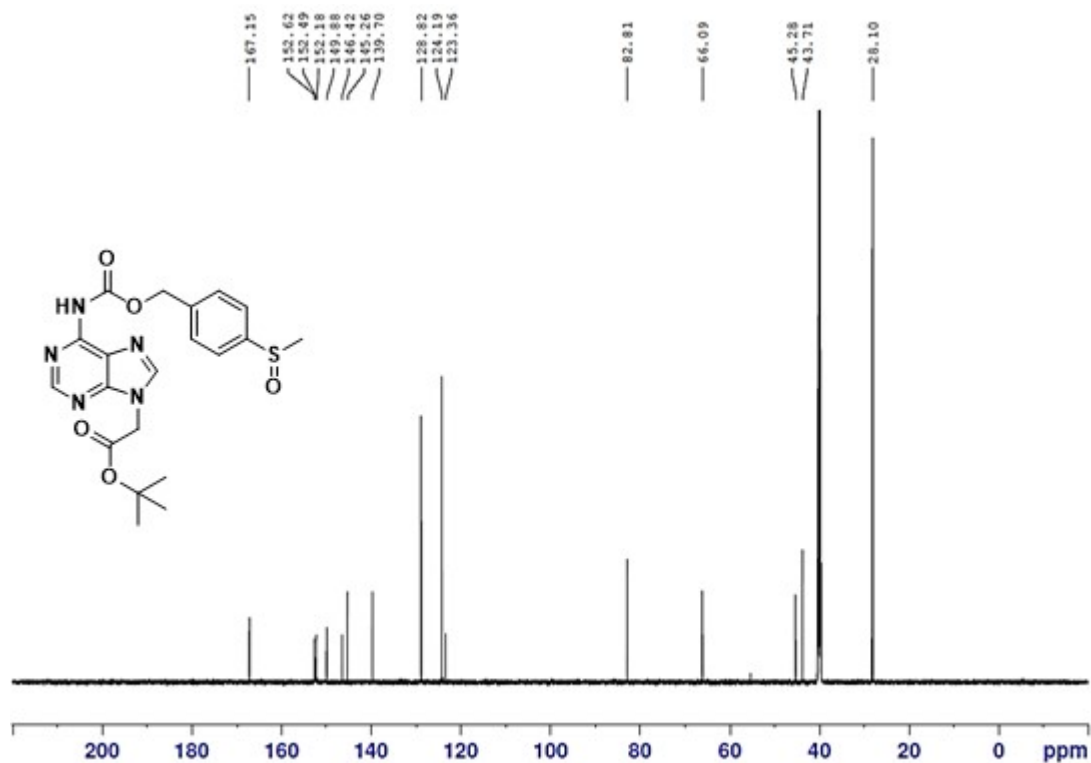
HRMS



<sup>1</sup>H NMR



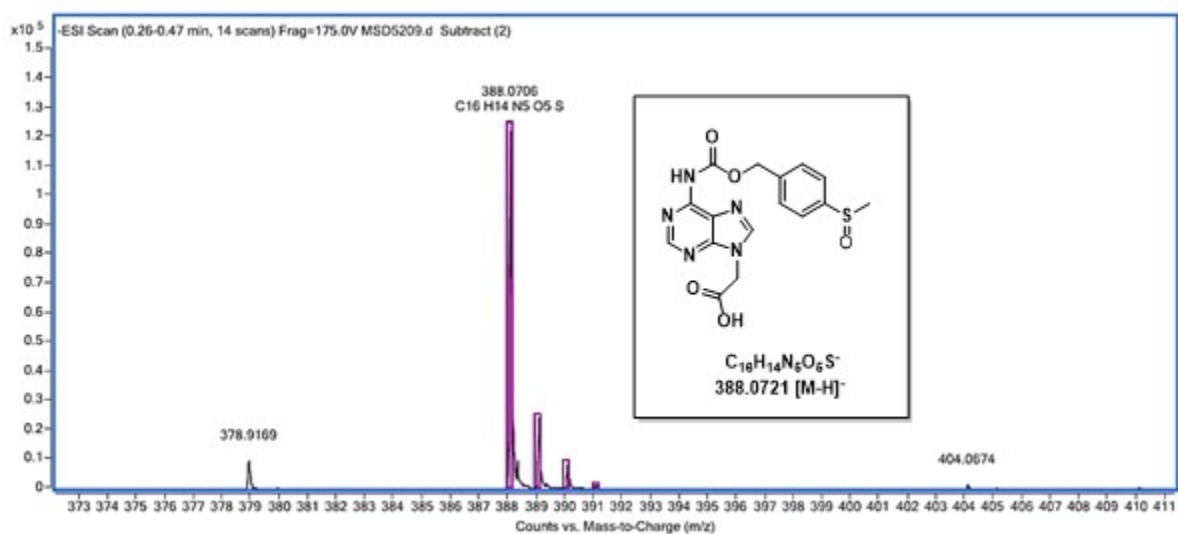
## <sup>13</sup>C NMR



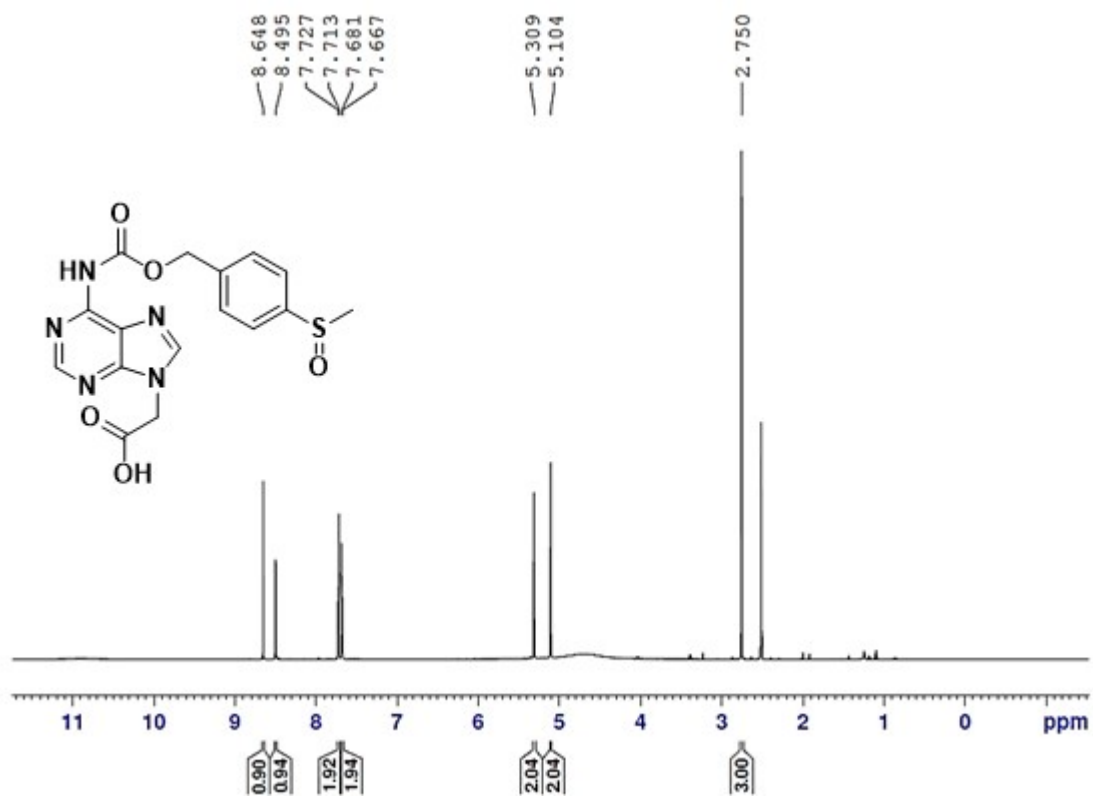
<sup>1</sup>H NMR ([d<sub>6</sub>] DMSO, 600 MHz, 25 °C): δ = 10.8 (s, NH, 1 H), 8.64 (s, 1 H), 8.44 (s, 1 H), 7.72 (d, *J* = 8.4 Hz, 2 H), 7.67 (d, *J* = 8.4 Hz, 2 H), 5.29 (s, 2 H), 5.08 (s, 2 H), 2.75 (s, 3 H), 1.42 (s, 9 H); <sup>13</sup>C NMR ([d<sub>6</sub>] DMSO, 100 MHz, 25 °C): δ = 167.2, 152.6, 152.5, 152.2, 149.9, 146.4, 145.3, 139.7, 128.8, 124.2, 123.4, 82.8, 66.1, 45.3, 43.7, 28.1.

## (N<sup>6</sup>-Msz-adenine-9-yl)acetic acid (8)

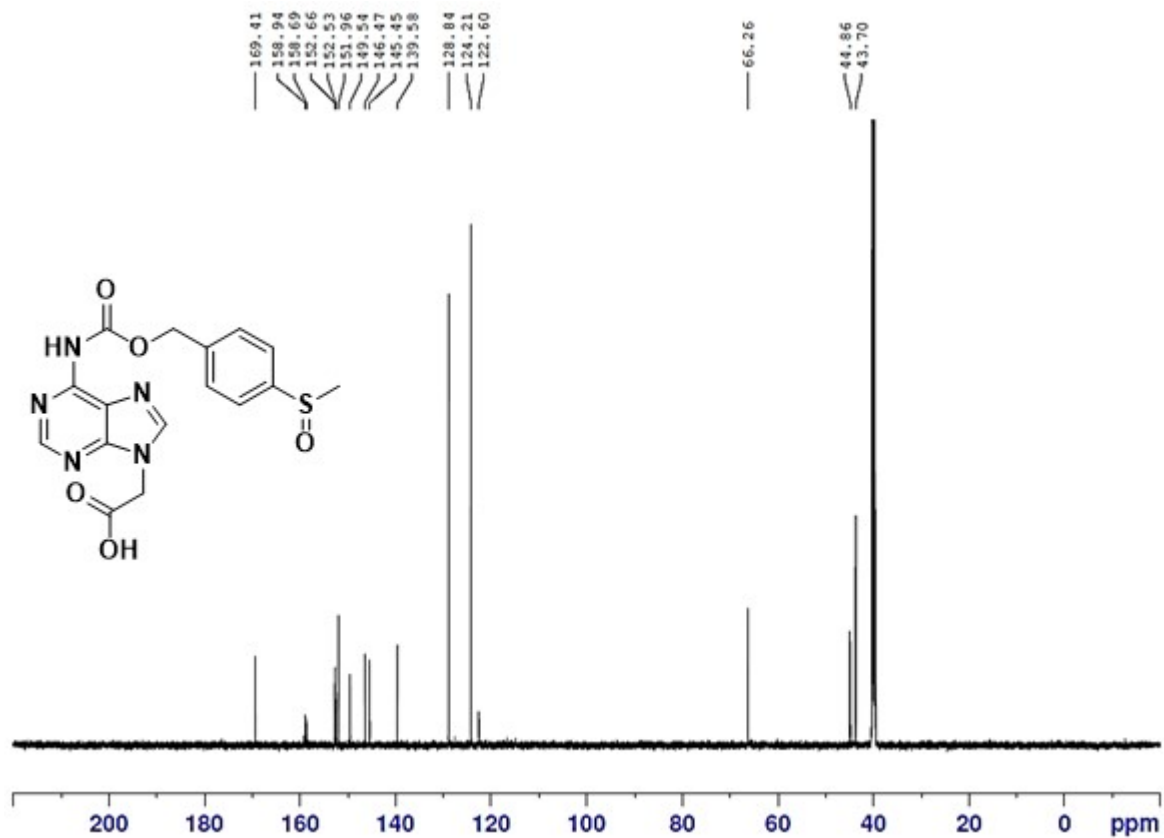
### HRMS



# <sup>1</sup>H NMR



# <sup>13</sup>C NMR

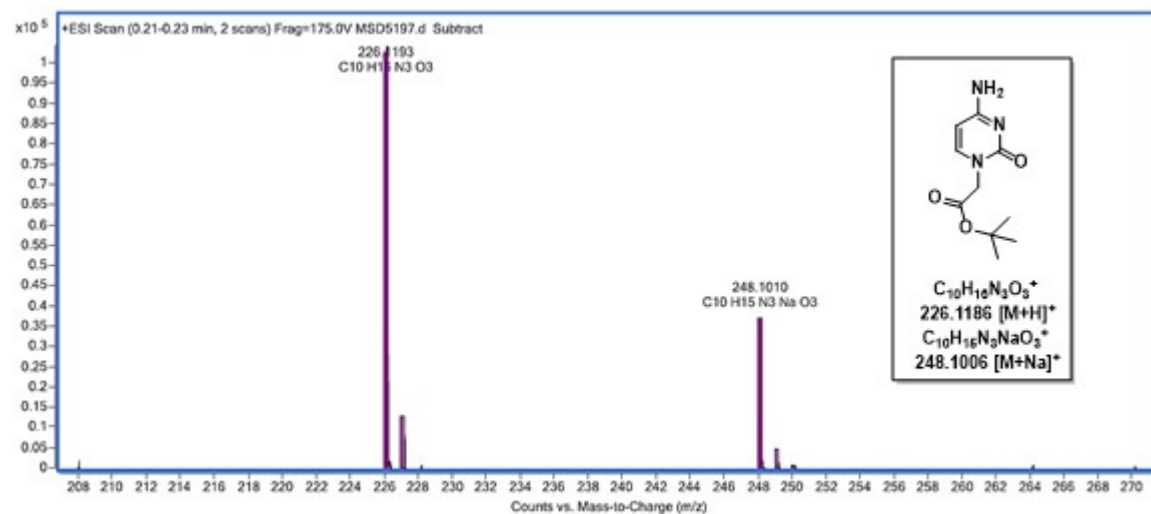


<sup>1</sup>H NMR ([d<sub>6</sub>] DMSO, 600 MHz, 25 °C):  $\delta$  = 8.65 (s, 1 H), 8.49 (s, 1 H), 7.72 (d,  $J$  = 8.4 Hz, 2 H), 7.67 (d,  $J$  = 8.4 Hz, 2 H), 5.31 (s, 2 H), 5.10 (s, 2 H), 2.75 (s, 3 H); <sup>13</sup>C NMR ([d<sub>6</sub>] DMSO, 100 MHz, 25 °C):  $\delta$  = 169.4, 158.9, 158.7, 152.7, 152.5, 151.2, 149.5, 146.5, 145.5, 139.6, 128.8, 124.2, 122.6, 66.3, 44.9, 43.7.

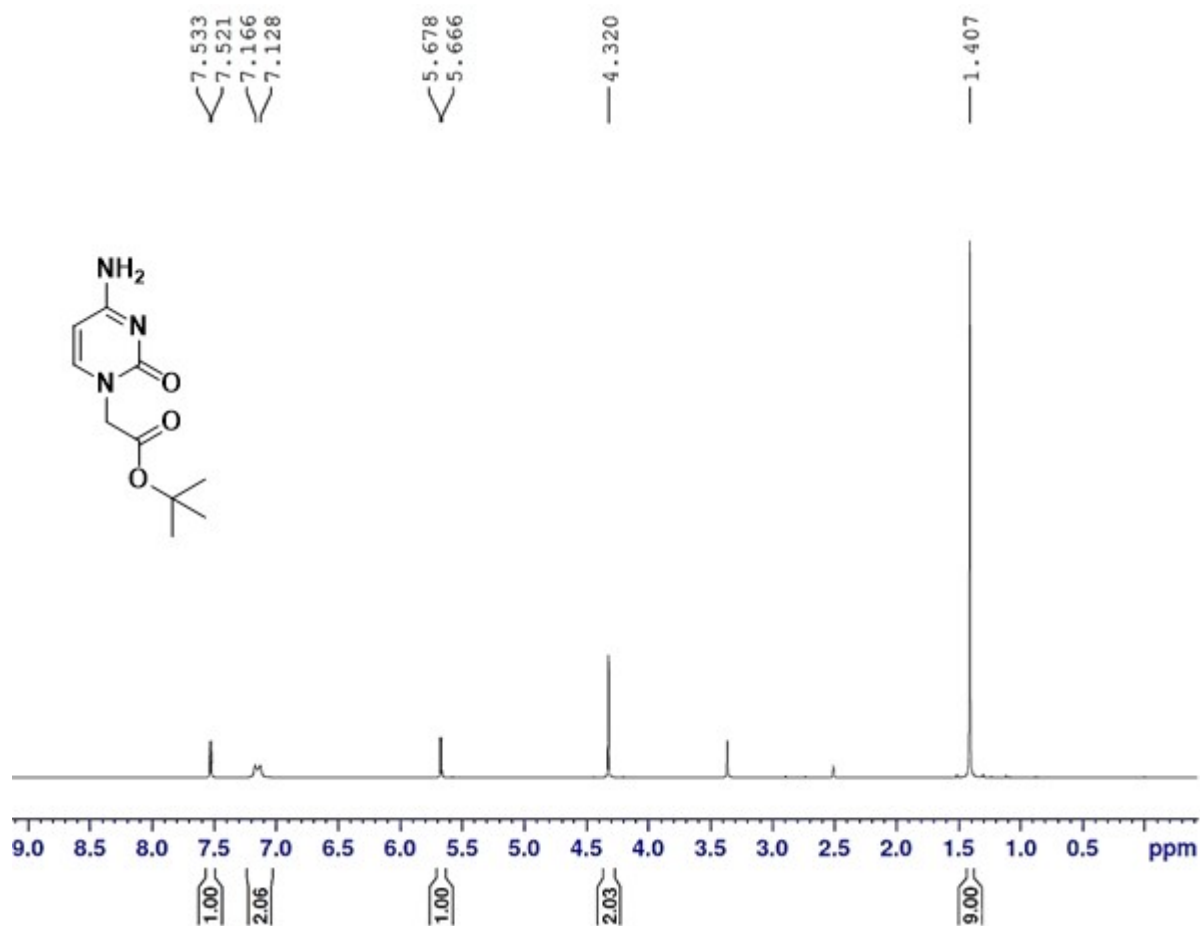


## *tert*-butyl cytosine-1-ylacetate (9)

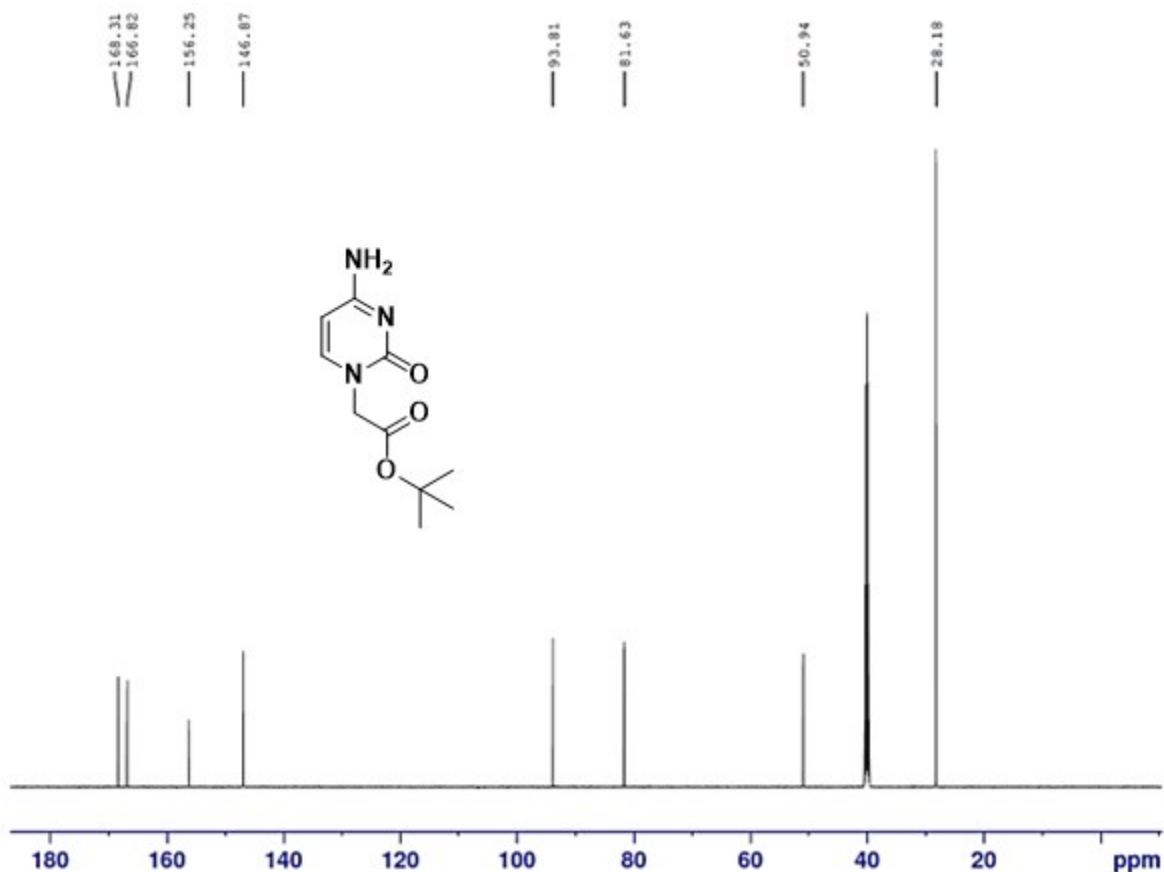
### HRMS



### <sup>1</sup>H NMR

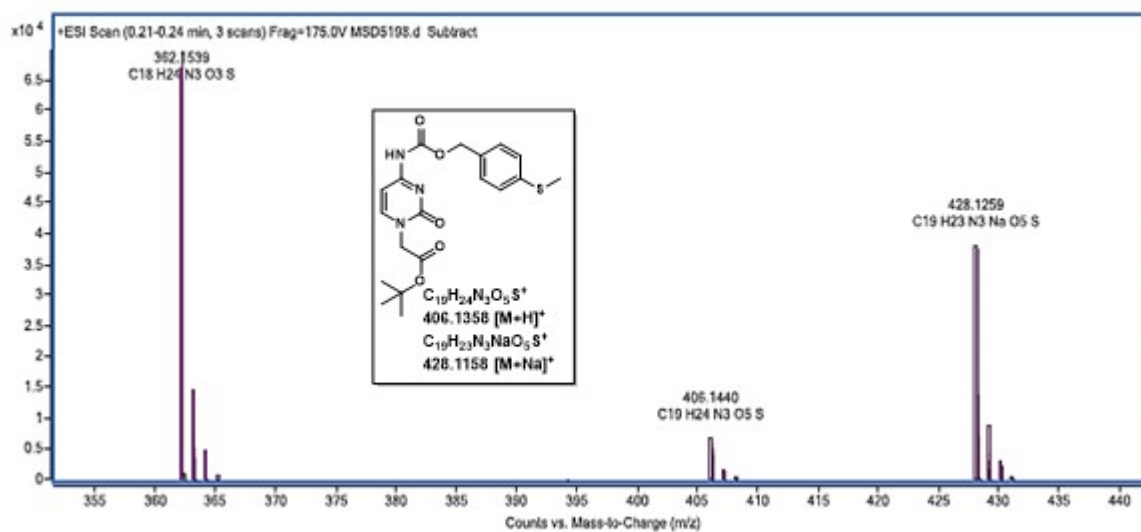


## <sup>13</sup>C NMR



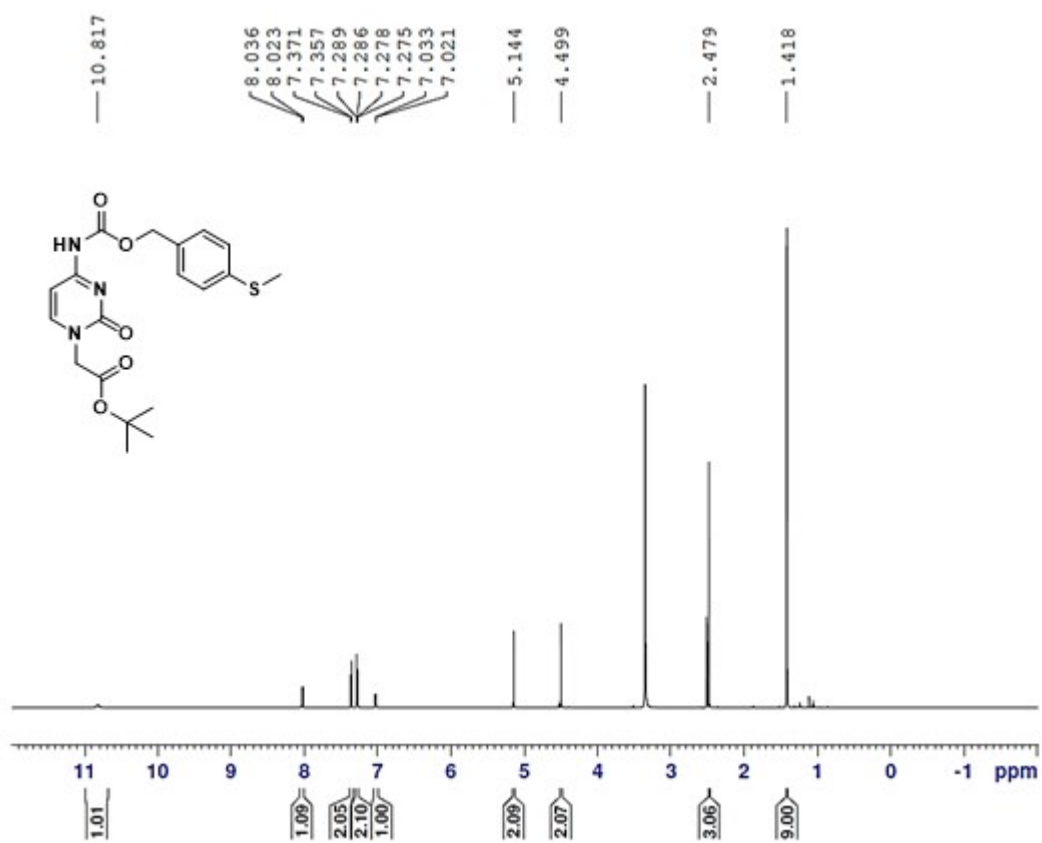
<sup>1</sup>H NMR ([d<sub>6</sub>] DMSO, 600 MHz, 25 °C):  $\delta$  = 7.52 (d,  $J$  = 7.18 Hz, 1 H), 7.17 - 7.12 (s, NH<sub>2</sub>, 2 H), 5.67 (d,  $J$  = 7.18 Hz, 1 H), 4.32 (s, 2 H), 1.41 (s, 9 H); <sup>13</sup>C NMR ([d<sub>6</sub>]DMSO, 100 MHz, 25 °C):  $\delta$  = 168.3, 166.8, 156.2, 146.8, 93.8, 81.6, 50.9, 28.2.

## tert-butyl (N<sup>4</sup>-Mtz-cytosine-1-yl)acetate (10)

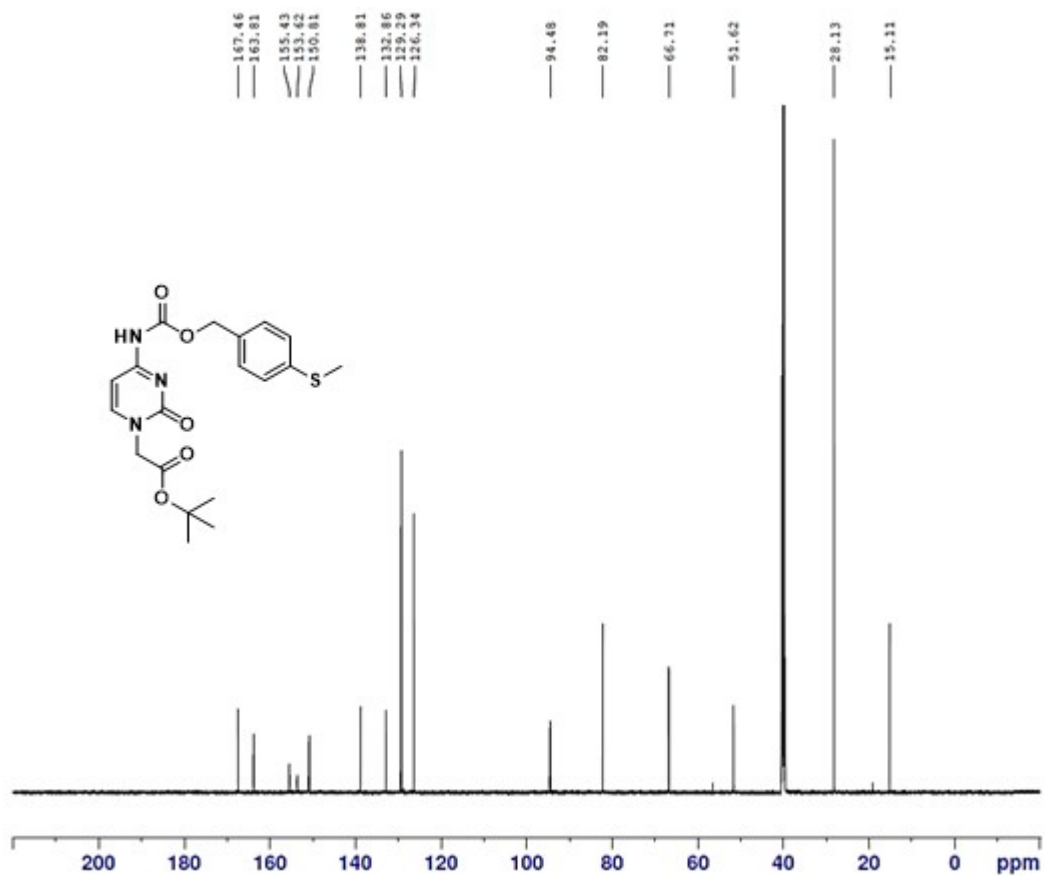


## HRMS

# <sup>1</sup>H-NMR

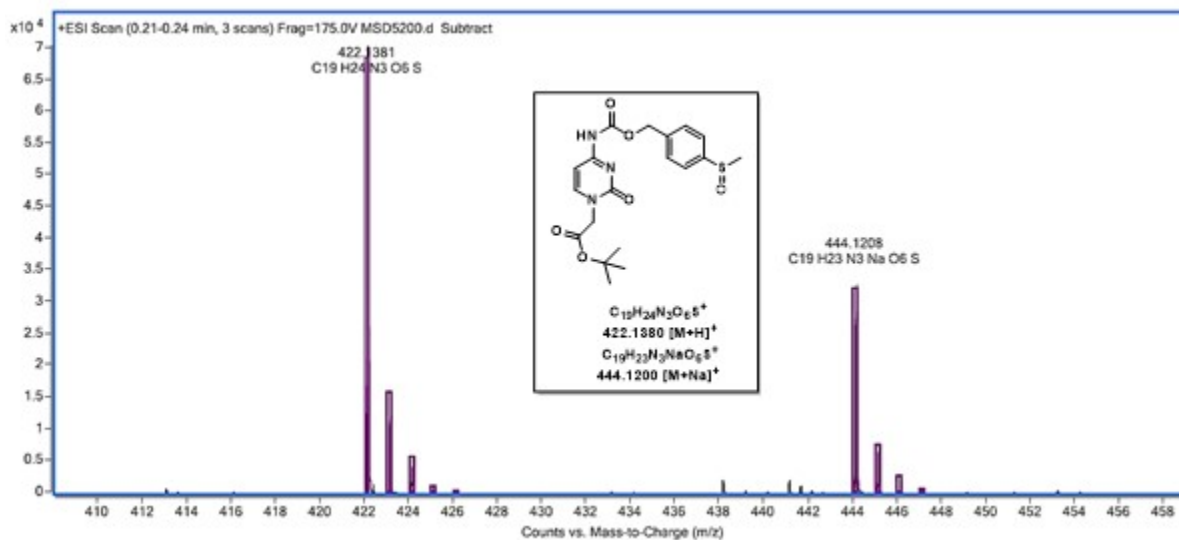


# <sup>13</sup>C-NMR



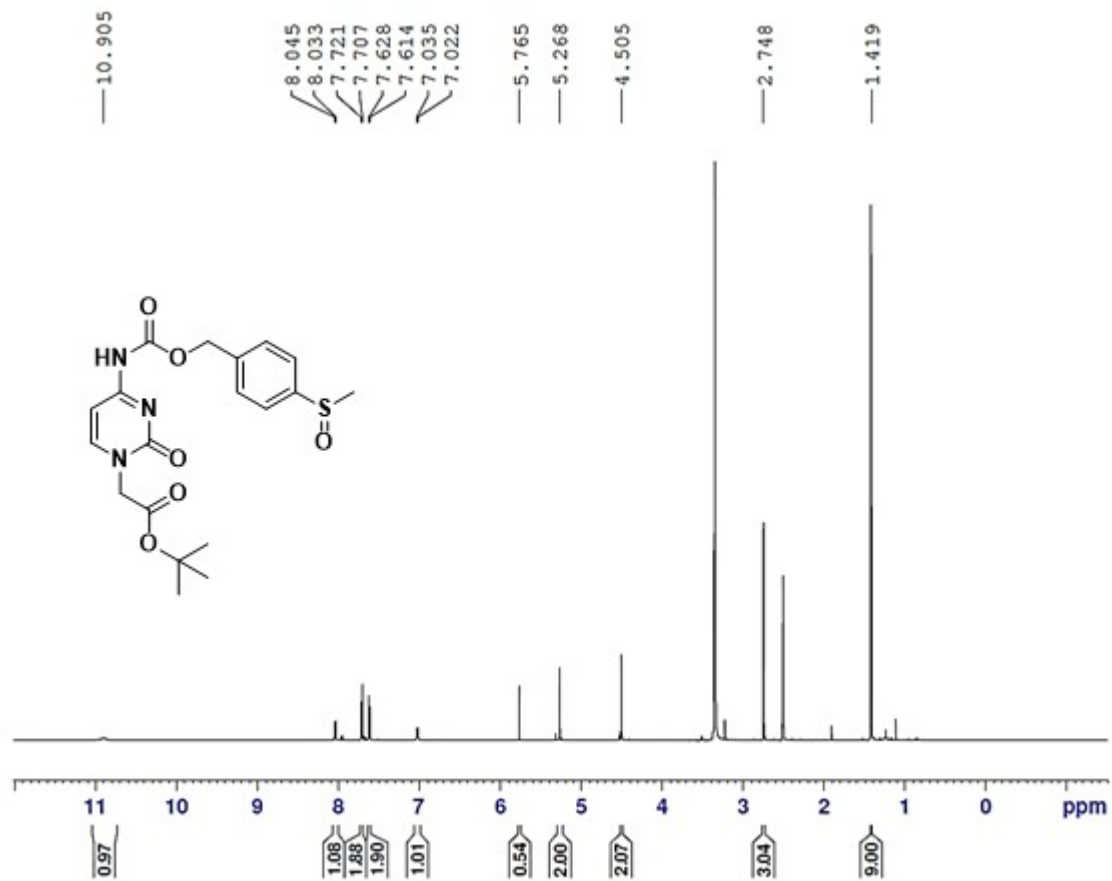
$^1\text{H}$  NMR ([d<sub>6</sub>]DMSO, 600 MHz, 25 °C):  $\delta$  = 10.8 (s, NH, 1 H), 8.02 (d,  $J$  = 7.3 Hz, 1 H), 7.36 (d,  $J$  = 8.3 Hz, 2 H), 7.28 (d,  $J$  = 8.3 Hz, 2 H), 7.02 (d,  $J$  = 7.3 Hz, 1 H), 5.14 (s, 2 H), 4.49 (s, 2 H), 2.41 (s, 3 H), 1.41 (s, 9 H);  $^{13}\text{C}$  NMR ([d<sub>6</sub>]DMSO, 100 MHz, 25 °C):  $\delta$  = 167.4, 163.8, 155.4, 153.6, 150.8, 138.8, 132.8, 129.3, 126.3, 94.4, 82.2, 66.7, 51.6, 28.1, 15.1.

### ***tert*-butyl (*N*<sup>4</sup>-Msz-cytosine-1-yl)acetate (11)**

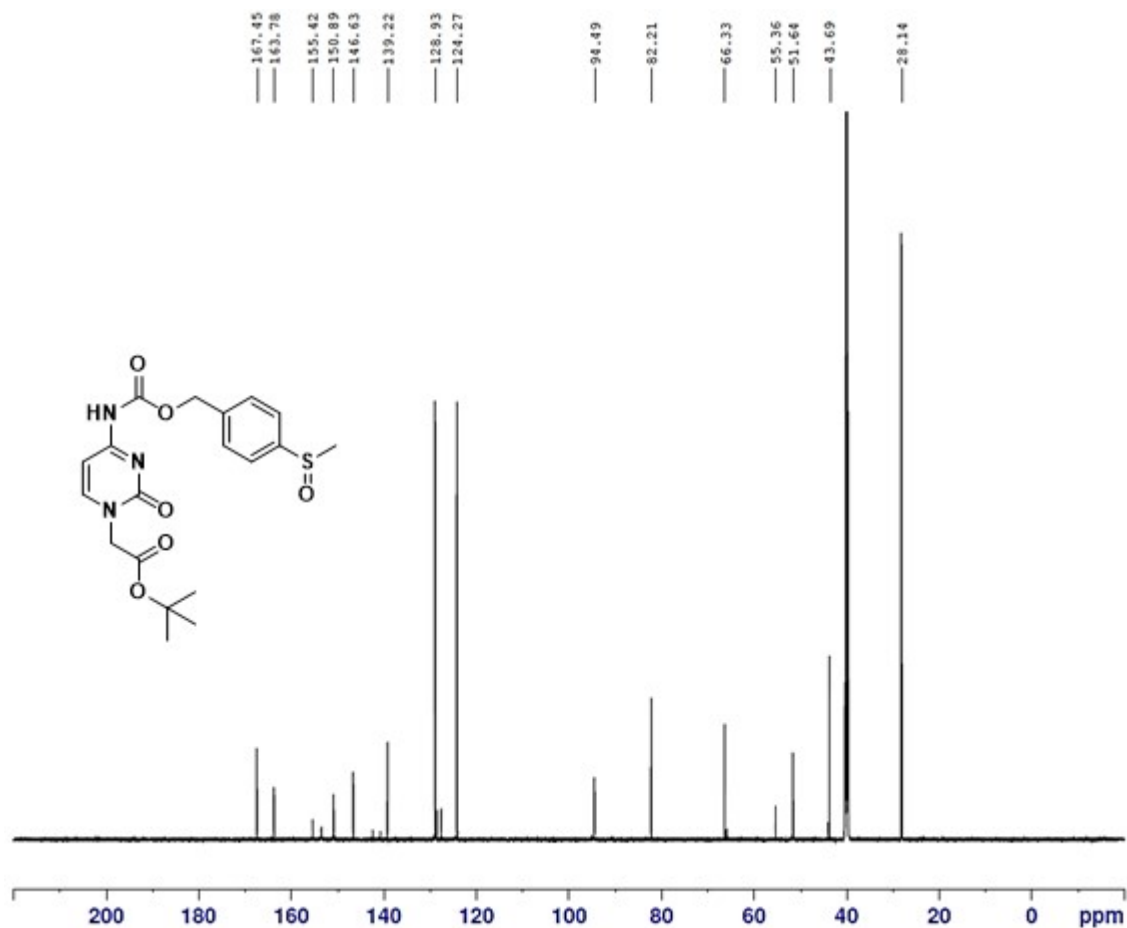


HRMS

$^1\text{H}$  NMR

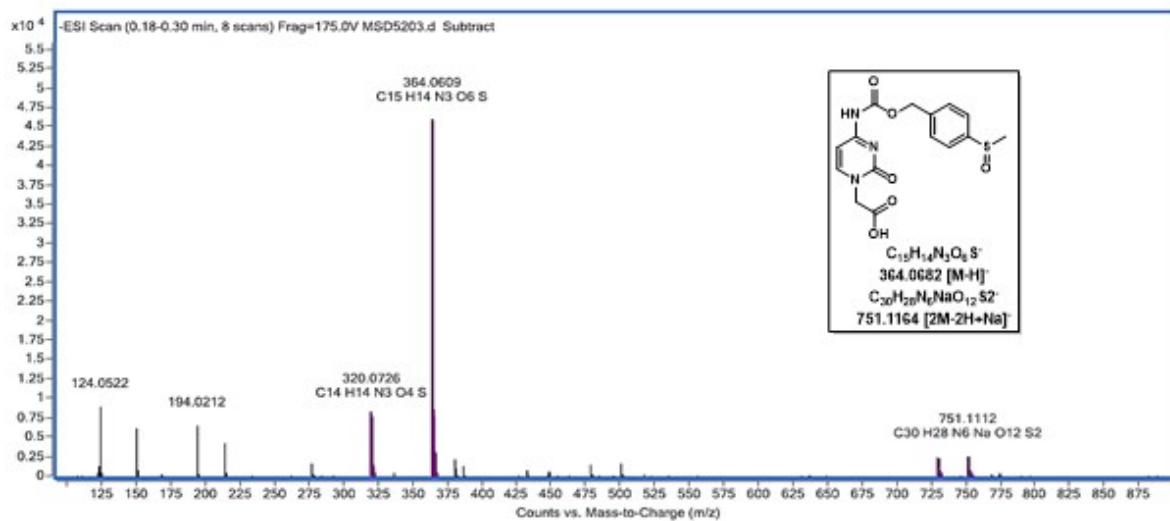


$^{13}\text{C}$  NMR



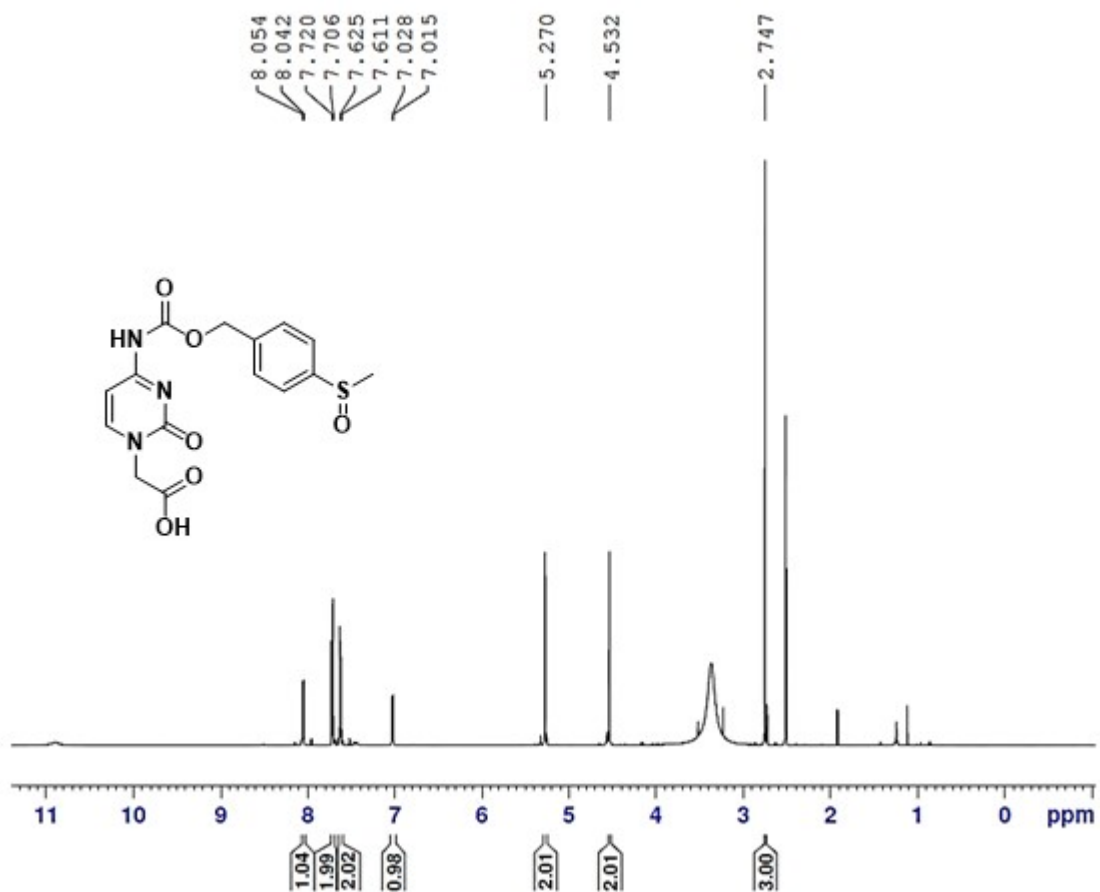
<sup>1</sup>H NMR ([d<sub>6</sub>]DMSO, 600 MHz, 25 °C): δ = 10.9 (s, NH, 1 H), 8.04 (d, *J* = 7.3 Hz, 1 H), 7.71 (d, *J* = 8.3 Hz, 2 H), 7.62 (d, *J* = 8.3 Hz, 2 H), 7.03 (d, *J* = 7.3 Hz, 1 H), 5.76 (s, 1H), 5.26 (s, 2 H), 4.50 (s, 2 H), 2.75 (s, 3 H), 1.42 (s, 9 H); <sup>13</sup>C NMR ([d<sub>6</sub>]DMSO, 100 MHz, 25 °C): δ = 167.4, 163.7, 155.4, 150.8, 146.6, 139.2, 128.9, 124.2, 94.5, 82.2, 66.3, 55.3, 51.6, 43.6, 28.1.

### (N<sup>4</sup>-Msz-cytosine-1-yl)acetic acid (12)

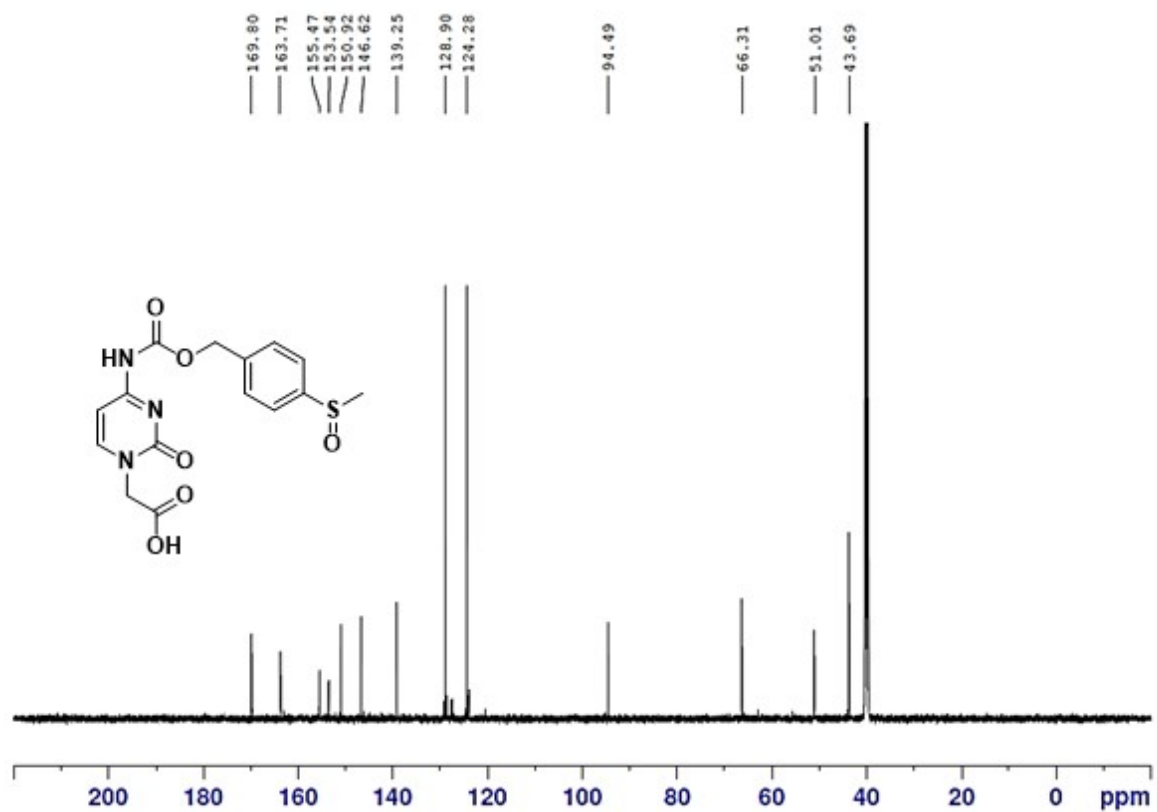


HRMS

# <sup>1</sup>H NMR

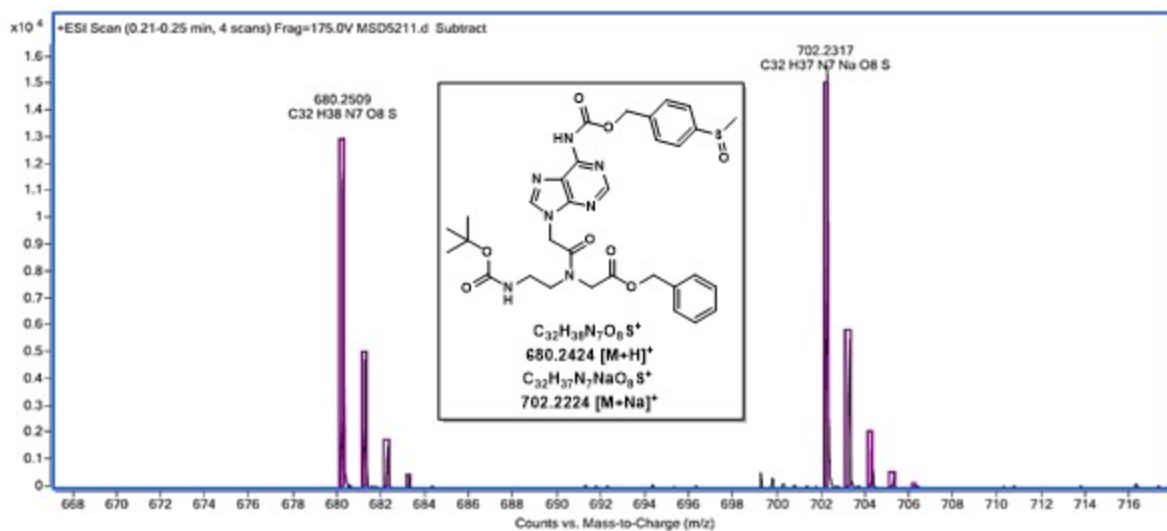


# <sup>13</sup>C NMR



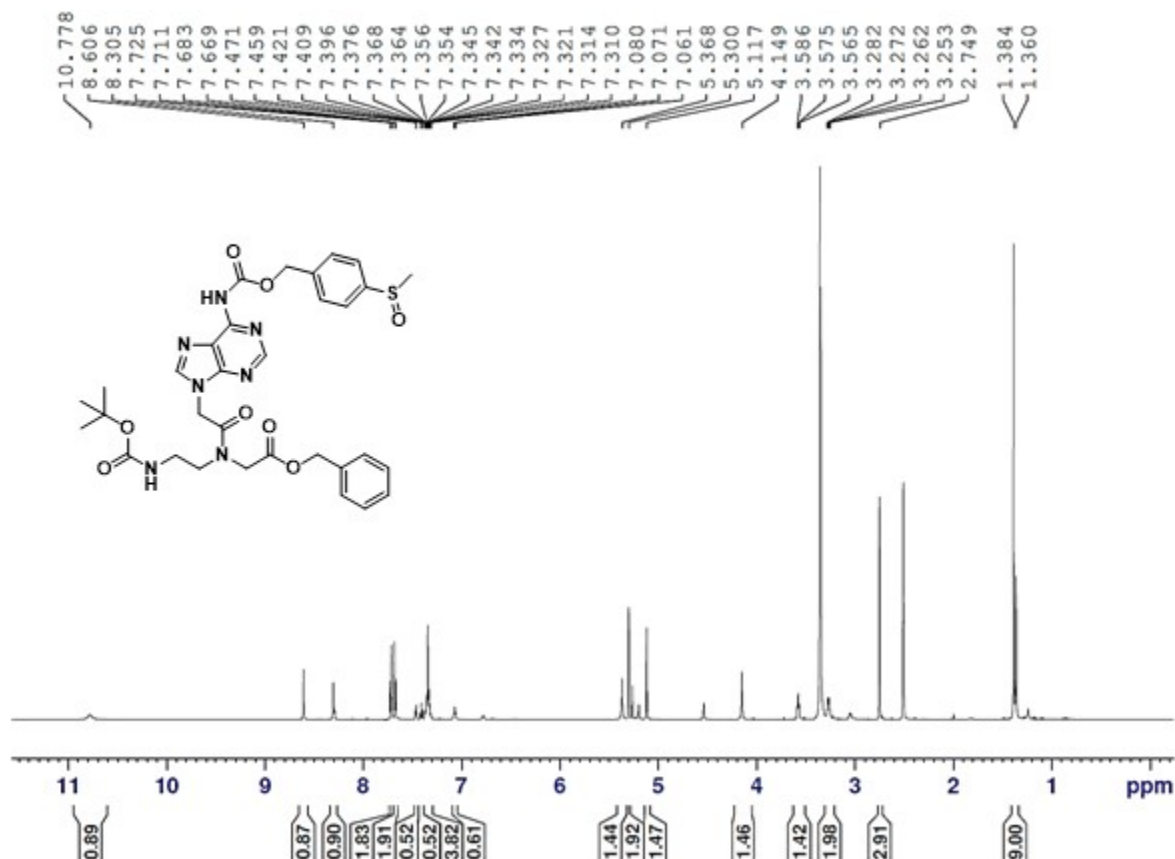
$^1\text{H}$  NMR ([d<sub>6</sub>]DMSO, 600 MHz, 25 °C):  $\delta$  = 8.04 (d,  $J$  = 7.3 Hz, 1 H), 7.71 (d,  $J$  = 8.3 Hz, 2 H), 7.61 (d,  $J$  = 8.3 Hz, 2 H), 7.02 (d,  $J$  = 7.3 Hz, 1 H), 5.26 (s, 2 H), 4.53 (s, 2 H), 2.74 (s, 3 H);  $^{13}\text{C}$  NMR ([d<sub>6</sub>]DMSO, 100 MHz, 25 °C):  $\delta$  = 169.8, 163.7, 155.4, 153.5, 150.9, 146.6, 139.2, 128.9, 124.2, 94.5, 66.3, 51.0, 43.7.

### Boc-PNA-A(Msz)-OBzl (13)



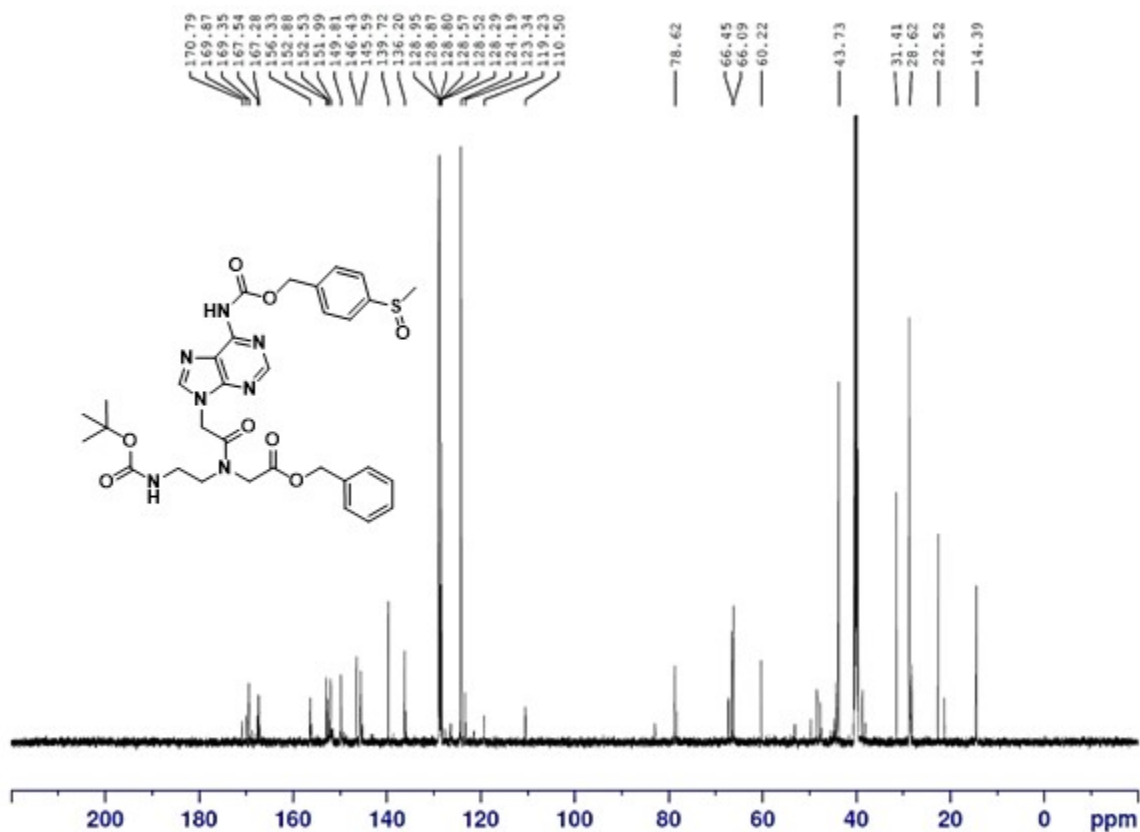
HRMS

### $^1\text{H}$ NMR





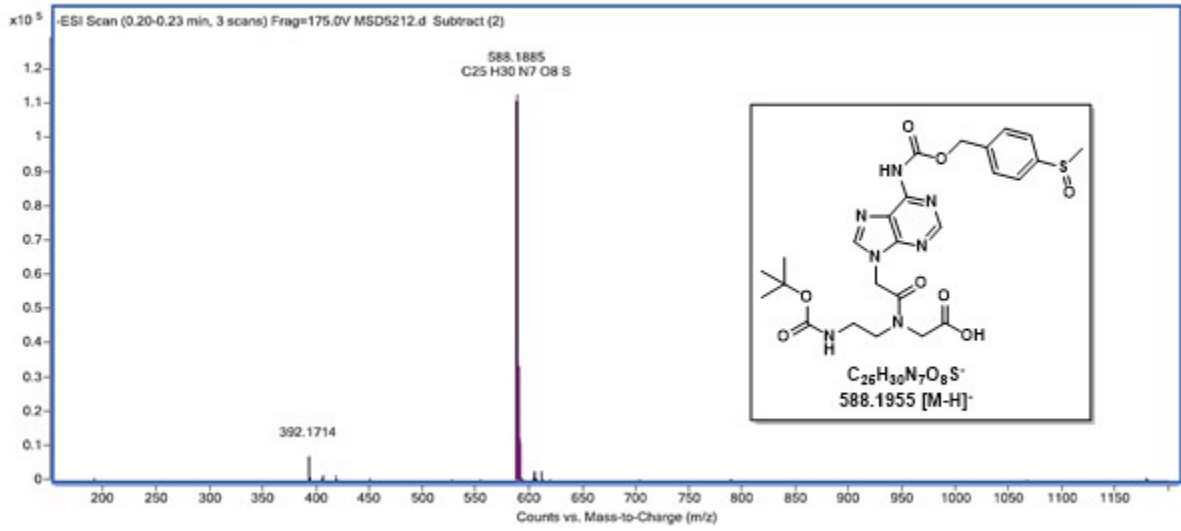
## <sup>13</sup>C NMR



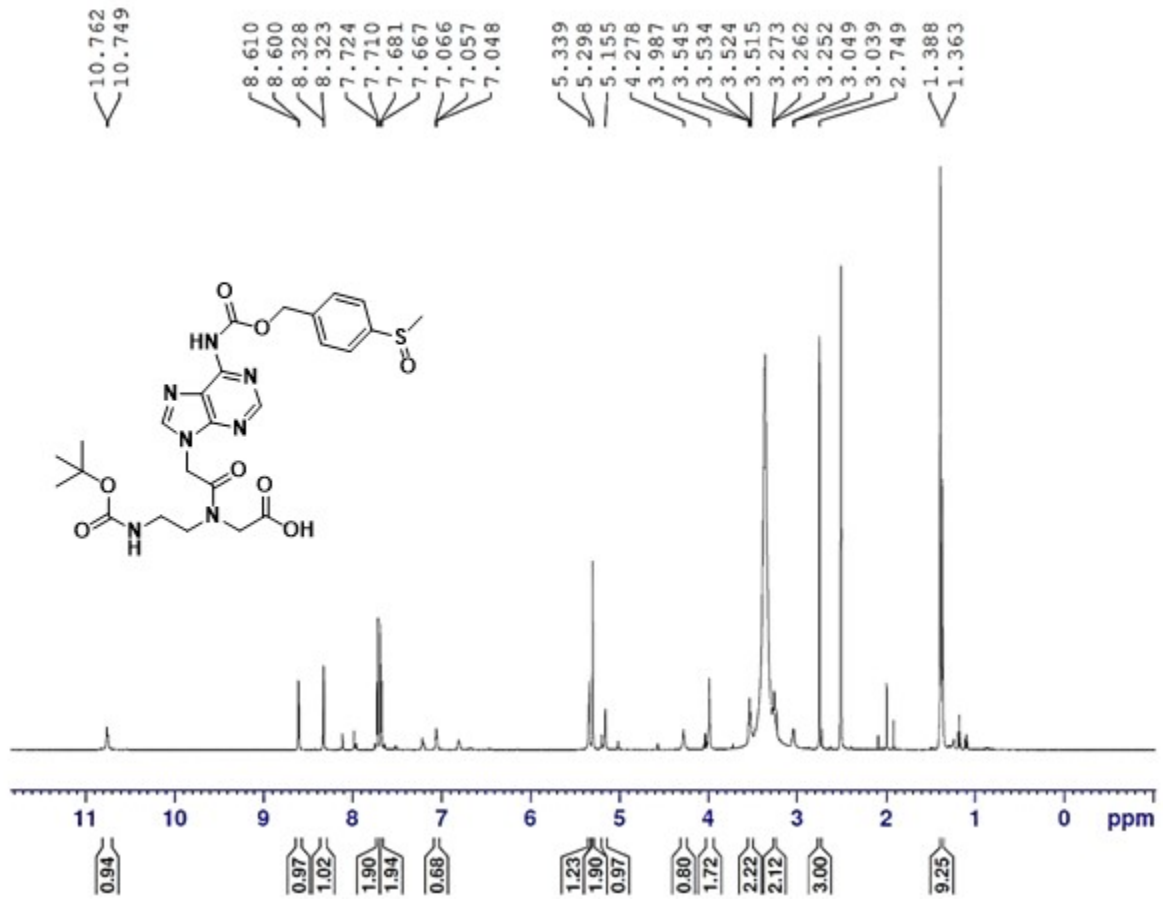
<sup>1</sup>H NMR ([d<sub>6</sub>]DMSO, 600 MHz, 25 °C):  $\delta$  = 10.78 (s, 1 H, NH), 8.61 (s, 1H), 8.31 (s, 1H), 7.72 (d,  $J$  = 8.29 Hz, 2 H), 7.67 (d,  $J$  = 8.29 Hz, 2 H), 7.46 (d,  $J$  = 7.23 Hz, 1 H), 7.41 (t,  $J$  = 7.34 Hz, 1 H), 7.35 (m, 4H), 7.07 (t,  $J$  = 5.82 Hz, 1 H), 5.37 (s, 2 H), 5.29 (s, 2 H), 5.11 (s, 2 H), 4.15 (s, 2 H), 3.58 (t,  $J$  = 6.34 Hz, 2 H), 3.27 (q,  $J$  = 6.05 Hz, 2 H), 2.75 (s, 3 H), 1.38-1.36 (s, 9H); <sup>13</sup>C NMR ([d<sub>6</sub>]DMSO, 100 MHz, 25 °C):  $\delta$  = 170.8, 169.8, 169.3, 167.5, 167.2, 156.3, 152.8, 152.5, 151.9, 149.8, 146.4, 145.6, 139.7, 136.2, 128.9, 128.8, 128.6, 128.5, 128.3, 124.1, 123.4, 119.2, 78.6, 66.4, 66.1, 60.2, 43.7, 31.4, 28.6, 22.5, 14.4.

### Boc-PNA-A(Msz)-OH (14)

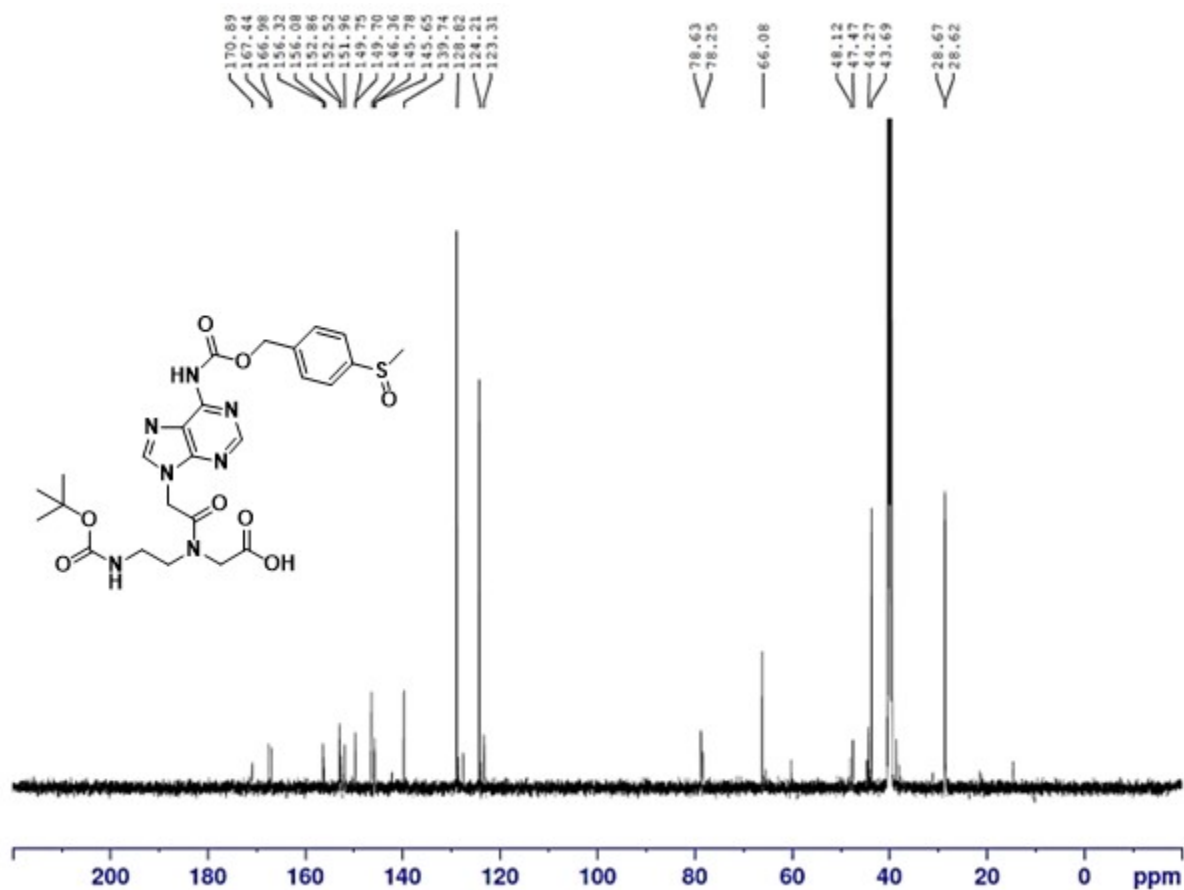
### HRMS



### <sup>1</sup>H NMR

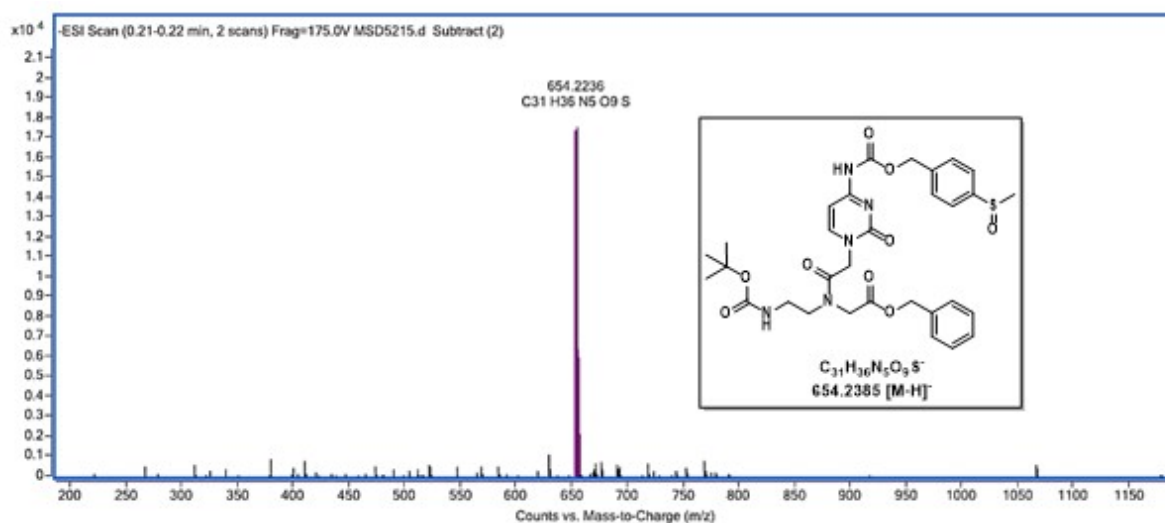


### <sup>13</sup>C NMR



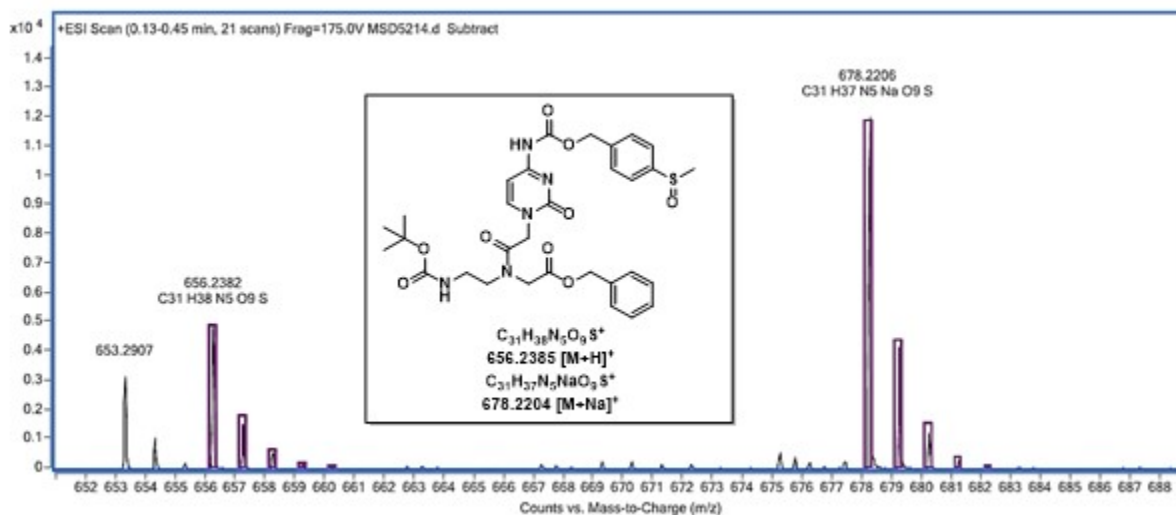
H), 3.26 (q,  $J = 6.17$  Hz, 2 H), 2.74 (s, 3 H), 1.38-1.36 (s, 9H); <sup>13</sup>C NMR ([d<sub>6</sub>]DMSO, 100 MHz, 25 °C):  $\delta = 170.8, 167.4, 166.9, 156.3, 152.9, 152.5, 151.9, 149.8, 149.7, 146.3, 145.8, 145.7, 139.7, 128.8, 124.2, 123.3, 78.6, 78.3, 66.1, 48.1, 47.5, 44.3, 43.6, 28.7, 28.6$ .

### Boc-PNA-C(Msz)-OBzl (15)

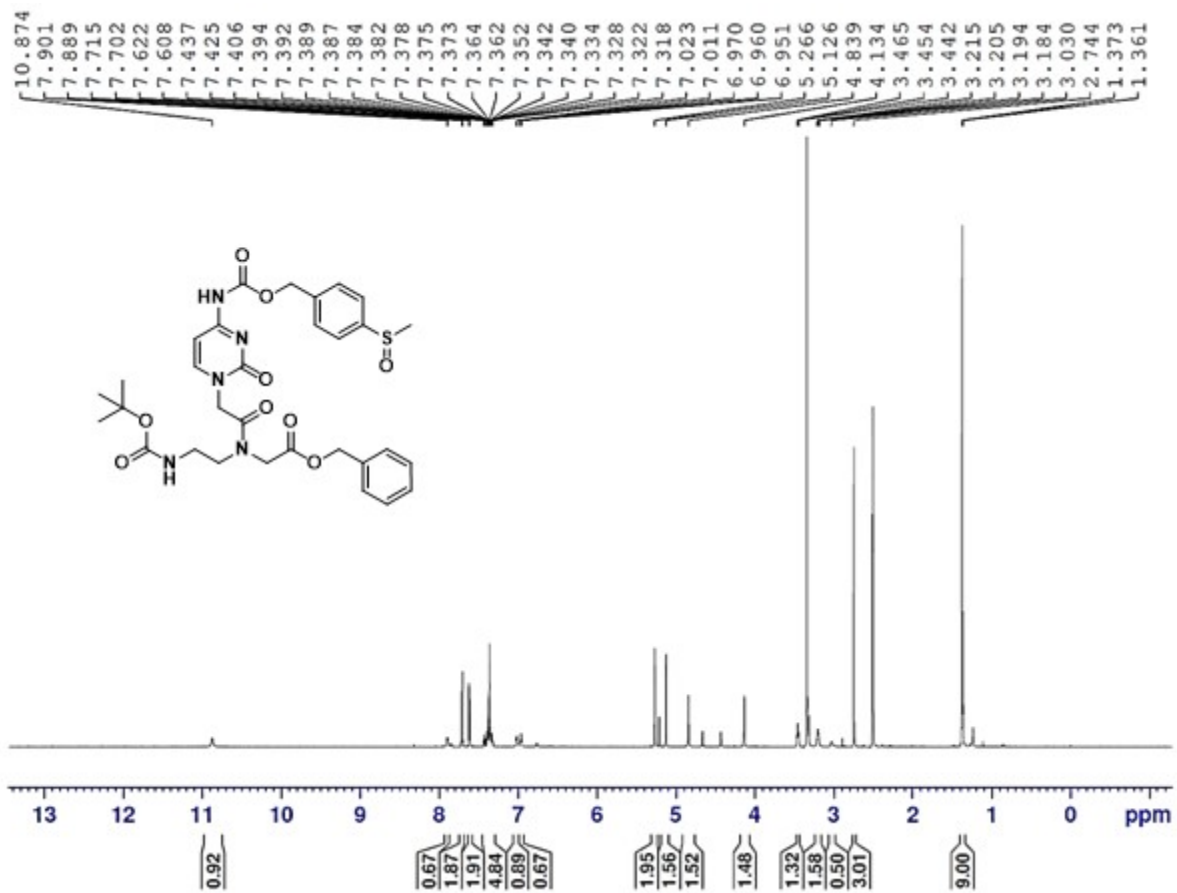


HRMS (+<sup>ve</sup> mode)

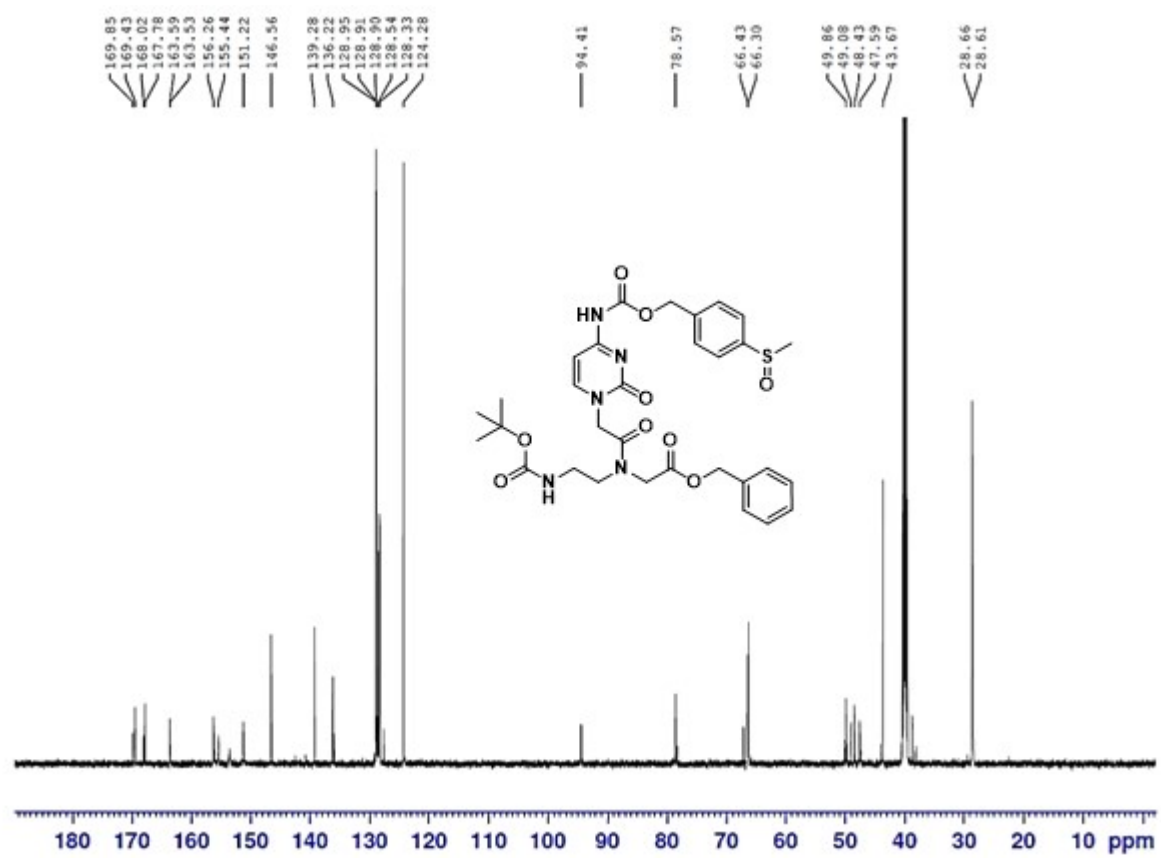
# HRMS (-ve mode)



# <sup>1</sup>H NMR

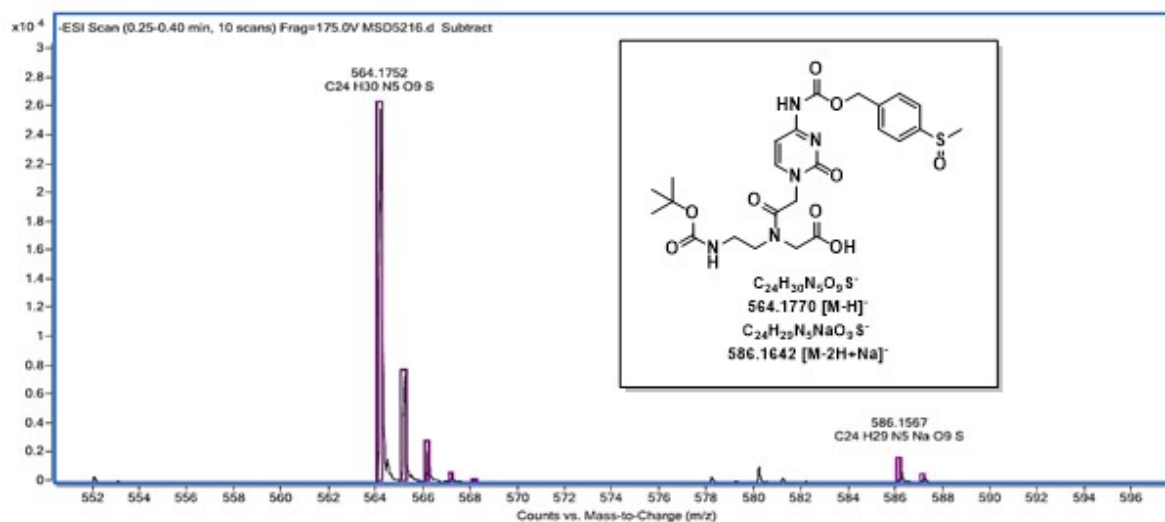


**<sup>13</sup>C NMR**



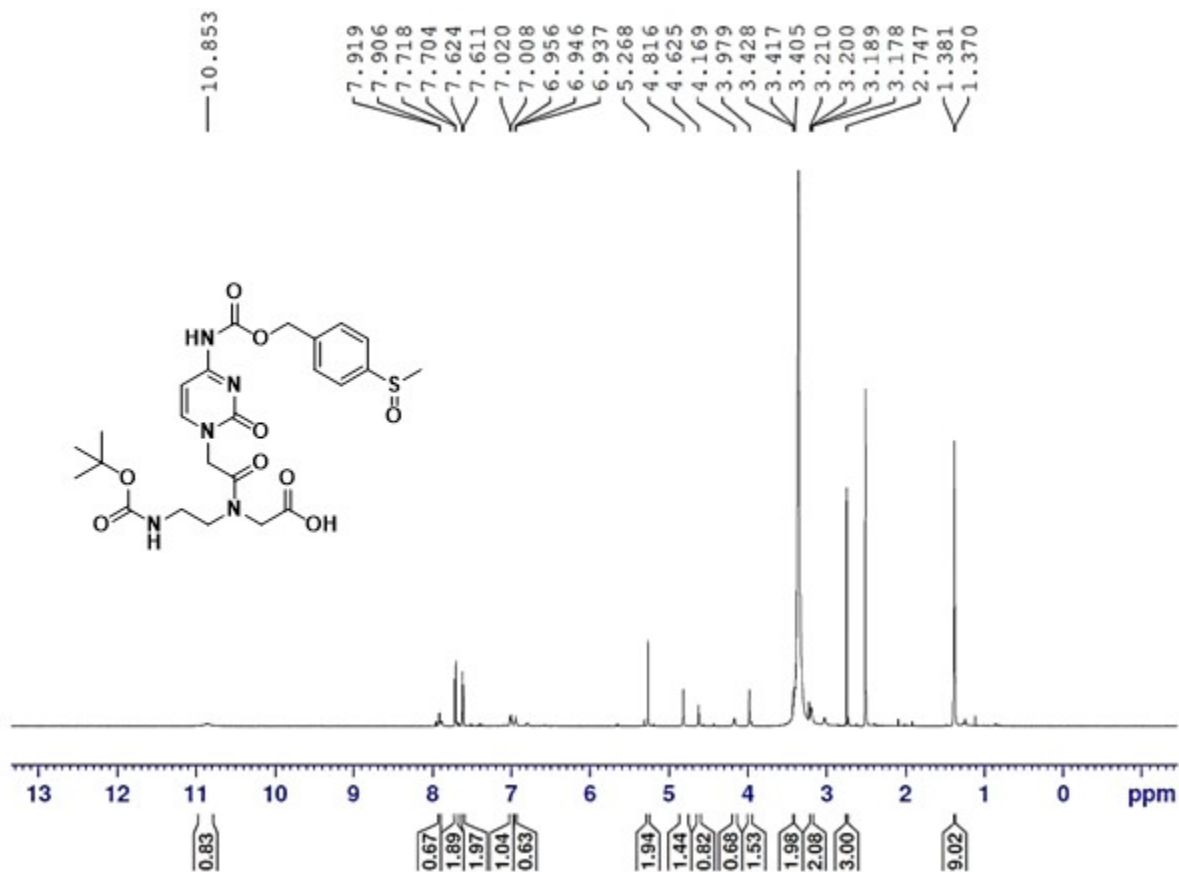
$^1\text{H}$  NMR ([d<sub>6</sub>]DMSO, 600 MHz, 25 °C):  $\delta$  = 10.87 (s, 1 H, NH), 7.89 (d,  $J$  = 7.10 Hz, 1 H), 7.71 (d,  $J$  = 8.26 Hz, 2 H, ArH), 7.67 (d,  $J$  = 8.22 Hz, 2 H, ArH), 7.36 (m, 5H), 7.02 (d,  $J$  = 7.22 Hz, 1 H), 6.96 (t,  $J$  = 5.74 Hz, 1 H), 5.26 (s, 2 H), 5.12 (s, 2 H), 4.84 (s, 2 H), 4.13 (s, 2 H), 3.45 (t,  $J$  = 6.77 Hz, 1 H), 3.19 (dd,  $J$  = 6.29 Hz, 2 H), 3.02 (dd,  $J$  = 6.24 Hz, 1 H), 2.74 (s, 3 H), 1.37-1.36 (s, 9H);  $^{13}\text{C}$  NMR ([d<sub>6</sub>]DMSO, 100 MHz, 25 °C):  $\delta$  = 169.8, 169.4, 168.0, 167.8, 163.6, 163.5, 156.2, 155.4, 151.2, 146.5, 139.2, 136.2, 128.9, 128.8, 128.5, 128.3, 124.3, 94.4, 78.5, 66.4, 66.2, 49.8, 49.1, 48.4, 47.6, 43.7, 28.7, 28.6.

## Boc-PNA-C(Msz)-OH (16)

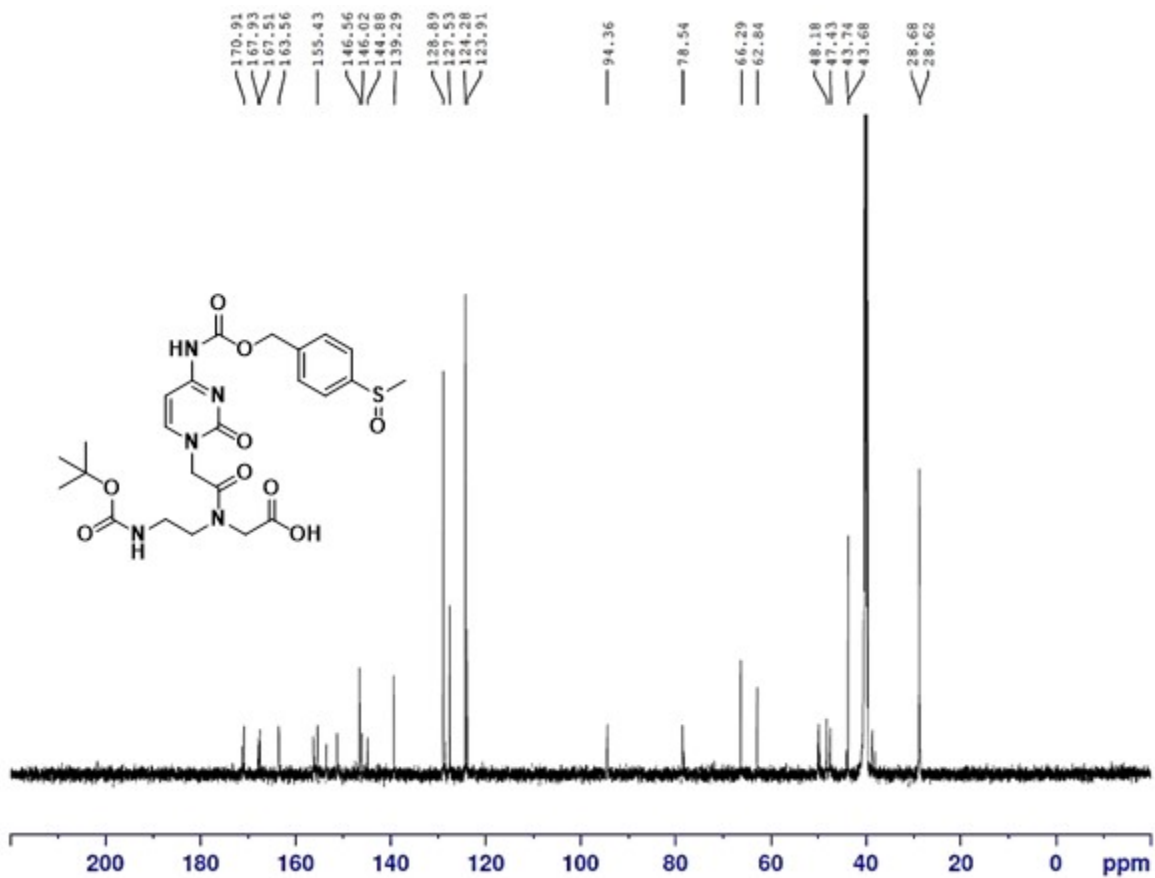


HRMS

$^1\text{H}$  NMR

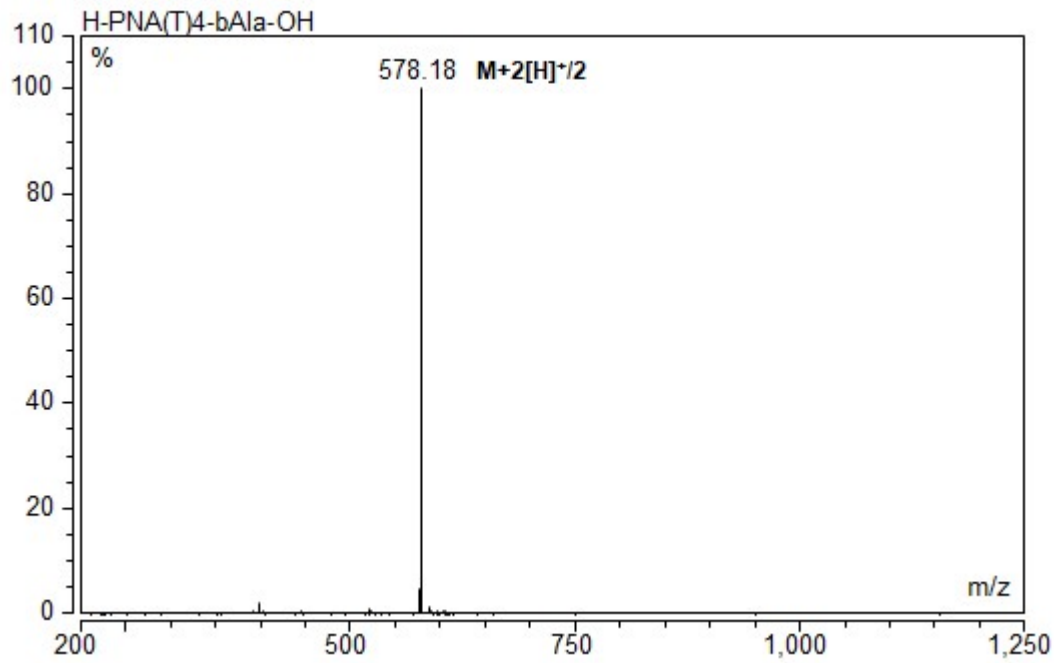


**<sup>13</sup>C NMR**



$^1\text{H}$  NMR ([d6]DMSO, 600 MHz, 25 °C):  $\delta$  = 10.85 (s, 1 H ), 7.91 (d,  $J$  = 7.37 Hz, 1 H), 7.71 (d,  $J$  = 8.23 Hz, 2 H, ArH), 7.62 (d,  $J$  = 8.22 Hz, 2 H, ArH), 7.01 (d,  $J$  = 7.28 Hz, 1 H), 6.94 (t,  $J$  = 5.71 Hz, 1 H), 5.27 (s, 2 H), 4.82 (s, 1 H), 4.62 (s, 1 H), 4.17 (s, 1 H), 3.98 (s, 2 H), 3.42 (t,  $J$  = 6.99 Hz, 2 H), 3.19 (q,  $J$  = 6.35 Hz, 2 H), 2.75 (s, 3 H), 1.38-1.36 (s, 9H);  $^{13}\text{C}$  NMR ([d6]DMSO, 100 MHz, 25 °C):  $\delta$  = 170.9, 167.9, 167.5, 163.6, 155.4, 146.6, 146.0, 144.8, 139.3, 128.8, 127.5, 124.3, 123.9, 94.4, 78.5, 66.3, 62.8, 48.2, 47.4, 43.7, 43.6, 28.7, 28.6.





### 3. LCMS of PNA oligomers

