Electronic Supplementary Information

Azide-alkyne cycloaddition for universal post-synthetic modifications of nucleic acids and effective synthesis of bioactive nucleic acid conjugates

Yu-Chih Su, Yu-Lun Lo, Chi-Ching Hwang, Li-Fang Wang, Min Hui Wu, Eng-Chi Wang, Yun-Ming Wang, Tzu-Pin Wang*

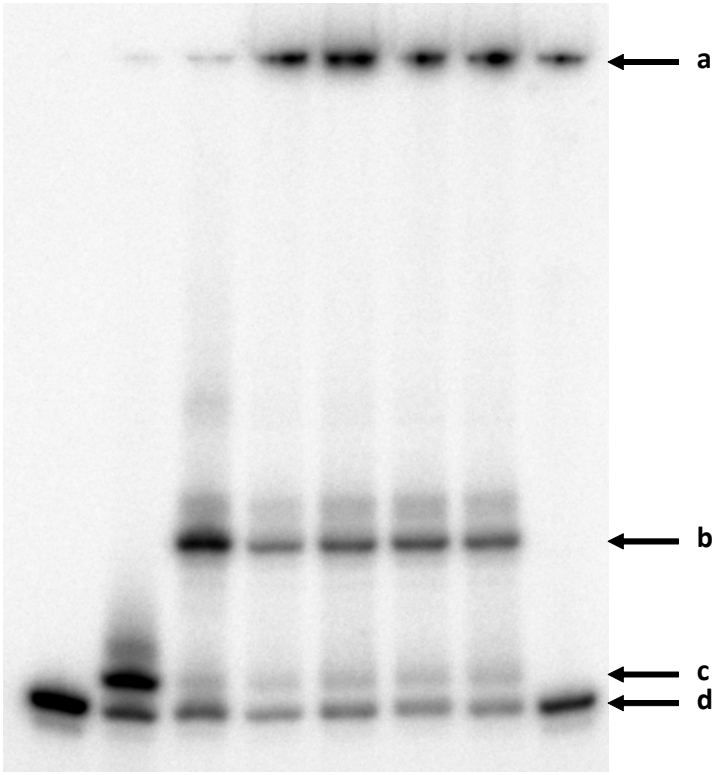
*Corresponding author: tzupinw@cc.kmu.edu.tw (T.-P. W.)

*Contact address: Department of Medicinal and Applied Chemistry, Kaohsiung Medical University, Kaohsiung, 80708, Taiwan

Contents

(A)

Maleimide Linker:	N/A	N/A	**3e**	**3a**	**3b**	**3c**	**3d**	N/A
Cystamine Linker:	-	+	+	+	+	+	+	-
POC (b) Yield (%):	N/A	N/A	72	27	27	37	30	0



(B)

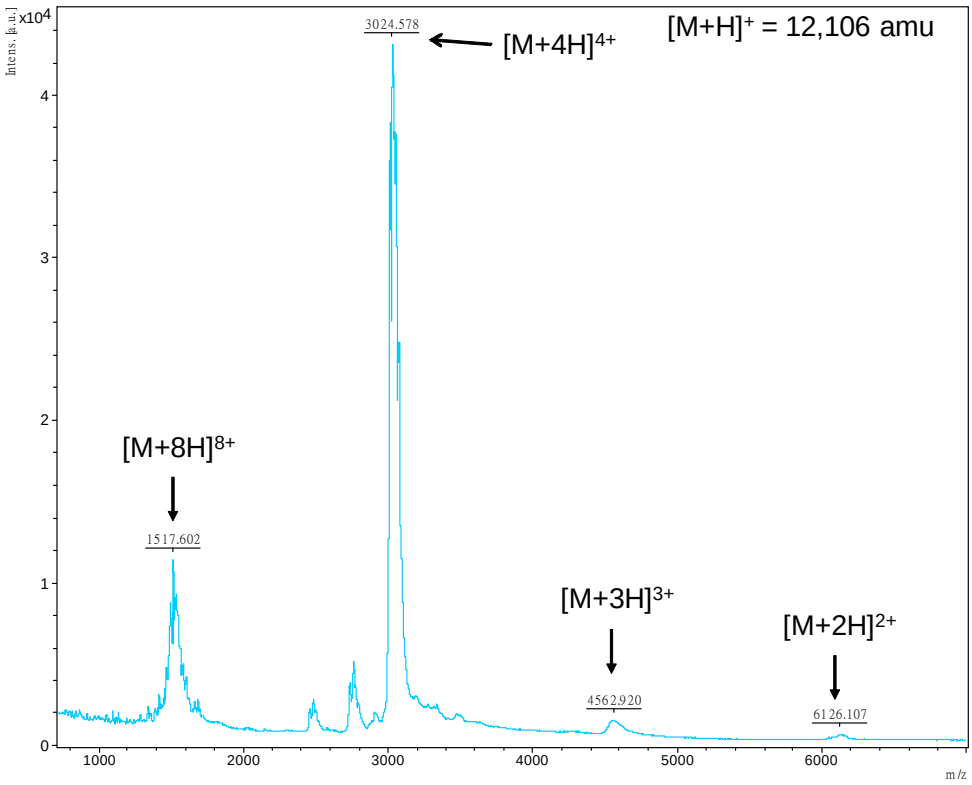


Fig. S1 Exploring potential of the *N*-maleoyl amino acid succinimidyl esters (3) for synthesis of POCs by Michael addition of a thiolate peptide. (A) The ^{32}P-labeled 3' primer DNA with or without previous modifications with cystamine and **3** were reacted with a cysteine-containing Tat peptide under the conditions of Michael addition and analyzed the reaction products by 20% urea-PAGE. **a**, the multiple peptide-linked DNA conjugate; **b**, the POC product; **c**, the cystamine-modified 3' primer DNA; **d**, the 3' primer DNA. (B) MALDI-TOF analysis of the denaturing PAGE-purified Product **a** from the far right reaction in Fig. S1A. The molecular mass for the 3' primer DNA is 6,038 amu and for the cysteine-containing Tat peptide is 1,517 amu. With four of the peptides covalently linked to the 3' primer DNA, the molecular mass of the DNA conjugate is equal to the 12,106 amu. It is noted that, even though the unusual MALDI-TOF peak distribution of Product **a** as shown here, the observed peaks could not be attributed to $[M+H]^+$, $[2M+H]^{2+}$, $[3M+H]^{3+}$, etc, because Product **a** had been purified by denaturing urea-PAGE prior to the MALDI-TOF analysis. We thus ruled out the possible presence of DNA conjugates with lower molecular mass as the identity of Product **a**.

CuAAC:	-	-	-	+	+	+	+	+	+
Alkyne : azido-DNA ratio:	-	-	-	0.33	1	10	20	40	20
DNA substrate:	-	1	2	3	3	3	3	3	3
Cu ligand:	-	-	-	4	4	4	4	4	5
Yield (%):	N/A	76	83	16	37	56	80	82	N/A

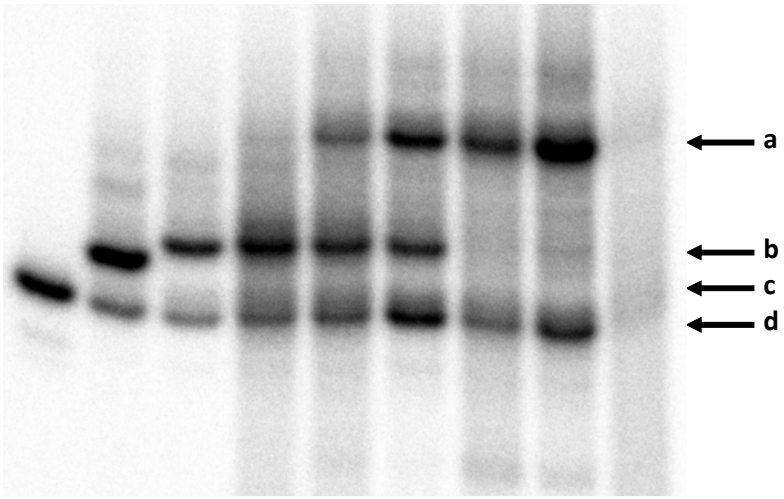


Fig. S2 Determination of the optimal alkyne : azido-DNA ratio and the appropriate copper ligand in the CuAAC reaction. The studied DNA was the 3' primer DNA and was labelled with 32-P at the 5' end before the reactions. The reaction products were analyzed by 20% urea-PAGE and visualized by an Amersham Typhoon PhosphorImager. 1, ethylenediamine; 2, ethylenediamine + **6d**; 3, ethylenediamine + **6d** + **10**; 4, THPTA; 5, NTB; **a**, the CuAAC reaction product between the azido 3' primer DNA and **10**; **b**, the **6d**-ethylenediamine-3' primer DNA conjugate; **c**, the ethylenediamine-modified 3' primer DNA; **d**, the 3' primer DNA.

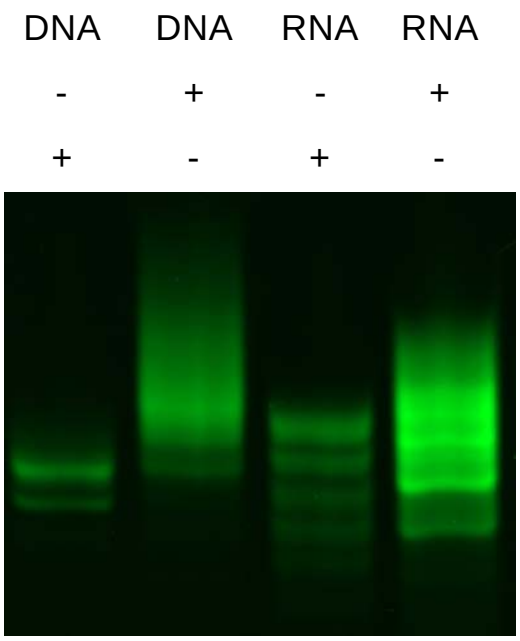


Fig. S3 The products of the CuAAC reactions between azido nucleic acids and the fluorophore 15, or between alkynyl nucleic acids and the other fluorophore 16 as analyzed by fluorescence imaging. The nucleic acid conjugates acquired from the CuAAC reactions were separated by 20% urea-PAGE and visualized by an Amersham Typhoon PhosphorImager through detection of fluorescein signals (Reference 14 in the text). DNA, the 3' primer DNA; RNA, the 17-mer RNA; alkynyl substrate, ethylenediamine + **19**; azido substrate, ethylenediamine + **6d**.

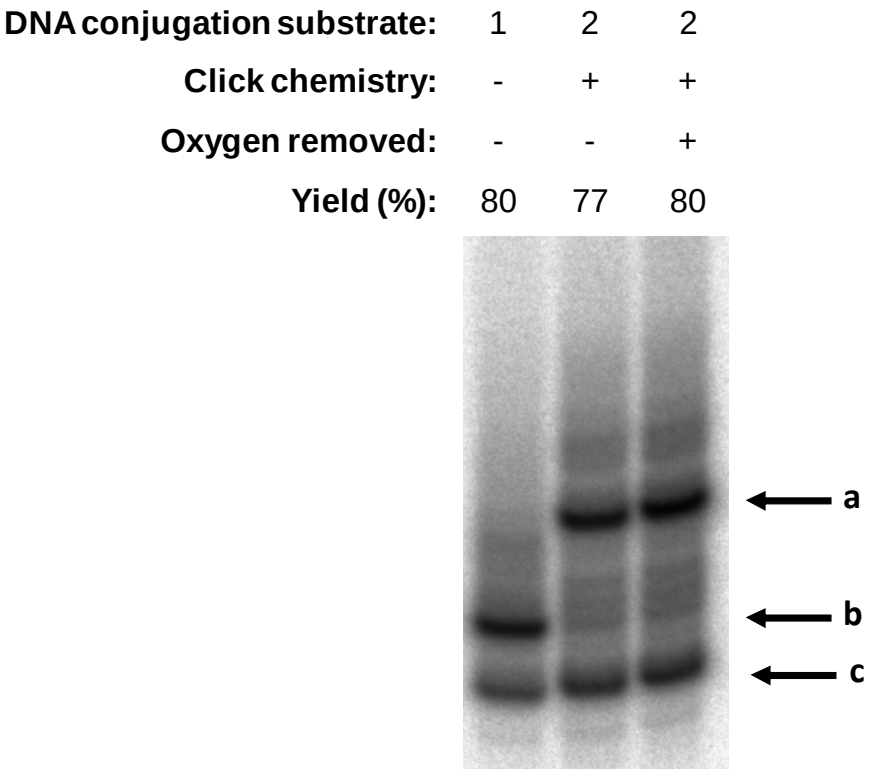


Fig. S4 Oxygen not the major cause of nucleic acid degradation in the CuAAC reactions. The DNA, the 3' primer DNA, was labeled with ^{32}P at the 5' end, and modified with ethylenediamine by the two-step phosphoramidation reaction and with **6d** by the amidation reaction, sequentially. The afforded DNA conjugate was further subjected to the CuAAC reaction with **10** in the presence/absence of oxygen to attain the final reaction products. The CuAAC reaction products were analyzed by 20% urea-PAGE and visualized by an Amersham Typhoon PhosphorImager. DNA conjugation substrate: 1, ethylenediamine + **6d**; 2, ethylenediamine + **6d** + **10**. a, the triazole product of the CuAAC reaction between the azido 3' primer DNA and **10**; b, the azido 3' primer DNA; c, the 3' primer DNA.

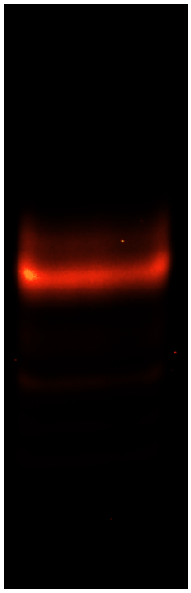


Fig. S5 Product analysis of the SPAAC reaction between the azido 3’ primer DNA and the cyclooctyne Alkyne MegaStokes dye 608 by fluorescence imaging.

The azido DNA was prepared by the two-step phosphoramidation reaction to modify the 3’ primer DNA with ethylenediamine and subsequently by amidation of the ethylenediamine-modified 3’ primer DNA with **6d**. The DNA conjugates obtained from the SPAAC reaction were purified by the first 20% urea-PAGE, analyzed by the second 20% urea-PAGE and finally visualized by an Amersham Typhoon PhosphorImager with the settings of the excitation wavelength at 488 nm and the emission wavelength at 580 nm.

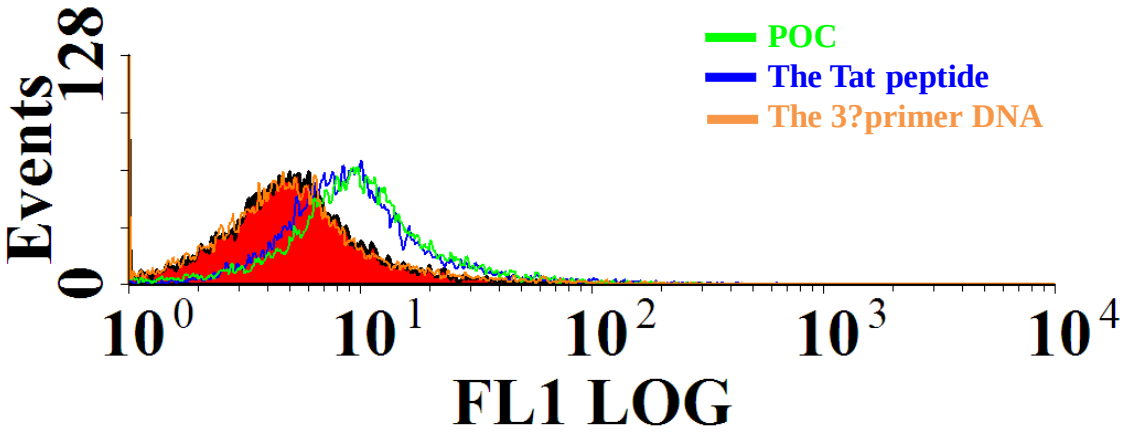
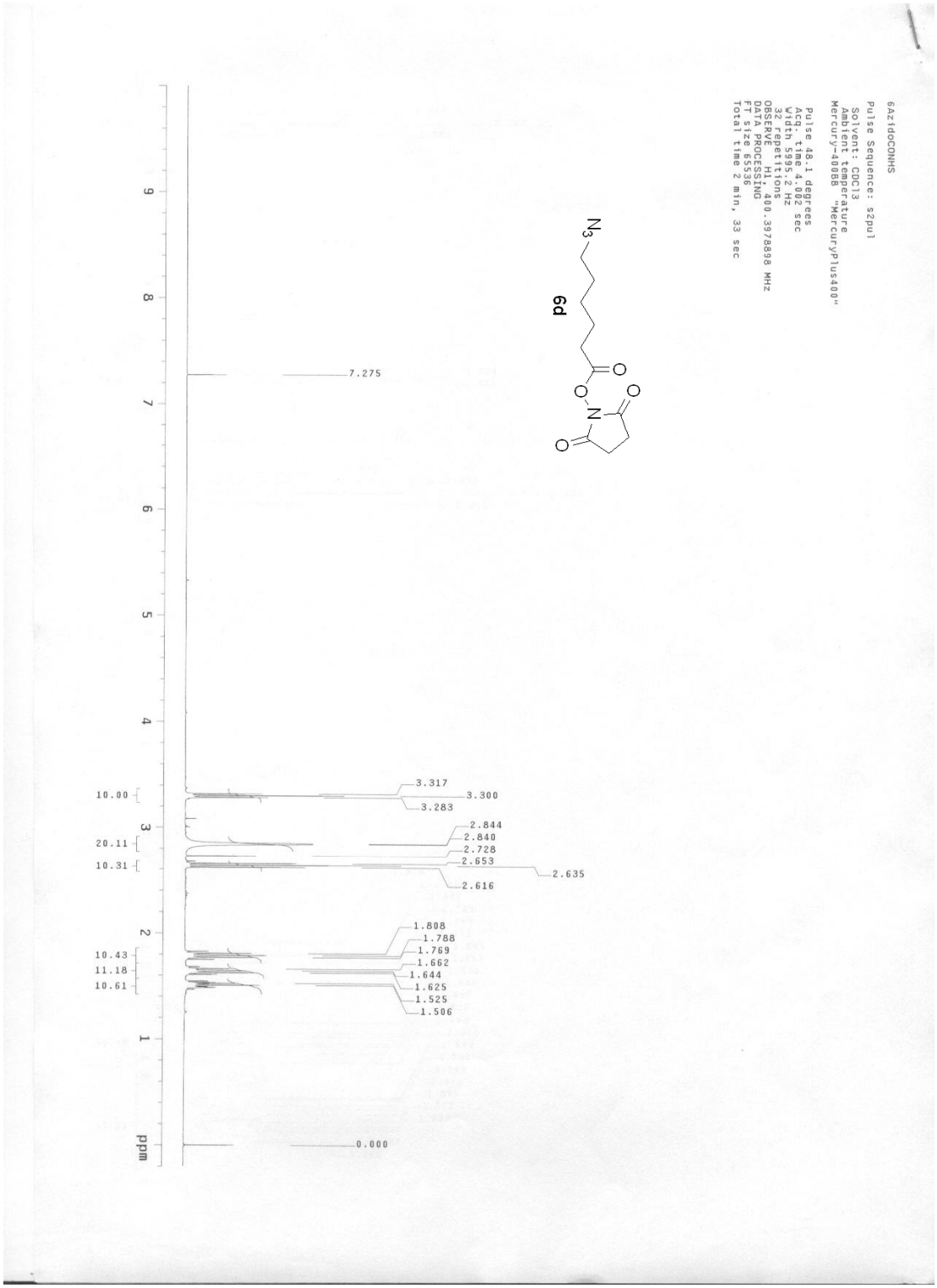
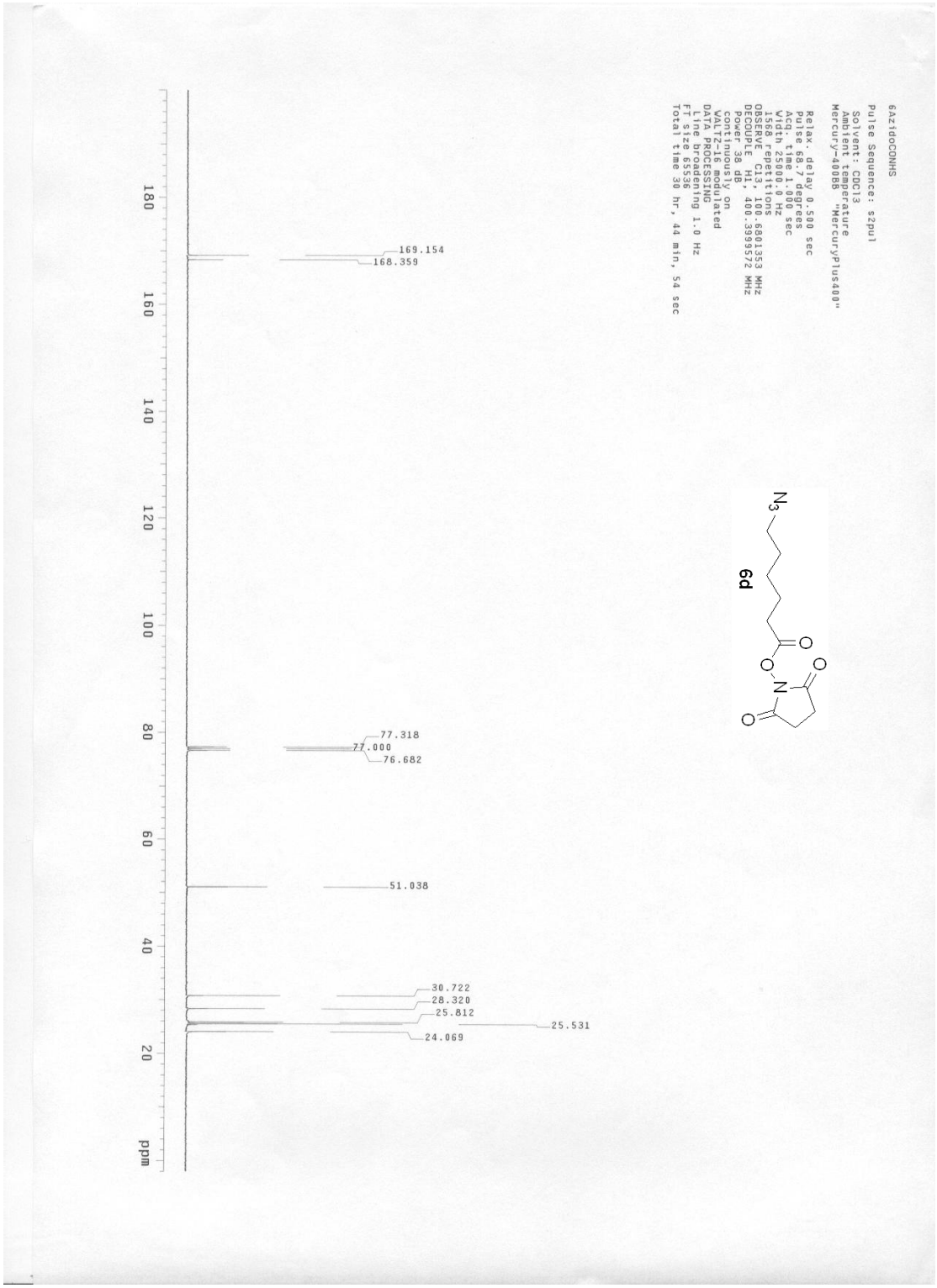


Fig. S6 Associations of the CuAAC-prepared POC with human A549 cells analyzed by flow cytometry. A549 cells were incubated with 5 μM of the FITC-labeled inoculates (the Tat peptide, the 3’ primer DNA and the CuAAC-prepared POC, respectively) for 24 h and followed by analysis of flow cytometry. Cells were washed and harvested with trypsin/EDTA before the analysis.

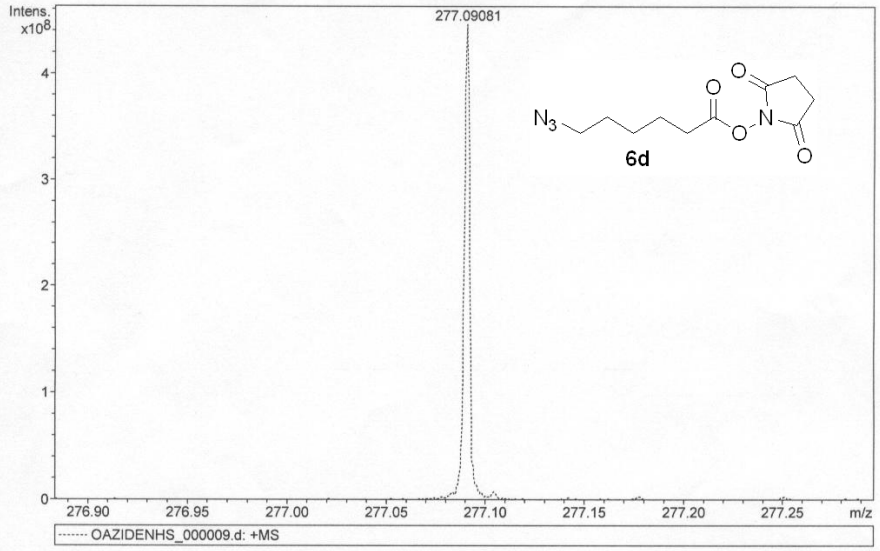

6Az1dOCONHS
Pulse Sequence: s2pul
Solvent: CDCl3
Ambient temperature
Mercury-400BB "MercuryPlus400"
Pulse 48.1 degrees
Acq. time 4.002 sec
Width 5995.2 Hz
32 repetitions
OBSERVE H1, 400.3978898 MHz
DATA PROCESSING
FT size 65536
Total time 2 min, 33 sec
6d
7.275
3.317
3.300
3.283
2.844
2.840
2.728
2.653
2.635
2.616
1.808
1.788
1.769
1.662
1.644
1.625
1.525
1.506
0.000
10.00
20.11
10.31
10.43
11.18
10.61
ppm


6Az1dOCONHS
Pulse Sequence: s2pul
Solvent: CDCl3
Ambient temperature
MercuryPlus400
Relax. delay 0.500 sec
Pulse 68.7 degrees
Acq. time 1.000 sec
Width 15060.0 Hz
1568 repetitions
OBSERVE C13, 100.6801353 MHz
DECOUPLE H1, 400.3999572 MHz
Power 38 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
Total time 30 hr, 44 min, 54 sec
169.154
168.359
77.318
77.000
76.682
51.038
30.722
28.320
25.812
25.531
24.069
180
160
140
120
100
80
60
40
20
ppm
6d
N3
O
O
O
N
O
O

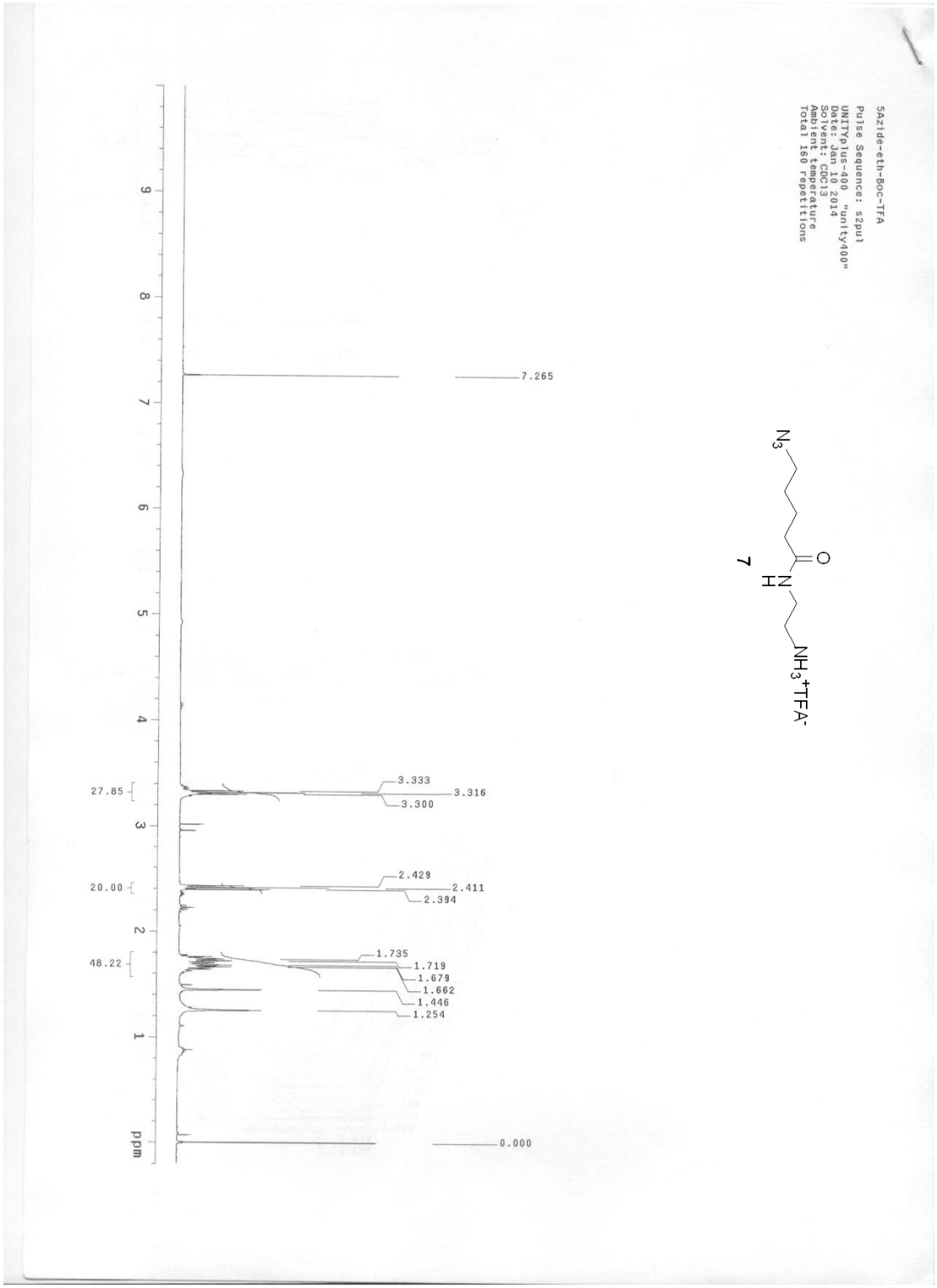
Mass Spectrum SmartFormula Report

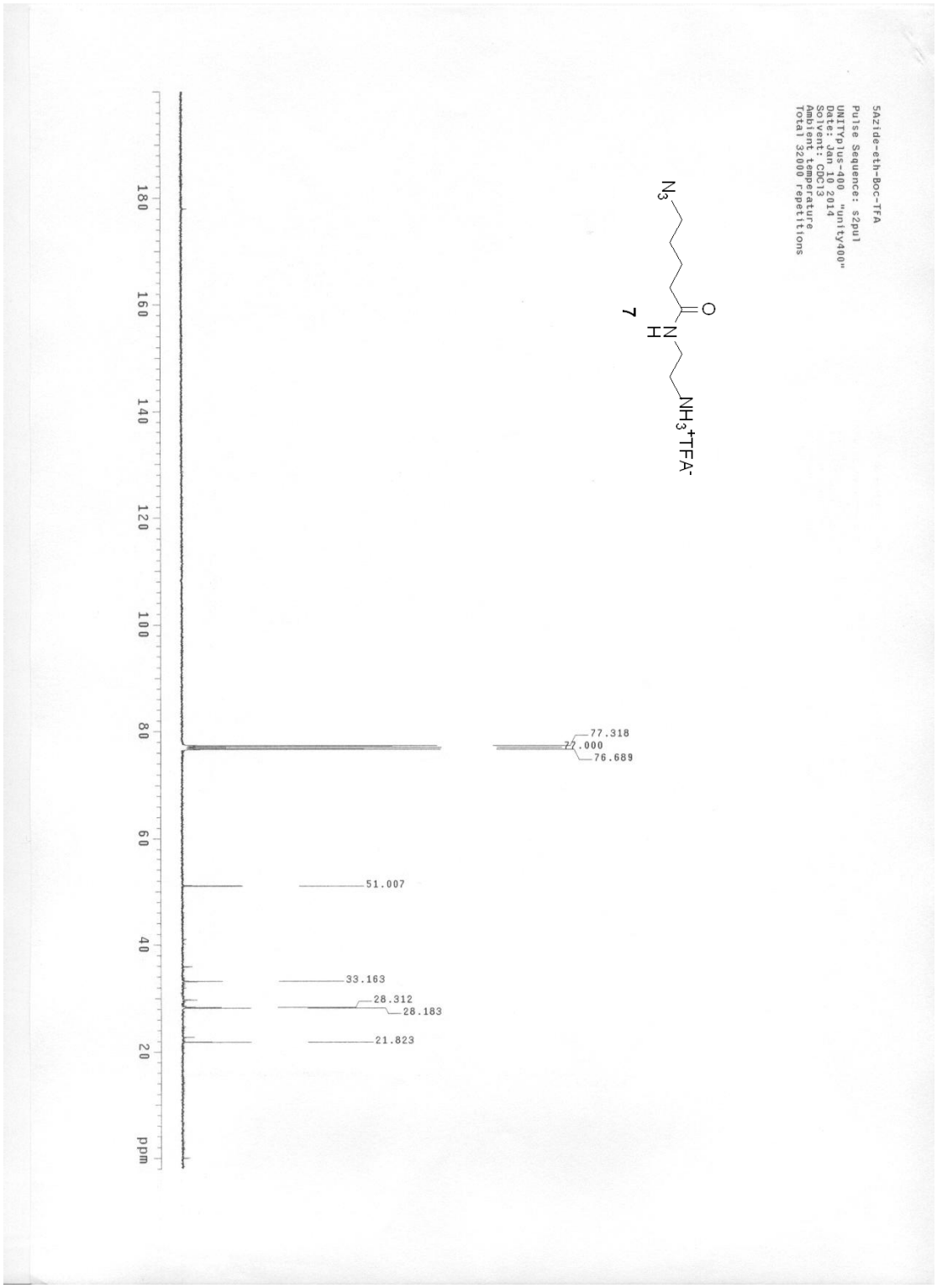
Analysis Info

Analysis Name	D:\Data\c5\OAZIDENHS_000009.d	1/13/2014 4:18:14 PM
Method	broadband first signal	
Sample Name	OAzide-NHS	Instrument: FT-MS solariX
Comment	ESI Positive	



Meas. m/z	#	Formula	Score	m/z	err [mDa]	err [ppm]	mSigma	rdb	e⁻ Conf	N-Rule
277.09081	1	C 10 H 14 N 4 Na O 4	100.00	277.09073	-0.09	-0.31	18.2	5.5	even	ok

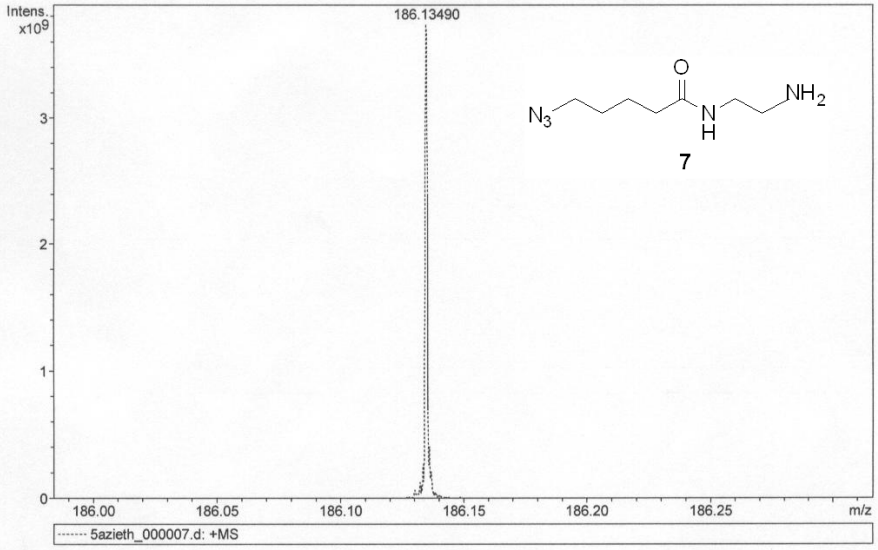

5Azide-eth-Boc-TFA
Pulse Sequence: s2pul
UNITYplus-400 "unity400"
Date: Jan 10 2014
Solvent: CDCl3
Ambient temperature
Total 160 repetitions
7.265
3.333
3.316
3.300
2.429
2.411
2.394
1.735
1.719
1.679
1.662
1.446
1.254
0.000
27.85
20.00
48.22
ppm
N3
O
N
H
NH3+TFA-
7


5Azide-eth-Boc-TFA
Pulse Sequence: s2pul
UNITYplus-400 "unity400"
Date: Jan 10 2014
Solvent: CDCl3
Ambient temperature
Total 32000 repetitions
7
77.318
77.000
76.689
51.007
33.163
28.312
28.183
21.823
ppm

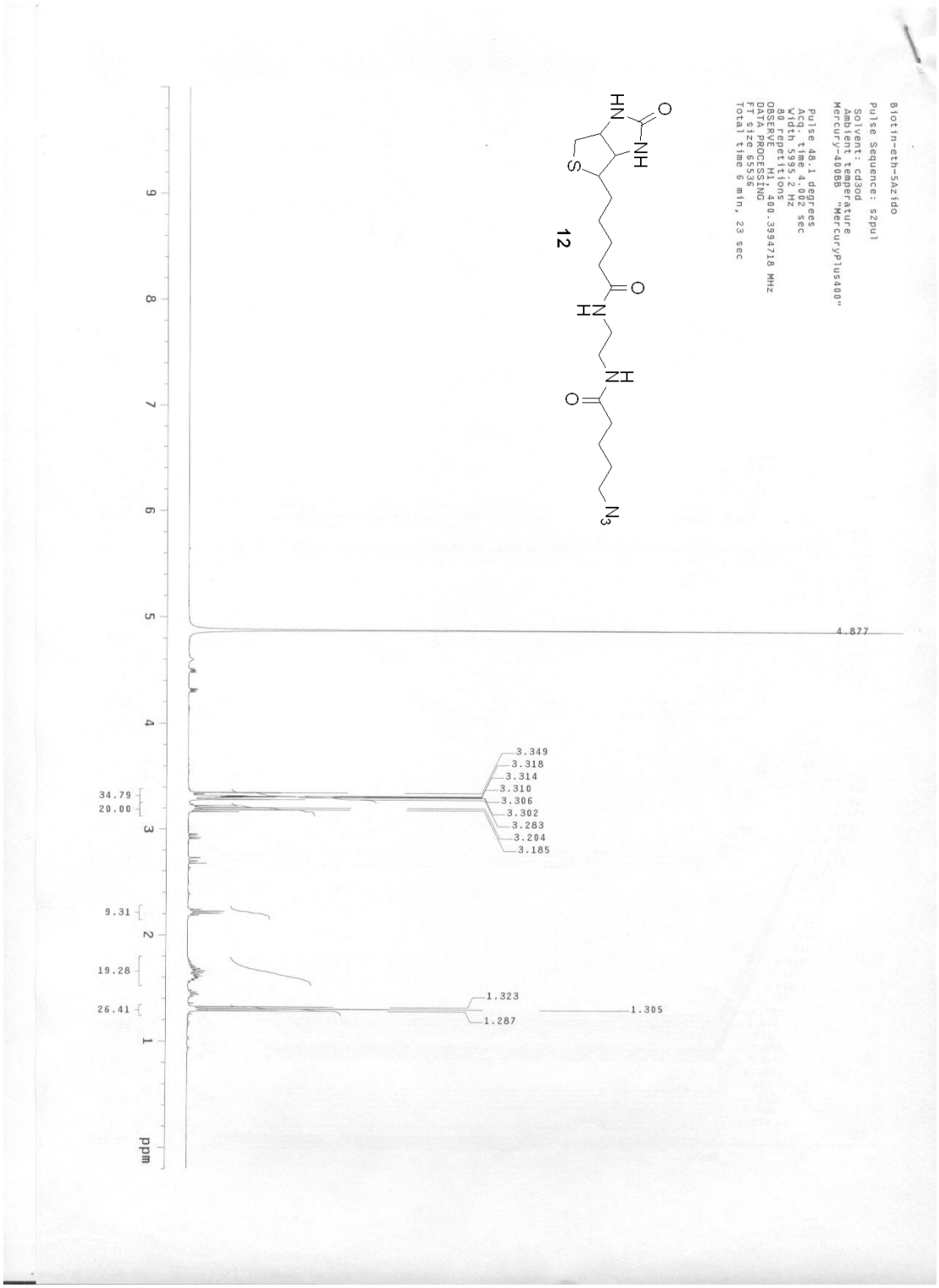
Mass Spectrum SmartFormula Report

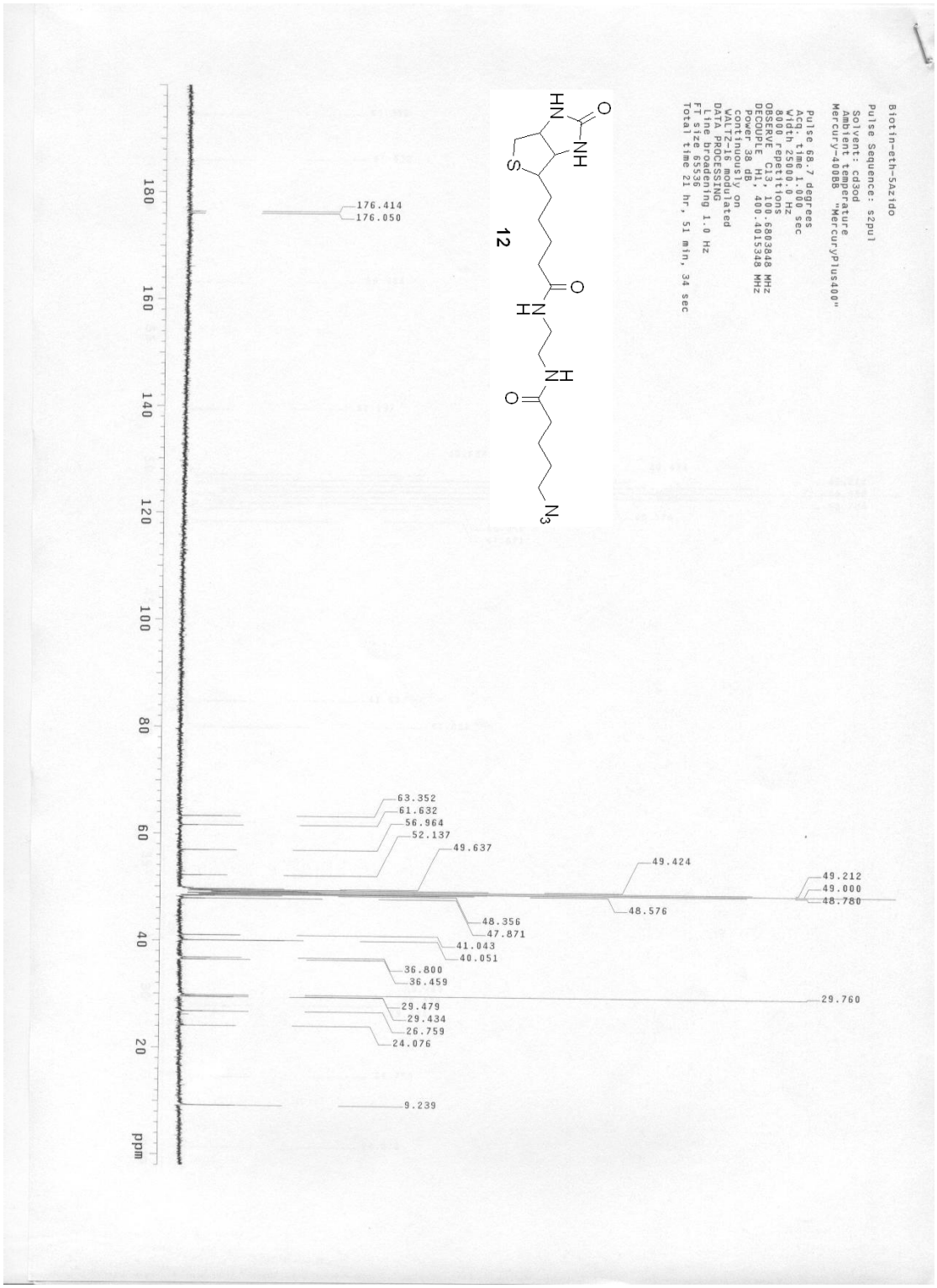
Analysis Info

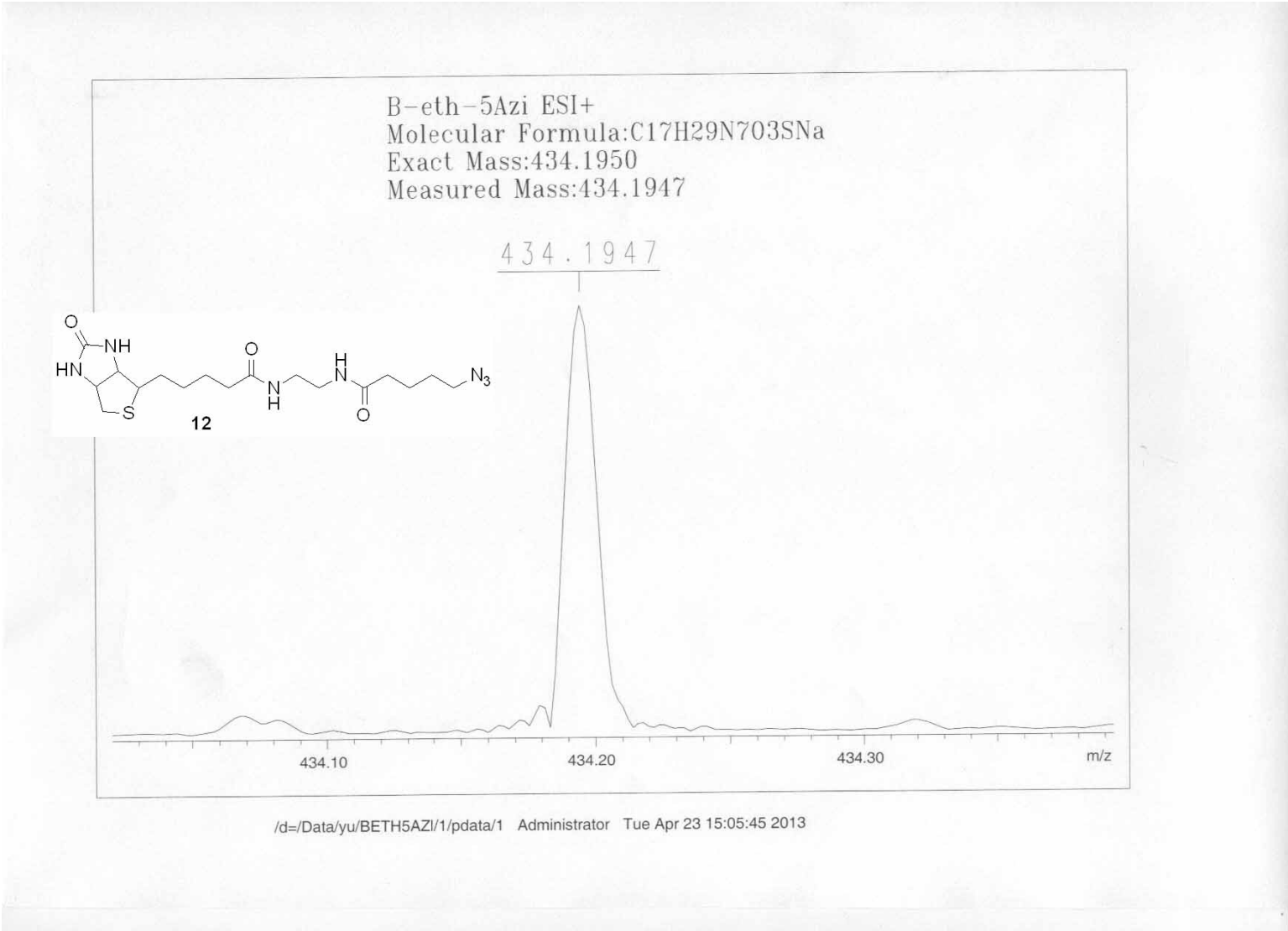
Analysis Name	D:\Data\c5\5azieth_000007.d		12/27/2013 2:42:32 PM
Method	broadband first signal		
Sample Name	5Azi-eth	Instrument:	FT-MS solariX
Comment	ESI Positive		

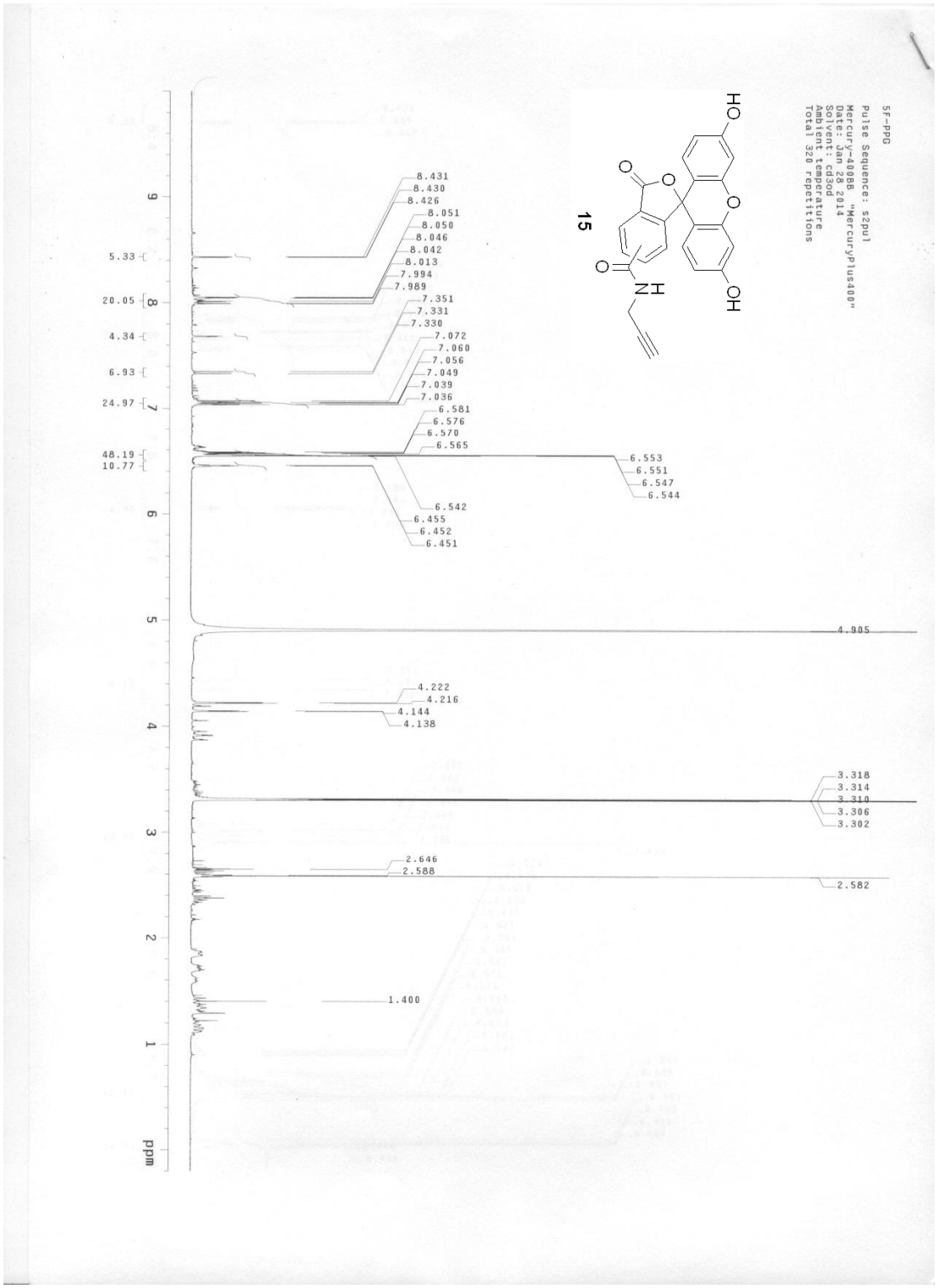


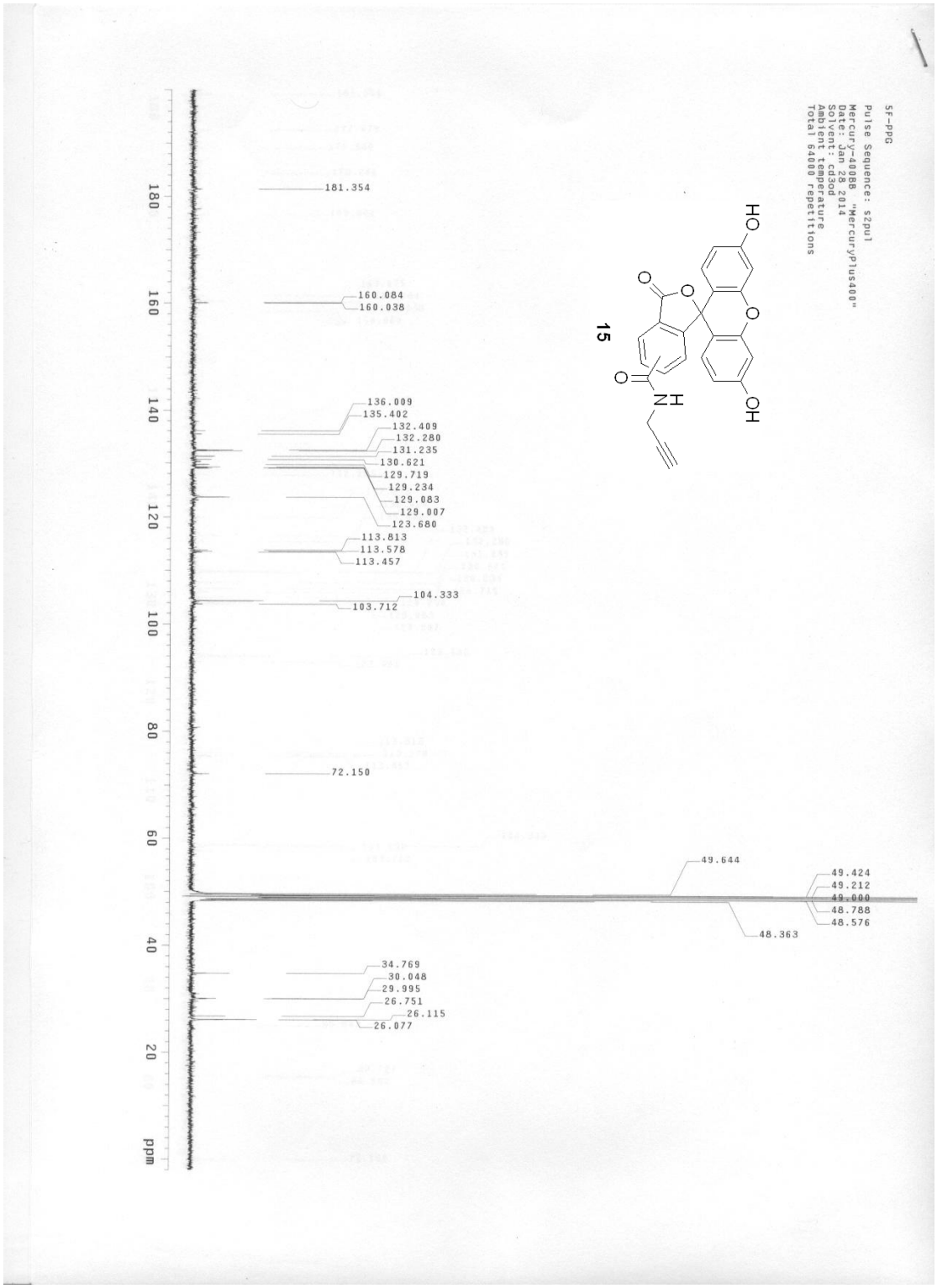
Meas. m/z	#	Formula	Score	m/z	err [mDa]	err [ppm]	mSigma	rdb	e⁻ Conf	N-Rule
186.13490	1	C 7 H 16 N 5 O	100.00	186.13494	0.04	0.22	14.9	2.5	even	ok


Biotin-eth-5Azido
Pulse Sequence: s2pul
Solvent: cd3od
Ambient temperature
Mercury-400BB "MercuryPlus400"
Pulse 48.1 degrees
Acq. time 4.002 sec
Width 5482.2 Hz
80 repetitions
OBSERVE H1, 400.3994718 MHz
DATA PROCESSING
FT size 65536
Total time 6 min, 23 sec
12
4.877
3.349
3.318
3.314
3.310
3.306
3.302
3.283
3.204
3.185
1.323
1.305
1.287
34.79
20.00
9.31
19.28
26.41
ppm

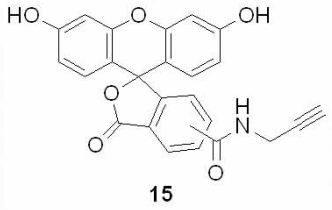

Biotin-eth-5Azido
Pulse Sequence: s2pul
Solvent: cd3od
Ambient temperature
Mercury-400BB "MercuryPlus400"
Pulse 68.7 degrees
Acq. time 1.000 sec
Width 25000.0 Hz
8000 repetitions
OBSERVE C13, 100.6803048 MHz
DECOUPLE H1, 400.4015348 MHz
Power 38 dB
continuously on
WALTZ-16 modulated
DATA PROCESSING
Line broadening 1.0 Hz
FT size 65536
Total time 21 hr, 51 min, 34 sec
12
176.414
176.050
63.352
61.632
56.964
52.137
49.637
49.424
49.212
49.000
48.780
48.576
48.356
47.871
41.043
40.051
36.800
36.459
29.760
29.479
29.434
26.759
24.076
9.239
ppm


B-eth-5Azi ESI+
Molecular Formula:C17H29N703SNa
Exact Mass:434.1950
Measured Mass:434.1947
434.1947
O
NH
HN
S
O
N
H
H
N
O
N3
12
434.10
434.20
434.30
m/z
/d=/Data/yu/BETH5AZI/1/pdata/1 Administrator Tue Apr 23 15:05:45 2013


5F-PPG
Pulse Sequence: s2pul
Mercury-400BB "MercuryPlus400"
Date: Jan 28 2014
Solvent: cd3od
Ambient temperature
Total 320 repetitions
15
HO
O
O
O
O
H
N
OH
8.431
8.430
8.426
8.051
8.050
8.046
8.042
8.013
7.994
7.989
7.351
7.331
7.330
7.072
7.060
7.056
7.049
7.039
7.036
6.581
6.576
6.570
6.565
6.553
6.551
6.547
6.544
6.542
6.455
6.452
6.451
4.905
4.222
4.216
4.144
4.138
3.318
3.314
3.310
3.306
3.302
2.646
2.588
2.582
1.400
5.33
20.05
4.34
6.93
24.97
48.19
10.77
ppm

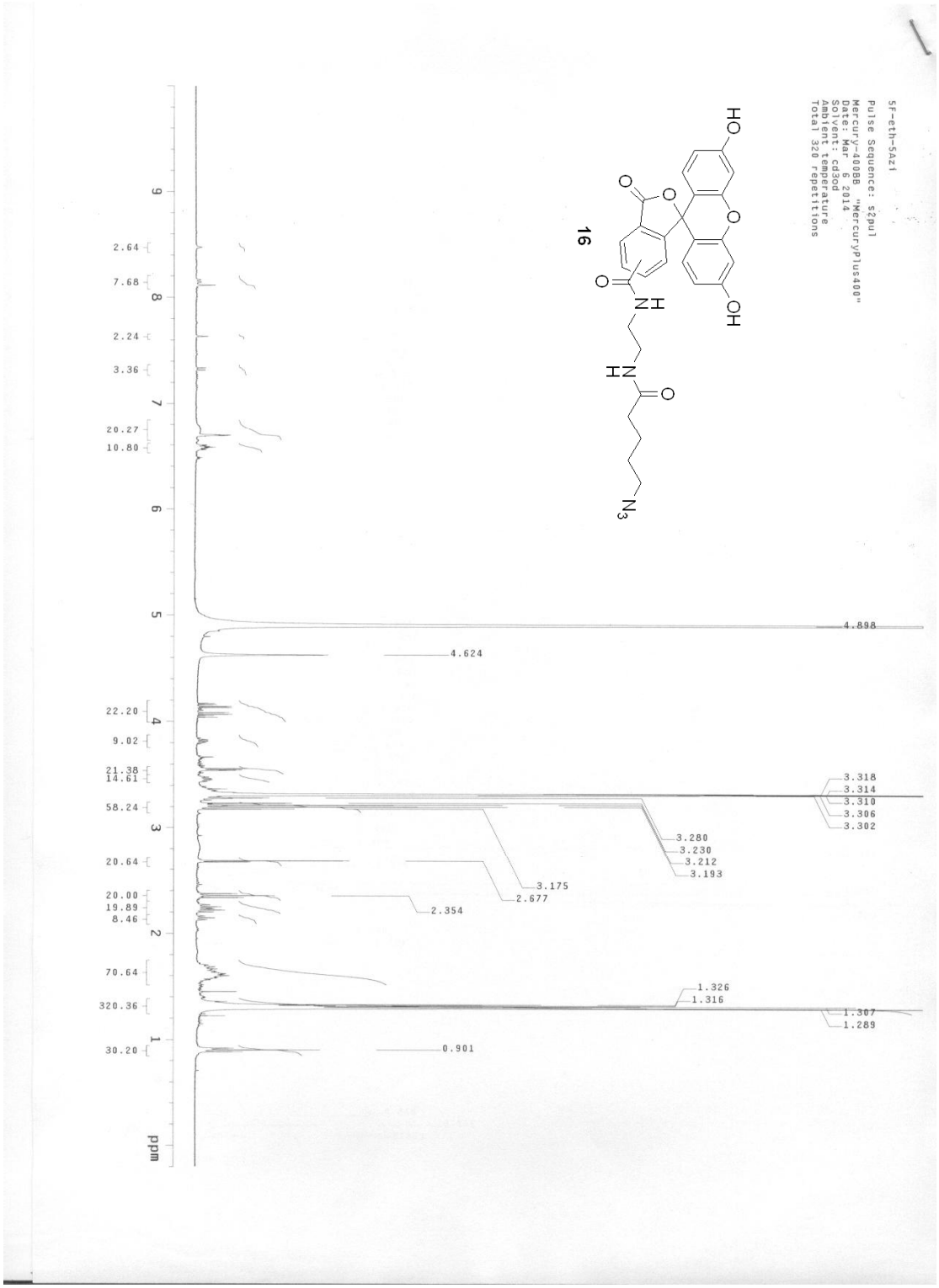

5F-PPG
Pulse Sequence: s2pul
Mercury-400BB "MercuryPlus400"
Date: Jan 28 2014
Solvent: cd3od
Ambient temperature
Total 6400 repetitions
15
HO
OH
O
NH
181.354
160.084
160.038
136.009
135.402
132.409
132.280
131.235
130.621
129.719
129.234
129.083
129.007
123.680
113.813
113.578
113.457
104.333
103.712
72.150
49.644
49.424
49.212
49.000
48.788
48.576
48.363
34.769
30.048
29.995
26.751
26.115
26.077
180
160
140
120
100
80
60
40
20
ppm

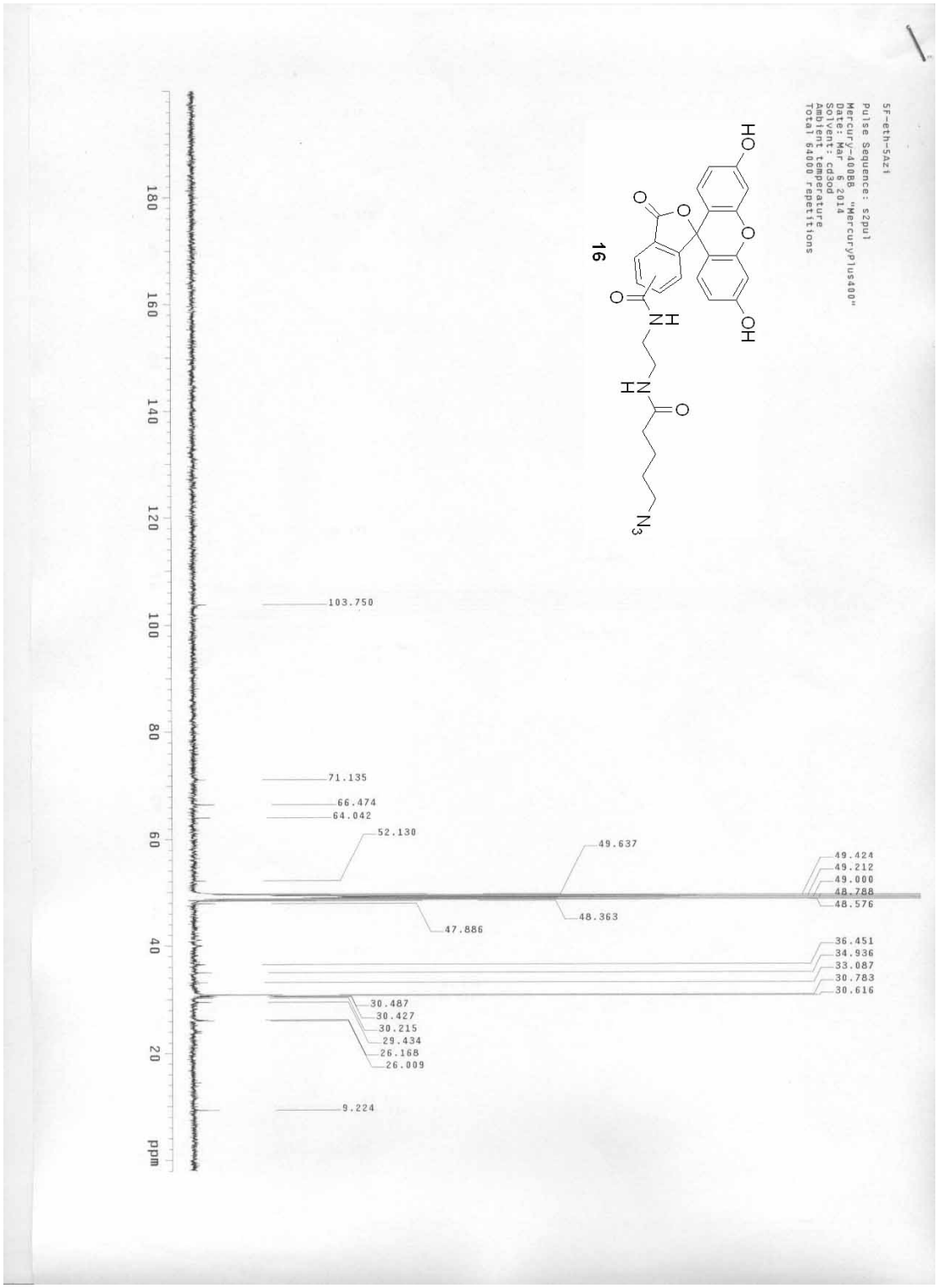
5F-PPG ESI+
Molecular Formula:C24H15N06Na
Exact Mass:436.0797
Measured Mass:436.0795



15

/d=/Data/yu/5fppg/1/pdata/1 Administrator Tue Apr 23 15:26:07 2013

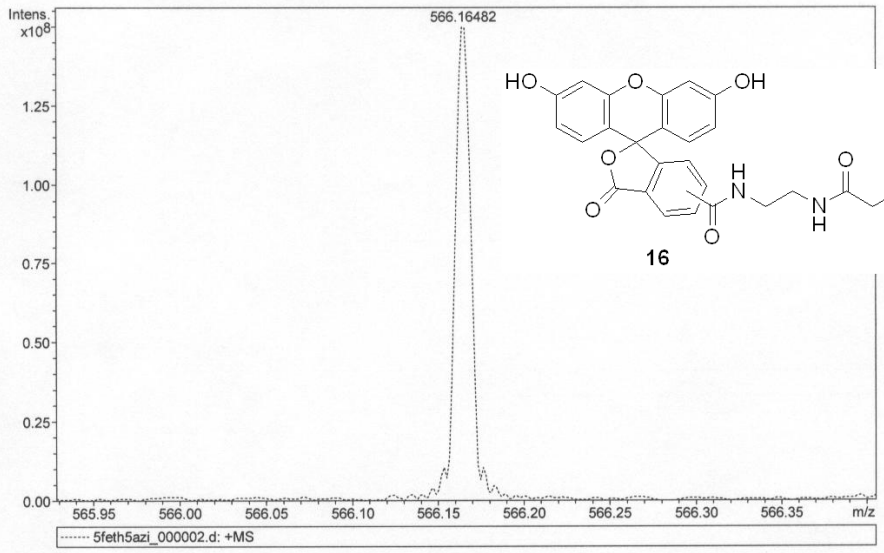

5F-eth-5Az1
Pulse Sequence: s2pul
Mercury-400BB "MercuryPlus400"
Date: Mar 6 2014
Solvent: cd3od
Ambient temperature
Total 320 repetitions
16
HO
O
O
OH
O
NH
HN
O
N3
4.898
4.624
3.318
3.314
3.310
3.306
3.302
3.280
3.230
3.212
3.193
3.175
2.677
2.354
1.326
1.316
1.307
1.289
0.901
2.64
7.68
2.24
3.36
20.27
10.80
22.20
9.02
21.38
14.61
58.24
20.64
20.00
19.89
8.46
70.64
320.36
30.20
ppm


5F-eth-5Azi
Pulse Sequence: s2pul
Mercury-400BB "MercuryPlus400"
Date: Mar 6 2014
Solvent: cd3od
Ambient temperature
Total 64000 repetitions
16
HO
OH
O
N
H
N3
103.750
71.135
66.474
64.042
52.130
49.637
49.424
49.212
49.000
48.788
48.576
48.363
47.886
36.451
34.936
33.087
30.783
30.616
30.487
30.427
30.215
29.434
26.168
26.009
9.224
ppm

Mass Spectrum SmartFormula Report

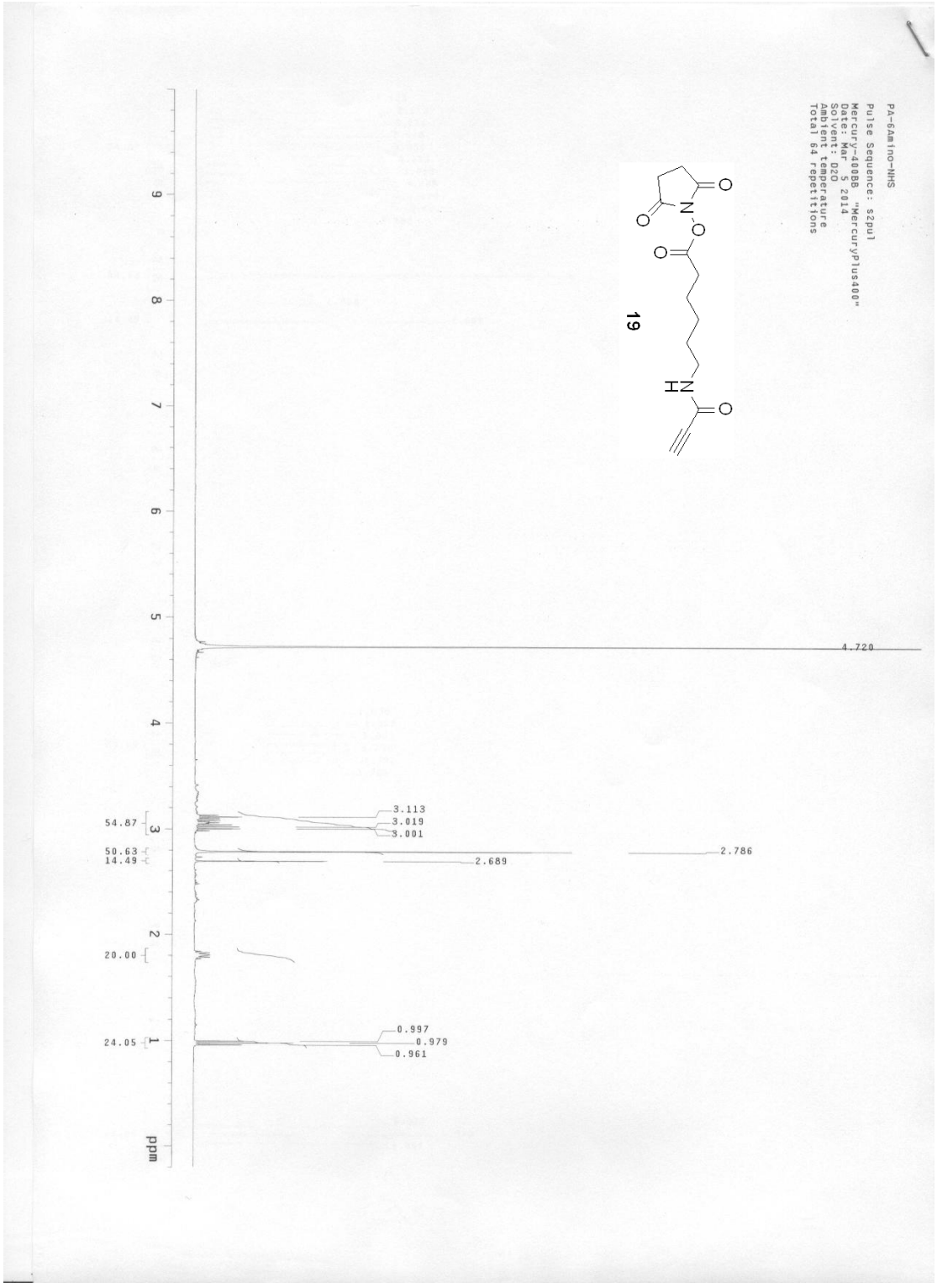
Analysis Info

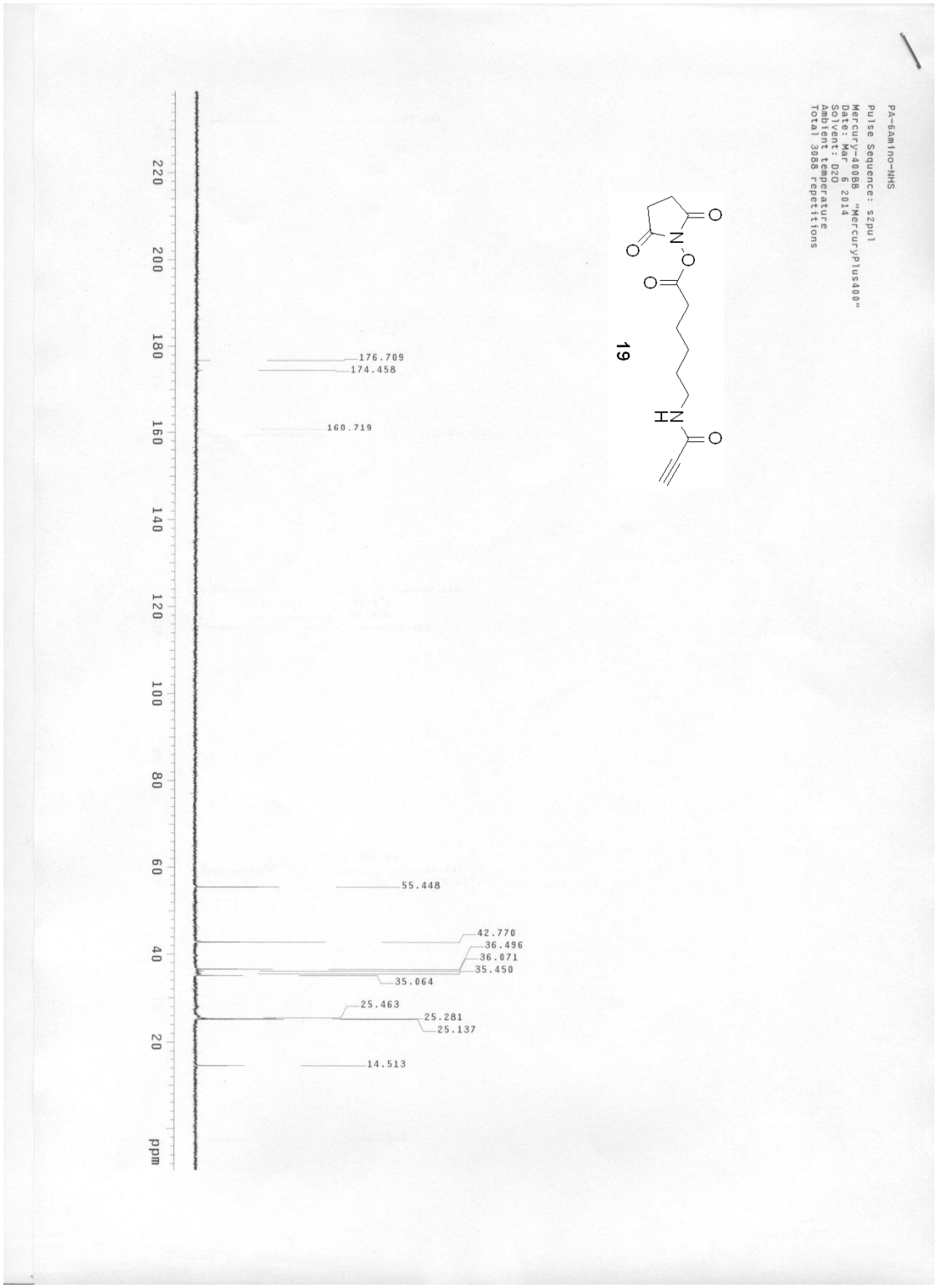
Analysis Name	D:\Data\c5\5feth5azi_000002.d	12/27/2013 3:26:30 PM	
Method	broadband first signal		
Sample Name	5F-eth-5Azi	Instrument: FT-MS	solariX
Comment	ESI Positive		



16

Meas. m/z	#	Formula	Score	m/z	err [mDa]	err [ppm]	mSigma	rdb	e⁻ Conf	N-Rule
566.16482	1	C 28 H 25 N 5 Na O 7	100.00	566.16462	-0.20	-0.35	11.6	18.5	even	ok

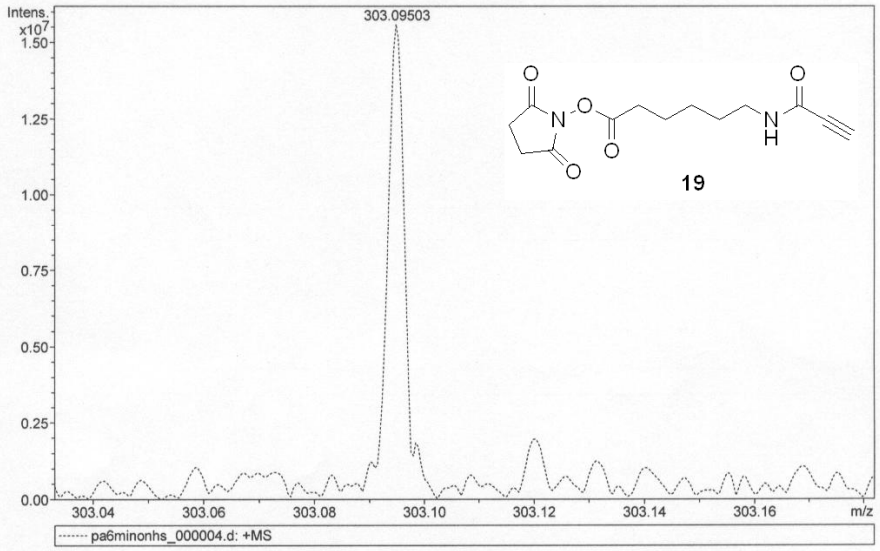

PA-6Amino-NHS
Pulse Sequence: s2pul
Mercury-400BB "MercuryPlus400"
Date: Mar 5 2014
Solvent: D2O
Ambient temperature
Total 64 repetitions
19
4.720
3.113
3.019
3.001
2.786
2.689
0.997
0.979
0.961
54.87
50.63
14.49
20.00
24.05
ppm


PA-6Amino-NHS
Pulse Sequence: s2pul
Mercury-400BB "MercuryPlus400"
Date: Mar 6 2014
Solvent: D2O
Ambient temperature
Total 3088 repetitions
19
176.709
174.458
160.719
55.448
42.770
36.496
36.071
35.450
35.064
25.463
25.281
25.137
14.513
220
200
180
160
140
120
100
80
60
40
20
ppm

Mass Spectrum SmartFormula Report

Analysis Info

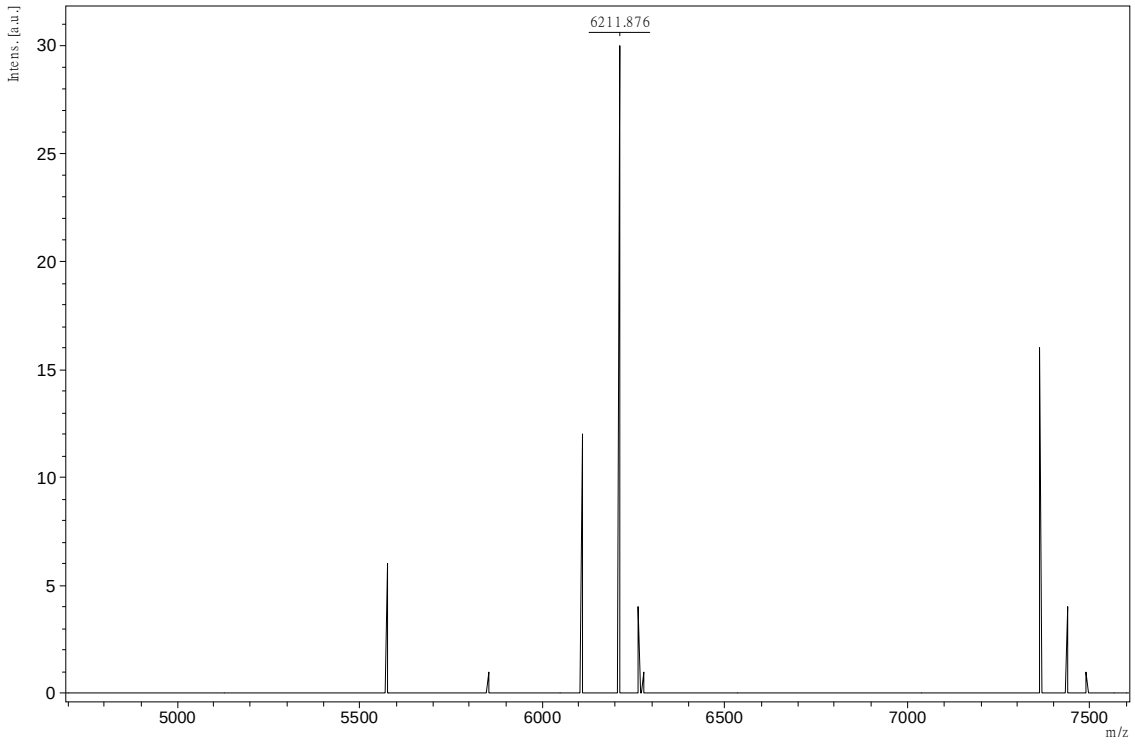
Analysis Name	D:\Data\c5\pa6minonhs_000004.d		12/27/2013 3:08:12 PM
Method	broadband first signal		
Sample Name	PA-6Amino NHS	Instrument:	FT-MS solariX
Comment	ESI Positive		



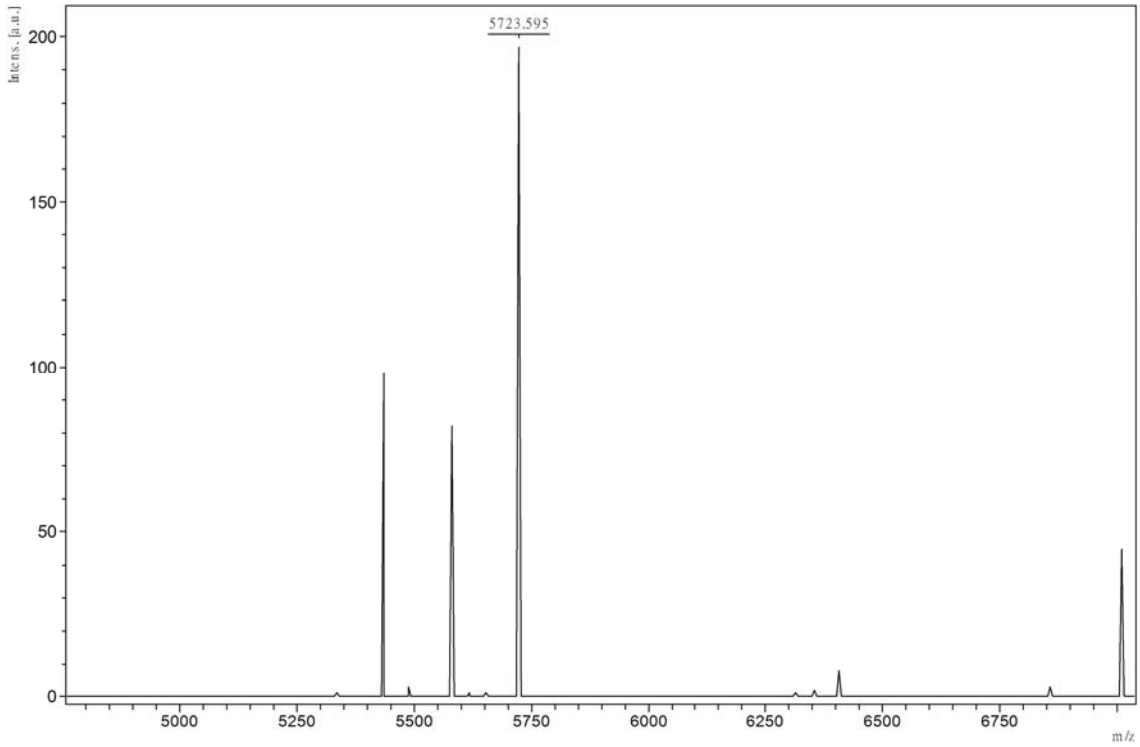
19

Meas. m/z	#	Formula	Score	m/z	err [mDa]	err [ppm]	mSigma	rdb	e⁻ Conf	N-Rule
303.09503	1	C 13 H 16 N 2 Na O 5	100.00	303.09514	0.11	0.36	187.8	6.5	even	ok

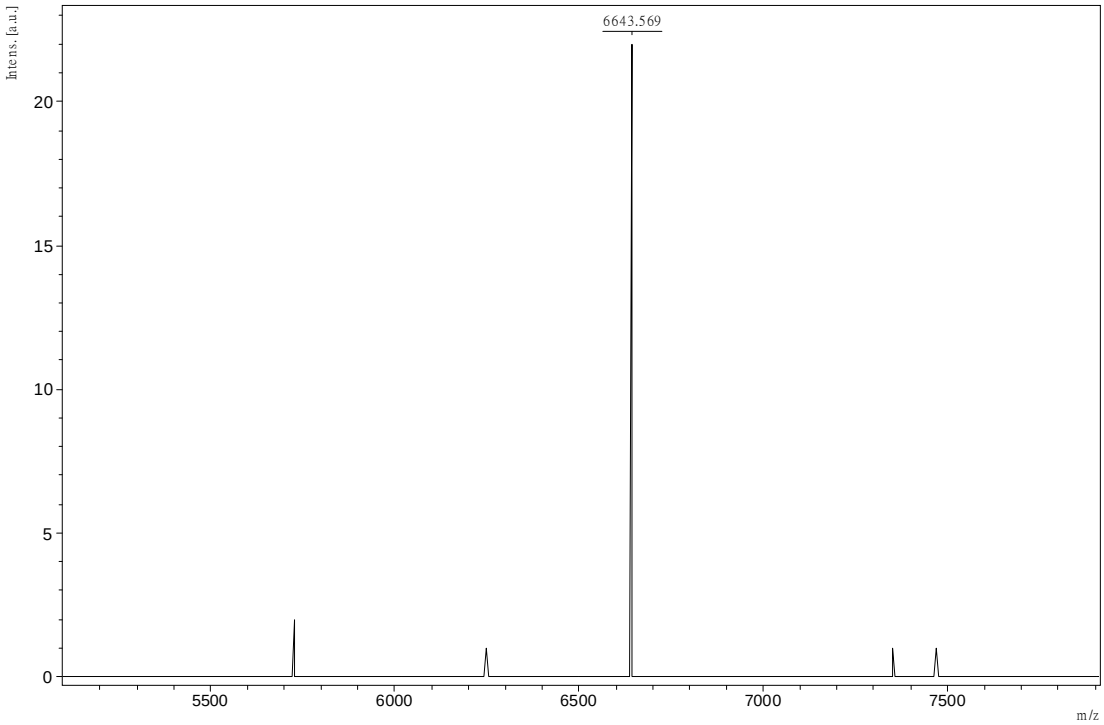
The 3’ primer DNA-ethylenediamine-**6d** conjugate.



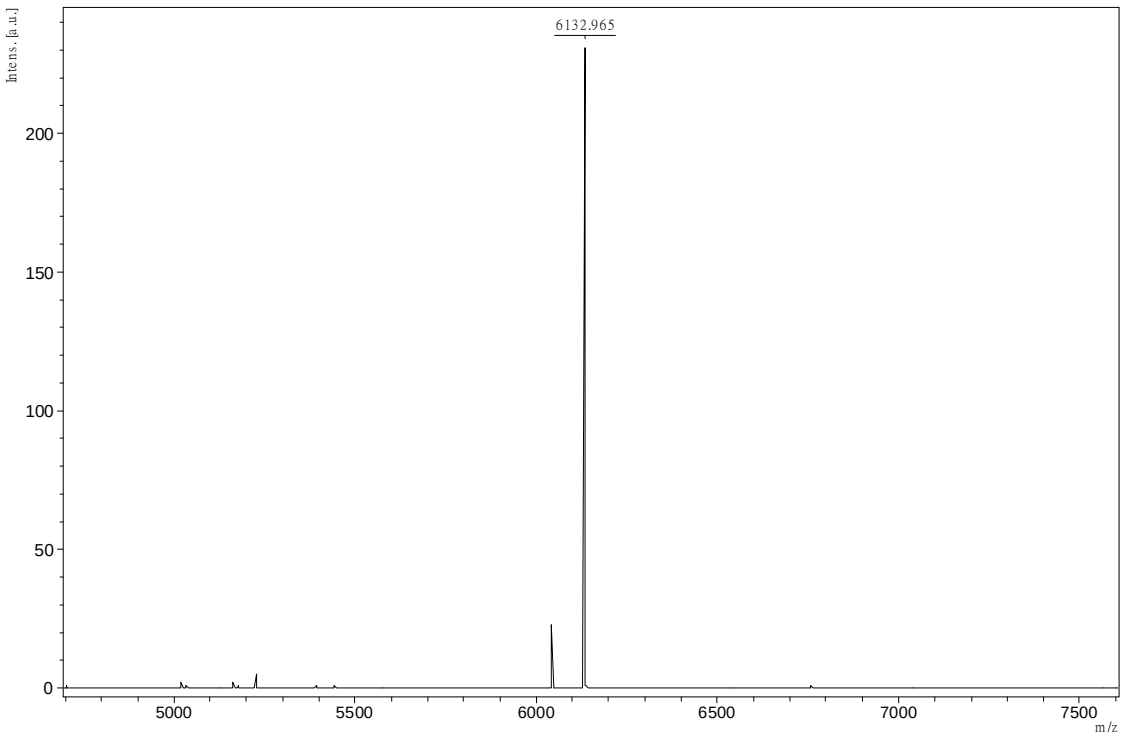
The 17-mer RNA-cystamine-**6d** conjugate.



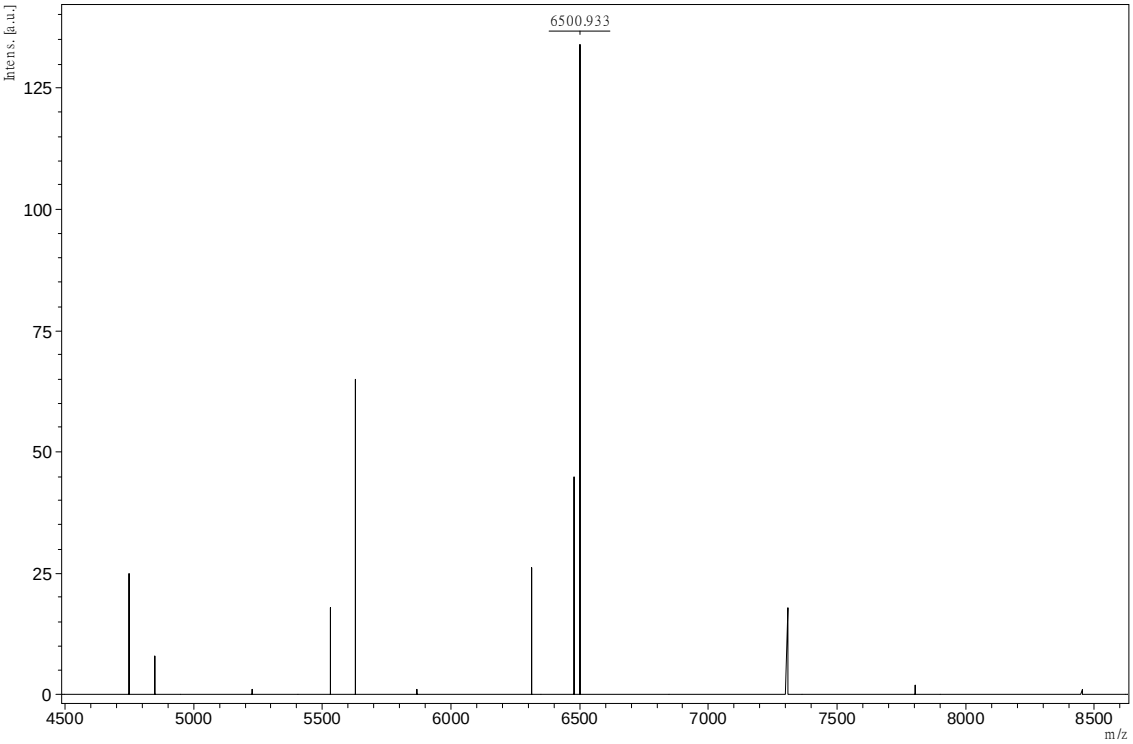
The CuAAC reaction product of the 3’ primer DNA-ethylenediamine-**6d** conjugate and **15**.



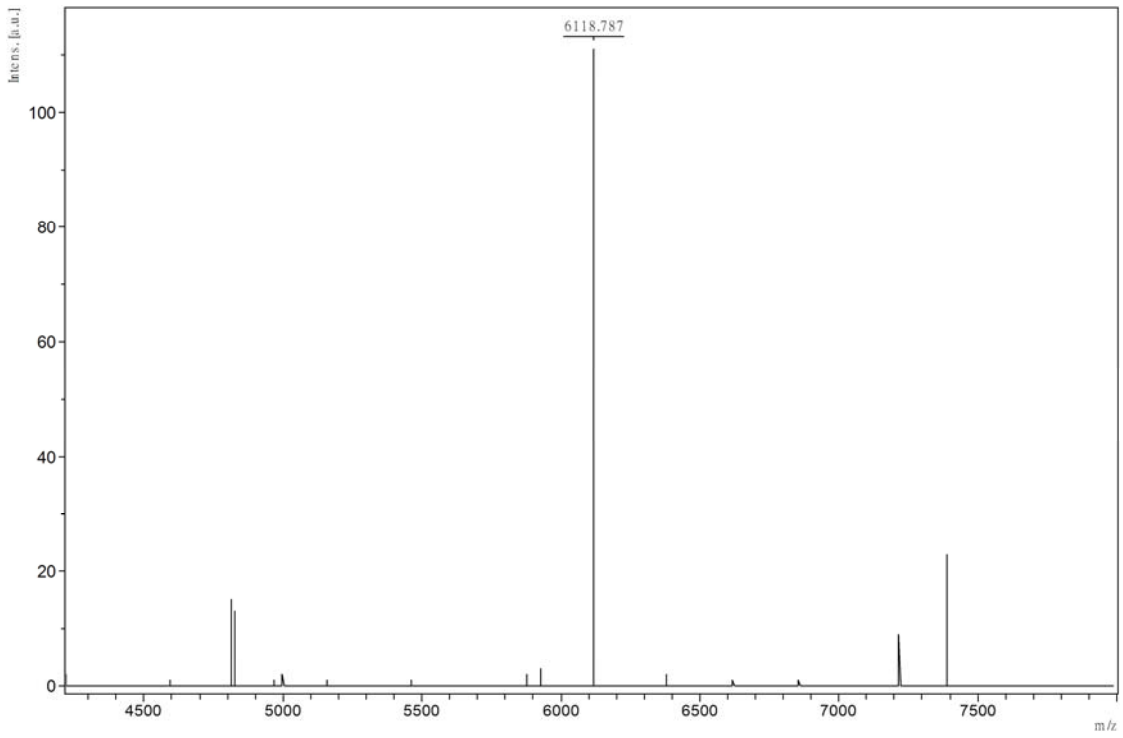
The CuAAC reaction product of the 17-mer RNA-cystamine-**6d** conjugate and **15**.



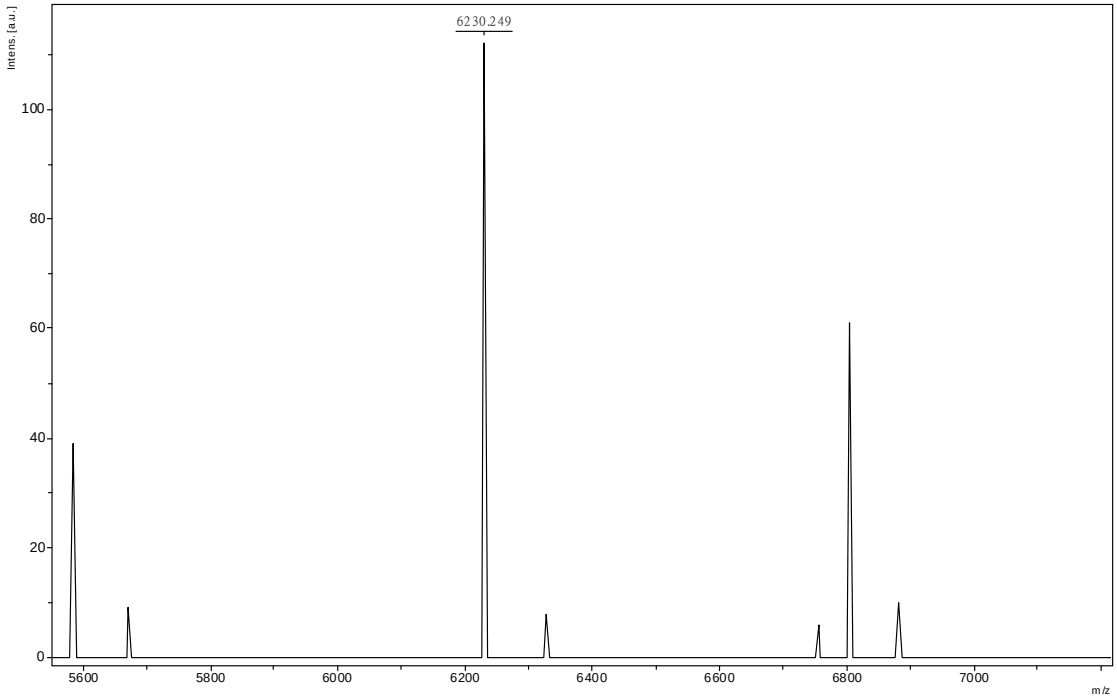
The CuAAC reaction product of the 3’ primer DNA-ethylenediamine-**6d** conjugate and **10**.



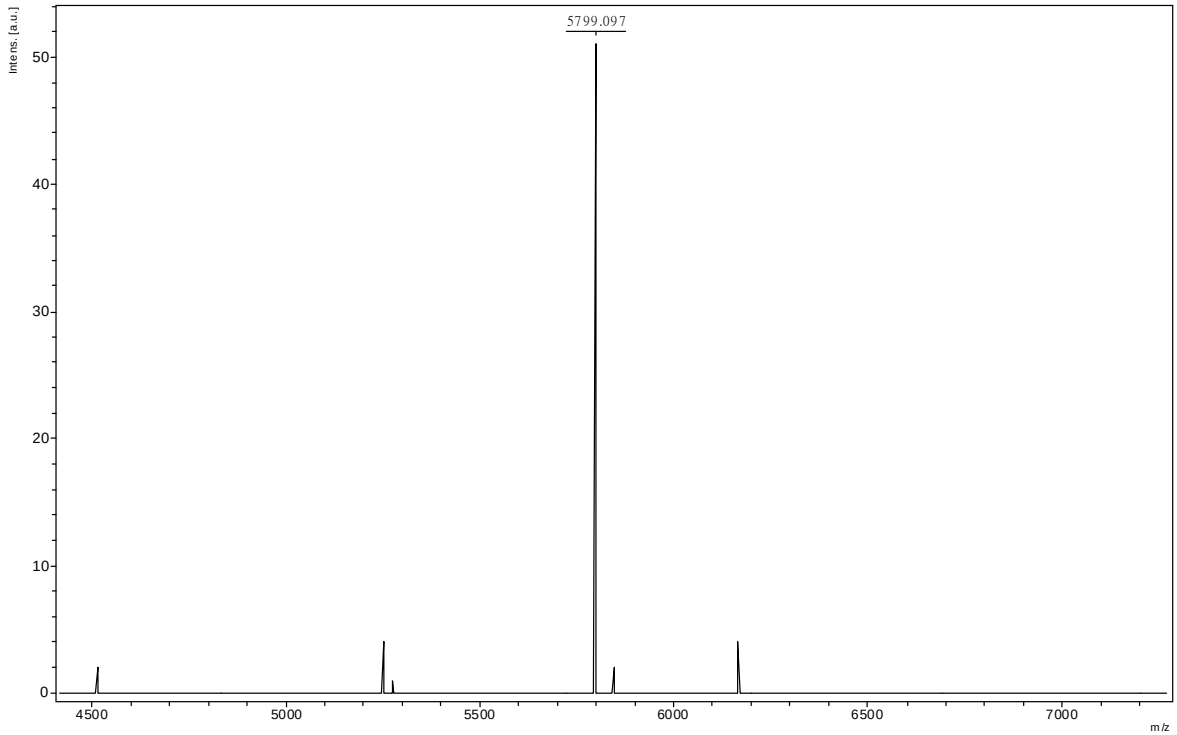
The CuAAC reaction product of the 17-mer RNA-cystamine-**6d** conjugate and **10**.



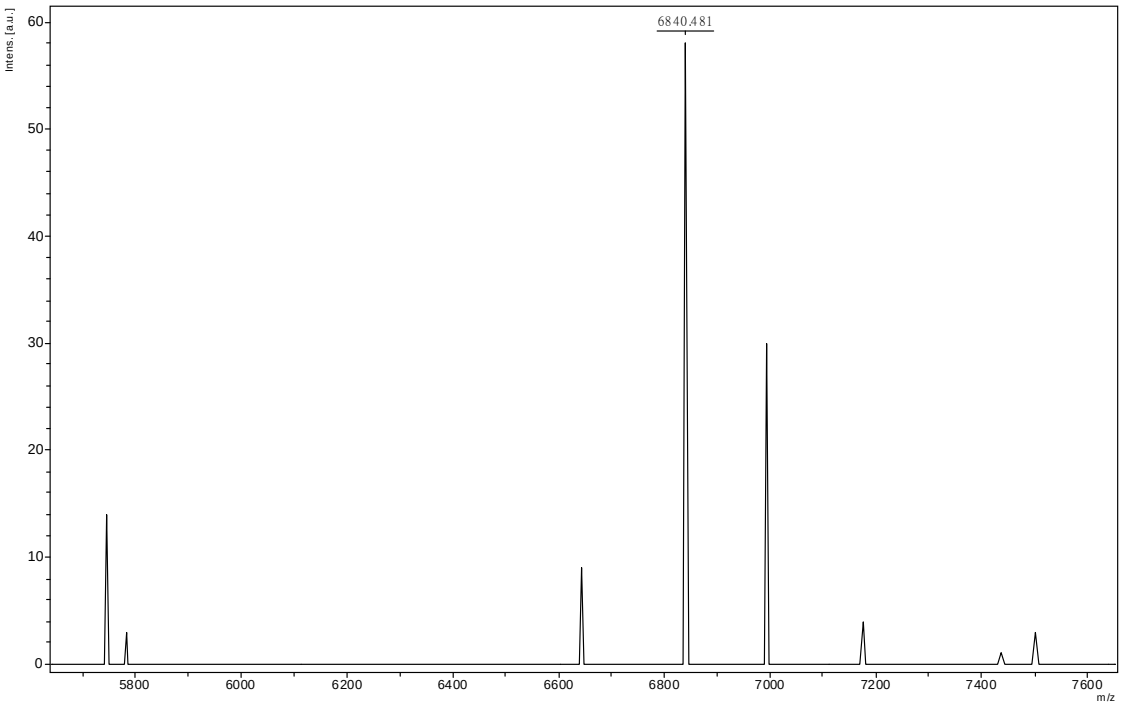
The 3’ primer DNA-ethylenediamine-**19** conjugate.



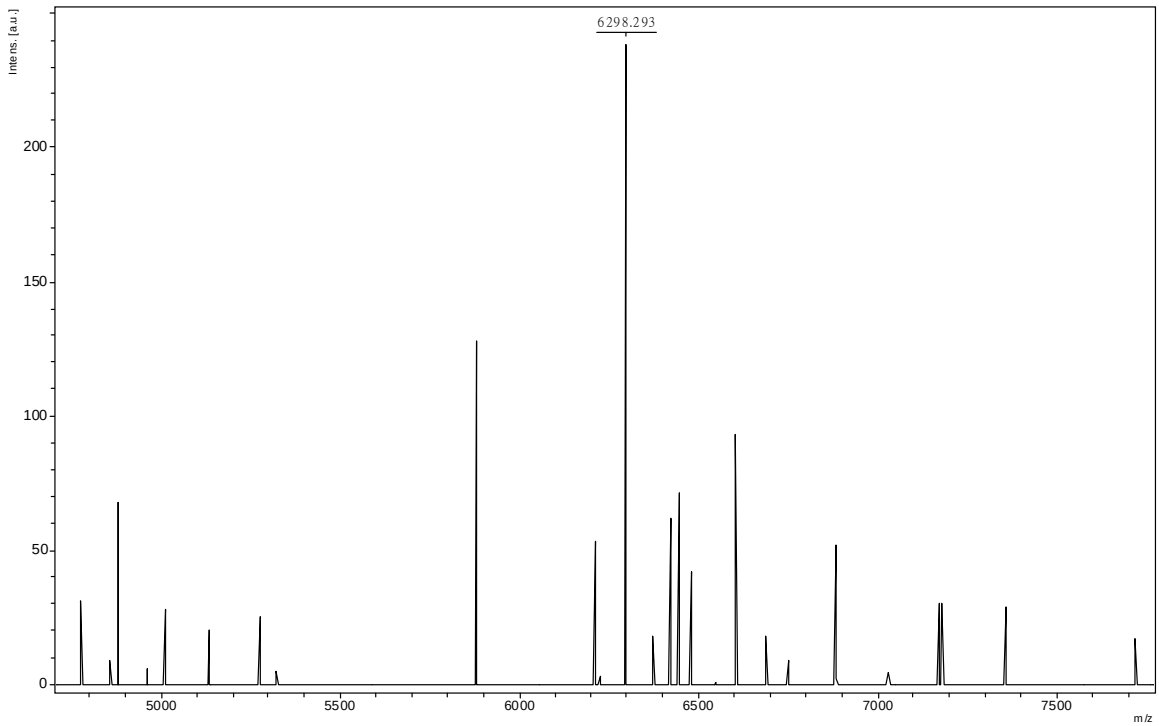
The 17-mer RNA-cystamine-**19** conjugate.



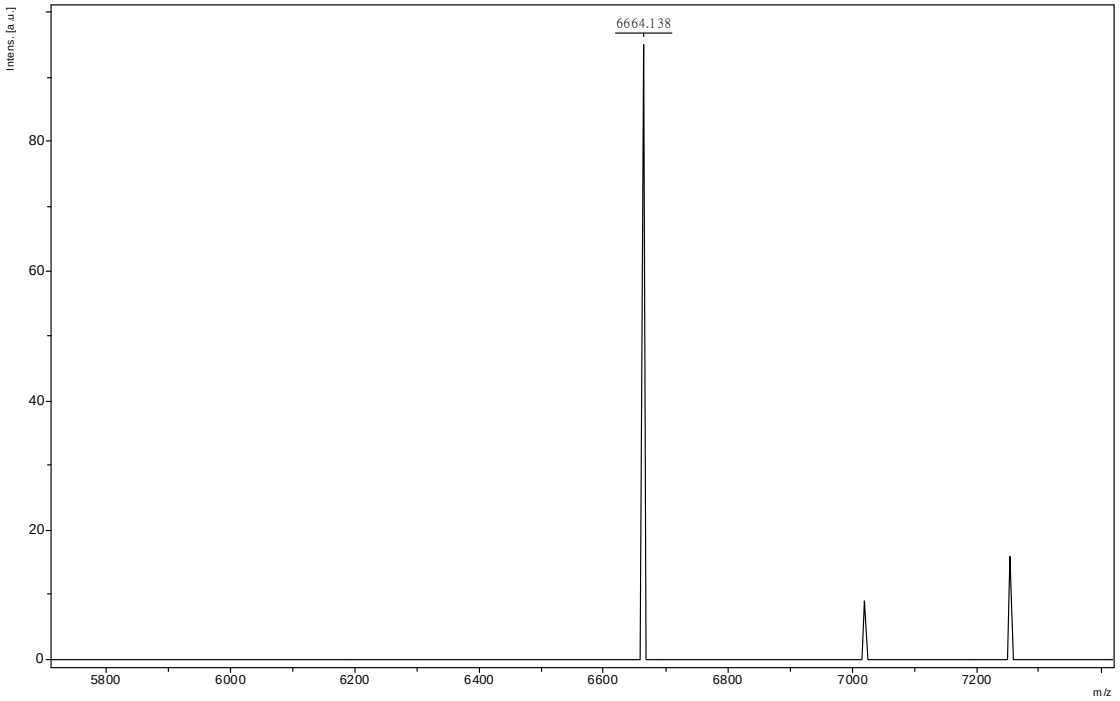
The CuAAC reaction product of the 3’ primer DNA-ethylenediamine-**19** conjugate and **16**.



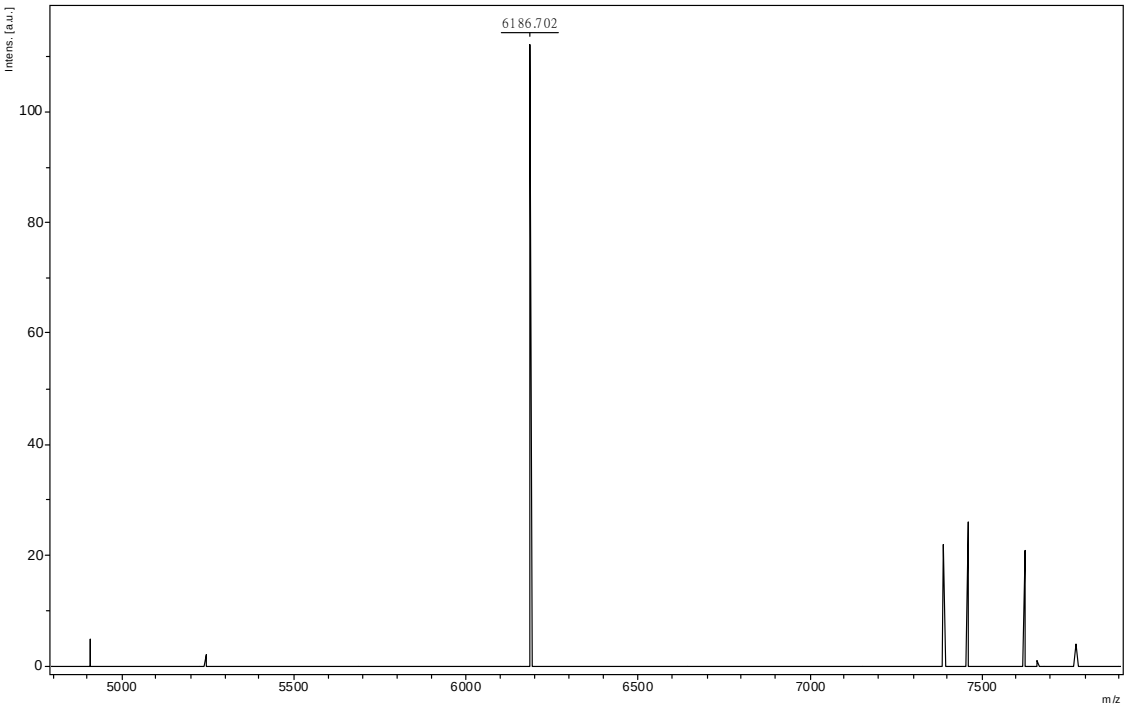
The CuAAC reaction product of the 17-mer RNA-cystamine-**19** conjugate and **16**.



The CuAAC reaction product of the 3’ primer DNA-ethylenediamine-**19** conjugate and **12**.



The CuAAC reaction product of the 17-mer RNA-cystamine-**19** conjugate and **12**.



The CuAAC reaction product of the 3’ primer DNA-ethylenediamine-**6d** conjugate and the alkynyl Tat peptide.