SUPPORTING INFORMATION

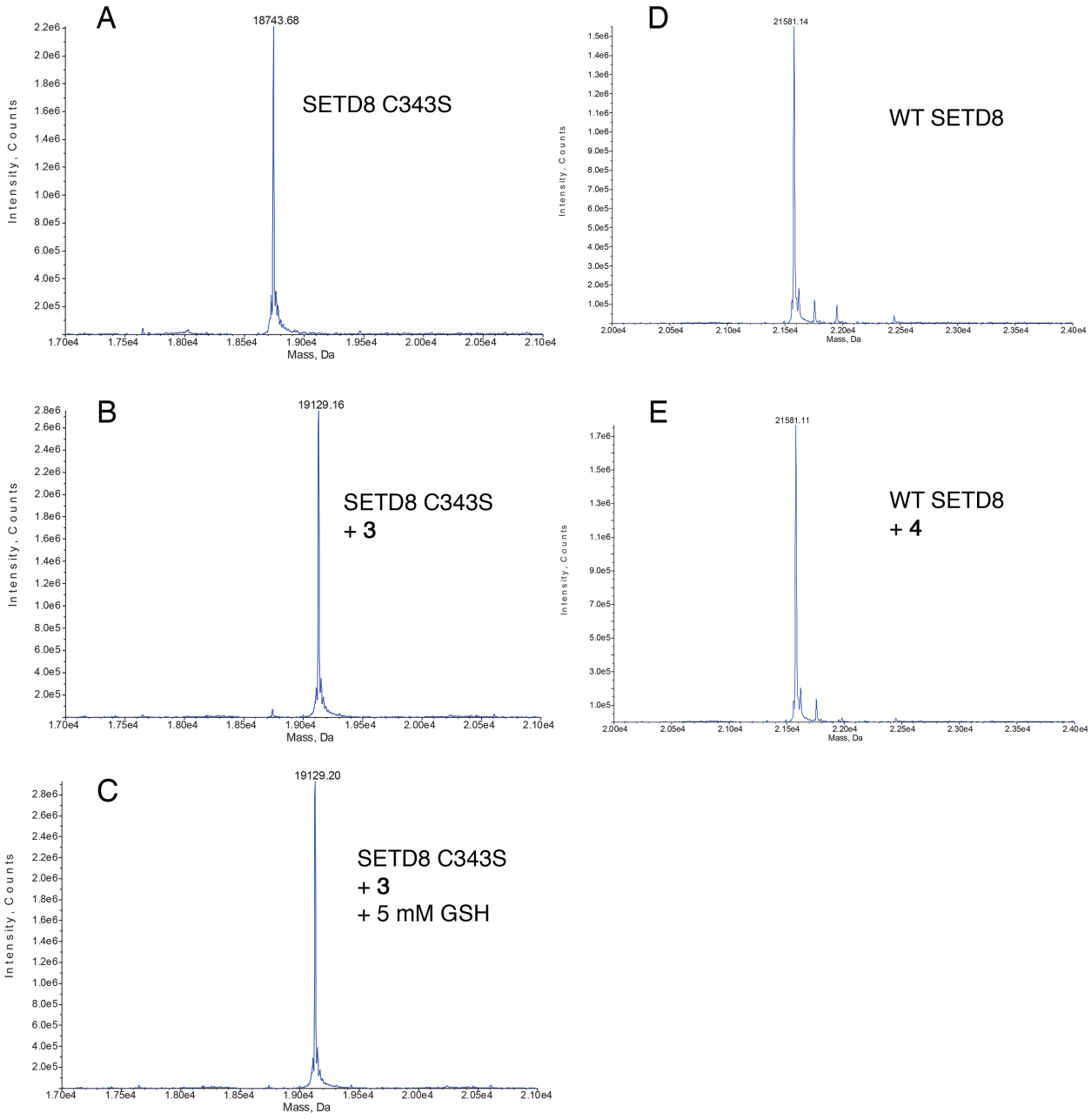
Structure Based Design of a Covalent Inhibitor of the SET Domain-Containing Protein 8 (SETD8) Lysine Methyltransferase

Kyle V. Butler[1#], Anqi Ma[1#], Wenyu Yu[2#], Fengling Li[2], Fabio Pittella-Silva[3], Shawna L. Organ[2], Wolfram Tempel[2], Nicolas Babault[1], Fabio Pittella-Silva[3], Jason Shao[1], Junyi Wang[3], Dalia Barsyte-Lovejoy[2], Masoud Vedadi[2,4], Minkui Luo[3], Peter J. Brown[2], Cheryl H. Arrowsmith[2,4], Jian Jin[1]*

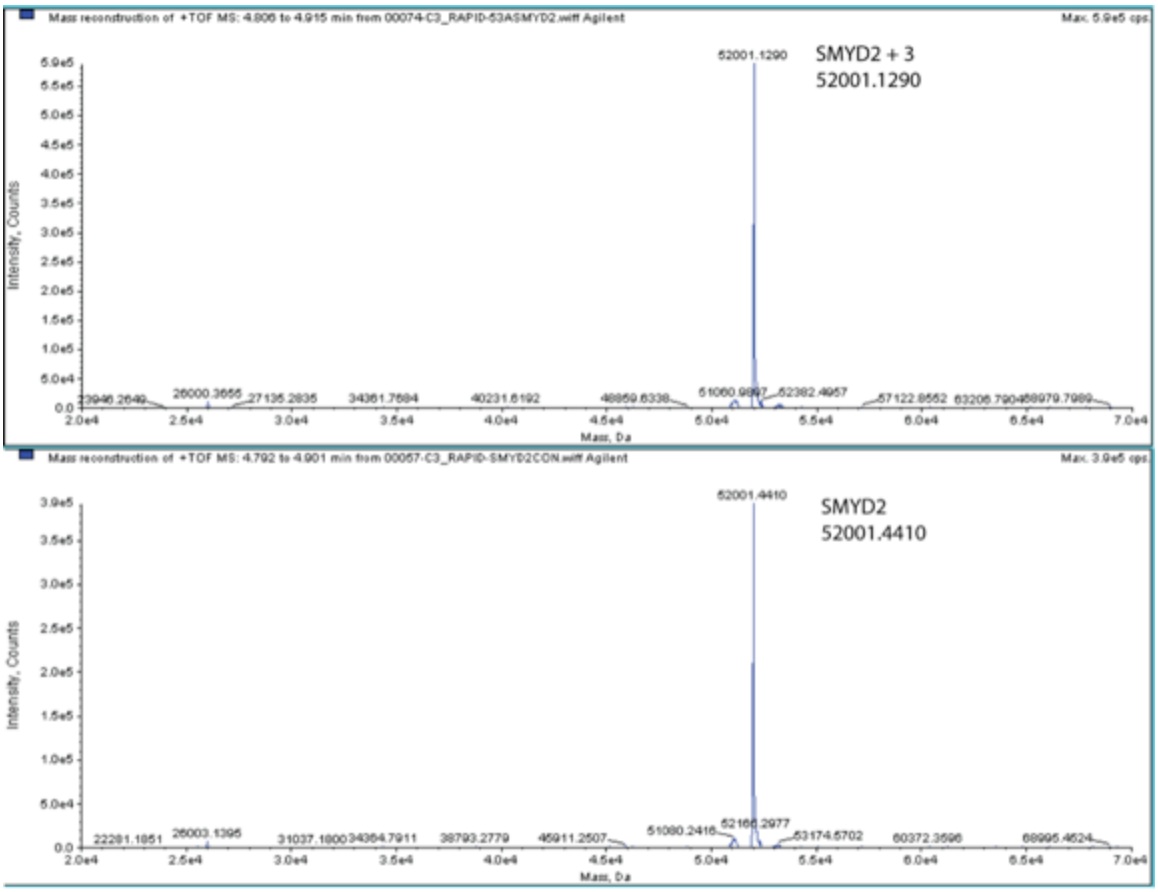
Table of Contents

A.	**Supplementary Figure 1**	**S2**
B.	**Supplementary Figure 2**	**S3**
C.	**Supplementary Figure 3**	**S4**
D.	**Supplementary Figure 4**	**S5**
E.	**Supplementary Figure 5**	**S6**
G.	**NMR Spectra**	**S7**
H.	**Extended Data Table 1**	**S11**

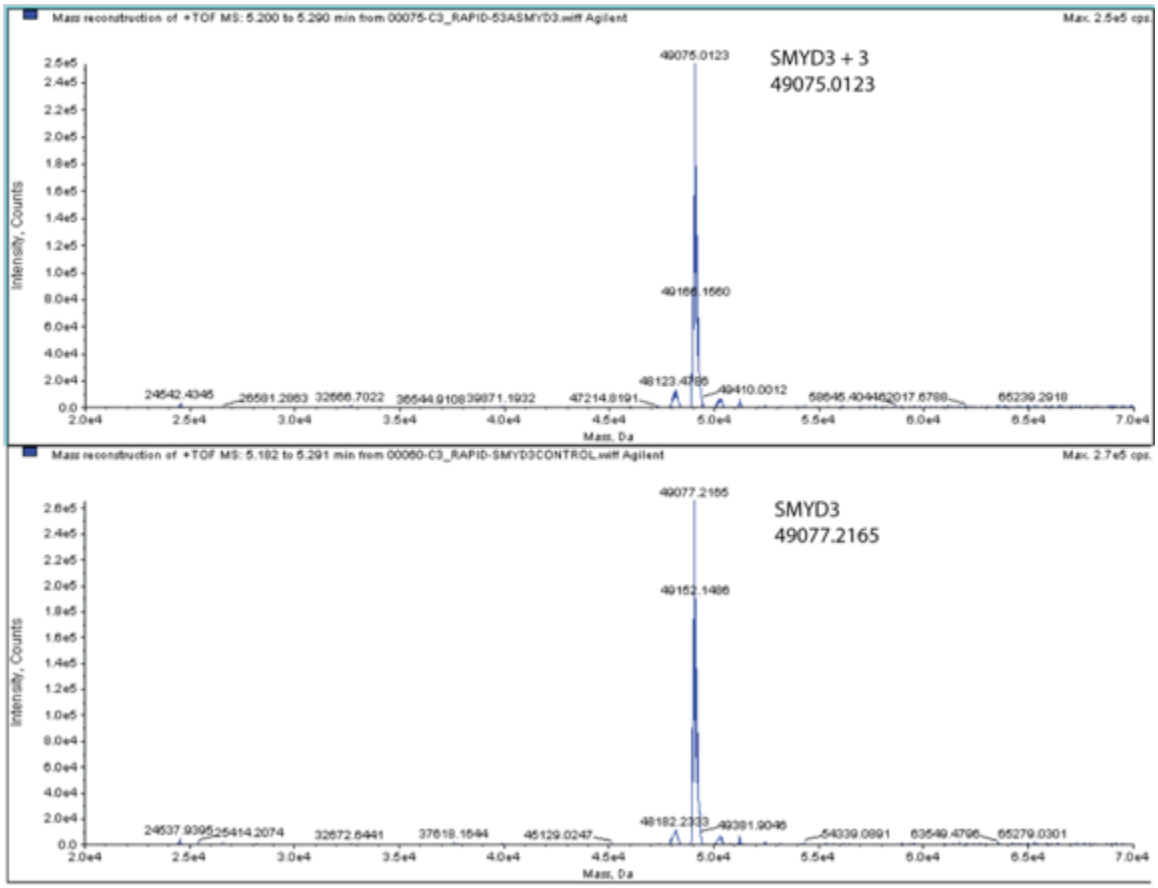
Supplementary Figures



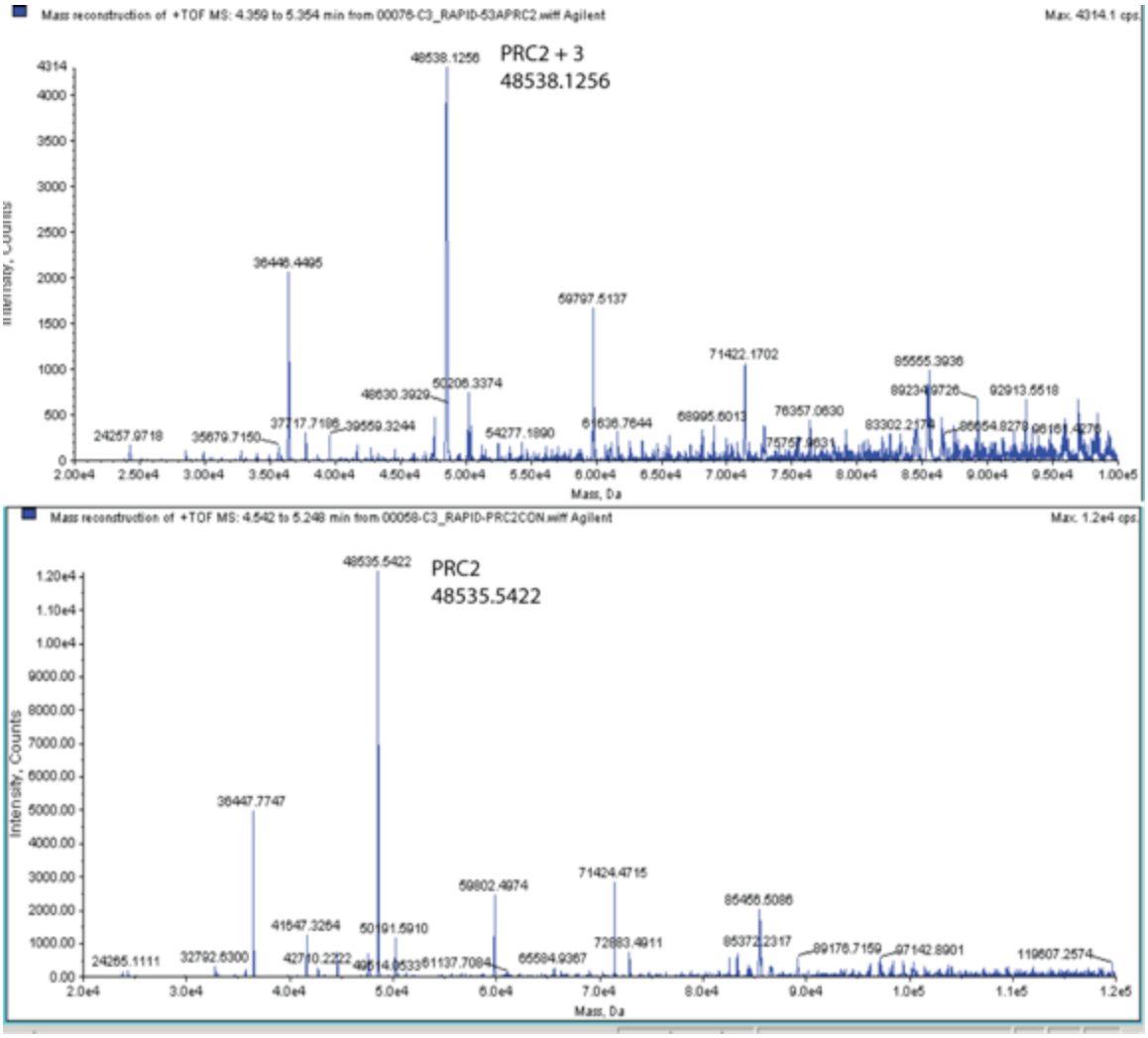
Supplementary Figure 1. Mass Spectrometry results of covalent modification of SETD8. (A) SETD8 C343S. (B) SETD8 C343S + **3**. (C) SETD8 C343S + **3** + 5 mM glutathione (GSH). (D) SETD8. (E) SETD8 + **4**.



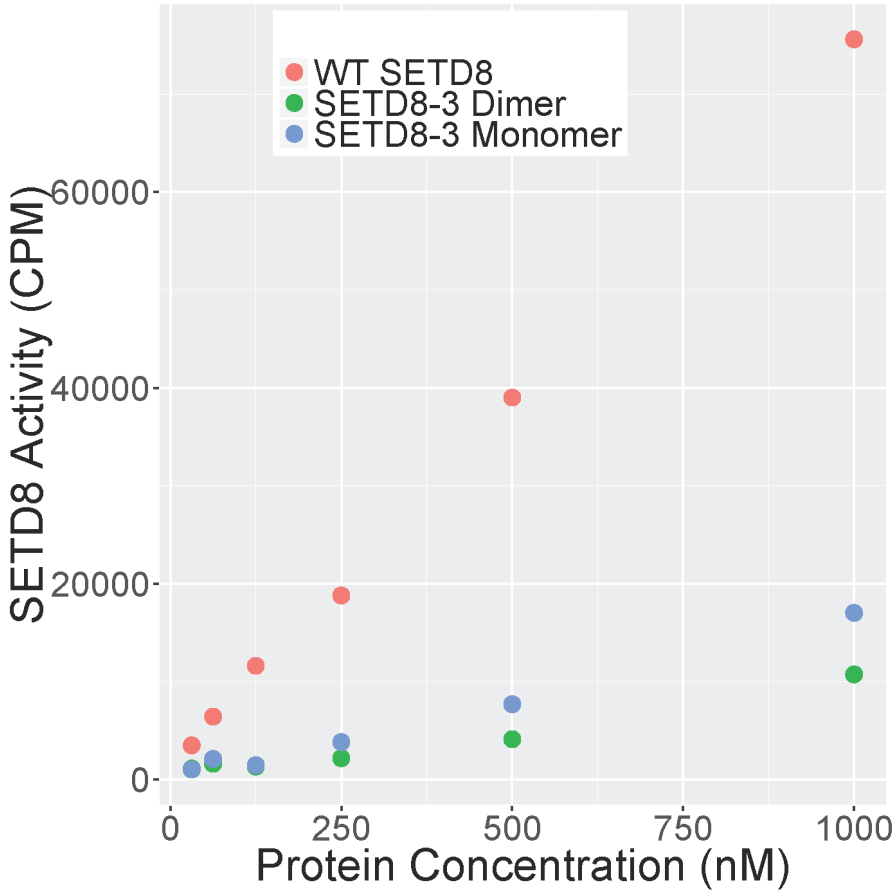
Supplementary Figure 2. Mass Spectrometry of SMYD2 with and without addition of **3**.



Supplementary Figure 3. Mass Spectrometry of SMYD3 with and without addition of **3**.

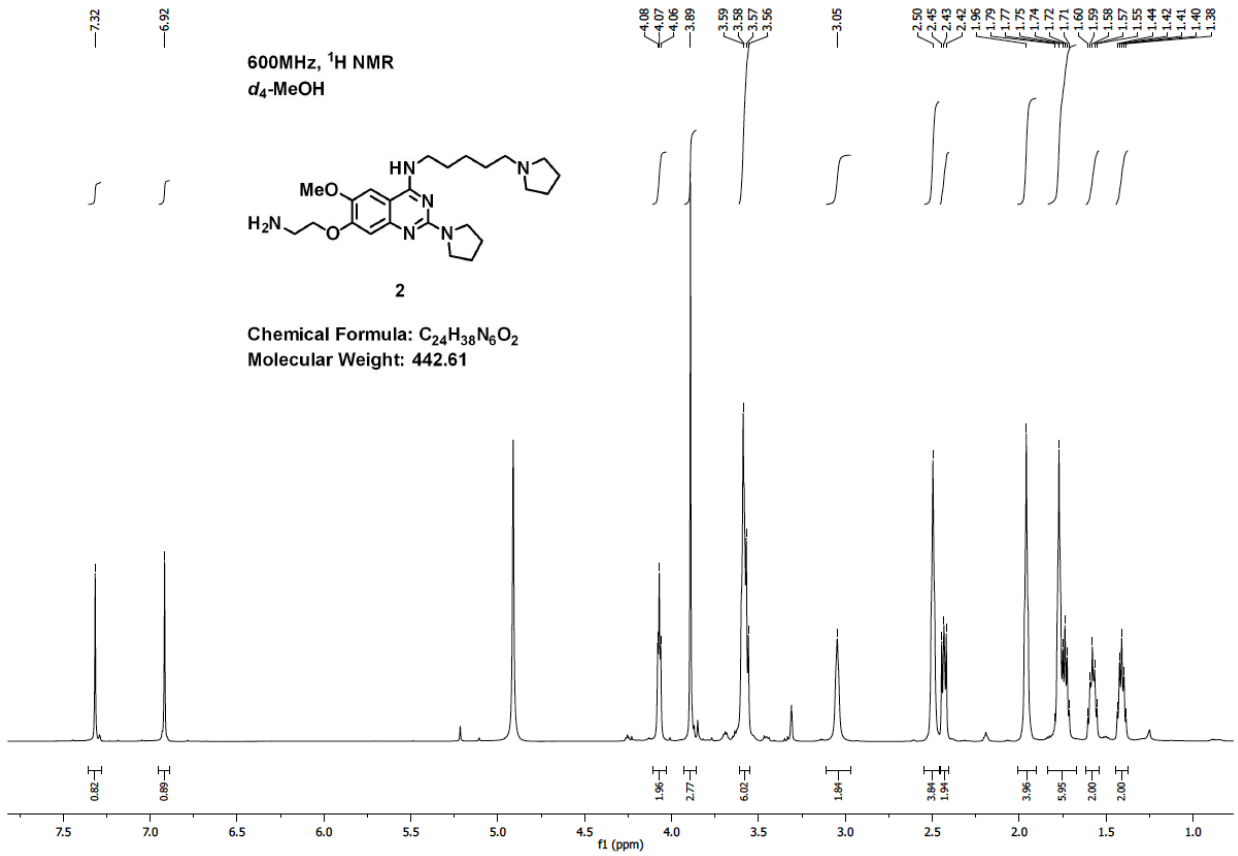


Supplementary Figure 4. Mass Spectrometry of PRC2 with and without addition of **3**.

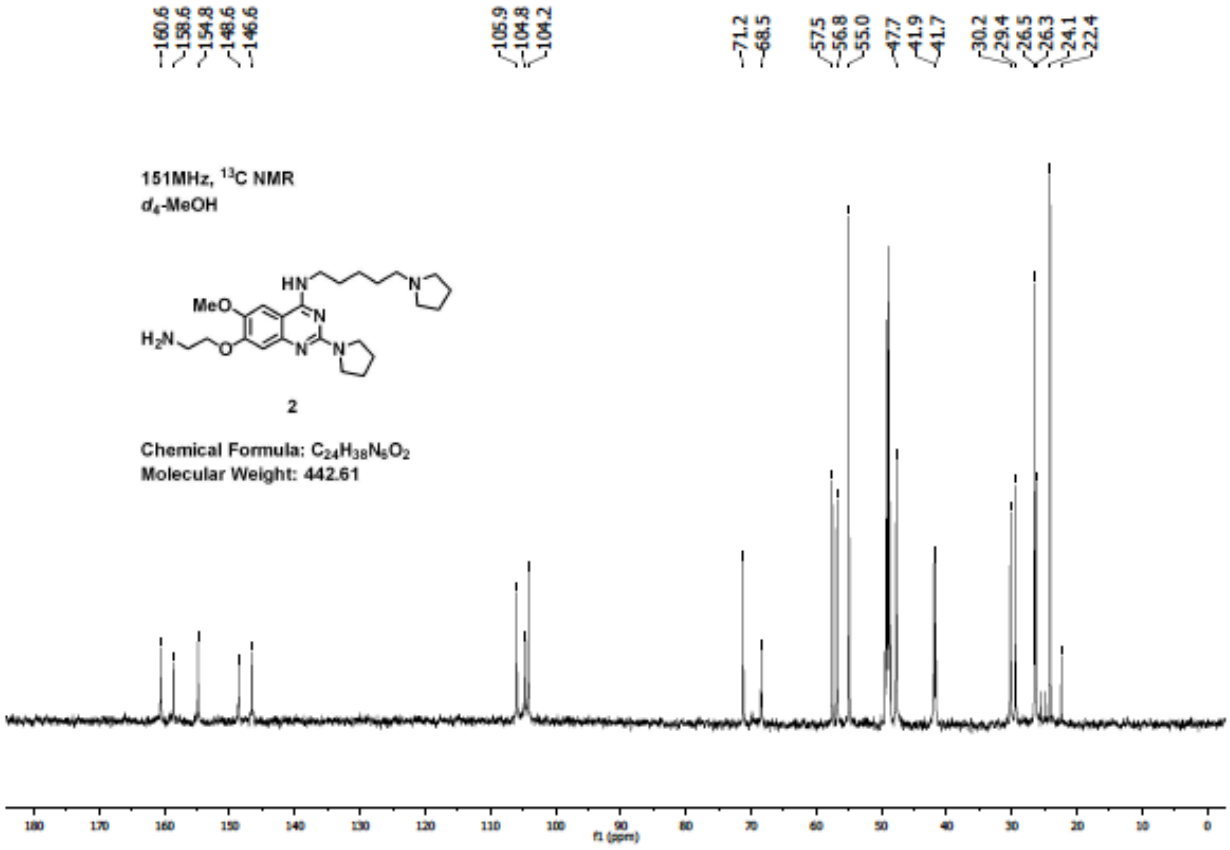


Supplementary Figure 5. Residual activity of the SETD8-**3** adduct in both monomer and dimer form compared to that of wild type SETD8.

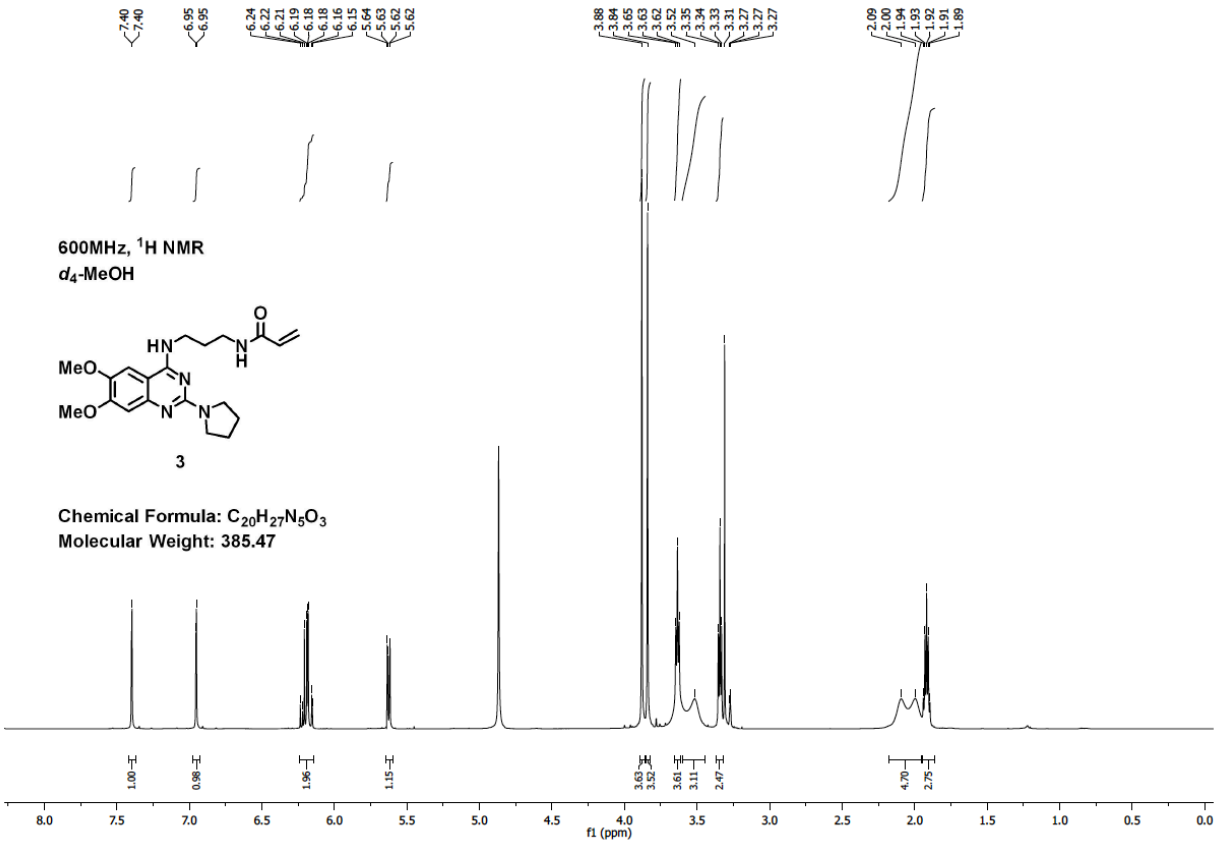
NMR Spectra



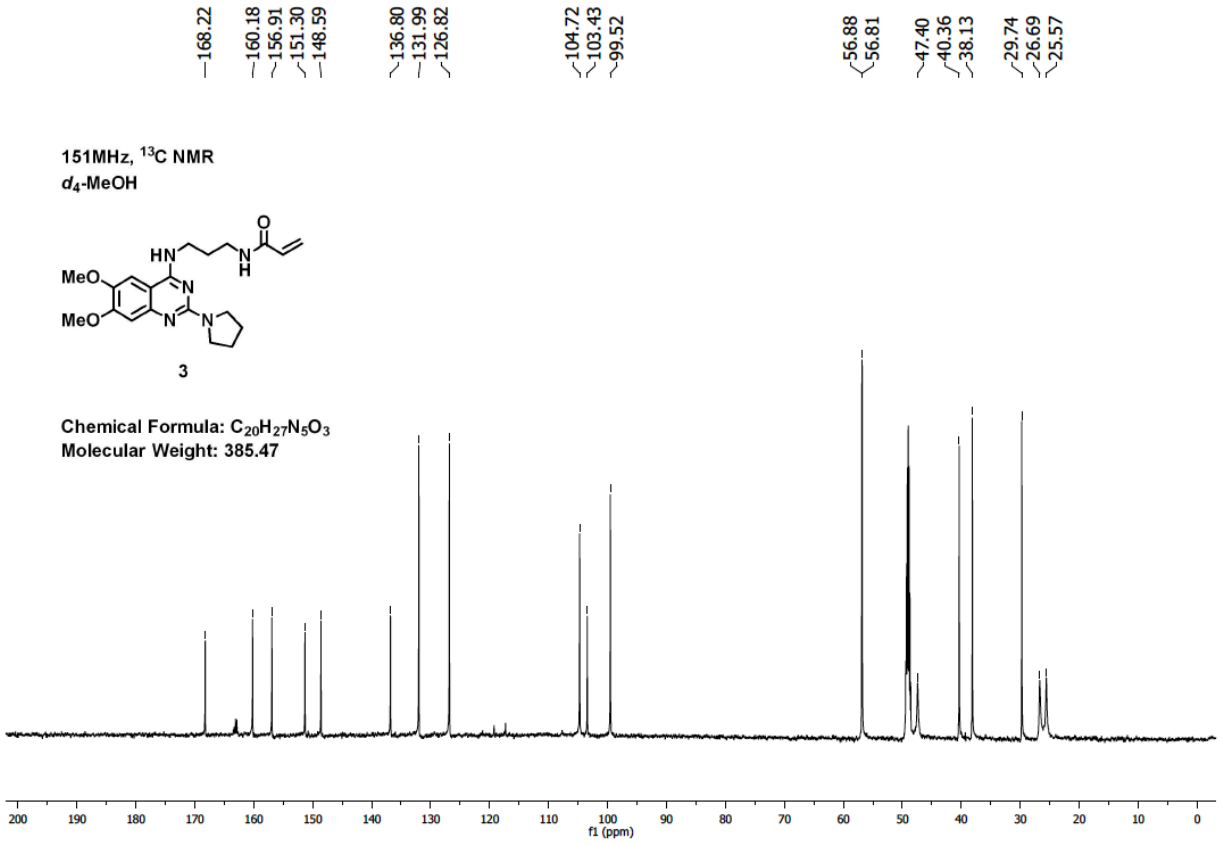
^{1}H NMR of **2**.


160.6
158.6
154.8
148.6
146.6
105.9
104.8
104.2
71.2
68.5
57.5
56.8
55.0
47.7
41.9
41.7
30.2
29.4
26.5
26.3
24.1
22.4
151MHz, 13C NMR
d4-MeOH
HN
N
MeO
H2N
O
N
N
2
Chemical Formula: C24H38N6O2
Molecular Weight: 442.61
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

^{13}C NMR of **2**.


600MHz, ^{1}H NMR
d_4-MeOH
MeO
MeO
HN
N
H
O
N
N
N
3
Chemical Formula: $C_{20}H_{27}N_5O_3$
Molecular Weight: 385.47
f1 (ppm)

^{1}H NMR of **3**.


168.22
160.18
156.91
151.30
148.59
136.80
131.99
126.82
104.72
103.43
99.52
56.88
56.81
47.40
40.36
38.13
29.74
26.69
25.57
151MHz, 13C NMR
d4-MeOH
HN
N
H
O
MeO
MeO
N
N
3
Chemical Formula: C20H27N5O3
Molecular Weight: 385.47
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

^{13}C NMR of **3**.

Extended Data Table 1 Data collection and refinement statistics (**Molecular replacement**)

	SETD8/**2** (PDB: 5T5G)	SETD8/**3** (PDB: 5TH7)
Data collection		
Space group	$P3_221$	C2
Cell dimensions		
a, *b*, *c* (Å)	74.58, 74.58, 111.47	83.15,31.64,114.89
α, β, γ (°)	88.96,88.96,34.99	90.00,105.23,90.00
Molecule per asymmetric unit	1	2
Resolution (Å)	44.48-2.10 (2.16-2.10)*	37.51-1.95 (2.00-1.95)*
R_{sym} or R_{merge}	0.081 (0.982)	0.084 (0.842)
I/s*I*	19.9 (3.1)	11.2 (1.3)
Completeness (%)	99.9 (100.0)	99.0 (94.5)
Redundancy	10.7 (11.1)	3.4 (2.6)
Refinement		
Resolution (Å)	2.10	1.95
No. unique reflections	9530 (775)	21342 (1397)
$R_{work/}$ R_{free}	0.204/0.227	0.189/0.237
No. atoms /average B-factor [Å2]	1145/54.8	2366/38.3
Protein	1077/55.2	2189/38.9
Ligand/ion	43/51.4	56/22.4
Water	20/40.2	96/34.8
Other	5/48.9	25/34.9
R.m.s deviations		
Bond lengths (Å)	0.014	0.010
Bond angles (º)	1.5	1.1
Ramachandran plot		
Most favoured (%)	97.87	97.23
Additional allowed (%)	2.13	2.78
Outliers (%)	0	0
Protein Data Bank entry	--	--

*The values in parentheses refer to statistics in the highest bin.