Supporting Information

Sequence, Cloning, and Analysis of the Fluvirucin B_1 Polyketide Synthase from *Actinomadura vulgaris*

Tsung-Yi Lin, Lawrence S. Borketey, Gitanjeli Prasad, Stephanie A. Waters, and Nathan A. Schnarr*

Department of Chemistry, University of Massachusetts, 710 N. Pleasant Street, Amherst, Massachusetts 01003

schnarr@chem.umass.edu

Figure S1. **Expected AT Specificities :** Results from SEARCHPKS

Mod 1

Matching all known AT domains with your sequence

Query is the potential *AT* in your sequence

DOMAIN	E-VALUE	%IDENTITY	%POSITIVES	Substrate	ActivesiteMotif
Your_AT_sequence					QQGHSQGRSHTNV
nidda_mod04_AT	7e-85	53	65	Methylmalonate	QQGHSQGRSHTNV
ampho_mod01_AT	1e-81	49	62	Methylmalonate	QQGHSQGRSHTSV
nidda_mod05_AT	3e-81	48	63	Ethylmalonate	QQGHSQGRGHTNV
rifam_mod07_AT	7e-81	50	62	Methylmalonate	QQGHSQGRSHTNV
rifam_mod04_AT	2e-80	50	63	Methylmalonate	QQGHSQGRSHTNV
ampho_mod11_AT	1e-79	48	63	Methylmalonate	QQGHSQGRSHTSV
pimar_mod07_AT	3e-79	50	62	Methylmalonate	QQGHSQGRSHTNV
averm_mod01_AT	4e-79	50	63	Methylmalonate	QQGHSQGRSHTNV
averm_mod07_AT	4e-79	50	63	Methylmalonate	QQGHSQGRSHTNV
nysta_mod01_AT	4e-78	49	61	Methylmalonate	QQGHSQGRSHTNV

Mod2

Matching all known AT domains with your sequence

Query is the potential *AT* in your sequence

DOMAIN	E-VALUE	%IDENTITY	%POSITIVES	Substrate	ActivesiteMotif
Your_AT_sequence					QQGHSLGRFHNHV
averm_mod08_AT	4e-87	56	63	Malonate	QQGHSLGRFHAQV
averm_mod04_AT	9e-87	56	63	Malonate	QQGHSLGRFHAQV
averm_mod02_AT	1e-86	57	64	Malonate	QQGHSLGRFHAQV
averm_mod03_AT	1e-86	56	63	Malonate	QQGHSLGRFHAQV
averm_mod05_AT	3e-86	56	63	Malonate	QQGHSLGRFHAQV
ampho_mod13_AT	3e-82	49	65	Malonate	QQGHSIGRFHTHV
nysta_mod13_AT	6e-79	48	63	Malonate	QQGHSIGRFHTHV
ampho_mod03_AT	8e-79	47	63	Malonate	QQGHSIGRFHNHV
nysta_mod04_AT	3e-78	50	65	Malonate	QQGHSIGRFHNHV
nysta_mod03_AT	9e-78	47	63	Malonate	QQGHSIGRFHNHV

Mod3

Matching all known AT domains with your sequence

Query is the potential *AT* in your sequence

DOMAIN	E-VALUE	%IDENTITY	%POSITIVES	Substrate	ActivesiteMotif
Your_AT_sequence					QQGHSQGRSHTNV
pimar_mod07_AT	e-104	61	73	Methylmalonate	QQGHSQGRSHTNV
ampho_mod11_AT	e-101	58	72	Methylmalonate	QQGHSQGRSHTSV
ampho_mod01_AT	1e-98	57	70	Methylmalonate	QQGHSQGRSHTSV
nysta_mod01_AT	4e-98	58	71	Methylmalonate	QQGHSQGRSHTNV
nysta_mod11_AT	7e-98	58	71	Methylmalonate	QQGHSQGRSHTNV
nidda_mod04_AT	9e-98	58	68	Methylmalonate	QQGHSQGRSHTNV
sorap_mod07_AT	3e-93	55	68	Glycerate	QQGHSQGRSHTNV
olean_mod03_AT	1e-92	55	66	Methylmalonate	QQGHSQGRSHTNV
rifam_mod03_AT	2e-92	54	69	Methylmalonate	QQGHSQGRSHTNV
olean_mod01_AT	3e-92	55	66	Methylmalonate	QQGHSQGRSHTNV

Mod4

Matching all known AT domains with your sequence

Query is the potential *AT* in your sequence

DOMAIN	E-VALUE	%IDENTITY	%POSITIVES	Substrate	ActivesiteMotif
Your_AT_sequence					QQGHSLGR-RTQV
averm_mod08_AT	2e-74	52	58	Malonate	QQGHSLGRFHAQV
averm_mod04_AT	3e-74	52	58	Malonate	QQGHSLGRFHAQV
averm_mod02_AT	7e-74	52	58	Malonate	QQGHSLGRFHAQV
averm_mod03_AT	7e-74	52	58	Malonate	QQGHSLGRFHAQV
averm_mod05_AT	2e-73	52	58	Malonate	QQGHSLGRFHAQV
ampho_mod13_AT	2e-68	45	60	Malonate	QQGHSIGRFHTHV
nysta_mod04_AT	5e-66	46	60	Malonate	QQGHSIGRFHNHV
ampho_mod03_AT	2e-65	43	58	Malonate	QQGHSIGRFHNHV
nysta_mod03_AT	7e-65	43	57	Malonate	QQGHSIGRFHNHV
pimar_mod05_AT	2e-63	44	57	Malonate	QQGHSIGRFHGHV

Mod5

Matching all known AT domains with your sequence

Query is the potential *AT* in your sequence

DOMAIN	E-VALUE	%IDENTITY	%POSITIVES	Substrate	ActivesiteMotif
Your_AT_sequence					QQGHSQGRSHTNV
nidda_mod04_AT	7e-85	53	65	Methylmalonate	QQGHSQGRSHTNV
ampho_mod01_AT	1e-81	49	62	Methylmalonate	QQGHSQGRSHTSV
nidda_mod05_AT	3e-81	48	63	Ethylmalonate	QQGHSQGRGHTNV
rifam_mod04_AT	3e-81	50	63	Methylmalonate	QQGHSQGRSHTNV
rifam_mod07_AT	7e-81	50	62	Methylmalonate	QQGHSQGRSHTNV
pimar_mod07_AT	5e-80	50	62	Methylmalonate	QQGHSQGRSHTNV
ampho_mod11_AT	1e-79	48	63	Methylmalonate	QQGHSQGRSHTSV
averm_mod01_AT	3e-79	50	63	Methylmalonate	QQGHSQGRSHTNV
averm_mod07_AT	3e-79	50	63	Methylmalonate	QQGHSQGRSHTNV
nysta_mod01_AT	4e-78	49	61	Methylmalonate	QQGHSQGRSHTNV

Sequence Comparison Results for Flu and DEBS Modules

All sequence comparison done by BLAST of NCBI website. All modules listed in the column are used as query, and the modules listed in the row are used as subject.

Note1: In cases where the compared modules are of dramatically different sizes (i.e Flu Mod 1 vs. Flu Mod2) identities are related only to the KS-AT regions (regions of max identity) of the modules.

Note2: Comparison results in parentheses are results from reverse query/subject pairs.

Sequence comparison (Identities/Similarities**)**	Flu_Mod1	Flu_Mod2	Flu_Mod3	Flu_Mod4	Flu_Mod5	DEBS_Mod4
Flu_Mod1	**100%**/100%	**61%**/72%	**81%**/86%	**60%**/70%	**80%**/86%	**49%**/61%
Flu_Mod2	**61%**/72%	**100%**/100%	**64%**/75%	**95%**/96%	**64%**/75%	**58%**/72%
Flu_Mod3	**81%**/86%	**64%**/75%	**100%**/100%	**63%**/73%	**74%**/82%	**50%**/63%
Flu_Mod4	**60%**/70%	**94%**/96%	**63%**/73%	**100%**/100%	**60%**/70%	**41%**/56%
Flu_Mod5	**80%**/86%	**64%**/75%	**75%**/82%	**60%/70%**	**100%**/100%	**49%**/62%

Sequence comparison (Identities/Similarities**)**	DEBS_Mod1	DEBS_Mod2	DEBS_Mod3	DEBS_Mod4	DEBS_Mod5	DEBS_Mod6
DEBS_Mod1	**100%**/100%	**49%**/60%	**48%**/61%	**56%**/68%	**49%**/61%	**59%**/68%
DEBS_Mod2	(**49%**/60%)	**100%**/100%	**48%**/59%	**52%**/64%	**50%**/61%	**49%**/61%
DEBS_Mod3	(**48%**/61%)	(**48%**/59%)	**100%**/100%	**56%**/69%	**48%**/62%	**51%**/64%
DEBS_Mod4	(**56%**/68%)	(**52%**/64%)	(**56%**/69%)	**100%**/100%	**57%**/68%	(**52%**/64%)
DEBS_Mod5	(**49%**/61%)	(**50%**/61%)	(**48%**/62%)	**57%**/69%	**100%**/100%	(**50%**/61%)
DEBS_Mod6	(**59%**/68%)	(**49%**/61%)	(**51%**/64%),	**52%**/64%	**50%**/61%	**100%**/100%

Protein homology analysis of Flu-KR Sequence

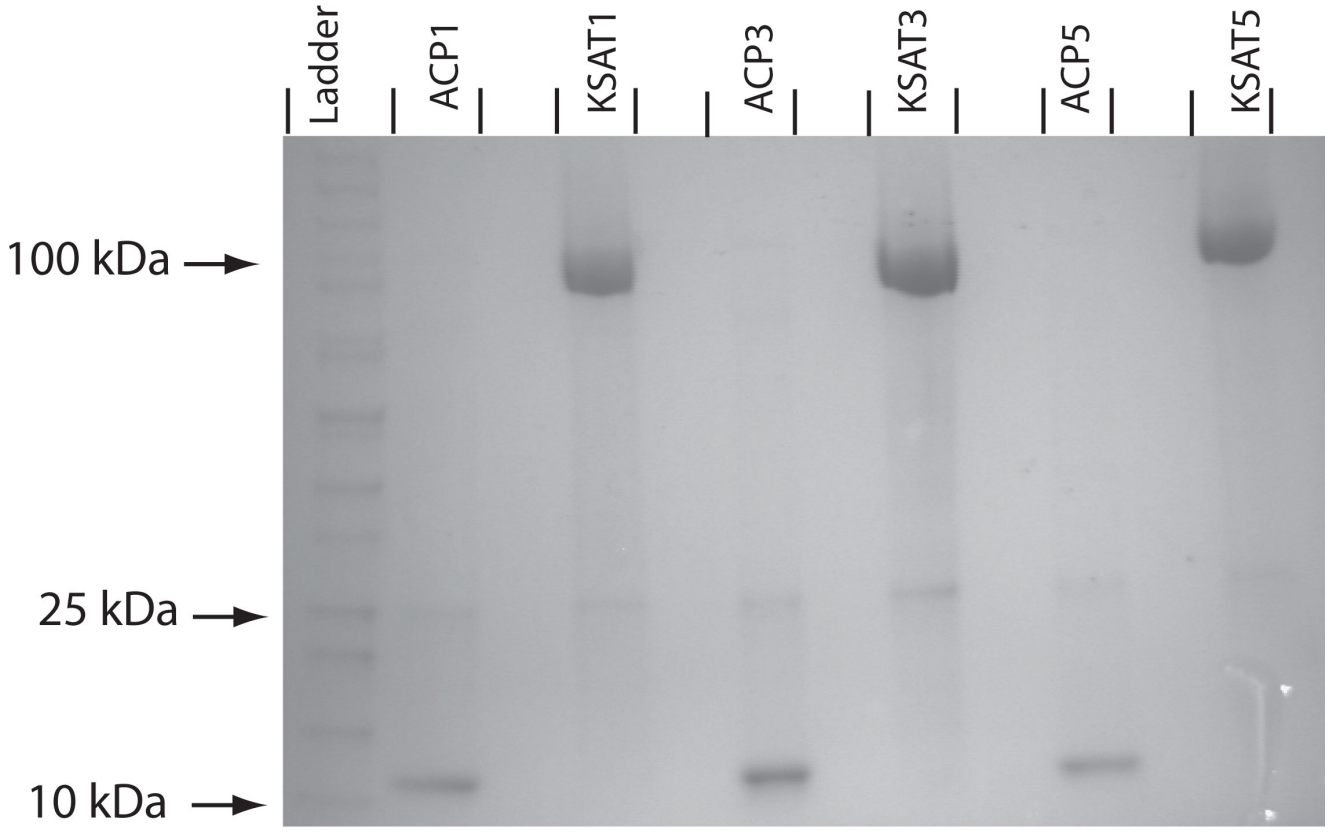
KR domain comparison between Fluvirucin modules and other B1 type KR domains. The homology analyses are done by Clustal Omega method of EMBL-EBI website, compared with the results in the reference.

Ref: Keatinge-Clay, A. T. *Chemistry & Biology*, **2007**, *14*, 898-908

Module	Loop	Catalytic region	Lid
Ave1	HTAGILDDAT-L	SSAAATFGAPGQANYAAANA	WGTWQGNG--LA--DSDKARAYLDRRG
Tyl1	HTAGILDDAV-I	SSVTGTWGNAGQGAYAAANA	WGLWGGG--GM---AAGAGEESLSRRG
Asc8	HTAATLDDGI-L	SSAAAVLGSPGQGNYAAANA	WGMWHTT-STLTGQLDDADRDRIRRGG
Ave7	HAAGVLDDAT-I	SSAAGILGSAGQGNYAAANA	dh
Ave9	HAAGVLDDAT-I	SSAAGILGSAGQGNYAAANA	WGLWEEA-SGMTGHLAGTDHRRIIRSG
Rap10	HTAGVLDDGV-V	SSAAGVLGSAGQGNYAVANA	dh
Flu1	HTAGVLDDGV-V	SSSAATLDSAGQGNYSAANA	WGLWEEA-SGMTGHLAGTDHRRIIRSG
Flu2	HAAGVGHDDVLV	SSGAAVWGSSGQASYAAANA	DH
Flu3	HTAGVLDDGV-V	SSSAATLDSAGQGNYSAANA	WGLWEDA-SGLTAKLTGTDHDRIRRSG
Flu4	HAAGVVDDGV-I	SSSAATLDSAGQGNYSAANA	DH
Flu5	HTAGVLDDGV-V	SSSAGTLDSAGQGNYSAANA	WGLWHQP-SGMSAHLTTTDITRIEQAG
	:. .* :	** :. **. *:.***	WGGWGGG--GM---MEDAGE-QLGRSG
			WGLWAPETGGMTNQLSELDLERLSRSG
			WGVWAPETGGMTRQLGDIDLERLSRSG
			WGLWHQP-SGMSAHLTTTDITRIEQAG
			** * : : : *

Note: DH, functional; dh, nonfunctional
PKS abbreviations:
Ave: Avermectin
Tyl: Tylosin
Asc: Ascomycin
Rap: Rapamycin
Flu: Fluvirucin B_1

Figure S2. PAGE of Fluvirucin KSATs and ACPs on 4-20% gel



Flu-ACP loading experiment by AT:

To a mixture of ACP (50 μM) and KS-AT (2 μM) in 100mM pH 7.0 phosphate buffer (50 μL total volume) containing 2.5 mM TCEP at 4^{o}C was added appropriate SNAc thioester substrate (500 μM) . The mixture was incubated 4 ^{o}C for 30 min to achieve AT assisted acylation of the ACP. Sequence grade modified trypsin was added to prepare samples with final ratio trypsin:ACP to be 1:10 (w/w). The mixture was incubated for 60 min at 37 ^{o}C. Digestion was quenched by addition of equal volume of 10% formic acid. Digests were stored at -80 ^{o}C until analysis.

ACP loading:

Protein	Probe	Mass expected (m/z, z=+3)
ACP1/3/5	-	896.8
ACP1/3/5	Ethyl-malonyl-SNAc	954.3
ACP1/3/5	Methyl-malonyl-SNAc	947.3
ACP1/3/5	Malonyl-SNAc	940.4

Figure S3. FluAT1 substrate selectivity

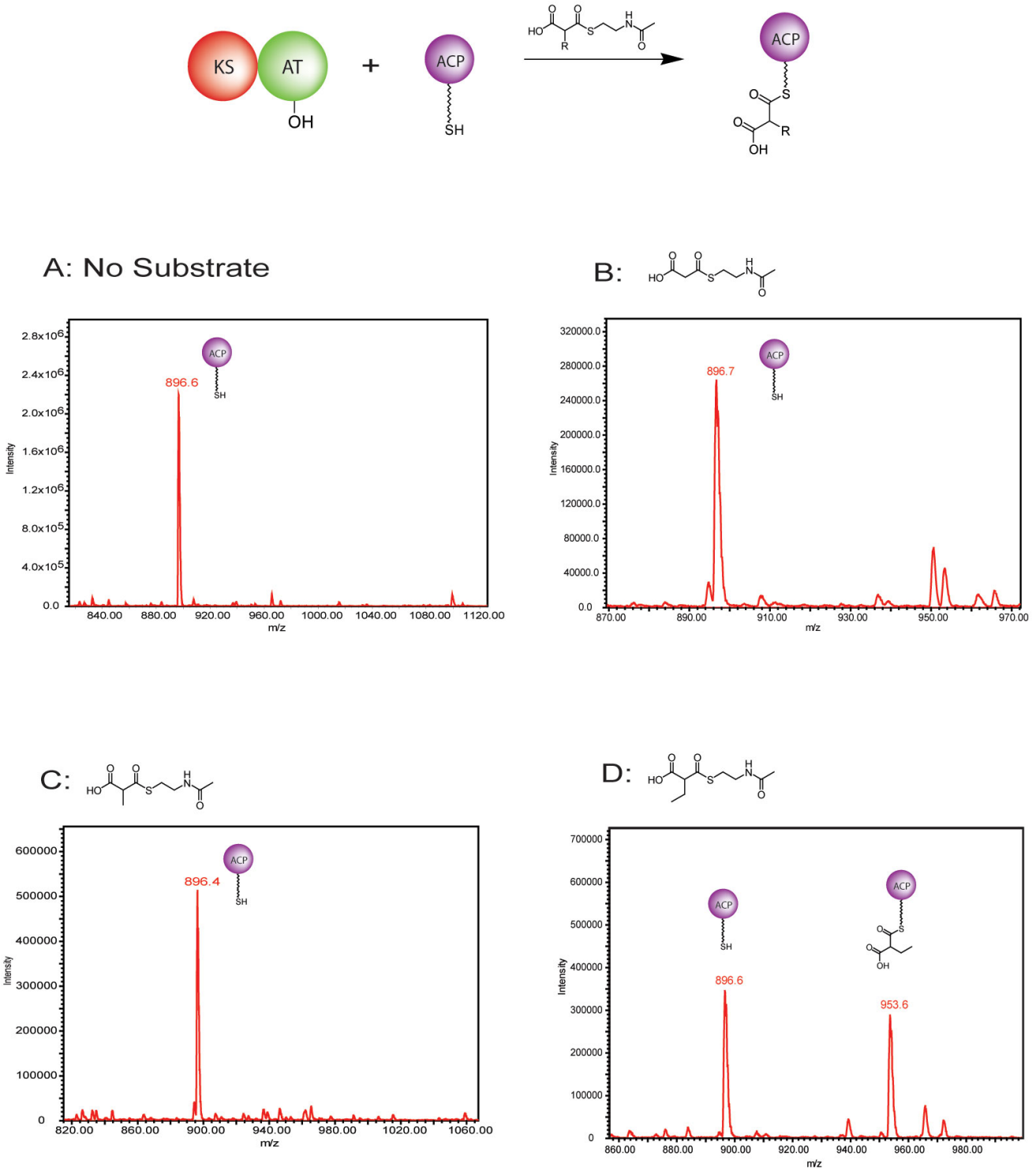


Figure S4. FluAT3 substrate selectivity

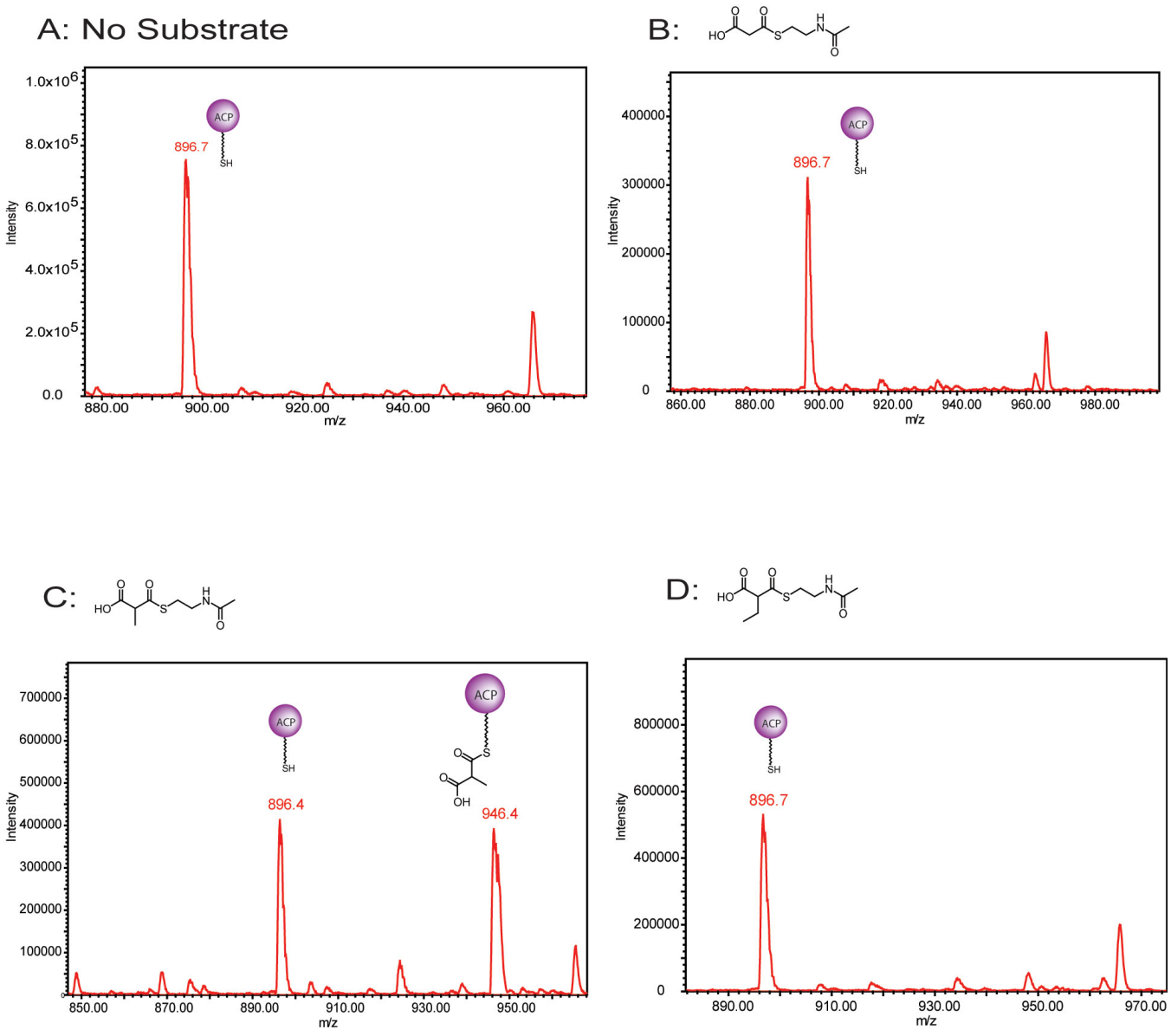
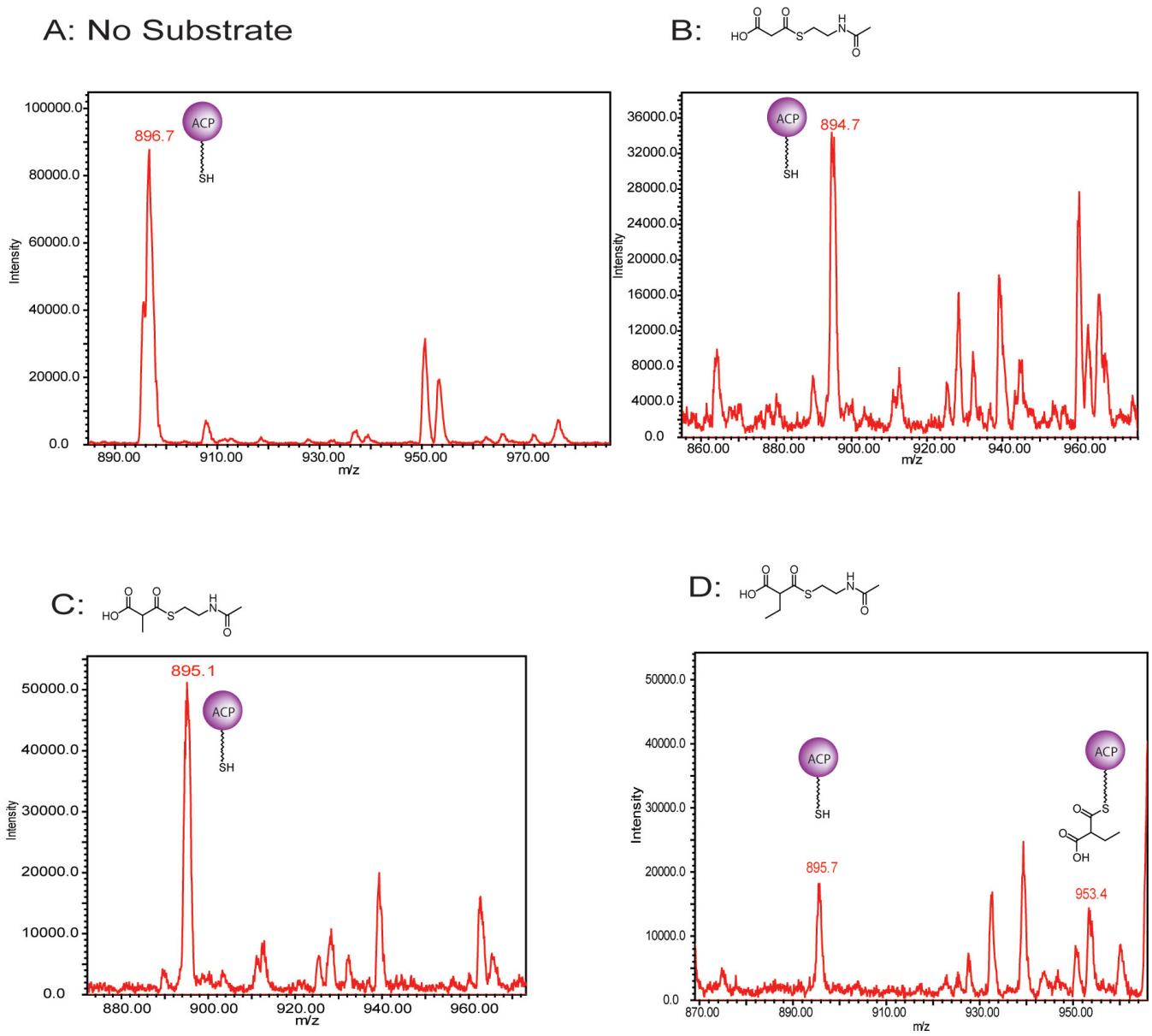


Figure S5. FluAT5 substrate selectivity



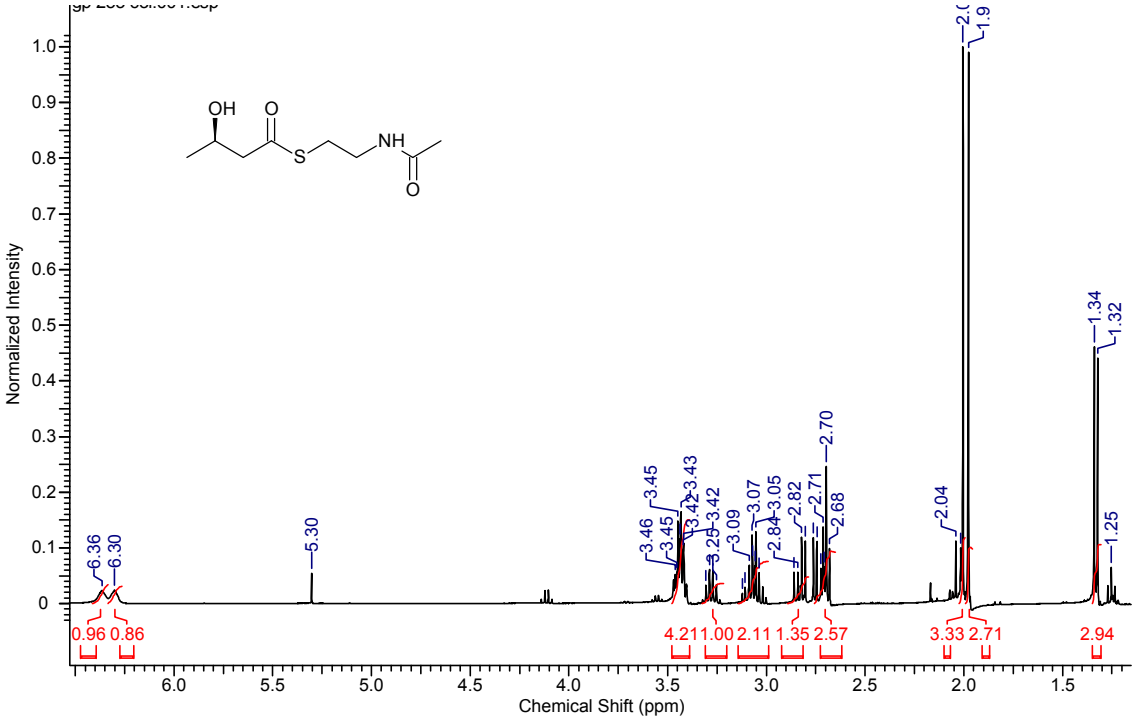
Flu-KSAT3 loading:

To a mixture of KSAT3 (25 μM) in 100 mM pH 7.0 phosphate buffer (50 μL total volume) containing 2.5 mM TCEP at 4°C was added appropriate SNAc thioester substrate (5mM). The mixture was incubated 4 °C for 60 min to achieve KS acylation. Sequence grade modified trypsin was added to prepare samples with final ratio trypsin:ACP to be 1:10 (w/w). The mixture was incubated for 60 min at 37 °C. Digestion was quenched by addition of equal volume of 10% formic acid. Digests were stored at -80 °C until analysis.

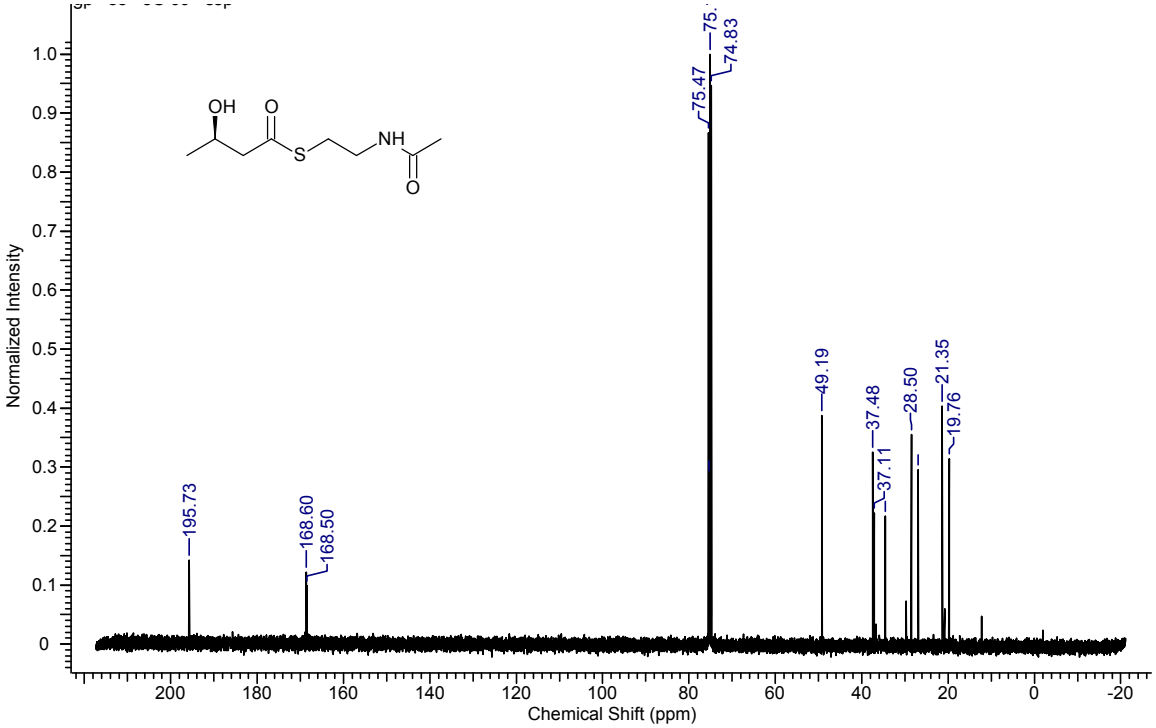
Protein	Probe	Mass expected(m/z, z=+2)
KSAT3	-	1667.6
KSAT3	3-hydroxy-butyric-SNAc	1710.1

(*R*)-3-hydroxybutyryl- SNAc

^{1}H NMR

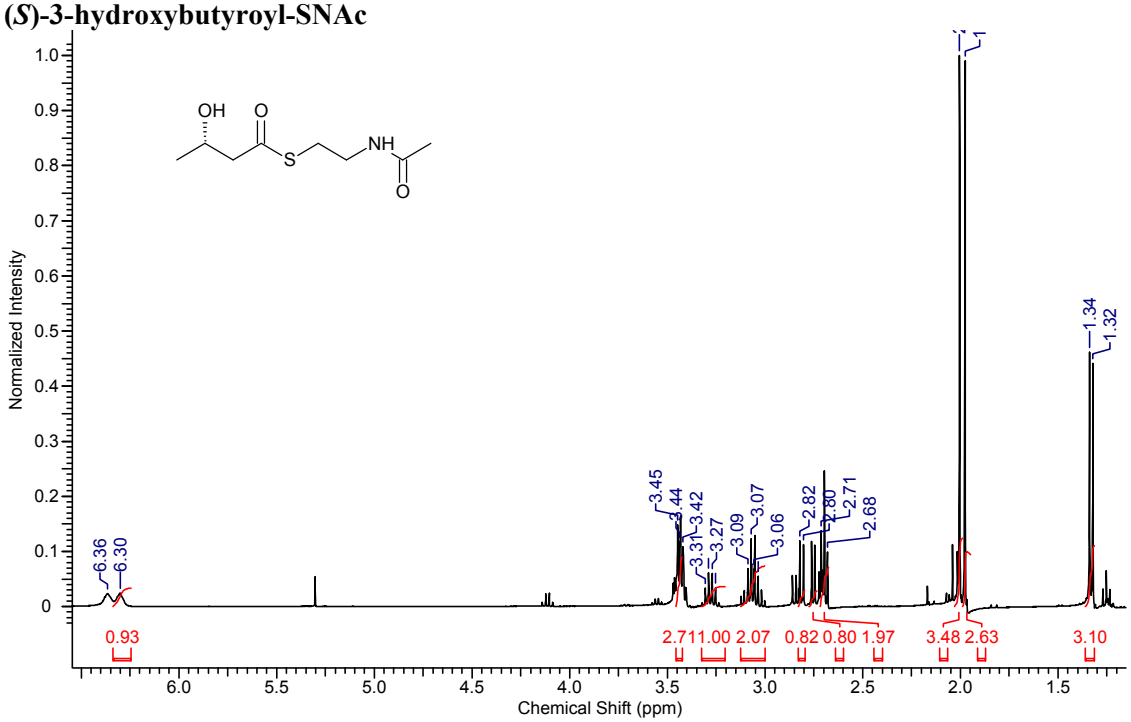


^{13}C NMR

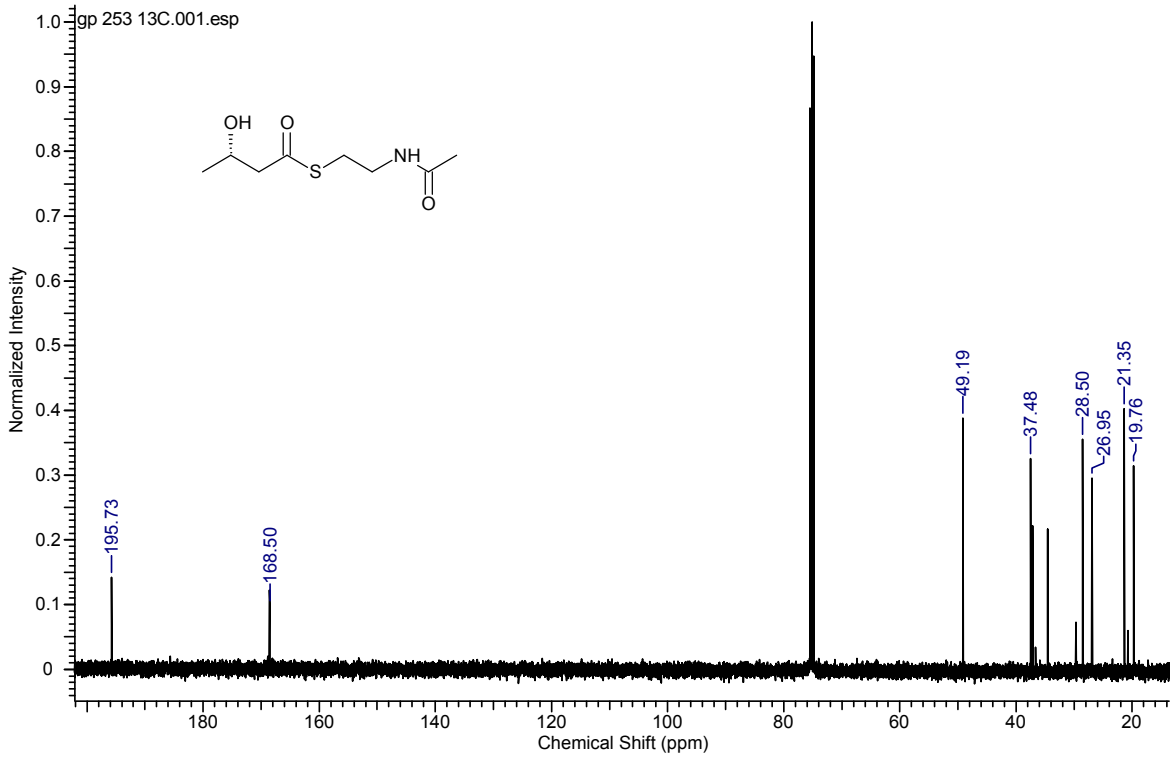


^{13}C NMR (101 MHz, CH_3Cl -*d*) ppm 19.76, 21.35, 26.95, 28.50, 34.57, 37.48, 49.19, 168.60 , 195.73

(*S*)-3-hydroxybutyroyl-SNAc

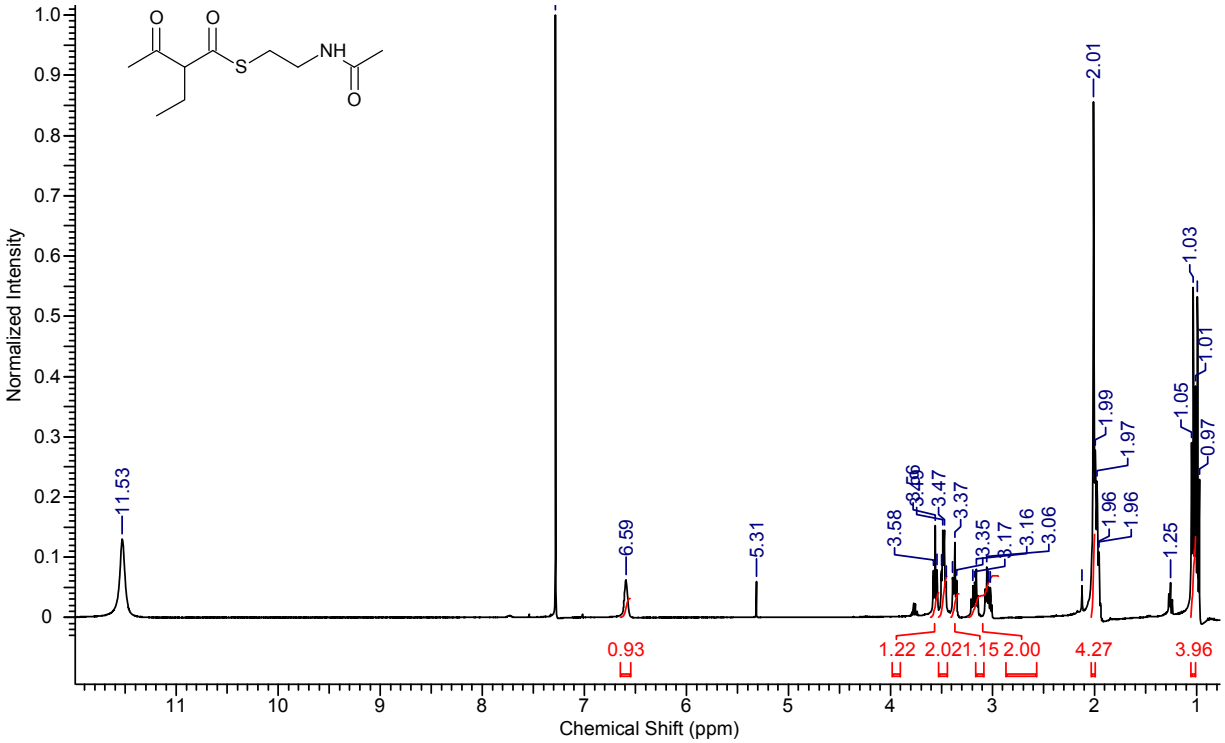


^{13}C NMR

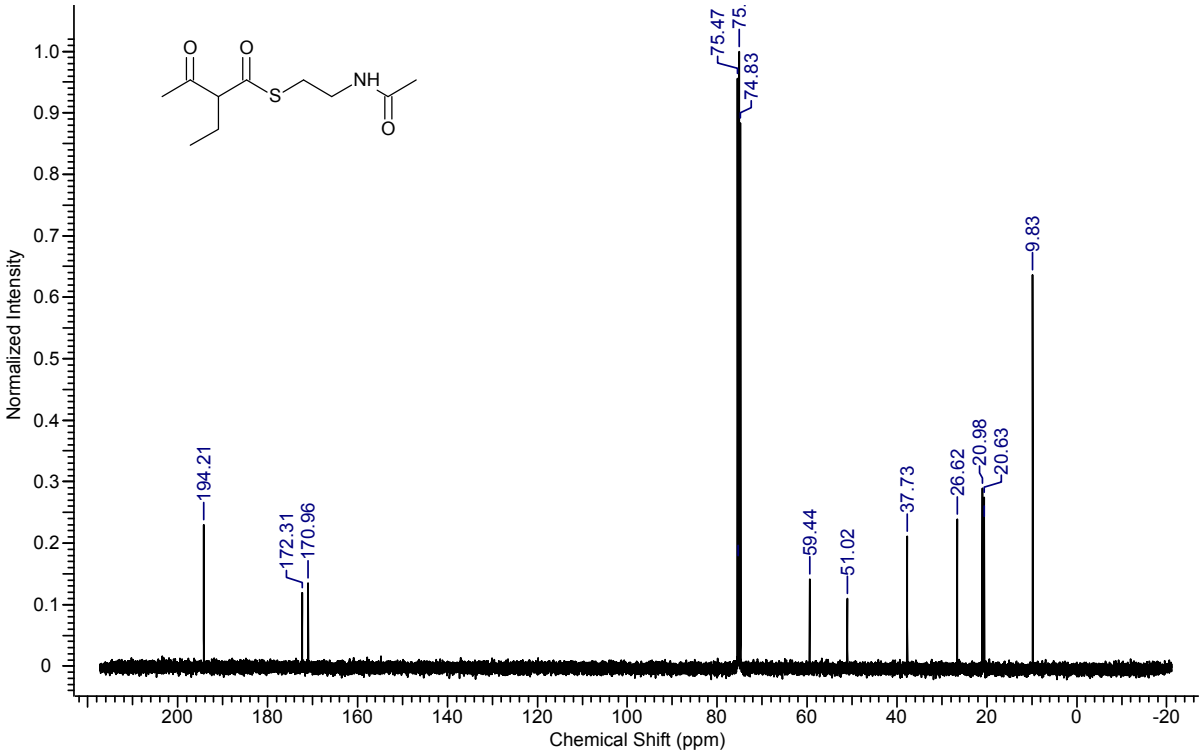


Ethylmalonyl-SNAc

^{1}H NMR

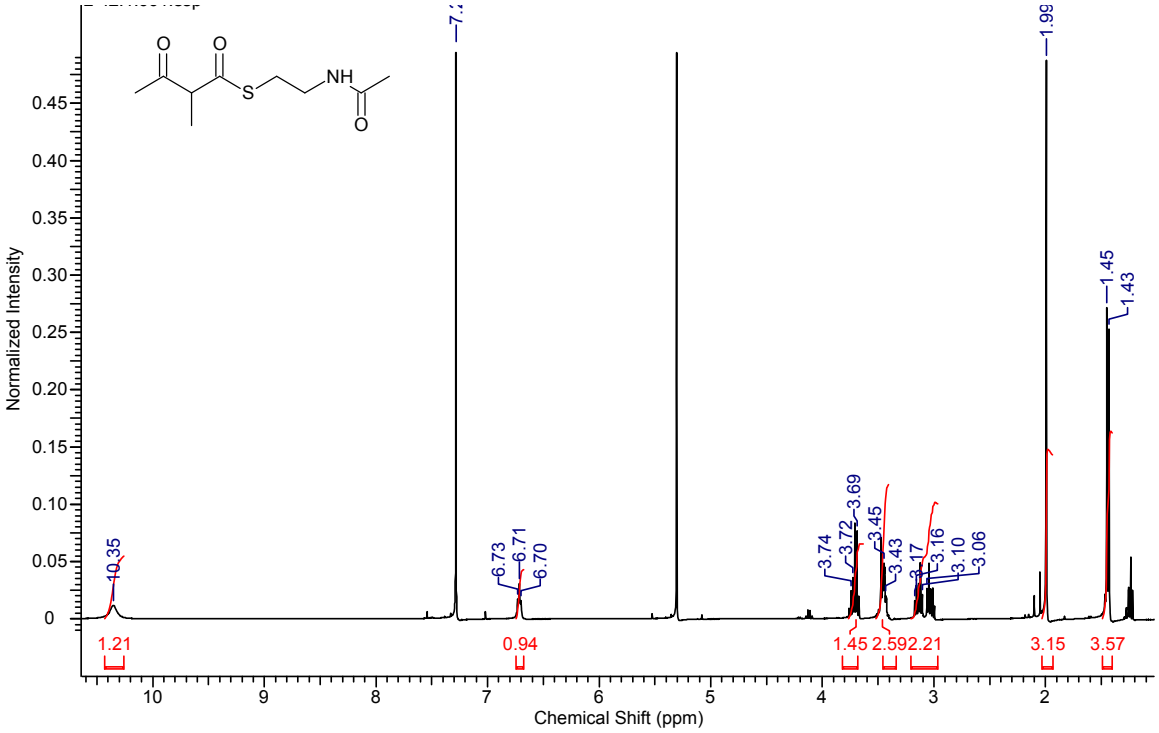


^{13}C NMR

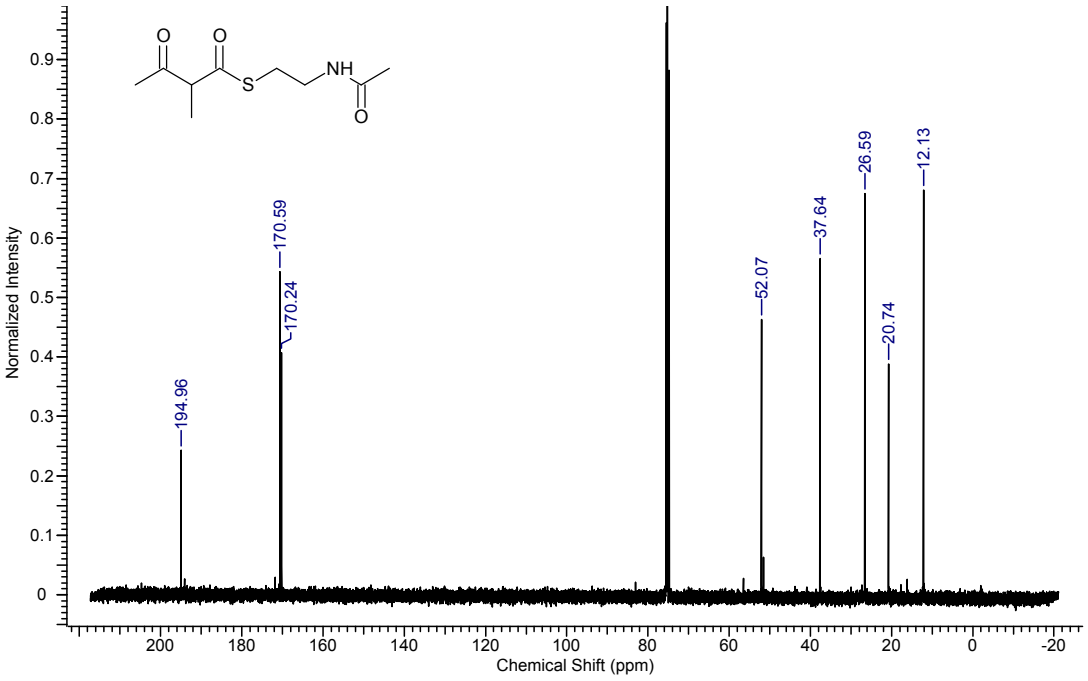


Methylmalonyl-SNAc

^{1}H NMR

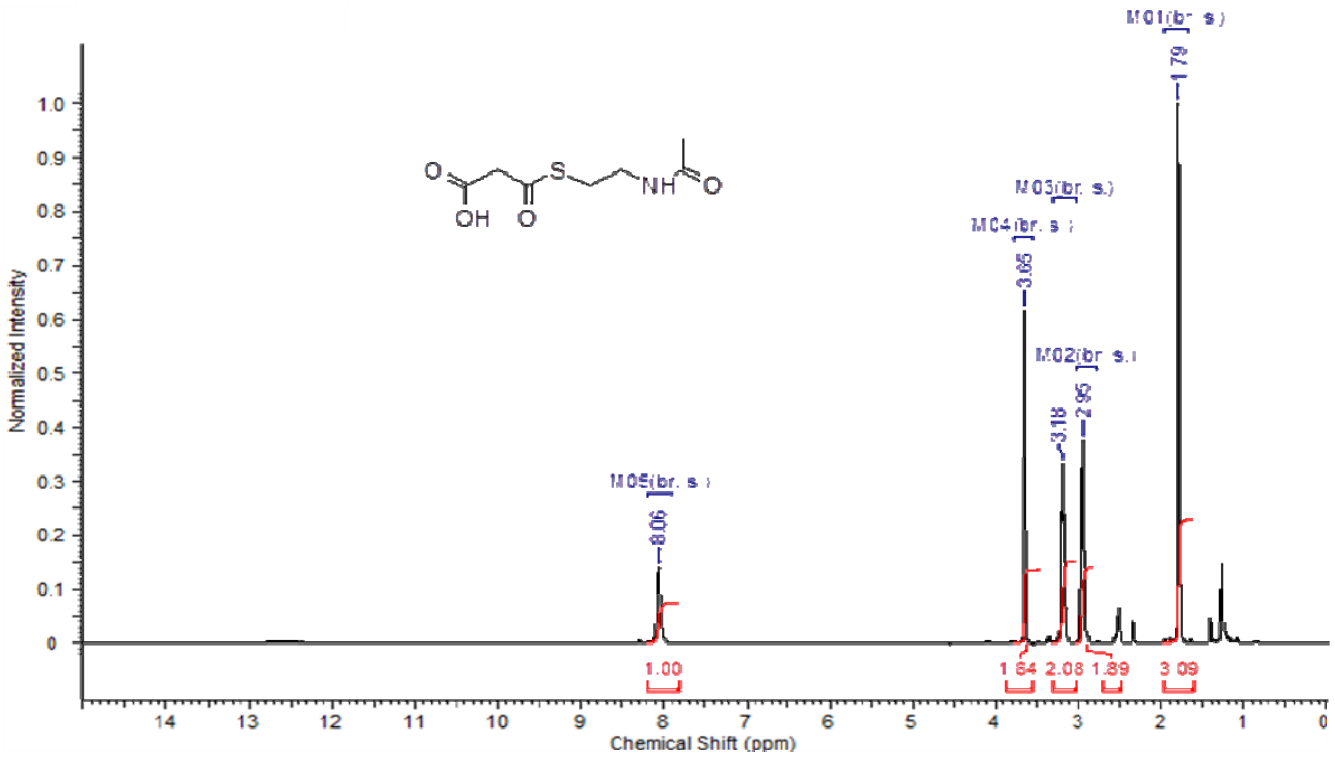


^{13}C NMR



Malonyl-SNAc

^{1}H NMR



^{13}C NMR