

Supporting Information

Grassystatins A–C from Marine Cyanobacteria, Potent Cathepsin E Inhibitors that Reduce Antigen Presentation

Jason C. Kwan, Erika A. Eksioglu, Chen Liu, Valerie J. Paul and Hendrik Luesch

Department of Medicinal Chemistry, University of Florida, Gainesville, Florida 32610, Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, Florida 32610, Smithsonian Marine Station, Fort Pierce, Florida 34949

Contents	Page Number
Table S1. NMR Spectral Data for Grassystatin A (1) in DMSO- <i>d</i> ₆	S2
Table S2. Conditions Employed for Protease Screen and IC ₅₀ Assays	S3
Figure S1. Assay Progress Curves (Cathepsin E and TACE)	S5
Figure S2. Docked Structures of Pepstatin A in Cathepsins D and E	S5
¹ H NMR Spectrum of Grassystatin A (1) in CDCl ₃	S6
¹ H NMR Spectrum of Grassystatin A (1) in CDCl ₃ , diluted	S7
¹ H NMR Spectrum of Grassystatin A (1) in CDCl ₃ , after D ₂ O exchange	S8
¹³ C NMR Spectrum of Grassystatin A (1) in CDCl ₃	S9
APT Spectrum of Grassystatin A (1) in CDCl ₃	S10
COSY Spectrum of Grassystatin A (1) in CDCl ₃	S11
HMQC Spectrum of Grassystatin A (1) in CDCl ₃	S12
HMBC Spectrum of Grassystatin A (1) in CDCl ₃	S13
ROESY Spectrum of Grassystatin A (1) in CDCl ₃	S14
TOCSY Spectrum of Grassystatin A (1) in CDCl ₃	S15
¹ H NMR Spectrum of Grassystatin A (1) in DMSO- <i>d</i> ₆	S16
COSY Spectrum of Grassystatin A (1) in DMSO- <i>d</i> ₆	S17
Edited HSQC Spectrum of Grassystatin A (1) in DMSO- <i>d</i> ₆	S18
HMBC Spectrum of Grassystatin A (1) in DMSO- <i>d</i> ₆	S19
ROESY Spectrum of Grassystatin A (1) in DMSO- <i>d</i> ₆	S20
¹ H NMR Spectrum of Grassystatin B (2) in CDCl ₃	S21
COSY Spectrum of Grassystatin B (2) in CDCl ₃	S22
HMQC Spectrum of Grassystatin B (2) in CDCl ₃	S23
HMBC Spectrum of Grassystatin B (2) in CDCl ₃	S24
ROESY Spectrum of Grassystatin B (2) in CDCl ₃	S25
TOCSY Spectrum of Grassystatin B (2) in CDCl ₃	S26
¹ H NMR Spectrum of Grassystatin C (3) in CDCl ₃	S27
COSY Spectrum of Grassystatin C (3) in CDCl ₃	S28
Edited HSQC Spectrum of Grassystatin C (3) in CDCl ₃	S29
HMBC Spectrum of Grassystatin C (3) in CDCl ₃	S30
ROESY Spectrum of Grassystatin C (3) in CDCl ₃	S31
TOCSY Spectrum of Grassystatin C (3) in CDCl ₃	S32

Table S1. NMR Spectral data for Grassystatin A (**1**) at 600MHz in DMSO-*d*₆

C/H no.	δ_{H} (J in Hz)	δ_{C} ^a mult.	¹ H- ¹ H COSY	HMBC ^b	ROESY
O-Me-Pro	1		172.4, s		
	2	4.27, dd (7.8, 7.8)	58.8, d	H-3a, H-3b	1, 3, 4
	3a	2.22, m	28.3, t	H-3b, H ₂ -4	4
	3b	1.72, m		H-3a, H ₂ -4	
	4	1.77 (2H)	24.7, t	H-3a, H-3b, H-5a, H-5b	3
	5a	3.34, m	46.4, t	H ₂ -4, H-5b	3
	5b	3.18, m		H ₂ -4, H-5a	
N-Me-Phe	6	3.61, s	51.5, q		6
	7		167.9, s		
	8	5.51, dd (9.7, 5.8)	55.1, d	H-9a, H-9b	7, 9, 10, 16, 17
	9a	2.99, dd (−14.0, 5.8)	33.8, t	H-8, H-9b	8, 10, 11/15
	9b	2.80, dd (−14.0, 9.7)		H-8, H-9a	7, 8, 10, 11/15
	10		137.8, s		
	11/15	7.20, m	129.4, d	H-12/14	9, 10, 11/15, 12/14, 13
	12/14	7.19, m	127.6, d	H-11/15, H-13	10, 11/15, 12/14, 13
	13	7.13, m	125.9, d	H-12/14	
	16	2.89, s	29.5, q		8, 17
Ala	17		172.1, s		
	18	4.54, dq (7.0, 6.9)	44.3, d	H ₃ -19, NH	19
	19	0.74, d (6.9)	16.1, q	H-18	17, 18
	NH	7.94, d (7.0)		H-18	21
Thr	21		169.7, s		
	22	4.15, dd (8.4, 4.6)	57.6, d	H-23, NH	21, 23, 24, 25
	23	3.85, qdd (6.3, 5.1, 4.6)	66.5, d	H-22, H ₃ -24, OH	21
	24	0.97, d (6.3)	19.2, q	H-23	22, 23
	OH	4.63, d (5.1)		H-23	22, 23, 24
	NH	7.61, d (8.4)		H-22	25
Sta	25		171.3, s		
	26	2.18, m (2H)	39.7, t	H-27	25, 27
	27	3.75, m	69.2, d	H ₂ -26, OH	
	28	3.76, m	50.3, d	H-29a, H-29b, NH	
	29a	1.32, m	39.5, t	H-28, H-29b, H-30	
	29b	1.17, m		H-28, H-29a, H-30	
	30	1.50, m	23.7, d	H-29a, H-29b, H ₃ -31, H ₃ -32	
	31	0.82, d (6.5)	23.0, q	H-30	29
	32	0.78, d (6.1)	21.5, q	H-30	29
	OH	4.82, d (4.8)		H-27	
	NH	7.13, m		H-28	
					H ₂ -26, OH (Sta), NH (Thr)
					H ₂ -26, OH (Sta), NH (Thr)
					H-29a, H-29b, H-30, H ₃ -31, H ₃ -32, NH (Sta)
Asn	33		171.6, ^c s		
	34	4.55, ddd (7.7, 7.5, 6.3)	49.5, d	H-35a, H-35b, NH	33, 35, 36
	35a	2.58, dd (−15.4, 7.5)	36.3, t	H-34, H-35b	33, 34
	35b	2.34, dd (−15.4, 6.3)		H-34, H-35a	33, 34, 36
	36		170.5, ^c s		
	NH ₂ a	7.38, s			
	NH ₂ b	6.88, s			
	NH	8.29, d (7.7)		H-34	35
Leu	37		172.0, s		
	38	4.37, m	50.1, d	H-39a, H-39b, NH	37, 39
	39a	1.47, m	40.7, t	H-38, H-39b, H-40	41, 42
	39b	1.35, m		H-38, H-39a, H-40	42
	40	1.54, m	23.6, d	H-39a, H-39b, H ₃ -41, H ₃ -42	
	41	0.82, d (6.5)	23.0, q	H-40	39
	42	0.78, d (6.1)	21.0, q	H-40	39
	NH	7.85, m		H-38	43
					H-38, H-39a, H-40, H-44, H ₃ -47, H-49 (w) ^f
Hiva-1	43		168.3, s		
	44	4.88, d (4.4)	77.9, d	H-45	43, 45, 46, 47
	45	2.13, qdd (7.0, 6.9, 4.4)	29.9, d	H-44, H ₃ -46, H ₃ -47	46, 47
	46	0.88, d (6.9)	18.5, q	H-45	44, 45, 47
	47	0.86, d (7.0)	16.5, ^d q	H-45	44, 45
Hiva-2	48		168.8, s		
	49	5.04, d (3.9)	75.7, d	H-50	48, 50, 51, 52, 53
	50	2.26, qdd (7.3, 6.8, 3.9)	29.3, d	H-49, H ₃ -51, H ₃ -52	51, 52
	51	0.99, d (6.8)	18.6, q	H-50	49, 50, 52
	52	0.92, d (7.3)	16.5, ^d q	H-50	49, 50, 51
					H-50
N,N-DiMe-Val	53		170.6, s		
	54	2.76, d (10.7)	73.2, d	H-55	53, 55, 57, 58/59
	55	1.91, dq (10.7, 7.3, 6.6)	26.9, d	H-54, H ₃ -56, H ₃ -57	54, 57
	56	0.91, d (7.3)	18.9, q	H-55	54, 55, 57
	57	0.84, d (6.6)	19.2, q	H-55	54, 55
	58/59	2.23, s	40.6, q		54, 58/59
					H-54

^aMultiplicity derived from edited HSQC. ^bProtons showing long-range correlation to indicated carbon. ^cThere is insufficient information to distinguish between carbons 33 and 36. ^dThese carbons have the same chemical shift. ^e“w” denotes a weak correlation.

Table S2. Conditions Employed for Protease Screen and IC₅₀ Assays

Protease	Substrate	Ex/Em	Buffer	[Sub] (μ M)	Positive control
Activated Protein C in 50% gly	Boc-DVLR-ANSNH-C ₄ H ₉	355/460	B	50	Gabexate mesylate
ACE1	MCA-RPPGFSAFK(Dnp),	320/405	A	10	Captopril
ADAM-9	MCA-PLAQAV-Dpa-RSSSR-NH ₃	320/405	I	10	GM6001
ADAM-10	MCA-PLAQAV-Dpa-RSSSR-NH ₃	320/405	I	10	GM6001
BACE 1	MCA-SEVNLDAEFRK(Dnp)-RR-NH ₂	320/405	J	10	Pepstatin A
Calpain 1	BioMol, N-Succinyl-Leu-Tyr-AMC	355/460	K	10	E-64
Caspase 1	Ac-LEHD-AMC	355/460	G	5	IETD-CHO
Caspase 2	Ac-LEHD-AMC	355/460	G	5	IETD-CHO
Caspase 3	Ac-DEVD-AMC	355/460	F	5	DEVD-CHO
Caspase 4	Ac-LEHD-AMC	355/460	G	5	IETD-CHO
Caspase 5	Ac-LEHD-AMC	355/460	G	5	IETD-CHO
Caspase 6	Ac-LEHD-AMC	355/460	G	5	IETD-CHO
Caspase 7	Ac-DEVD-AMC	355/460	F	5	DEVD-CHO
Caspase 8	Ac-LEHD-AMC	355/460	G	5	IETD-CHO
Caspase 9	Ac-LEHD-AMC	355/460	G	5	IETD-CHO
Caspase 10	Ac-LEHD-AMC	355/460	G	5	IETD-CHO
Caspase 11	Ac-LEHD-AMC	355/460	G	5	IETD-CHO
Caspase 14	Ac-LEHD-AMC	355/460	G	5	IETD-CHO
Cathepsin B	Z-FR-AMC (Biomol)	355/460	L	10	E-64
Activated Cathepsin C	Z-FR-AMC (Biomol)	355/460	L	10	E-64
Cathepsin D	MCA-KPILFFRLK(Dnp)-D-R-NH ₂	320/405	P	10	Pepstatin A
Cathepsin E	MCA-KPILFFRLK(Dnp)-D-R-NH ₂	320/405	P	10	Pepstatin A
Cathepsin H	R-AMC	355/460	M	10	E-64
Activated Cathepsin K/O/O2/X	Z-GPR-AMC	355/460	C	10	E-64
Cathepsin S	Z-FR-AMC (Biomol)	355/460	M	10	E-64
Activated Cathepsin V	Z-FR-AMC (Biomol)	355/460	C	10	E-64
Activated Cathepsin X/Z	MCA-RPPGFSAFK(Dnp),	320/405	D	10	E-64
Chymase	Suc-AAPF-AMC	355/460	B	10	Chymostatin
Chymotrypsin	Suc-AAPF-AMC	355/460	B	10	Chymostatin
DPP IV	H-GP-AMC	355/460	H	10	P32/98
DPP VIII	H-GP-AMC	355/460	H	10	P32/98
DPP IX	H-GP-AMC	355/460	H	10	P32/98
Elastase (Porcine Pancreatic)	AR-AMC	355/460	B	10	Gabexate mesylate
Factor Xa	CH ₃ SO ₂ -D-CHA-Gly-Arg-AMC-AcOH	355/460	N	10	Gabexate mesylate
Factor Xla	(Boc-Glu(OBzl)-Ala-Arg)-MCA	355/460	A	10	Gabexate mesylate
Activated Kallikrein 4	VPR-AMC	380/460	E	10	Gabexate mesylate
Kallikrein 5	Z-VVR-AMC	355/460	A	10	Gabexate mesylate
Activated Kallikrein 8	VPR-AMC	380/460	E	10	Gabexate mesylate
MMP1	(5-FAM/QXLTM) FRET peptide [QXL520-g-Abu-P-Cha-Abu-Smc-H-A-Dab(5-FAM)-A-K-NH ₂]	485/520	H	5	GM6001
MMP2	(5-FAM/QXLTM) FRET peptide	485/520	H	5	GM6001
MMP3	(5-FAM/QXLTM) FRET peptide	485/520	H	5	GM6001
MMP7	(5-FAM/QXLTM) FRET peptide	485/520	H	5	GM6001
MMP8 (Catalitic domain)	(5-FAM/QXLTM) FRET peptide	485/520	H	5	GM6001
MMP9 (Catalitic domain)	(5-FAM/QXLTM) FRET peptide	485/520	H	5	GM6001
MMP10 (Catalitic domain)	(5-FAM/QXLTM) FRET peptide	485/520	H	5	GM6001
MMP11 (Catalitic domain)	(5-FAM/QXLTM) FRET peptide	485/520	H	5	GM6001
MMP12 (Catalitic domain)	(5-FAM/QXLTM) FRET peptide	485/520	H	5	GM6001
MMP13 (Catalitic domain)	(5-FAM/QXLTM) FRET peptide	485/520	H	5	GM6001
MMP14 (Catalitic domain)	(5-FAM/QXLTM) FRET peptide	485/520	H	5	GM6001

Papain (Cysteine Protease)	Z-FR-AMC	355/460	M	10	E-64
Activated Plasma Kallikrein	Z-FR-AMC	380/460	A	10	Gabexate mesylate
Plasmin (Human)	H-D-CHA -Ala-Arg-AMC.2AcOH	355/460	A	10	Gabexate mesylate
Renin (Human)	RE(EDANS)IHPFHLVIHTK(DABCYL)R	340/520	A	50	Pepstatin A
TACE	MCA-PLAQAV-Dpa-RSSSR-NH ₂	320/405	I	10	GM6001
Thrombin alpha	H-D-CHA -Ala-Arg-AMC.2AcOH	355/460	O	10	Gabexate mesylate
Trypsin (Bovine Pancreatic)	H-D-CHA -Ala-Arg-AMC.2AcOH	355/460	A	10	Gabexate mesylate
Tryptase beta 2 (Serin Protease)	Z-GPR-AMC	355/460	A	10	Gabexate mesylate
Tryptase gamma 1 (Human lung)	Z-GPR-AMC,	355/460	A	10	Gabexate mesylate
Urokinase	Bz-b-Ala-Gly-Arg-AMC.AcOH	355/460	A	10	Gabexate mesylate

A	25 mM Tris pH 8.0, 100 mM NaCl, 0.01% Brij35
B	100 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10 mM CaCl ₂ , 0.025% CHAPS, 1.5 mM DTT
C	25 mM Sodium Acetate pH 5.5, 0.1 M NaCl, 5 mM DTT
D	25 mM Sodium Acetate pH 3.5, 5 mM DTT
E	25 mM Tris pH 9, 150 mM NaCl
F	50 mM HEPES pH 7.4, 100 mM NaCl, 0.01% CHAPS, 0.1 mM EDTA, 10 mM DTT
G	50 mM HEPES pH 7.4, 1 M Na Citrate, 100 mM NaCl, 0.01% CHAPS, 0.1 mM EDTA, 10 mM DTT
H	50 mM HEPES (pH7.5), 10 mM CaCl ₂ , 0.01% Brij-35, store at 4 °C; add 0.1 mg/ml BSA in buffer before use.
I	25 mM Tris pH 9, 2.5 μM ZnCl ₂ , 0.005% Brij
J	0.1 M Sodium Acetate, pH 4.0
K	75 mM Tris pH 7.0, 0.005% Brij35, 3 mM DTT, 0.5 mM CaCl ₂
L	25 mM MES pH 6, 50 mM NaCl, 0.005% Brij35, 5 mM DTT
M	75 mM Tris pH 7.0, 1 mM EDTA, 0.005% Brij35, 3 mM DTT
N	25 mM Tris pH 8.0, 100 mM NaCl, 0.01% Brij35, 0.25 mg/ml BSA
O	25 mM Tris pH 8.0, 100 mM NaCl, 0.01% Brij35, 2.5 mM CaCl ₂ , 1 mg/ml BSA
P	0.1 M Sodium Acetate, pH 3.5, 0.1 M NaCl

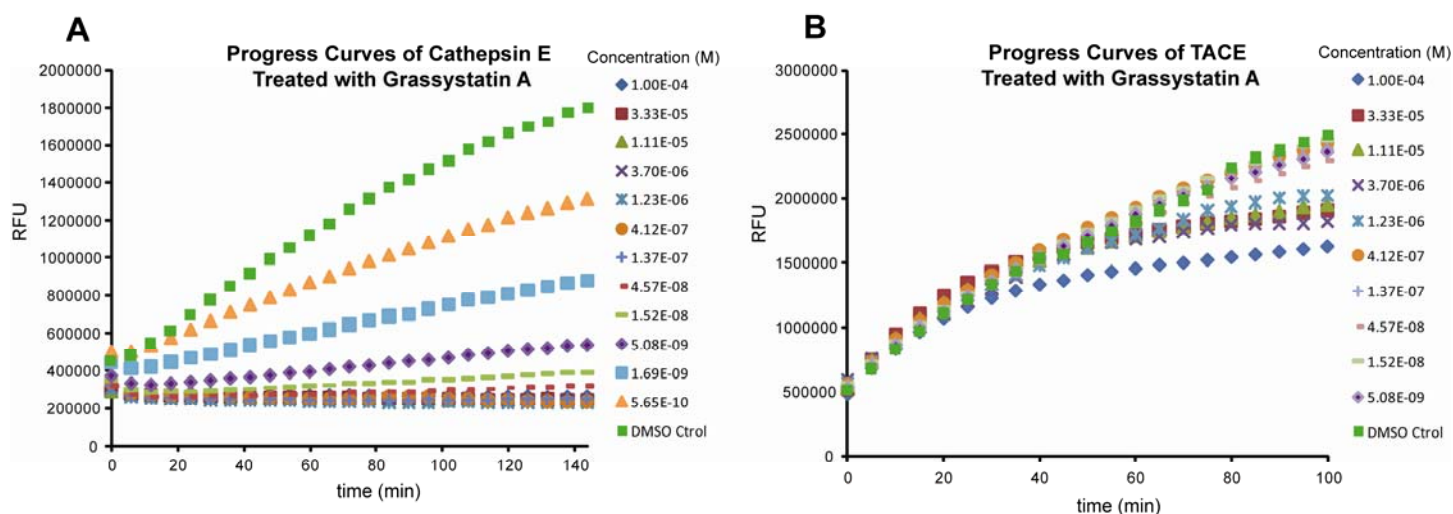


Figure S1. Progress curves of cathepsin E and TACE treated with grassystatin A (**1**). A) Inhibition of cathepsin E is not time-dependent, and initial rate is affected by **1**. B) Inhibition of TACE is time-dependent, and so initial rate is not affected by **1** as the onset of inhibition is slow.

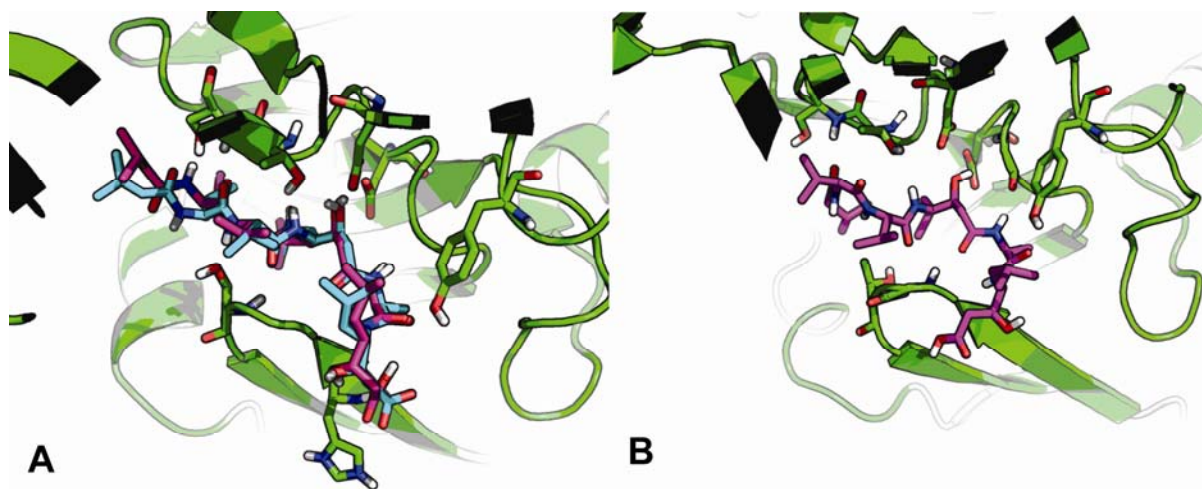
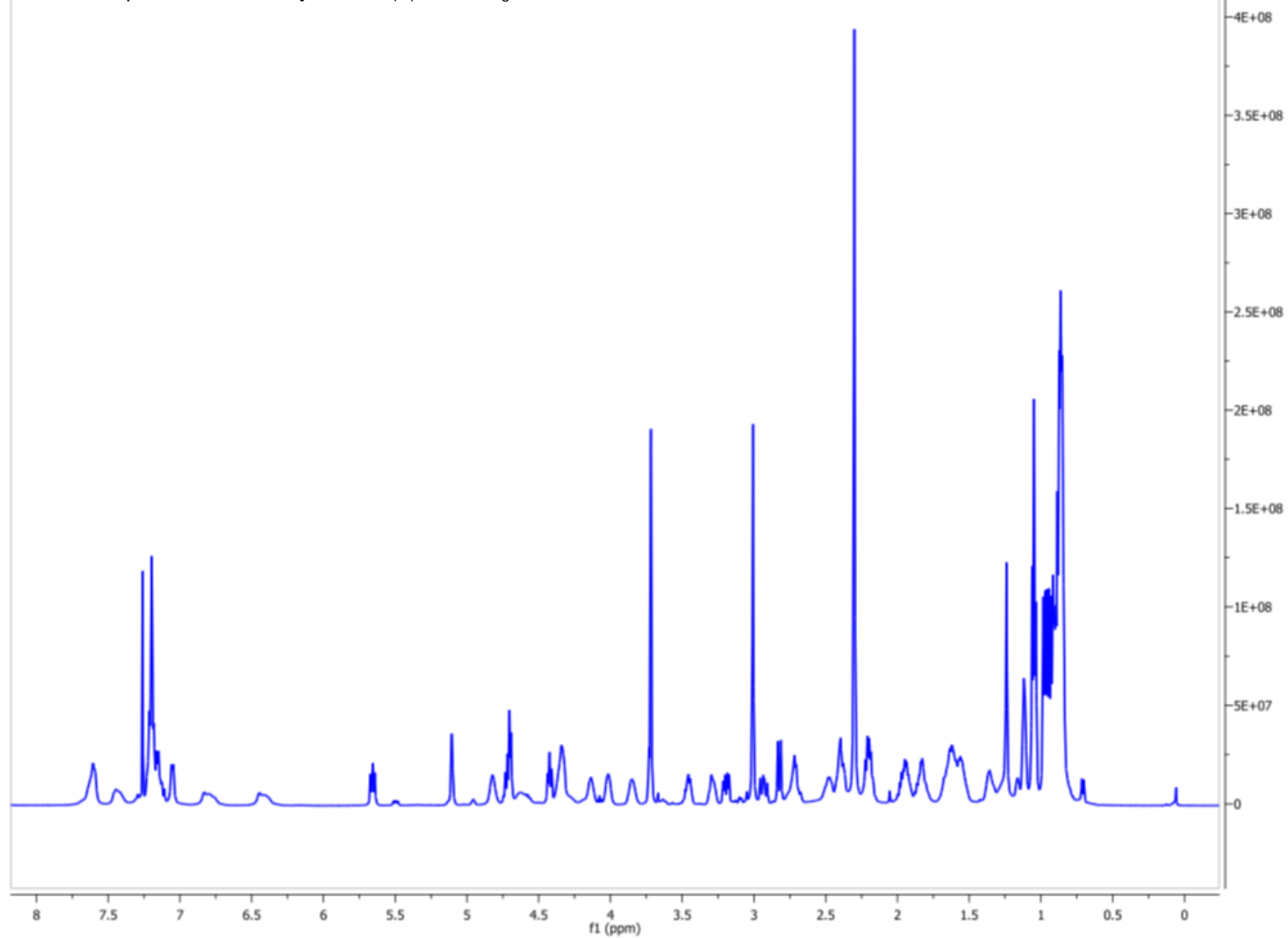
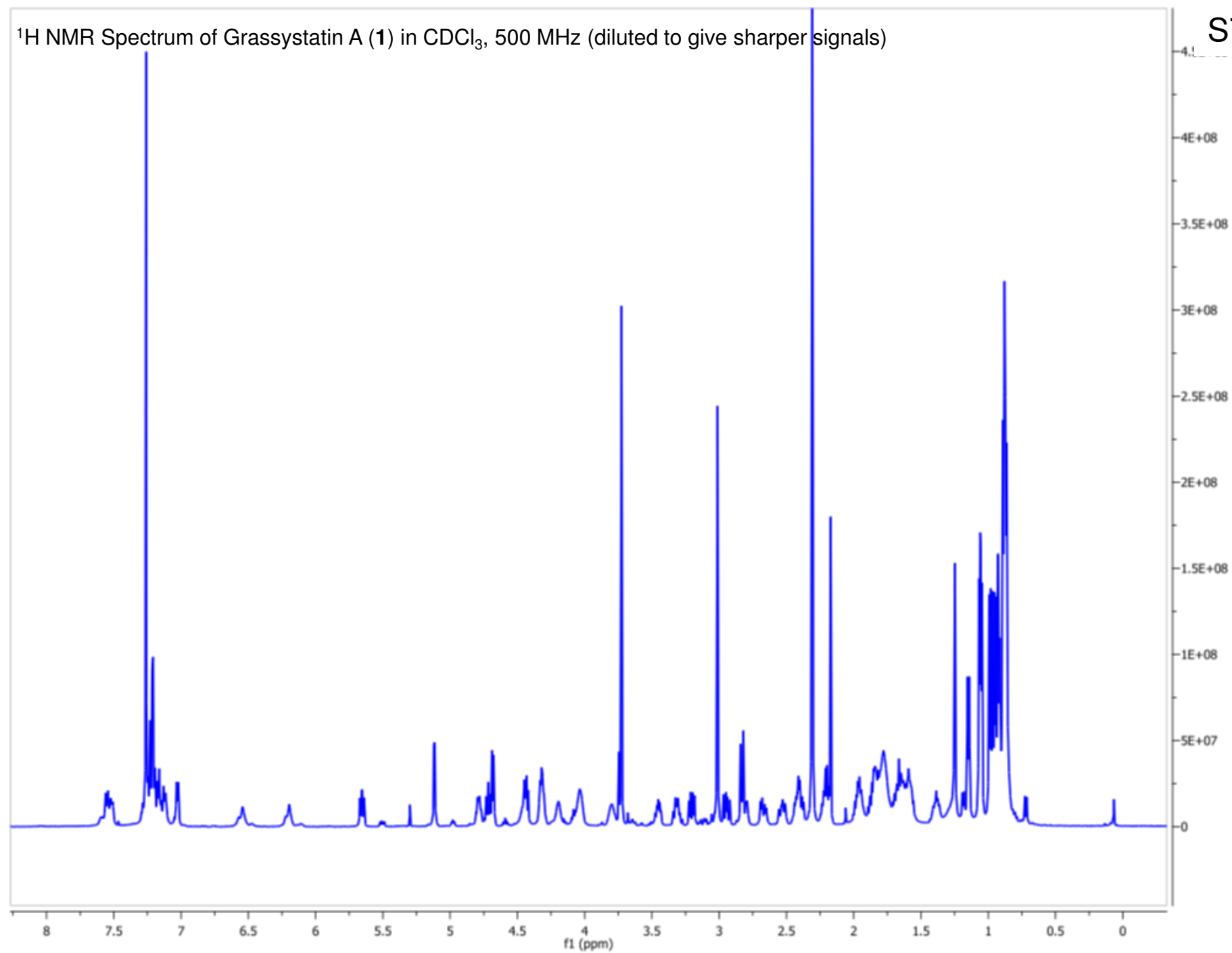


Figure S2. A) Docked structure of pepstatin A (magenta) overlaid with the X-ray structure of pepstatin A bound to cathepsin D (1LYB, cyan). Cathepsin D is shown in green. Residues for which the crystal structure suggests interact by hydrogen bonding to pepstatin A are shown. B) Docked conformation of pepstatin A (magenta) in Cathepsin E (green). Residues that are within hydrogen bonding distance to a donor/acceptor in pepstatin A are shown.

^1H NMR Spectrum of Grassystatin A (**1**) in CDCl_3 , 500 MHz

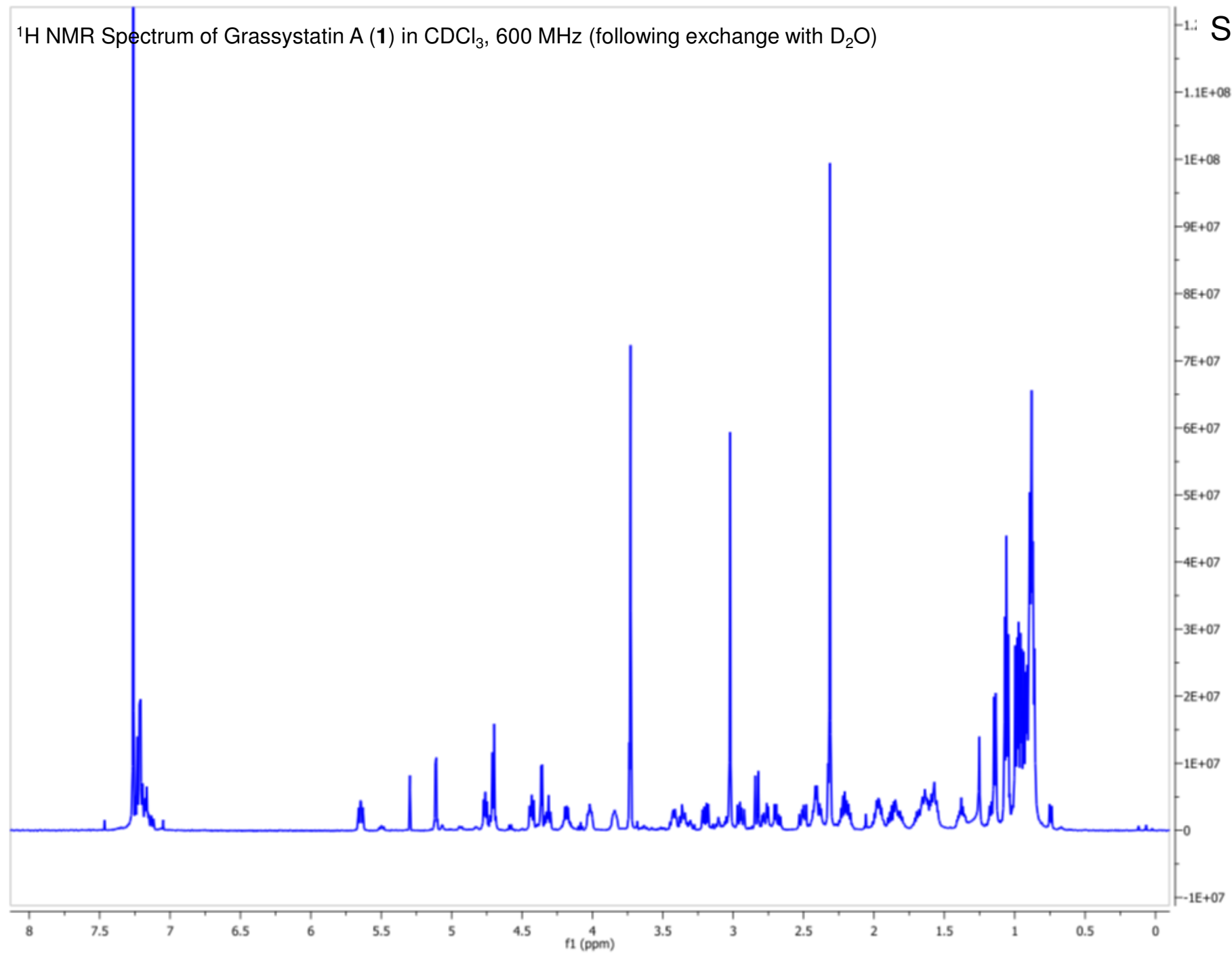


^1H NMR Spectrum of Grassystatin A (**1**) in CDCl_3 , 500 MHz (diluted to give sharper signals)

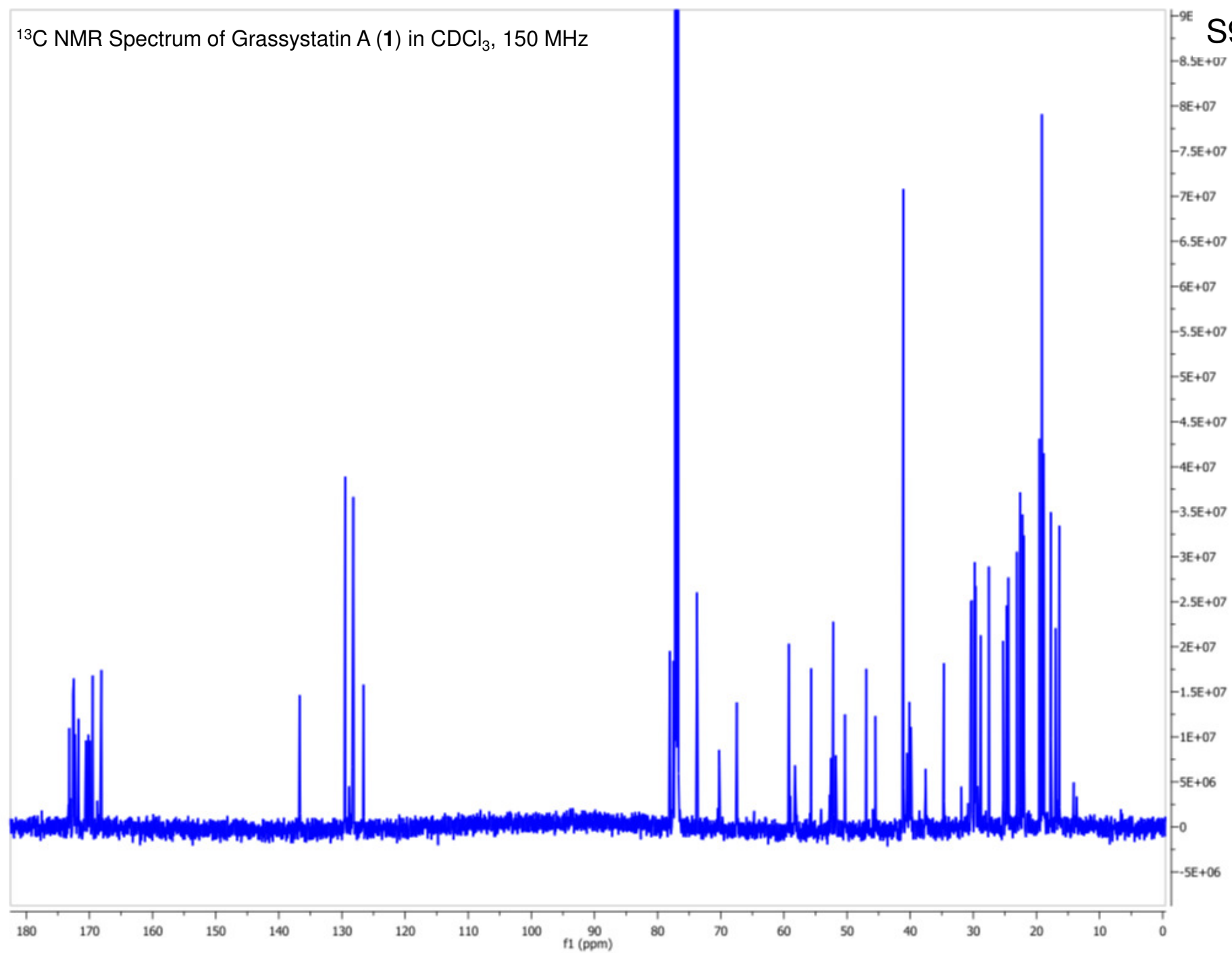


^1H NMR Spectrum of Grassystatin A (**1**) in CDCl_3 , 600 MHz (following exchange with D_2O)

S8

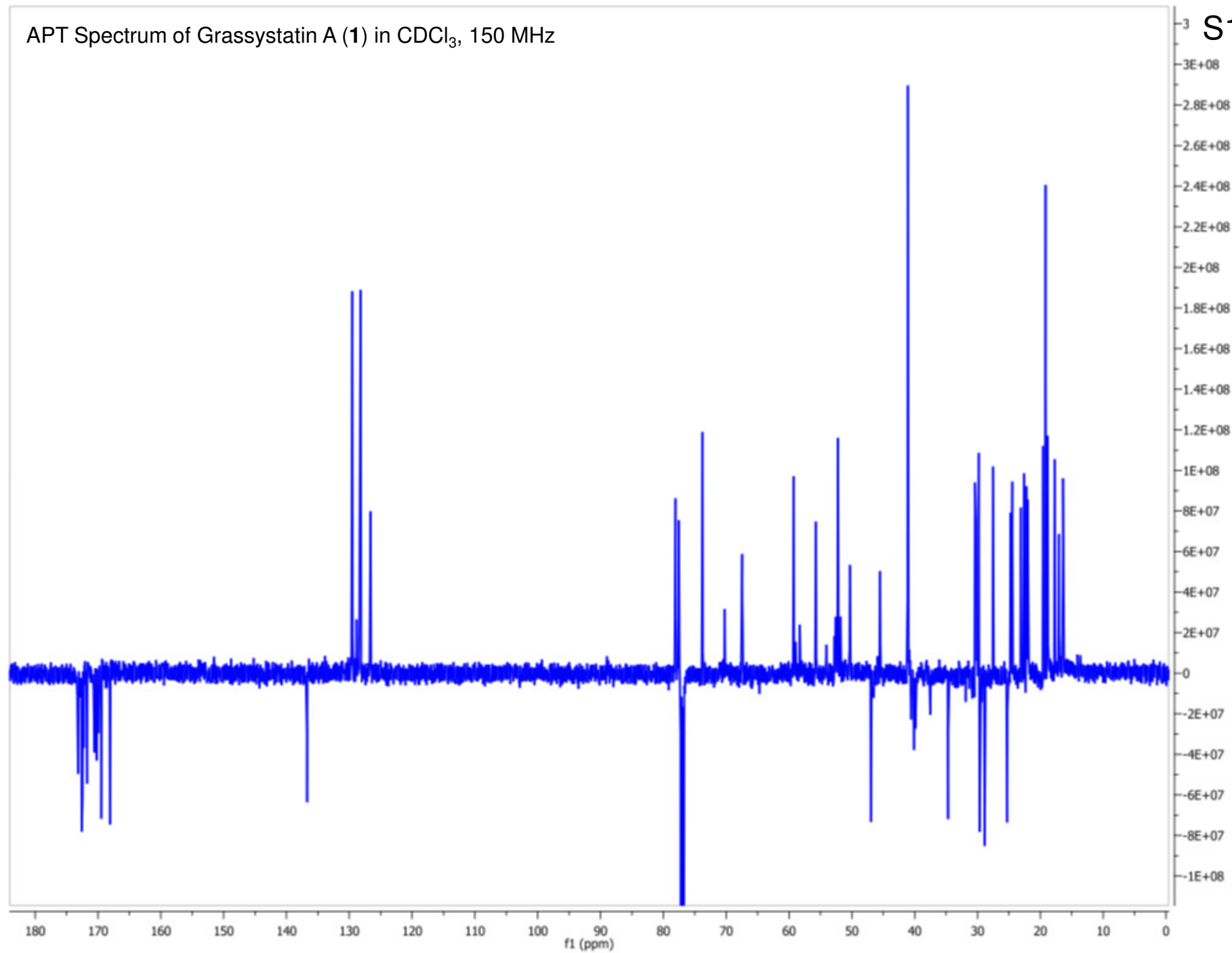


^{13}C NMR Spectrum of Grassystatin A (**1**) in CDCl_3 , 150 MHz



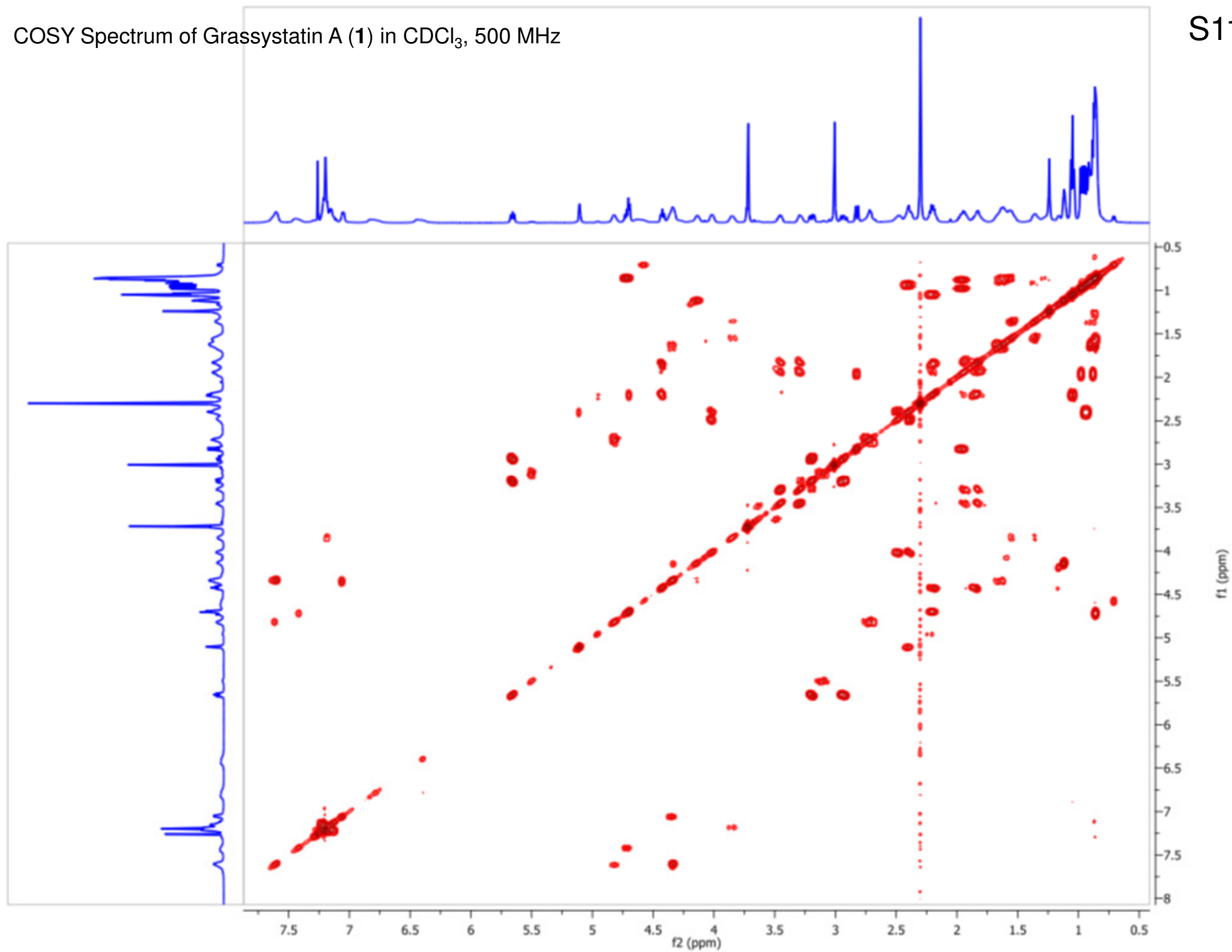
APT Spectrum of Grassystatin A (**1**) in CDCl₃, 150 MHz

S10



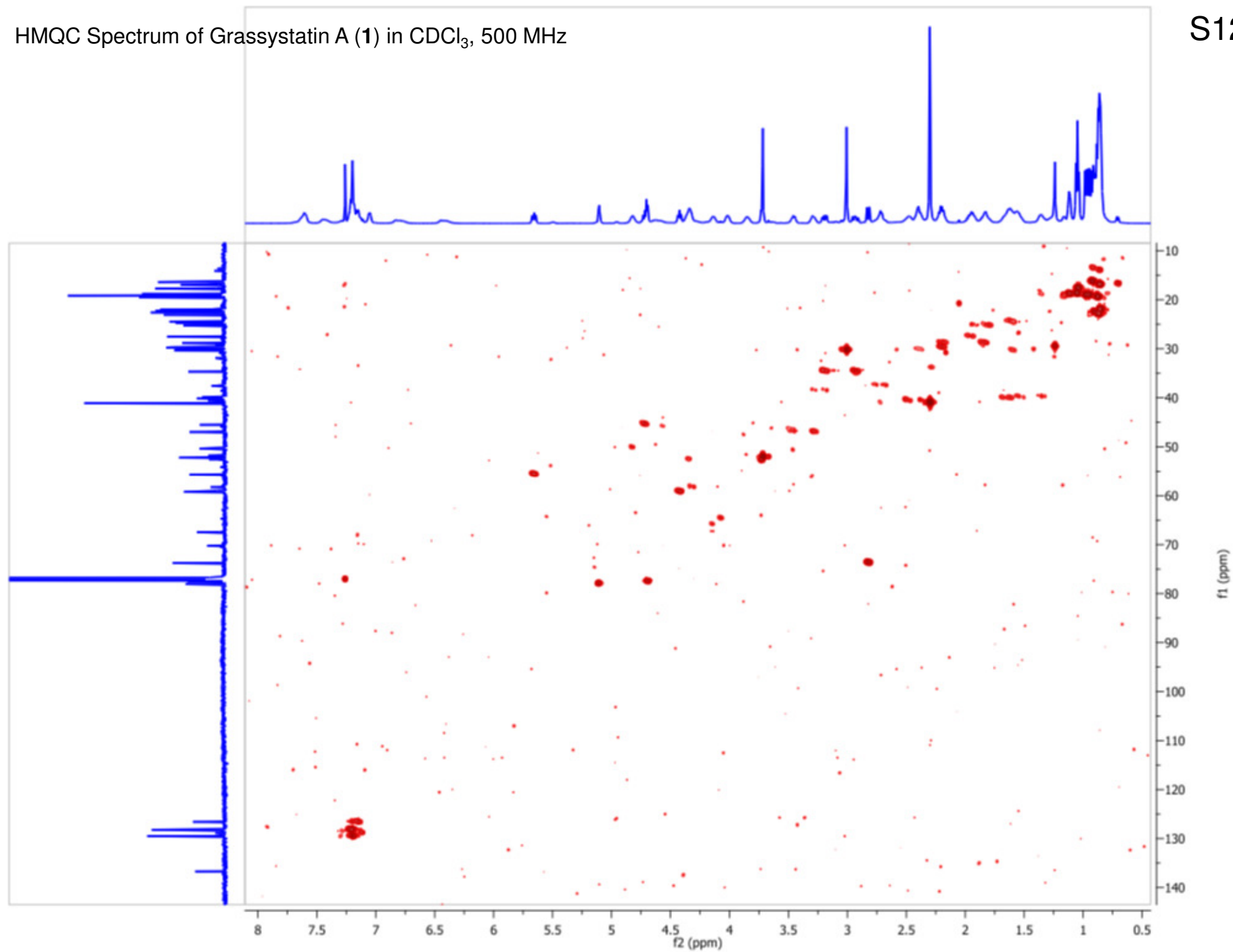
COSY Spectrum of Grassystatin A (**1**) in CDCl₃, 500 MHz

S11



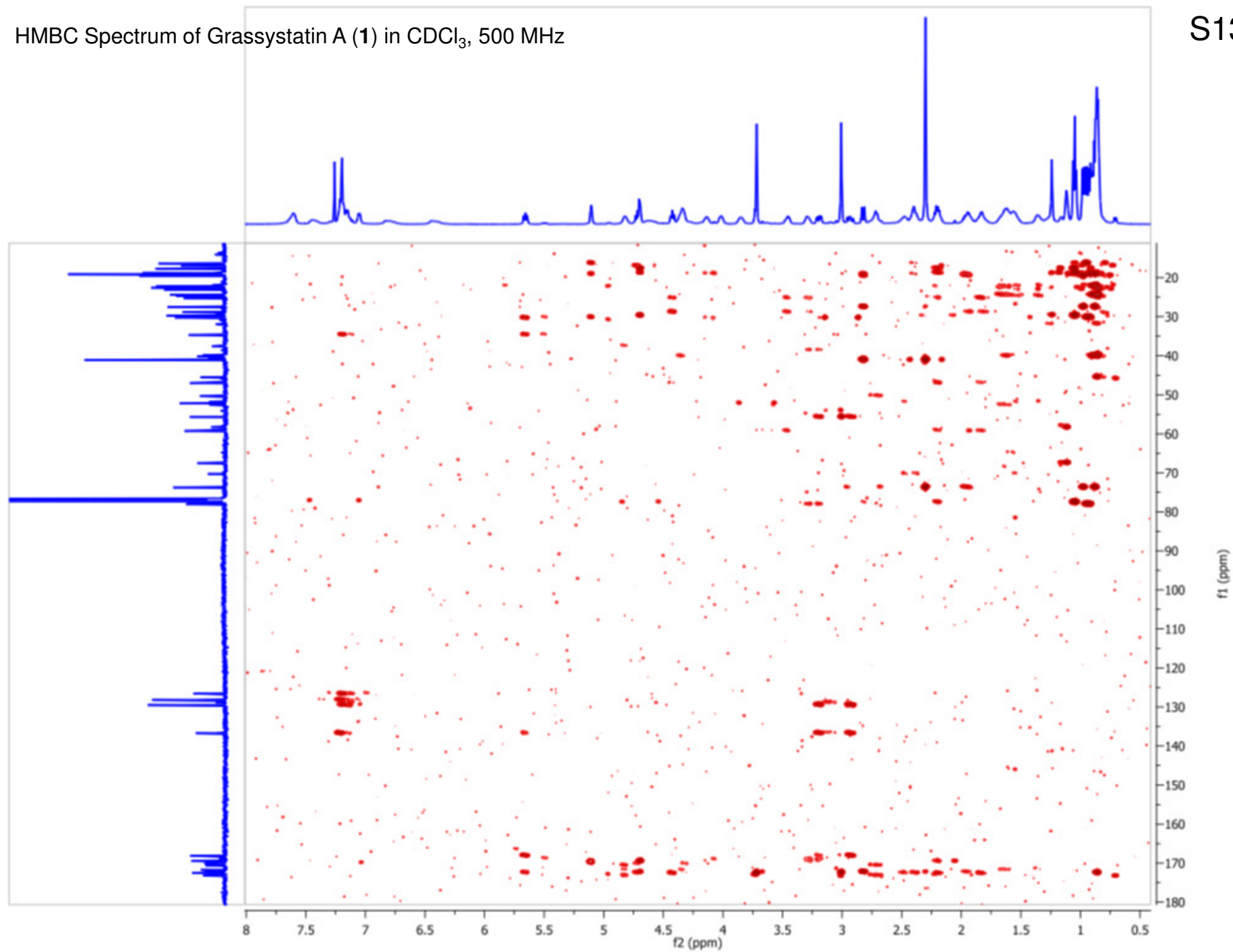
HMQC Spectrum of Grassystatin A (1) in CDCl₃, 500 MHz

S12



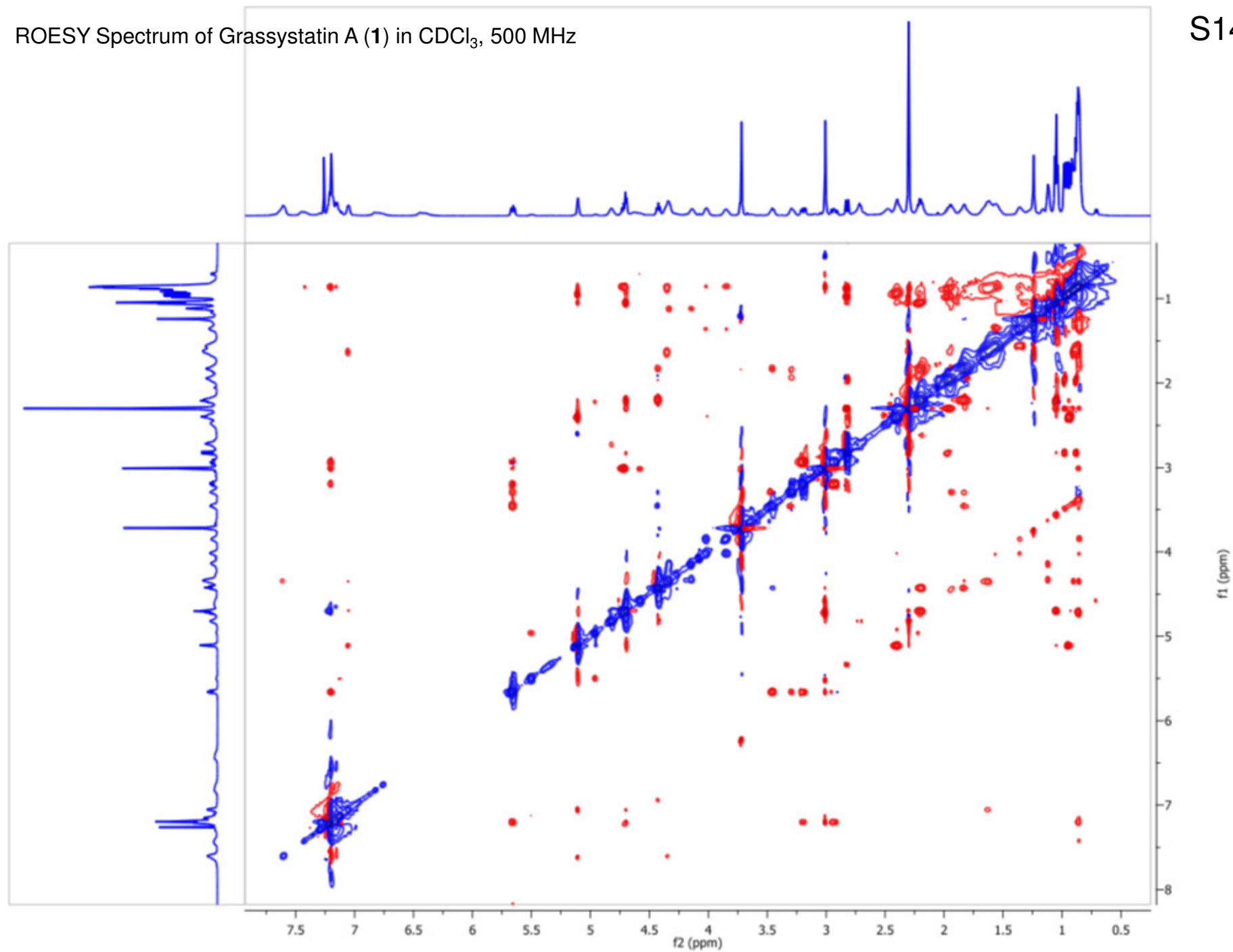
HMBC Spectrum of Grassystatin A (**1**) in CDCl₃, 500 MHz

S13



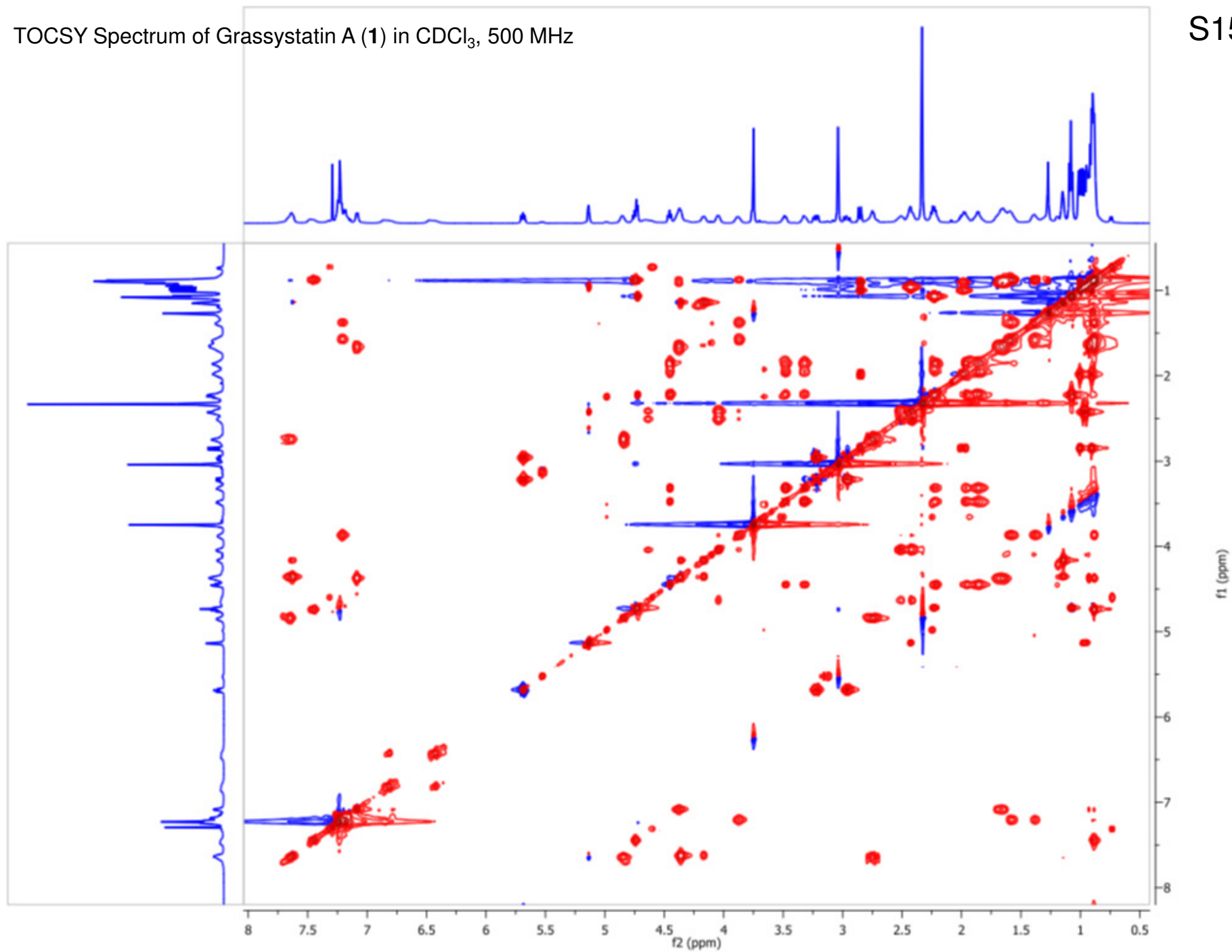
ROESY Spectrum of Grassystatin A (**1**) in CDCl₃, 500 MHz

S14



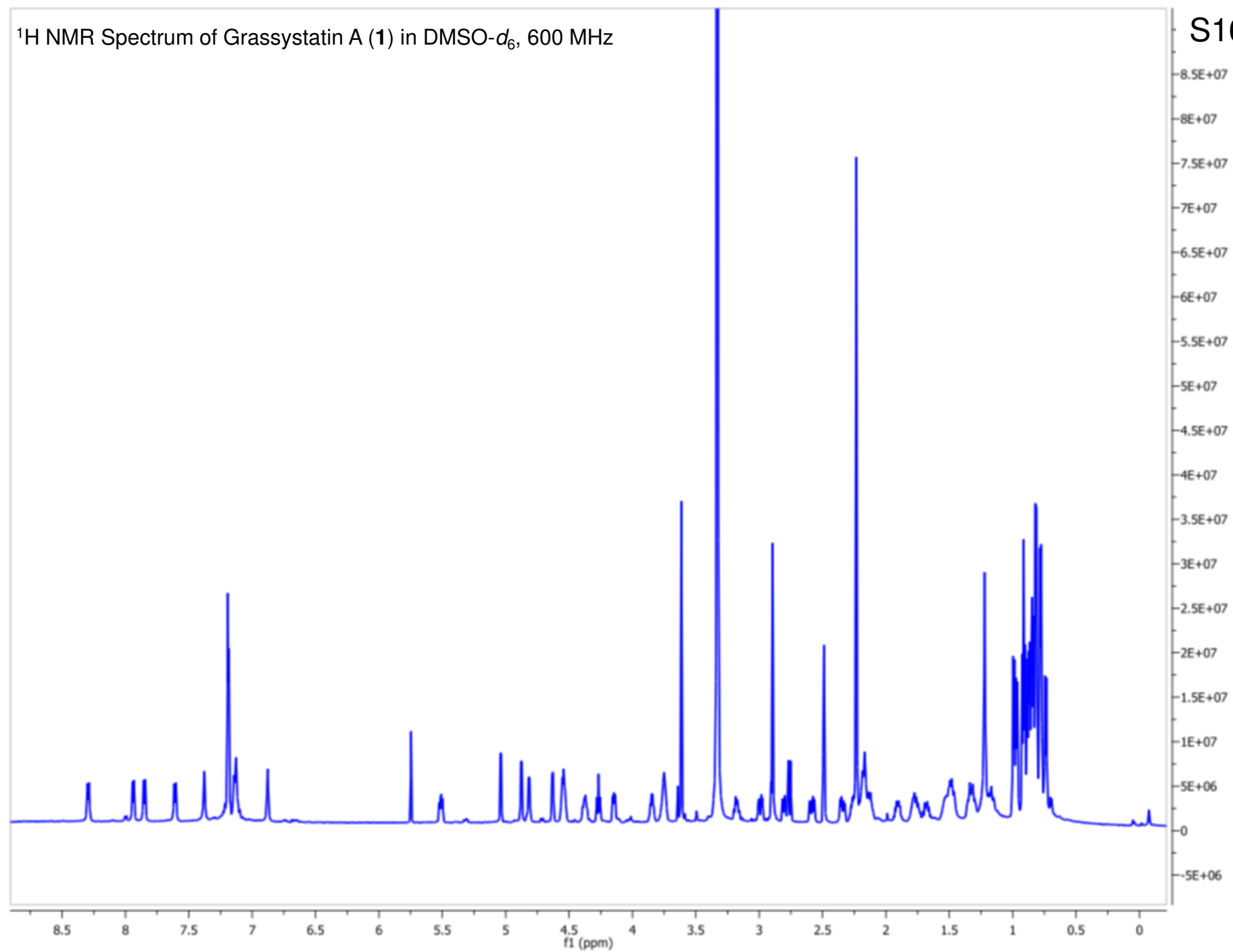
TOCSY Spectrum of Grassystatin A (**1**) in CDCl₃, 500 MHz

S15



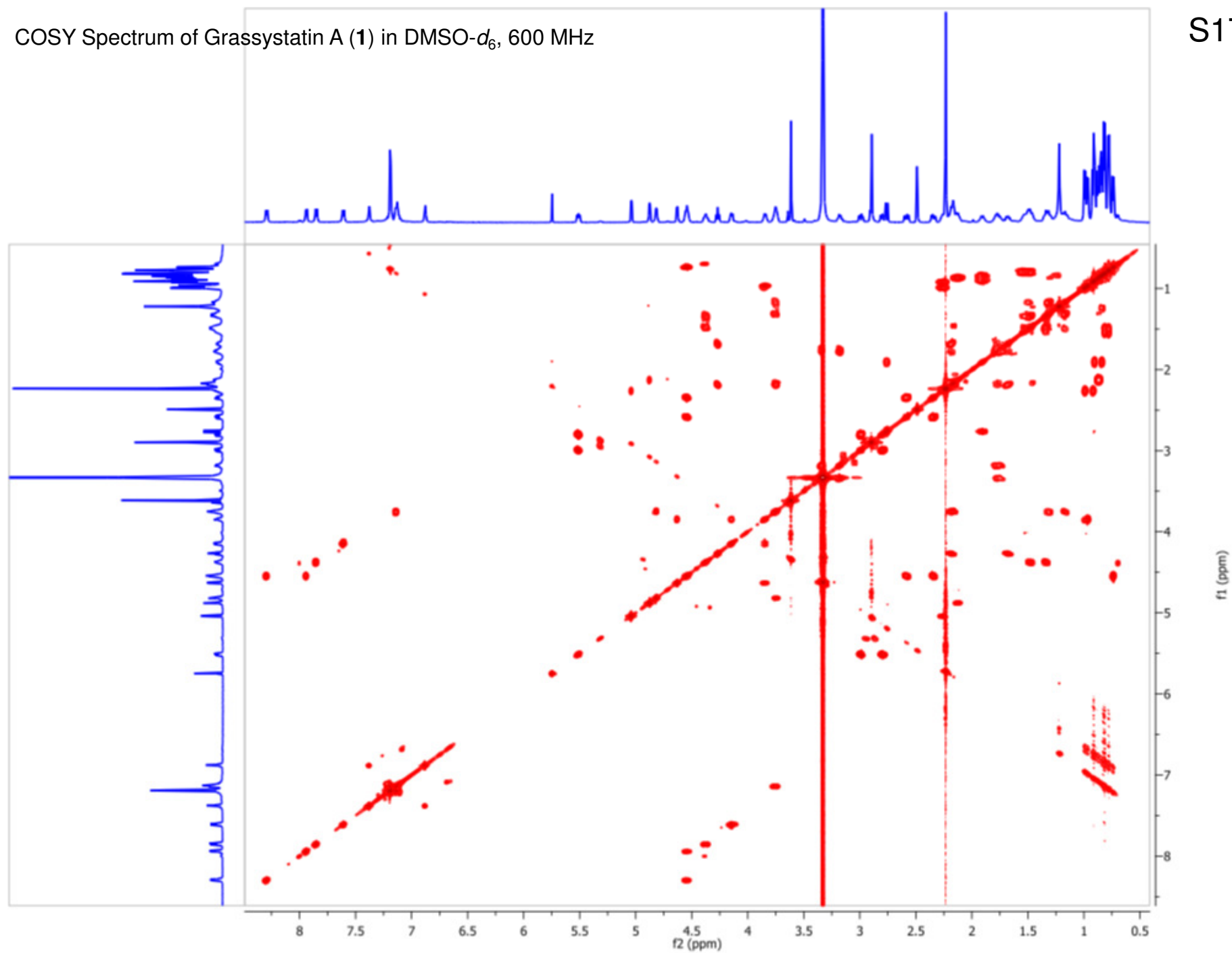
^1H NMR Spectrum of Grassystatin A (**1**) in $\text{DMSO}-d_6$, 600 MHz

S16



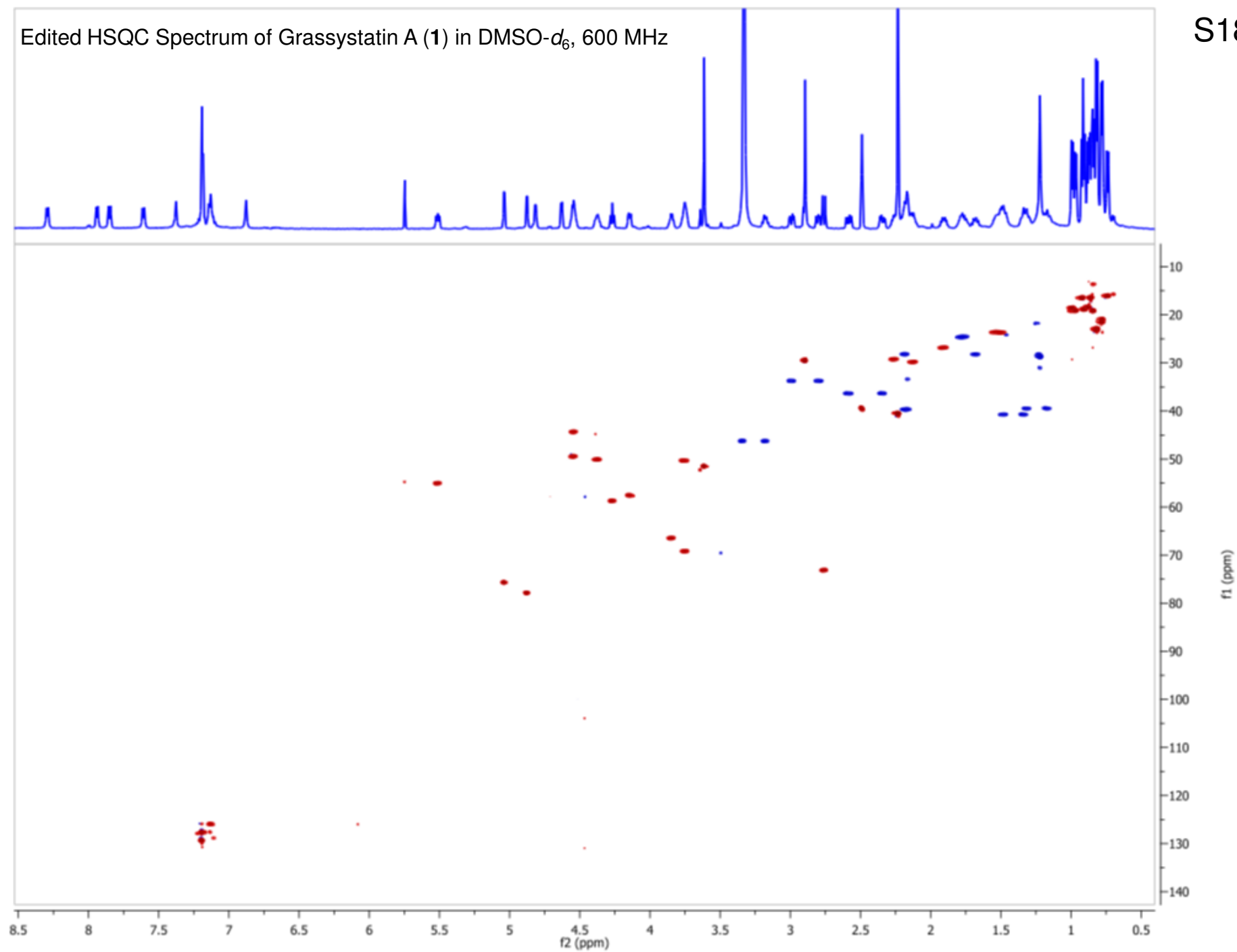
COSY Spectrum of Grassystatin A (**1**) in DMSO- d_6 , 600 MHz

S17



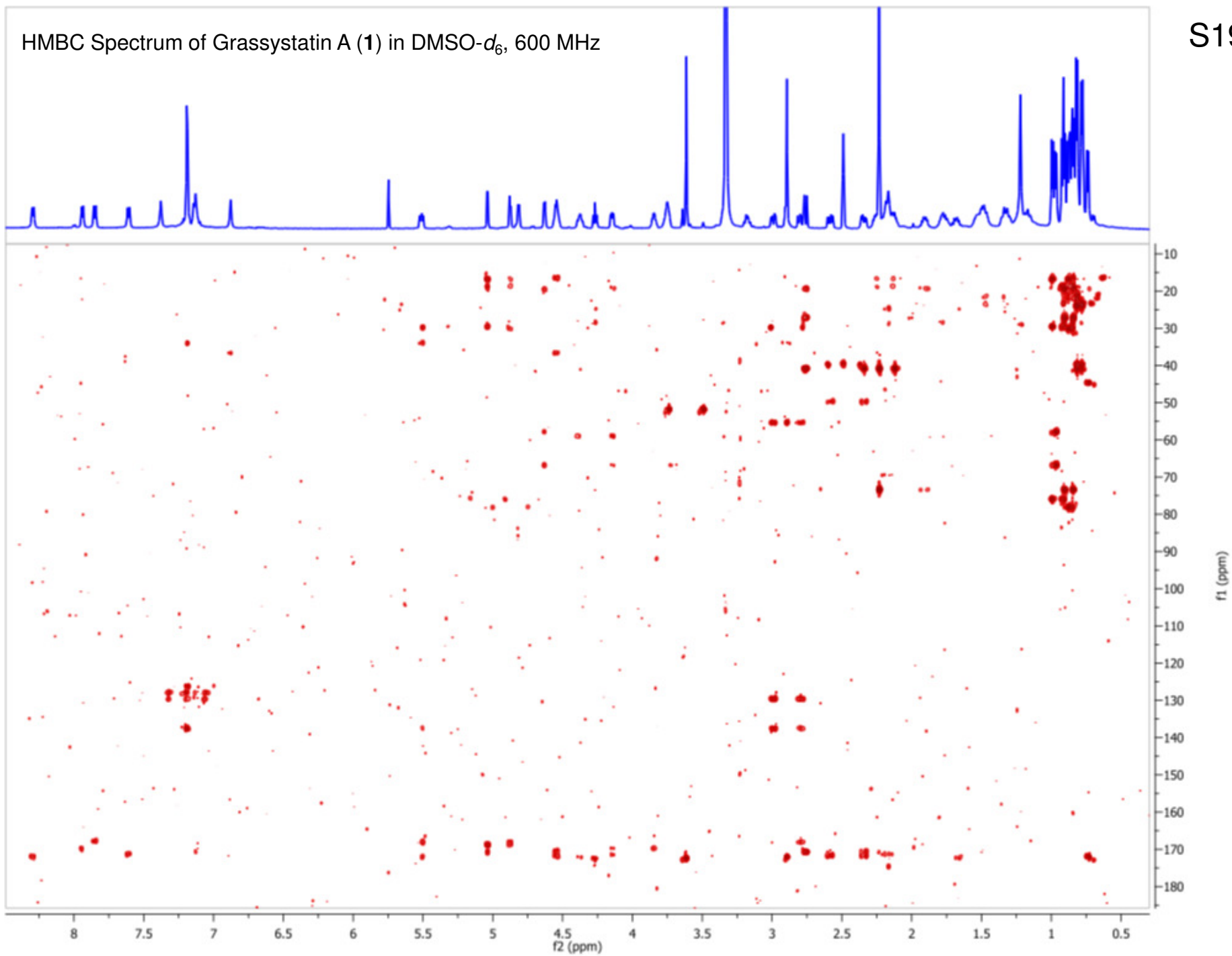
Edited HSQC Spectrum of Grassystatin A (**1**) in DMSO- d_6 , 600 MHz

S18



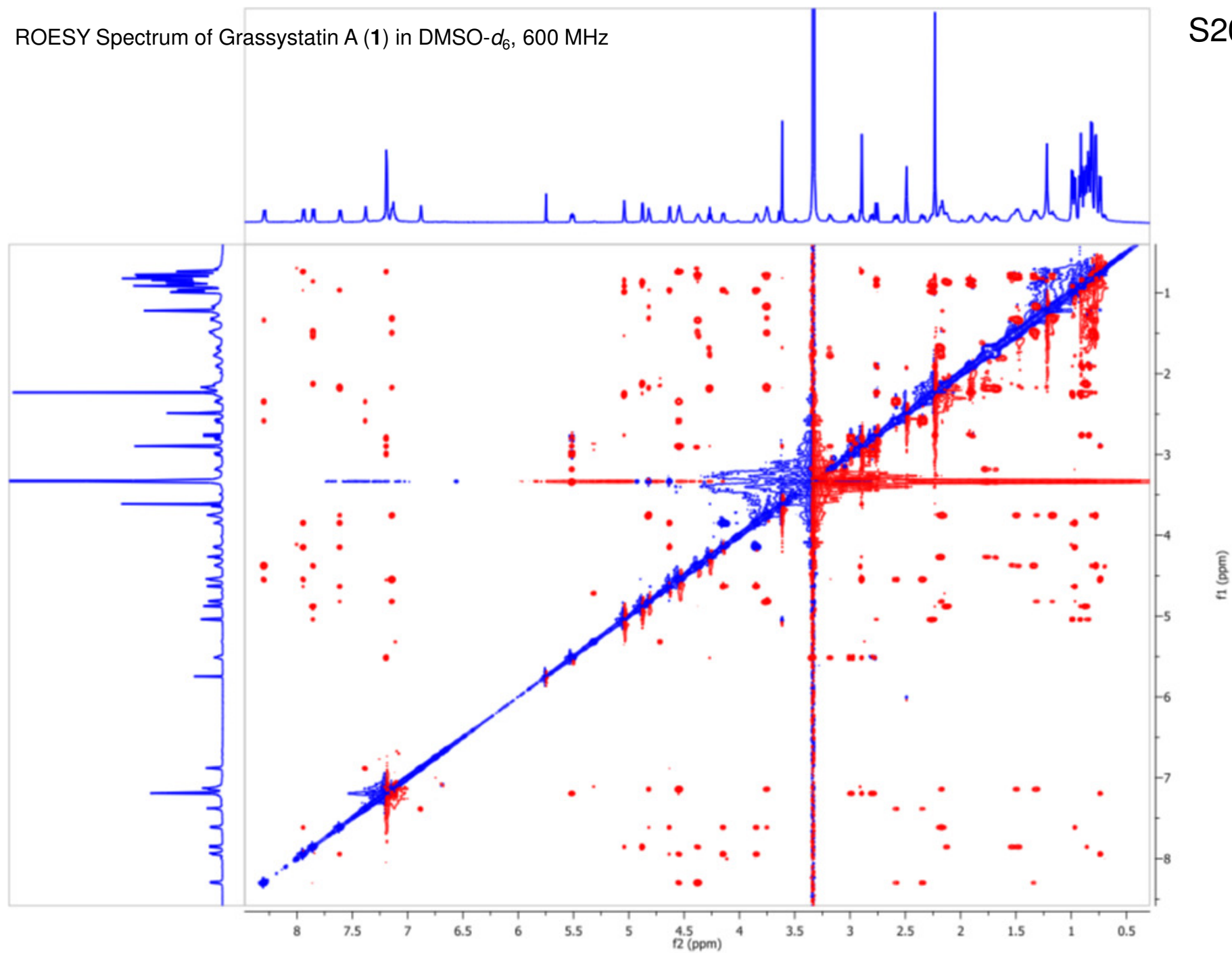
HMBC Spectrum of Grassystatin A (**1**) in DMSO- d_6 , 600 MHz

S19



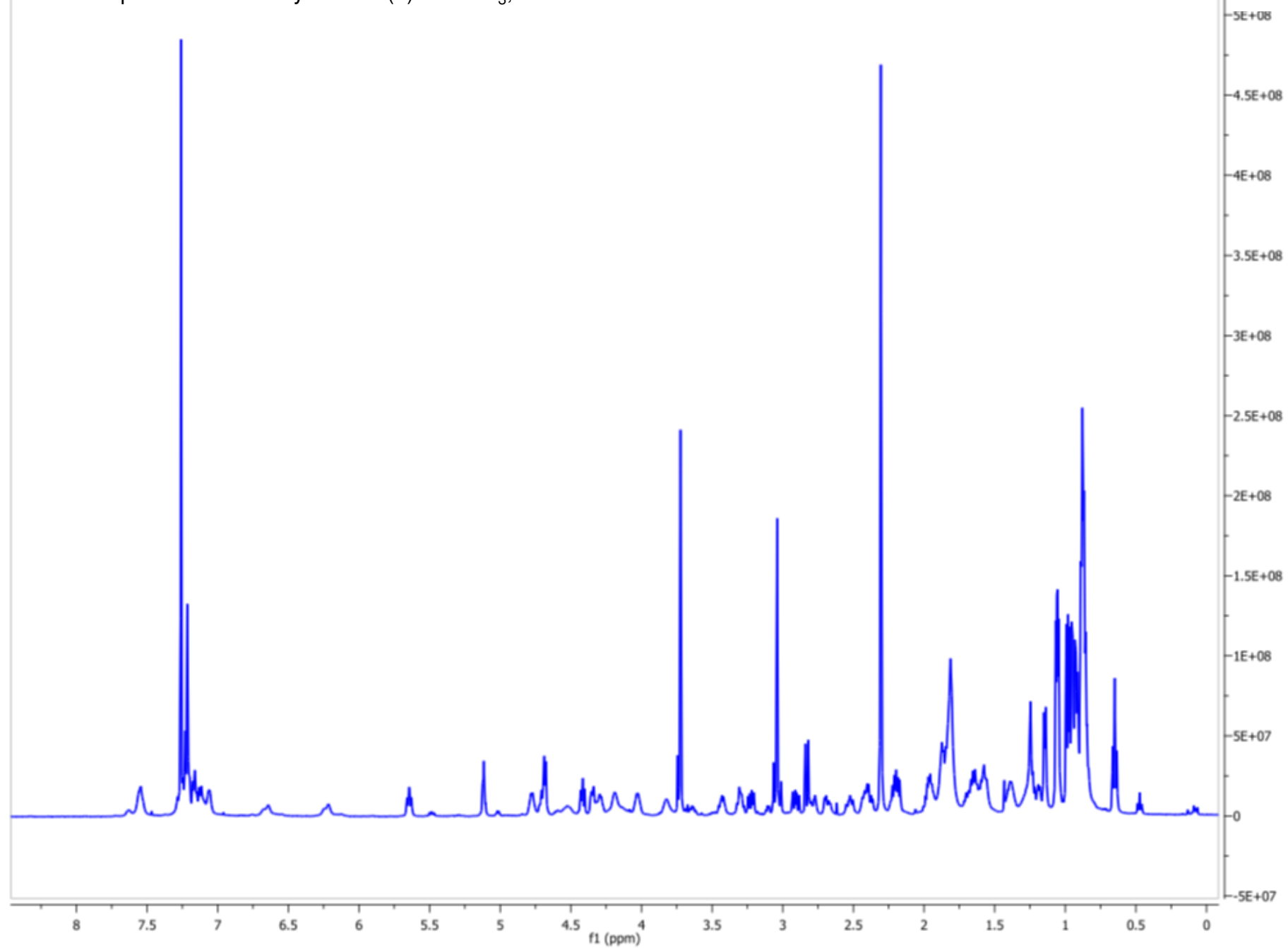
ROESY Spectrum of Grassystatin A (**1**) in DMSO- d_6 , 600 MHz

S20



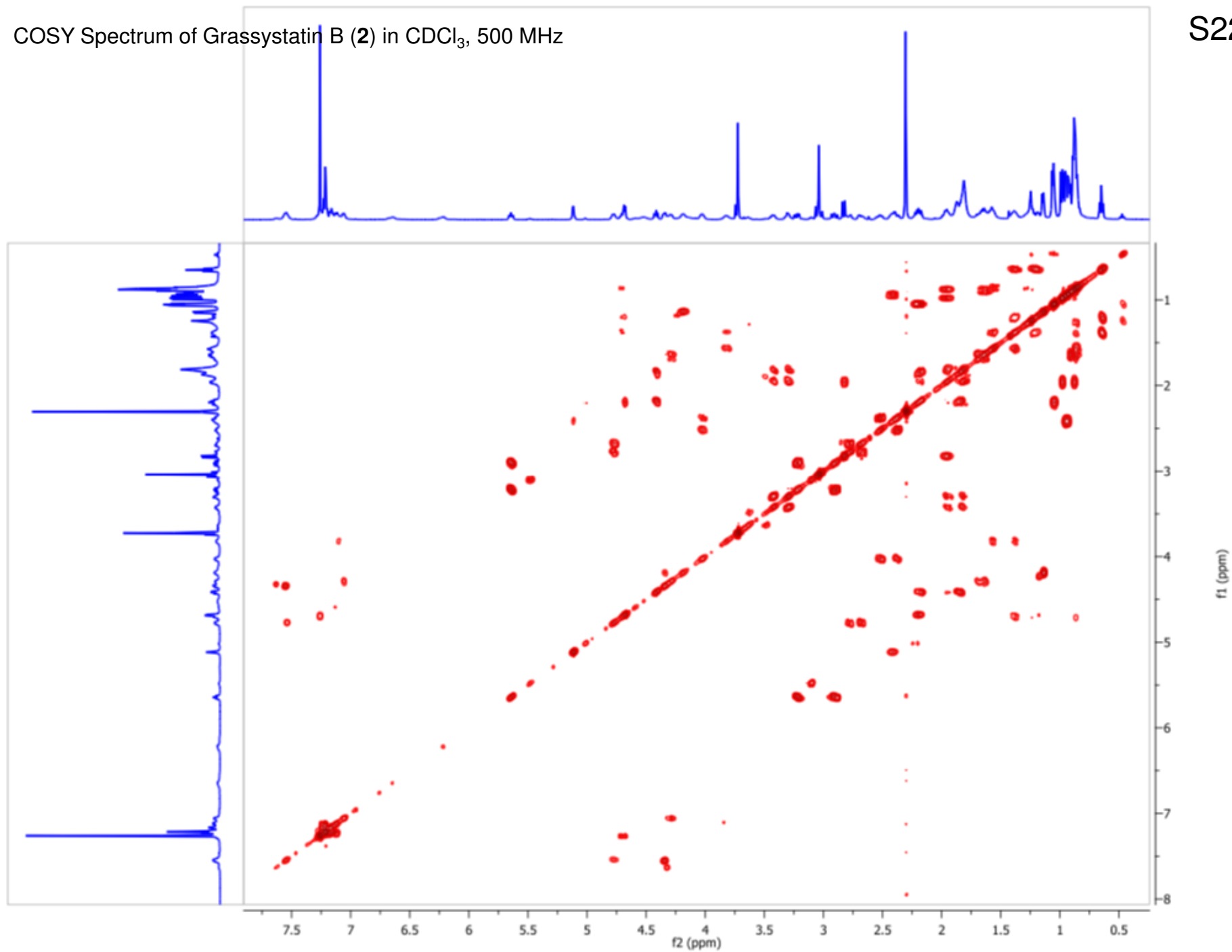
^1H NMR Spectrum of Grassystatin B (**2**) in CDCl_3 , 500 MHz

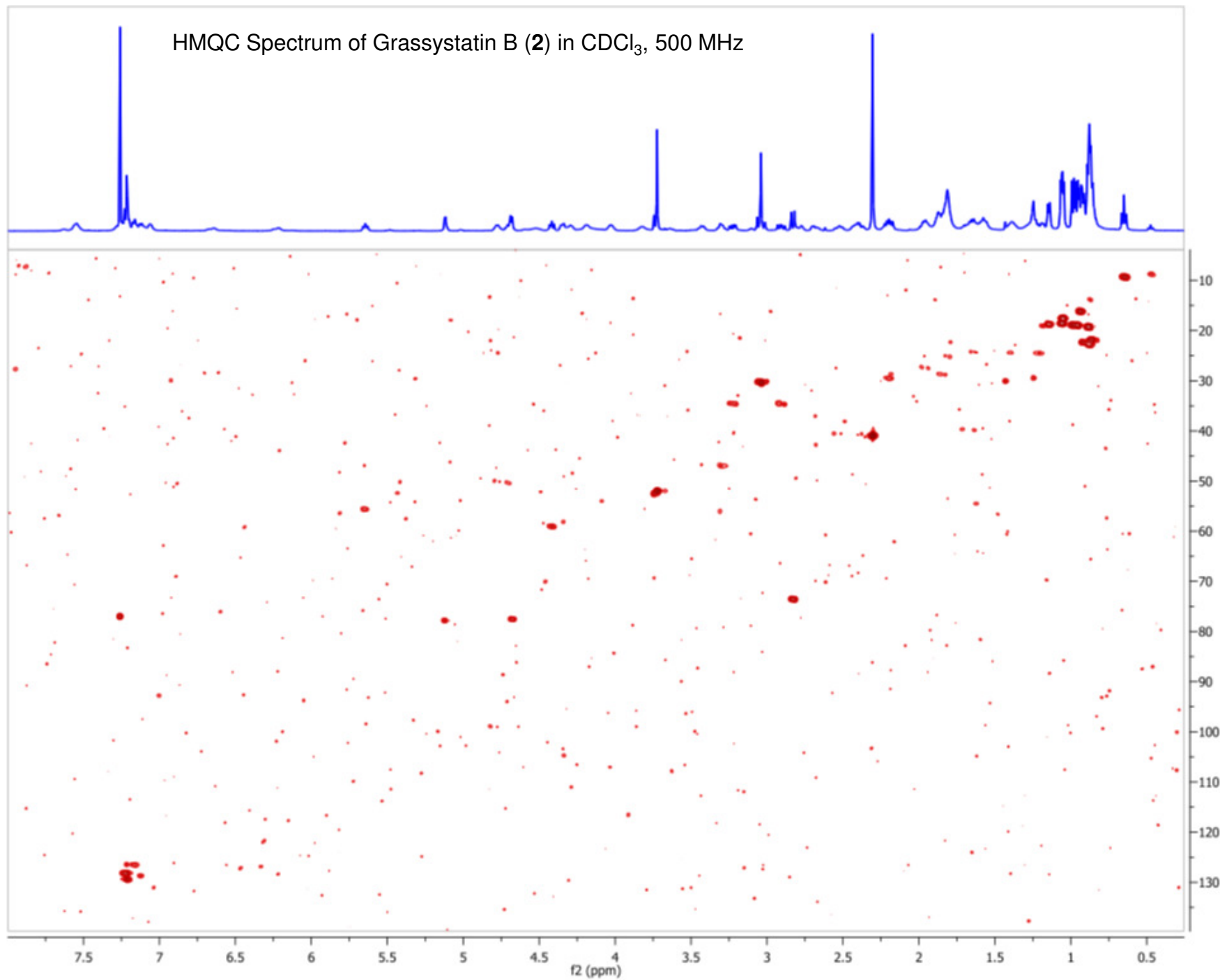
S21

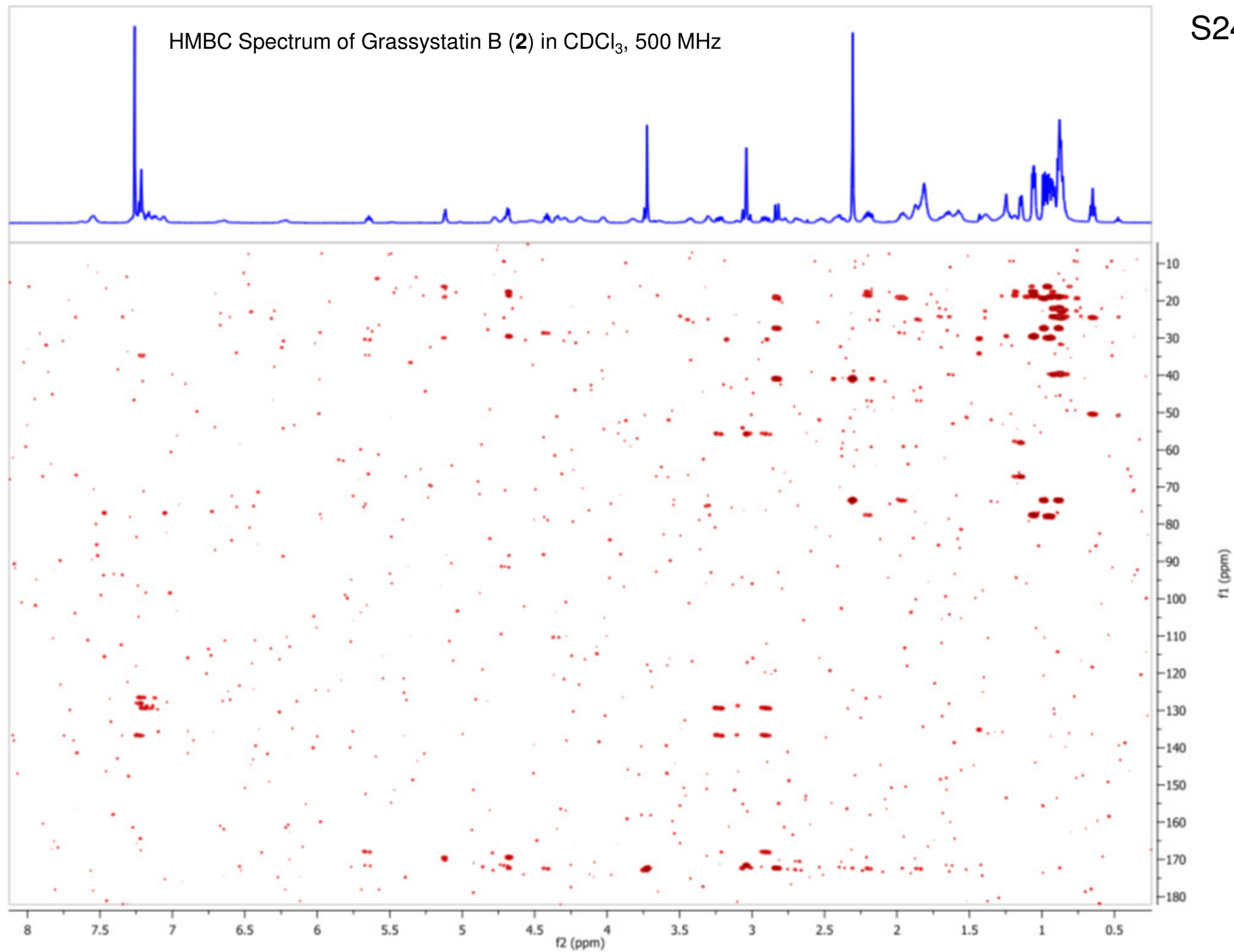


COSY Spectrum of Grassystatin B (**2**) in CDCl₃, 500 MHz

S22

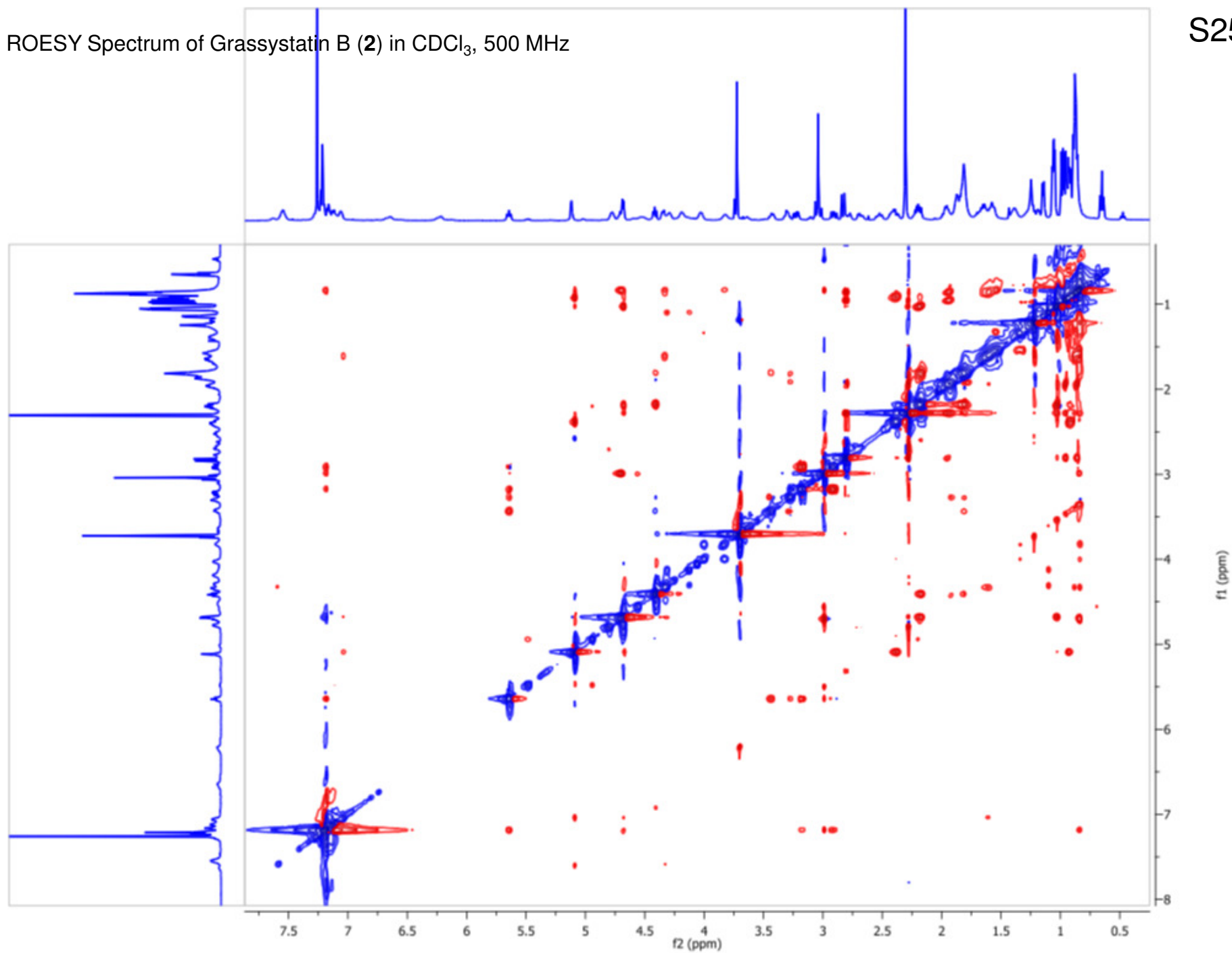






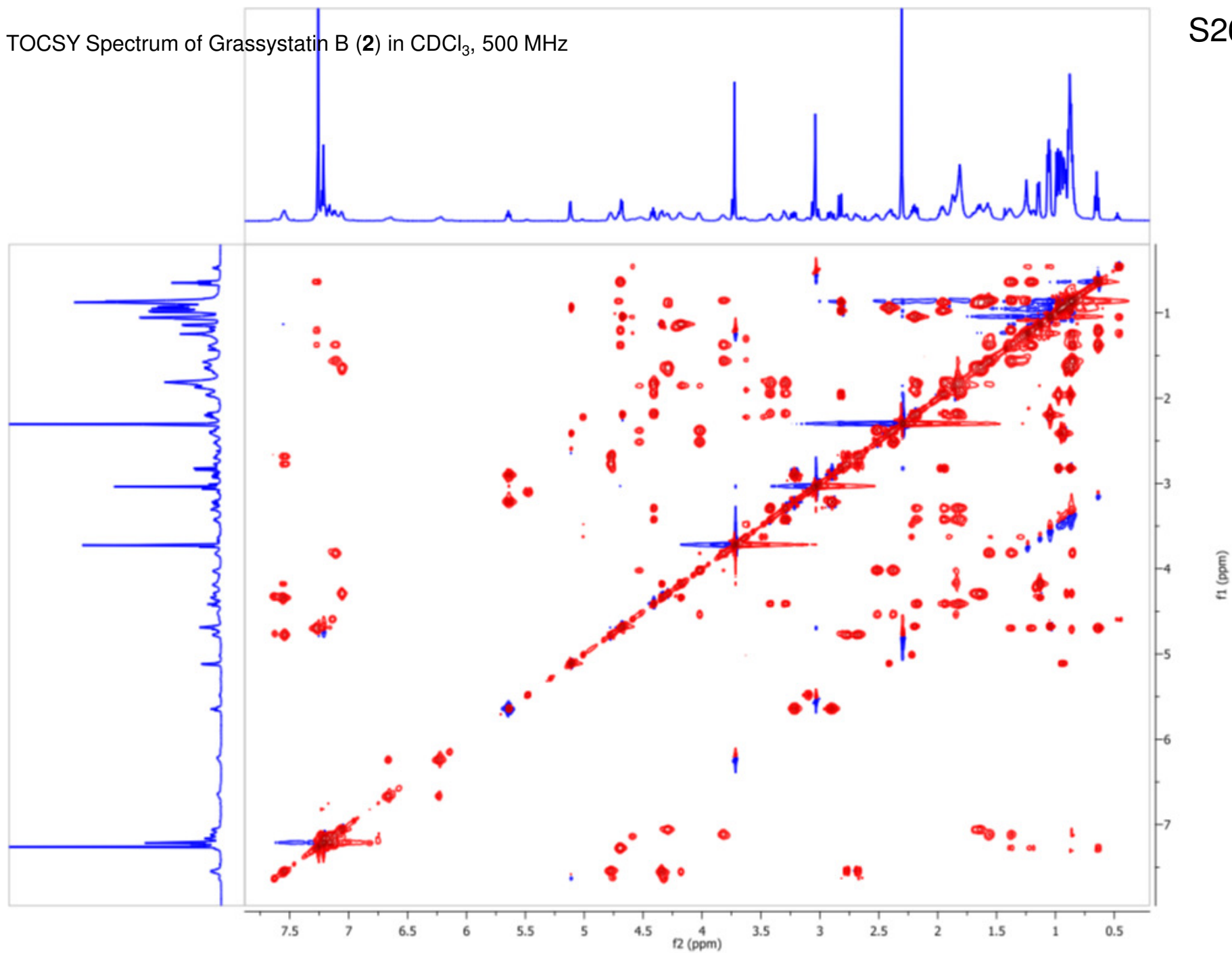
ROESY Spectrum of Grassystatin B (**2**) in CDCl₃, 500 MHz

S25



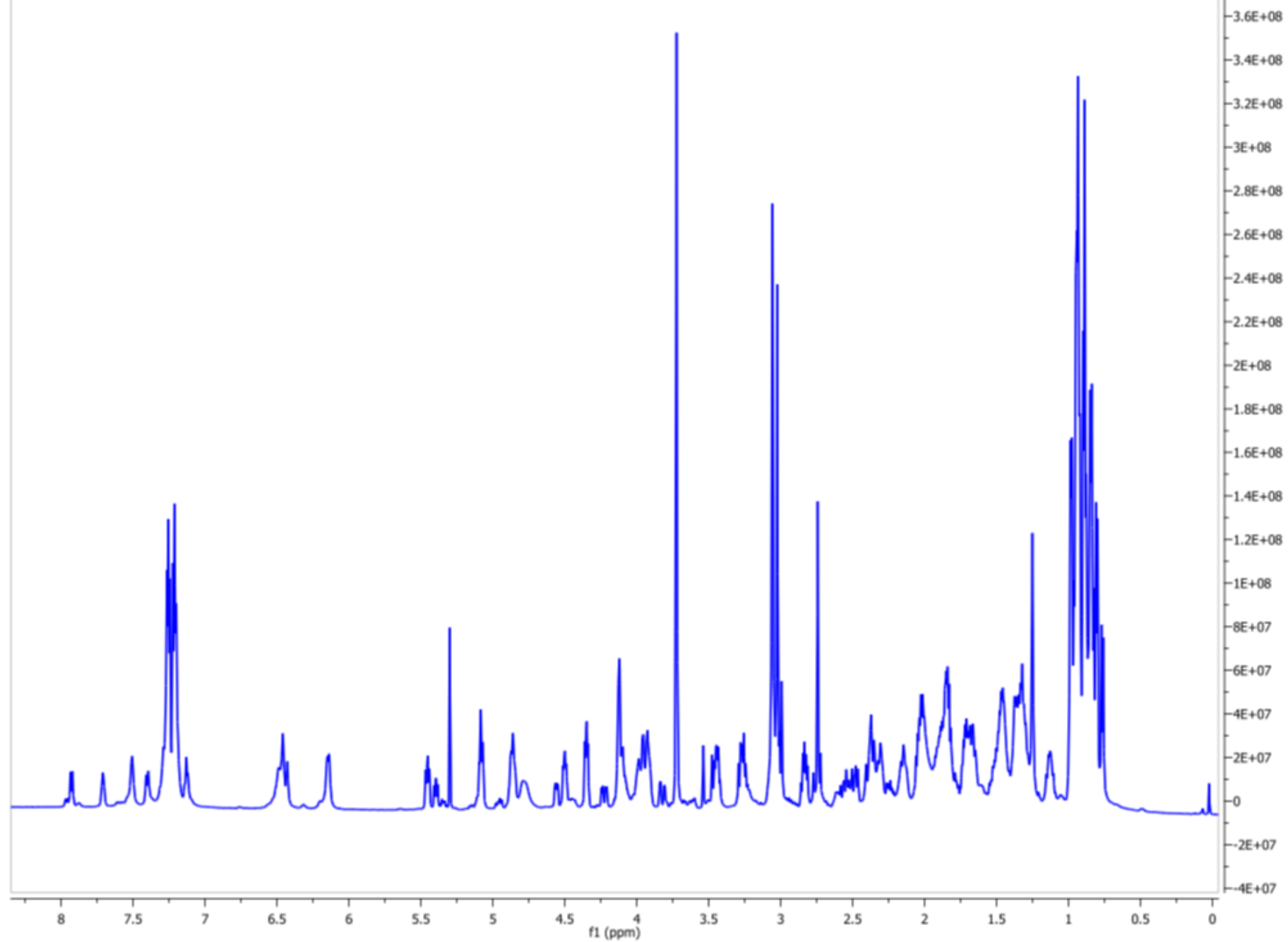
TOCSY Spectrum of Grassystatin B (**2**) in CDCl_3 , 500 MHz

S26



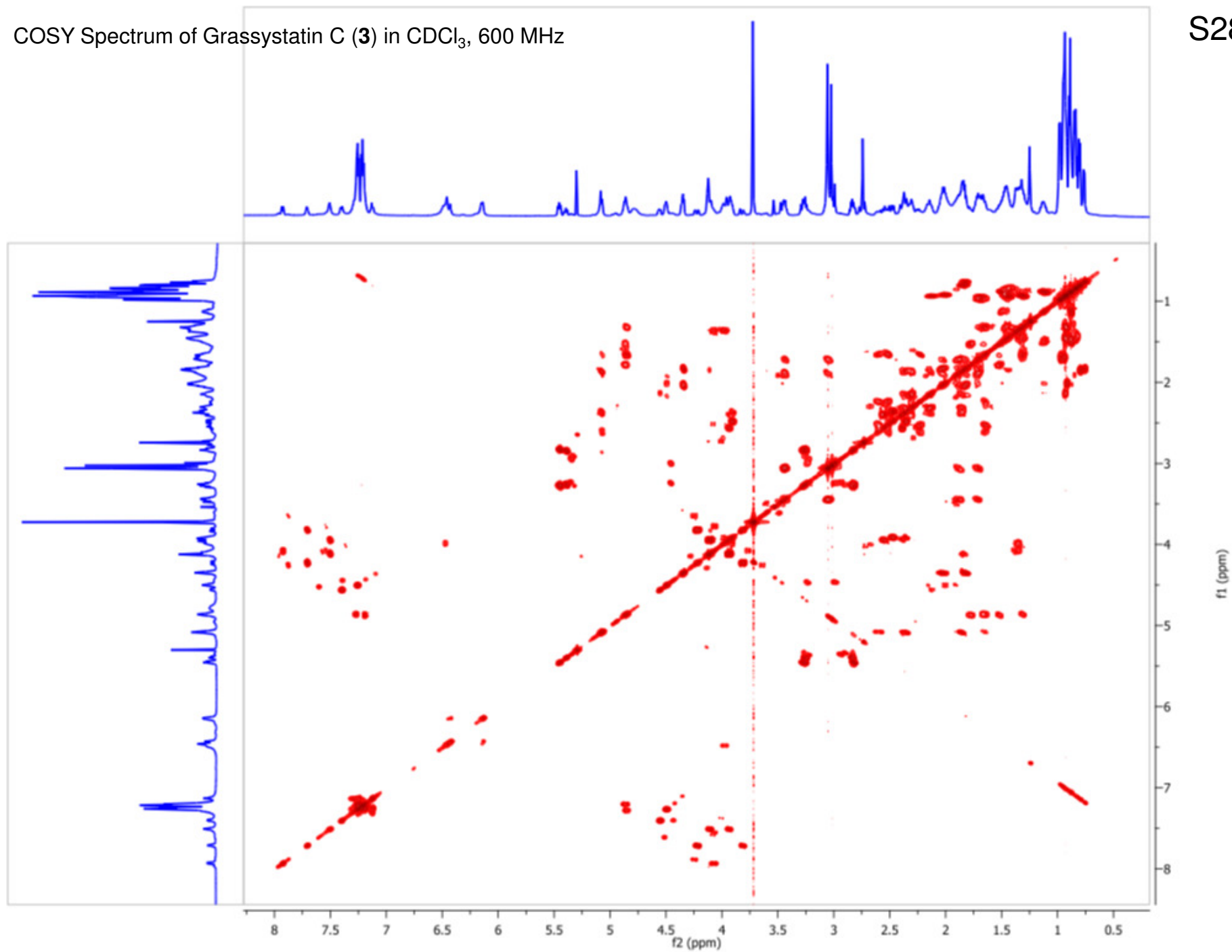
^1H NMR Spectrum of Grassystatin C (**3**) in CDCl_3 , 600 MHz

S27

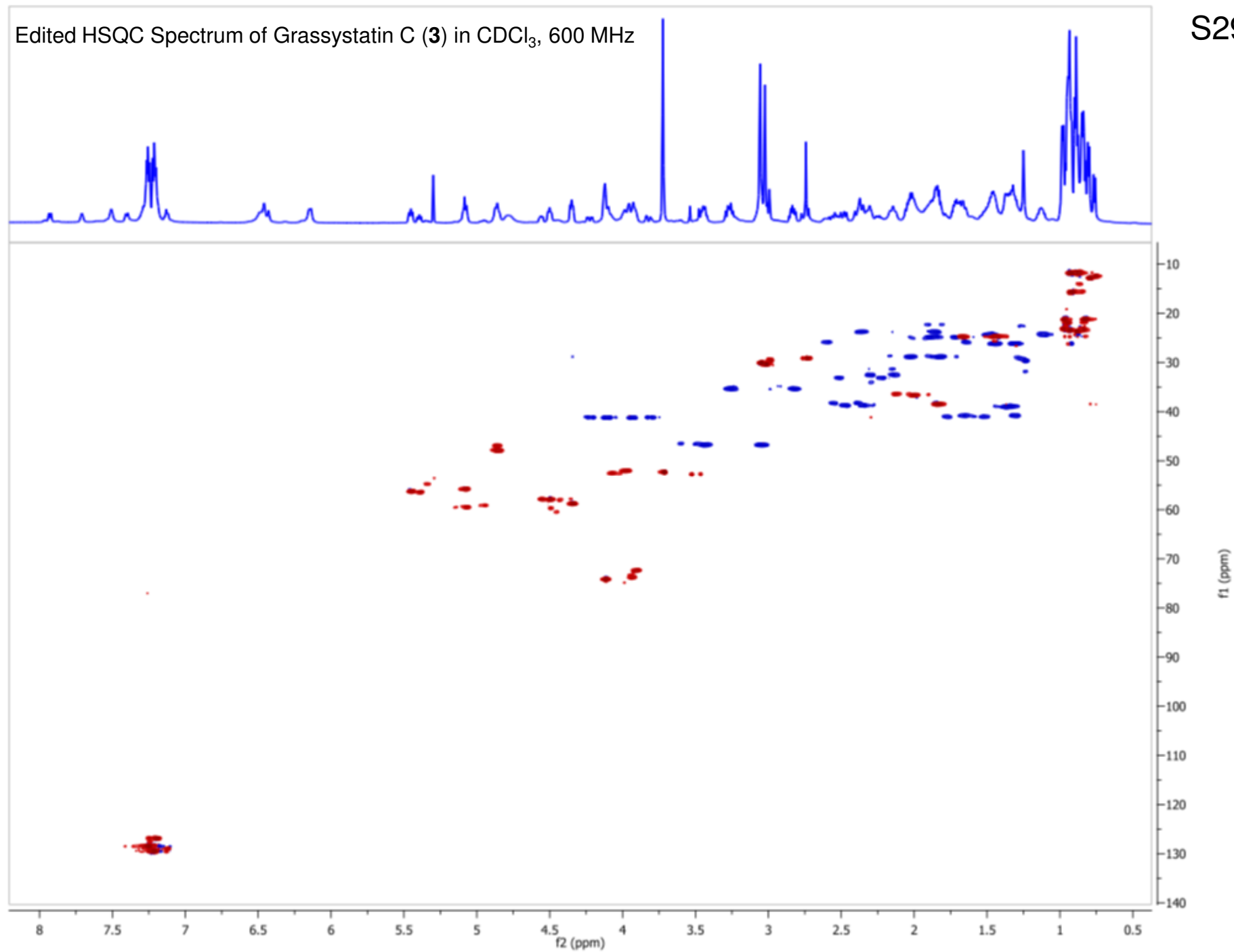


COSY Spectrum of Grassystatin C (**3**) in CDCl₃, 600 MHz

S28

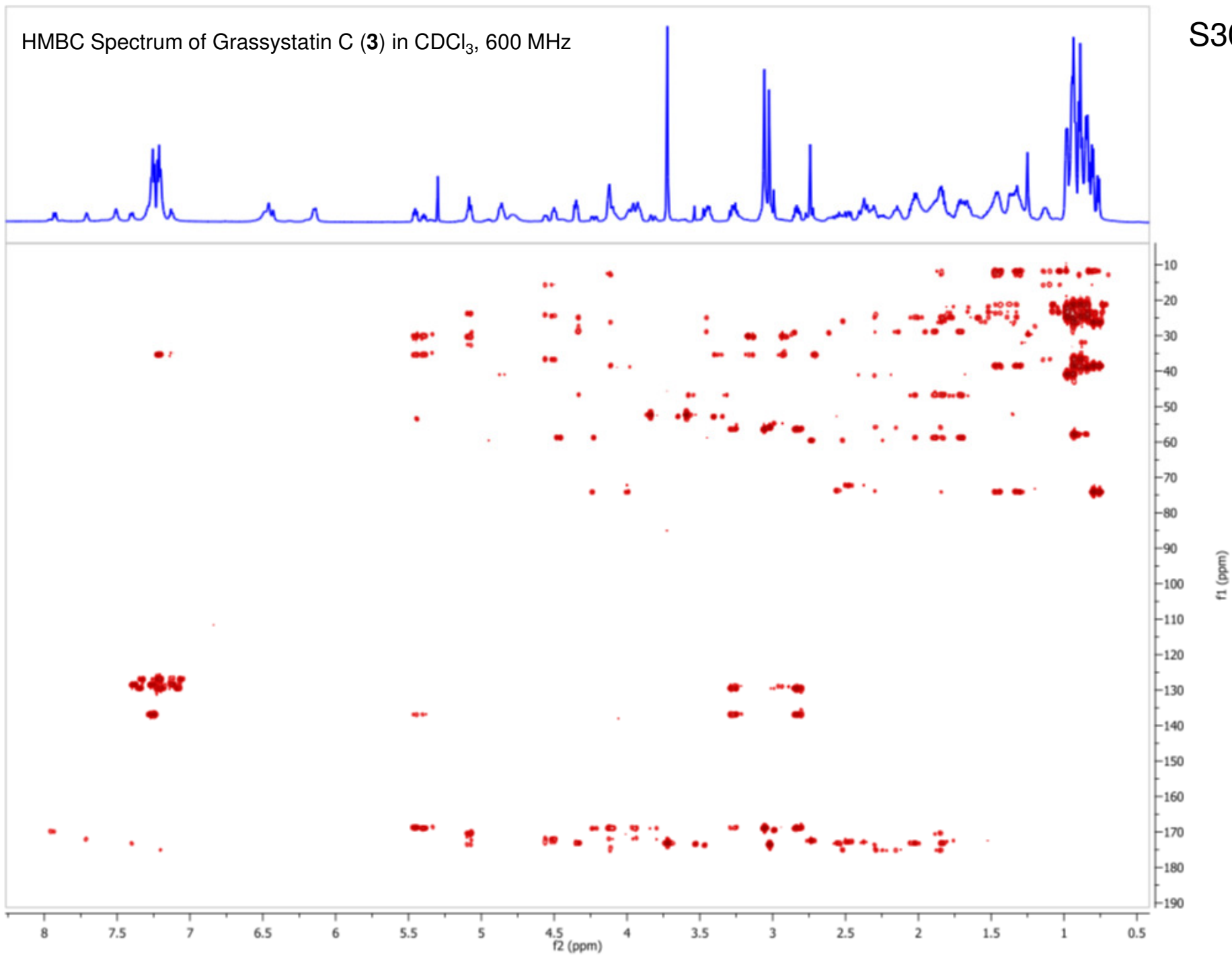


Edited HSQC Spectrum of Grassystatin C (**3**) in CDCl₃, 600 MHz



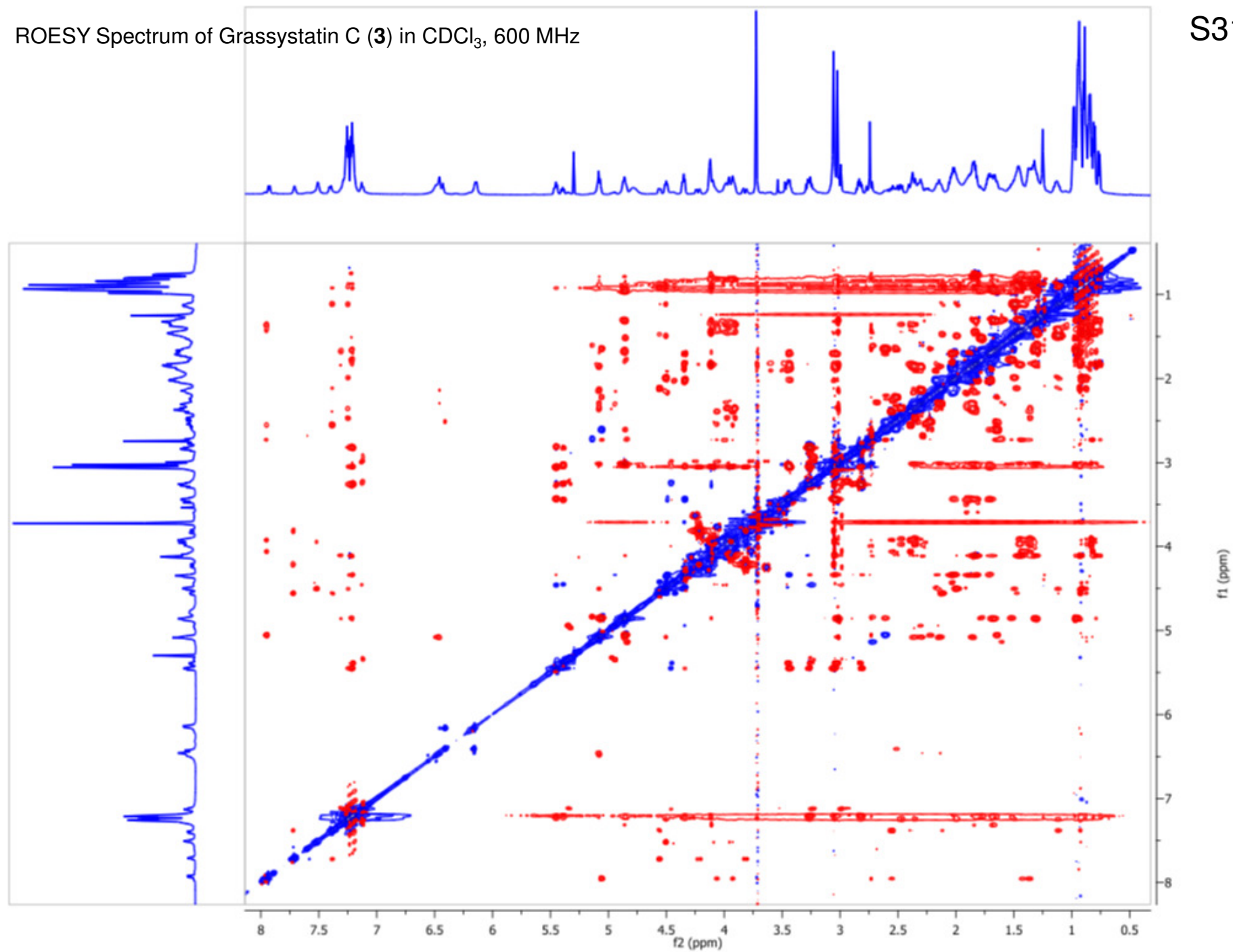
HMBC Spectrum of Grassystatin C (**3**) in CDCl₃, 600 MHz

S30



ROESY Spectrum of Grassystatin C (**3**) in CDCl₃, 600 MHz

S31



TOCSY Spectrum of Grassystatin C (**3**) in CDCl₃, 600 MHz

S32

