SUPPORTING INFORMATION

Grassystatin G, a New Cathepsin D Inhibitor from Marine Cyanobacteria: Discovery, Synthesis, and Biological Characterization

Lobna A. Elsadek,[a,b] Gustavo Seabra,[a,b] Valerie J. Paul,[c] and Hendrik Luesch*[,a,b,d]

[a] Department of Medicinal Chemistry and [b] Center for Natural Products, Drug Discovery and Development (CNPD3), University of Florida, 1345 Center Drive, Gainesville, FL 32610, United States; [c] Smithsonian Marine Station, 701 Seaway Drive, Ft. Pierce, FL 34949, United States; [d] Program in Cancer and Stem Cell Biology, Duke-NUS Medical School, Singapore, 169857, Singapore.

Table of Contents

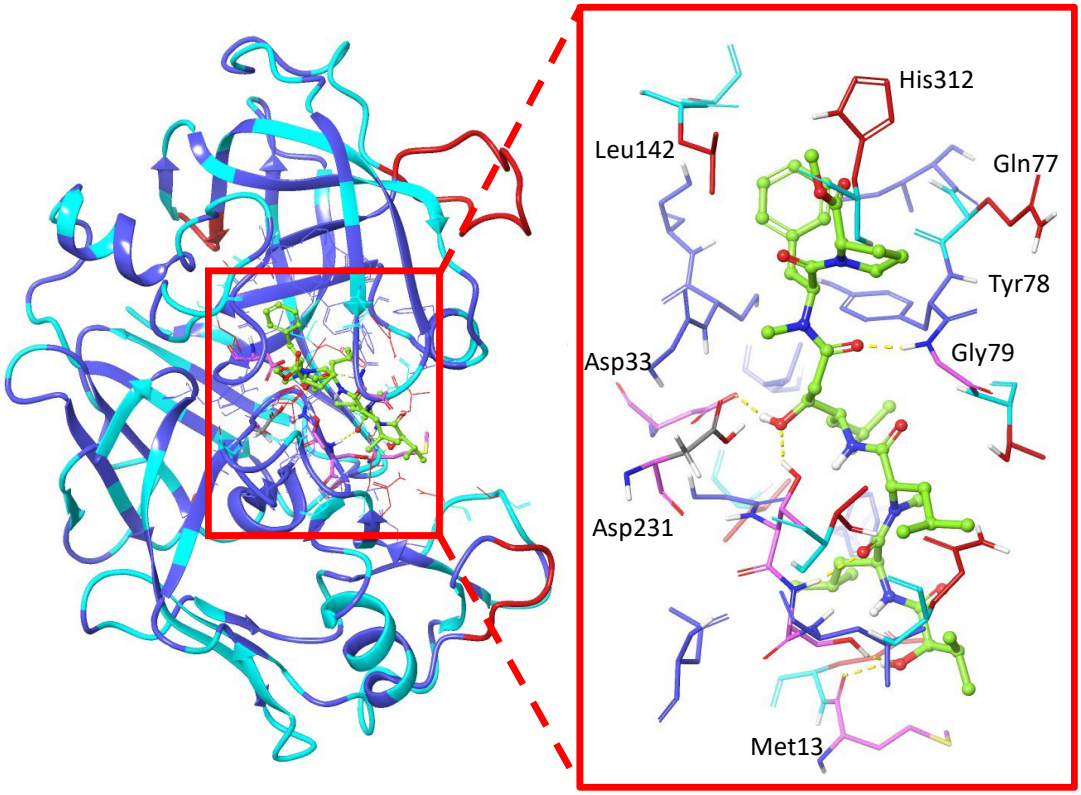

Fig. S1. Homology model of cathepsin-E showing the docked structure of grassystatin G. Residues in dark blue are the same as in the template (cathepsin-D, PDBID: 1LYB). Residues in light blue have been mutated, and residues in red were not present in the template and needed to be reconstructed.

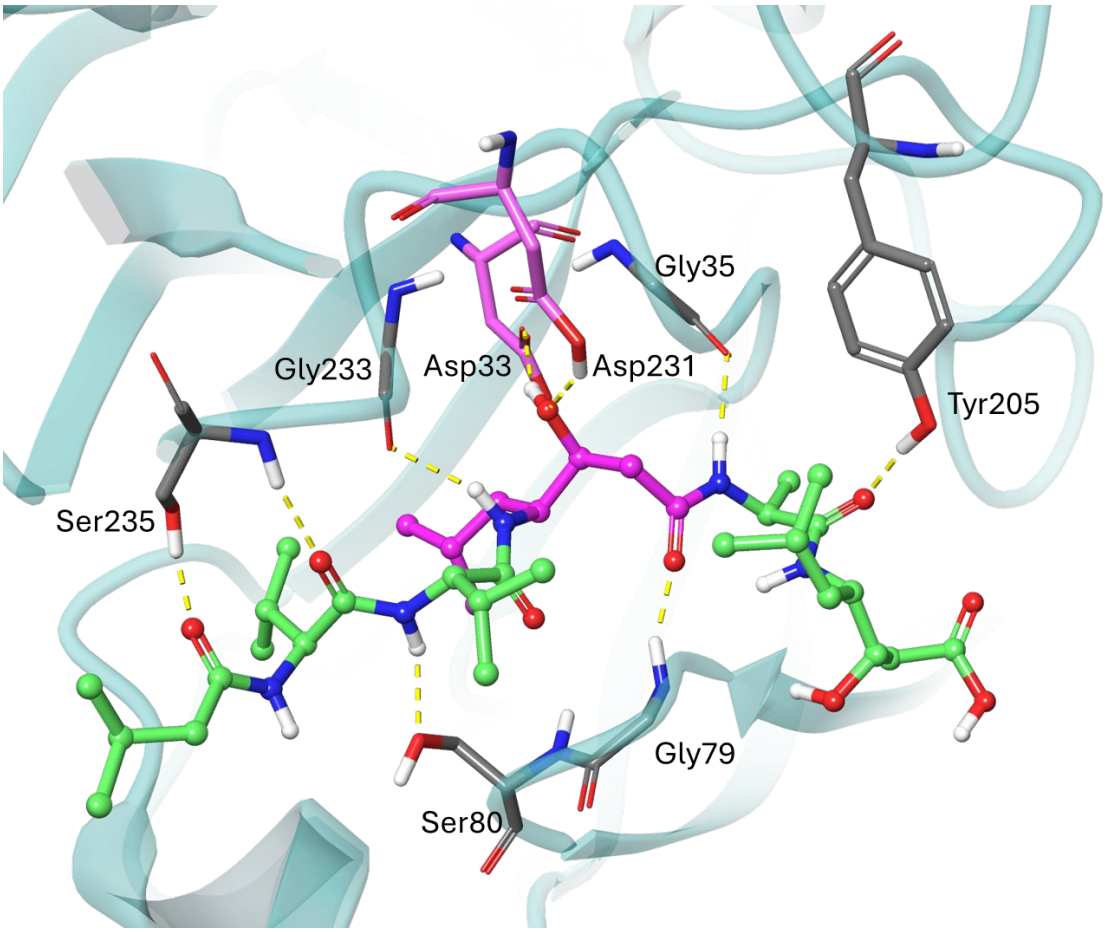

Fig. S2. Structure of Pepstatin A bound to Cathepsin D (PDBID: 1LYB). Pepstatin is shown in ball-and-stick model, and cathepsin residues in thick tubes. The statin unit and Asp33/231 are highlighted.

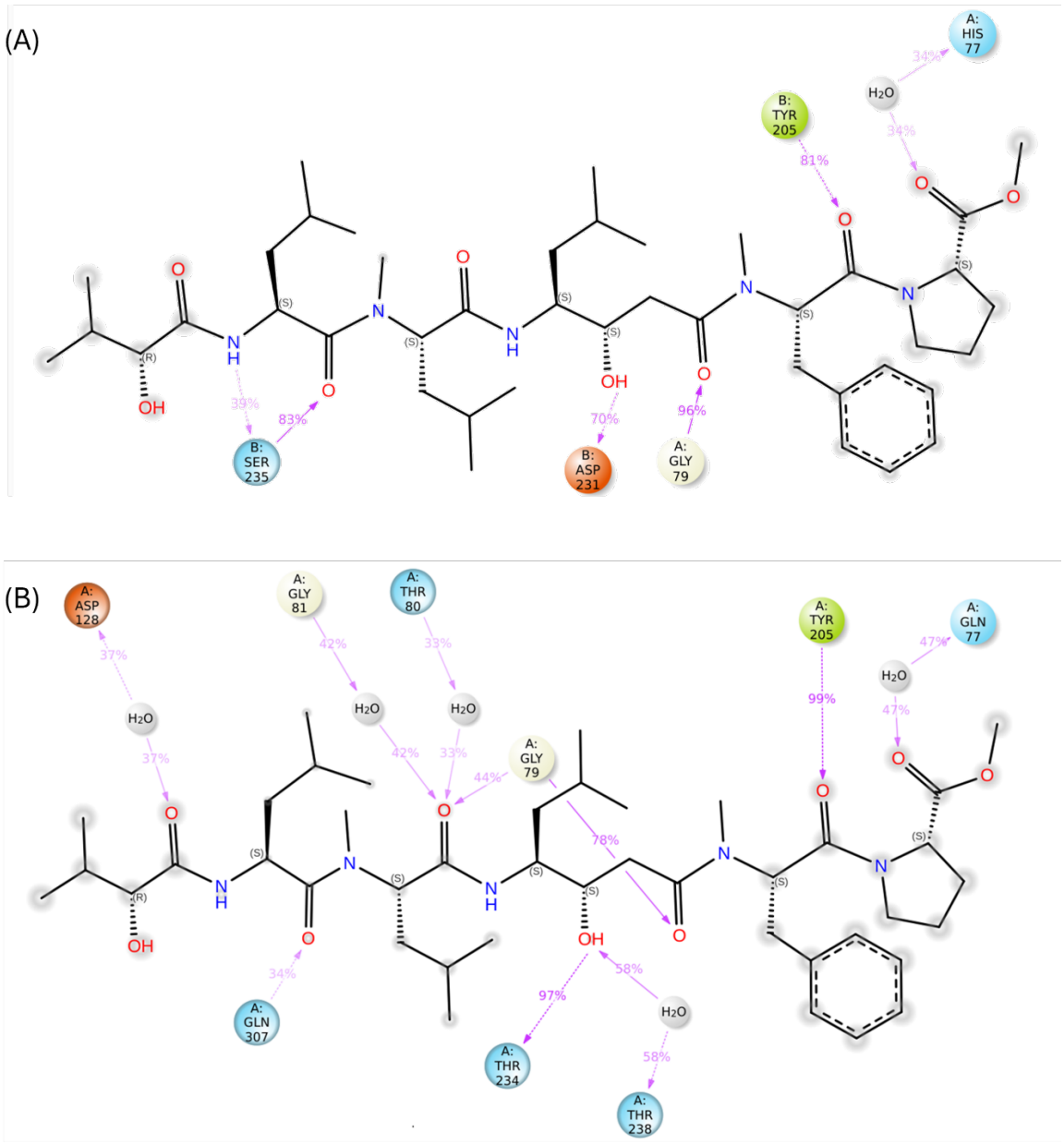

Fig. S3. Diagrams of persistent interactions between Grassystatin G and Cathepsin D (A), and Cathepsin E (B), during the last 400 ns of simulations. Only interactions present over 30% or more of the total time are shown. The residue colors represent: Green: hydrophobic; blue: polar, and red: negatively charged residues, and grey clouds indicate solvent exposure. Arrows indicate hydrogen bonds fro donor to acceptor.

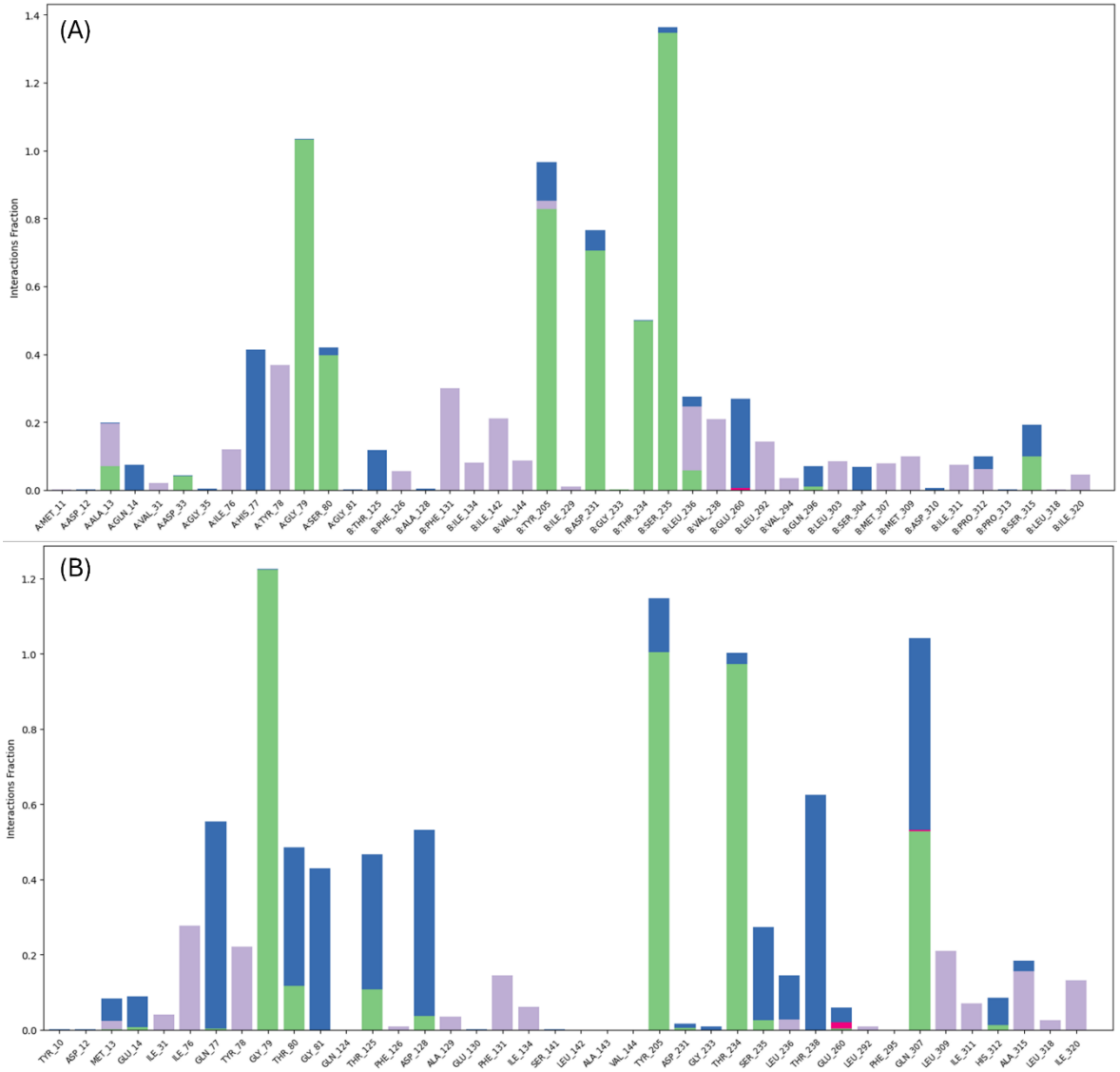

Fig. S4. Analysis of the interactions between Grassystatin G and Cathepsin D (A) and Cathepsin E (B), during the last 400 ns of simulations. The bar colors represent the types of interactions. Green: hydrogen bond; blue: water bridges, purple: hydrophobic interactions; and red: ionic interactions.

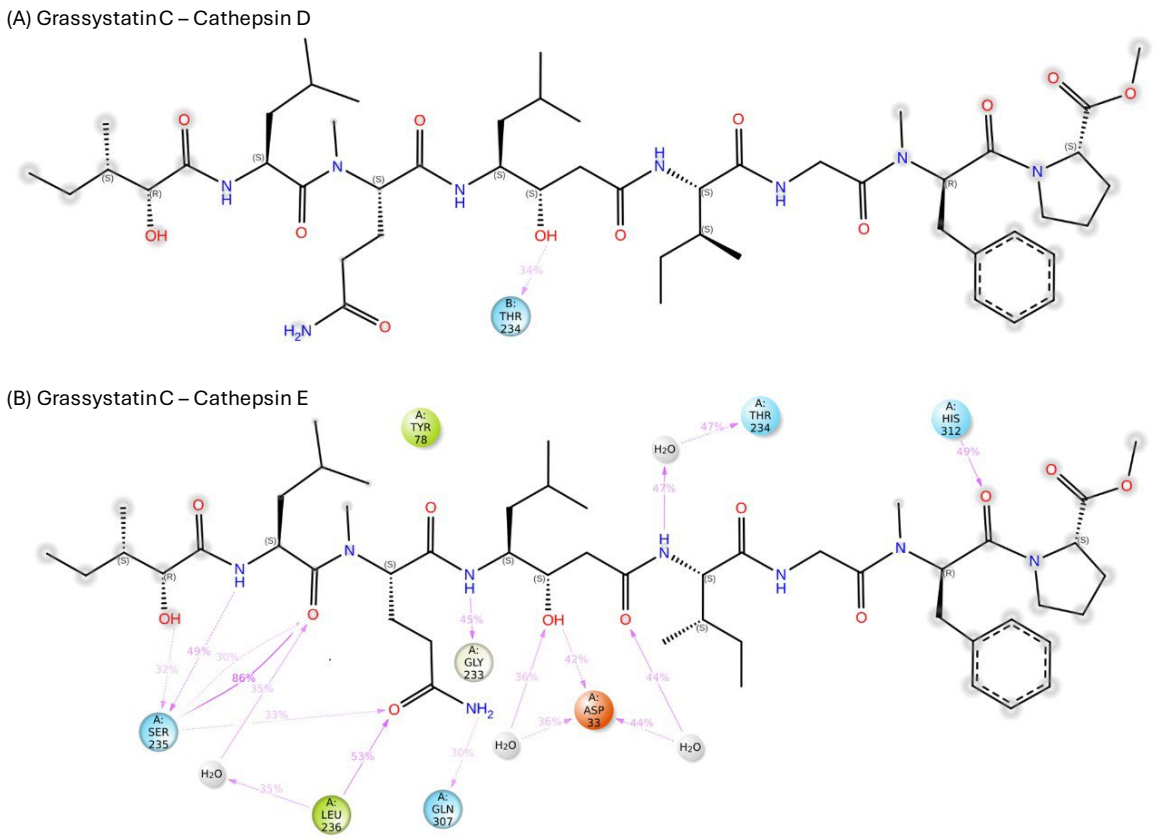

Fig. S5. Diagrams of persistent interactions between Grassystatin C and Cathepsin D (A), and Cathepsin E (B), during the last 400 ns of simulations. Only interactions present over 30% or more of the total time are shown. The residue colors represent: Green: hydrophobic; blue: polar, and red: negatively charged residues, and grey clouds indicate solvent exposure. Arrows indicate hydrogen bonds fro donor to acceptor.

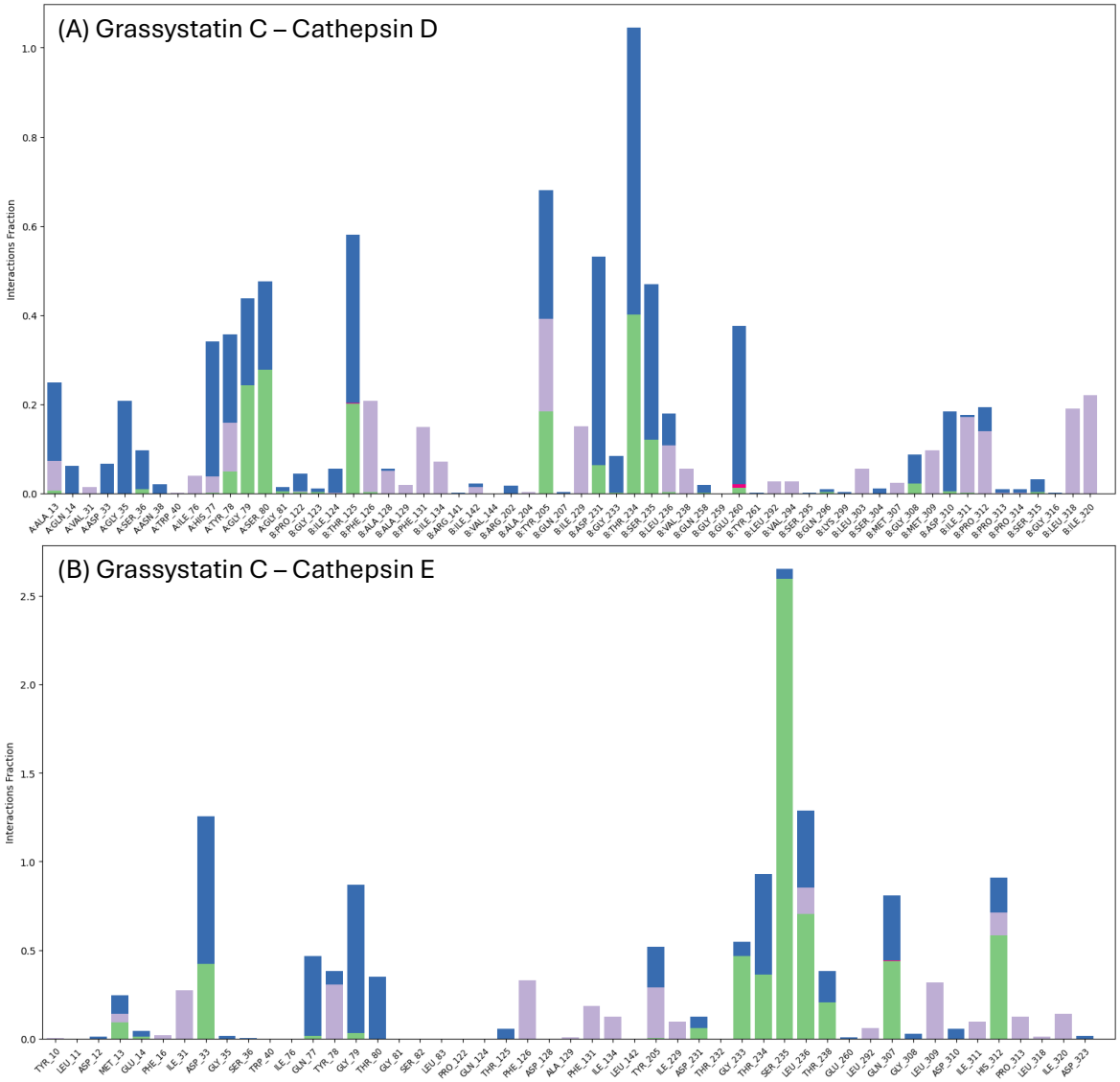

Fig. S6. Analysis of the interactions between Grassystatin C and Cathepsin D (A) and Cathepsin E (B), during the last 400 ns of simulations. The bar colors represent the types of interactions. Green: hydrogen bond; blue: water bridges, purple: hydrophobic interactions; and red: ionic interactions.

Table S1. Most persistent interactions between grassystatin G and Cathepsins D and E.*

		Interactions in CathD			Interactions in CathE			Difference (E-D)		
CathD Residue	**CathE Residue**	**HB**	**WB**	**HP**	**HB**	**WB**	**HP**	**HB**	**WB**	**HP**
Asp33	Asp	33	0	0	0	0	0	-33	0	0
His77	Gln	0	330		3	437	0	3	107	0
Tyr78	Tyr	0	0	267	0	0	174	0	0	-93
Gly79	Gly	792	3	0	785	0	0	-7	-3	0
Ser80	**Thr**	304	16	0	94	291	0	-210	275	0
Phe131	Phe	0	0	232			115	0	0	-117
Ile142	**Leu**	0	0	169	0	0	0	0	0	-169
Tyr205	Tyr	660	92	19	799	114	0	139	22	-19
Asp231	Asp	566	47	0	4	9	0	-562	-38	0
Thr234	Thr	336	0	0	780	18	0	444	18	0
Ser235	Ser	743	13	0	21	186	0	-722	173	0
Leu236	Leu	46	146	24	21	93	0	-25	-53	-24
Val238	Thr	0	0	168	0	0	495	0	0	327
Glu260	Glu	0	166	0	3	30	0	3	-136	0

* The numbers represent the number of frames in which the interaction was present, out of a total of 800 frames, representing the last 400 ns of simulation. HB = Hydrogen bonds, WB = water bridges, HP = hydrophobic interactions.

Table S2. Most persistent interactions between Grassystatin C and Cathepsins E and D.*

		Interactions in CathE			Interactions in CathD			Difference (D-E)		
CathE Residue	**CathD Residue**	**HB**	**WB**	**HP**	**HB**	**WB**	**HP**	**HB**	**WB**	**HP**
Asp33	Asp	339	654	0	0	50	0	-339	-604	0
Gln77	His	13	323	0	1	220	24	-12	-103	24
Tyr78	Tyr	0	56	246	39	155	86	39	99	-160
Gly79	Gly	26	428	0	194	127	0	168	-301	0
Thr80	Ser	0	248	0	175	146	0	175	-102	0
Phe126	Phe	0	0	234	2	0	146	2	0	-88
Phe131	Phe	0	0	145	0	0	114	0	0	-31
Tyr205	Tyr	4	176	227	147	217	159	143	41	-68
Asp231	Asp	49	51	0	50	352	0	1	301	0
Gly233	Gly	366	63	0	1	57	0	-365	-6	0
Thr234	Thr	284	429	0	321	358	0	37	-71	0
Ser235	Ser	776	34	0	69	236	0	-707	202	0
Leu236	Leu	427	293	118	2	58	79	-425	-235	-39
Thr238	Val	164	141	0	0	0	44	-164	-141	44
Gln307	Met	346	223	0	0	0	18	-346	-223	18
Leu309	Met	0	0	247	0	0	76	0	0	-171
His312	Pro	467	152	102	0	43	108	-467	-109	6

*The numbers represent the number of frames in which the interaction was present, out of a total of 900 frames, representing the last 400 ns of simulation. HB = Hydrogen bonds, WB = water bridges, HP = hydrophobic interactions.

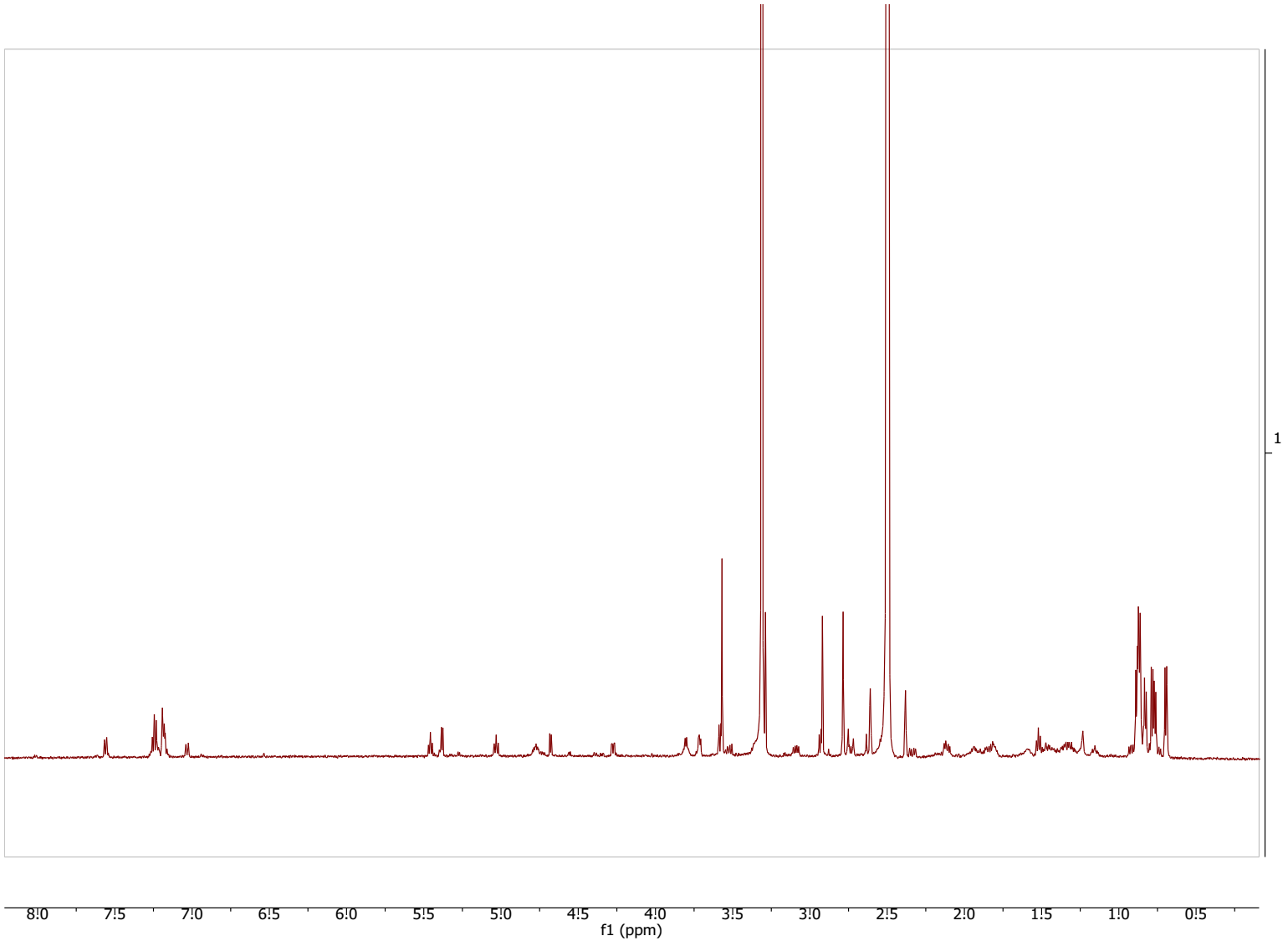

Fig. S7. ^{1}H NMR spectrum (600 MHz, DMSO-d_6) of natural product grassystatin G (**1**).

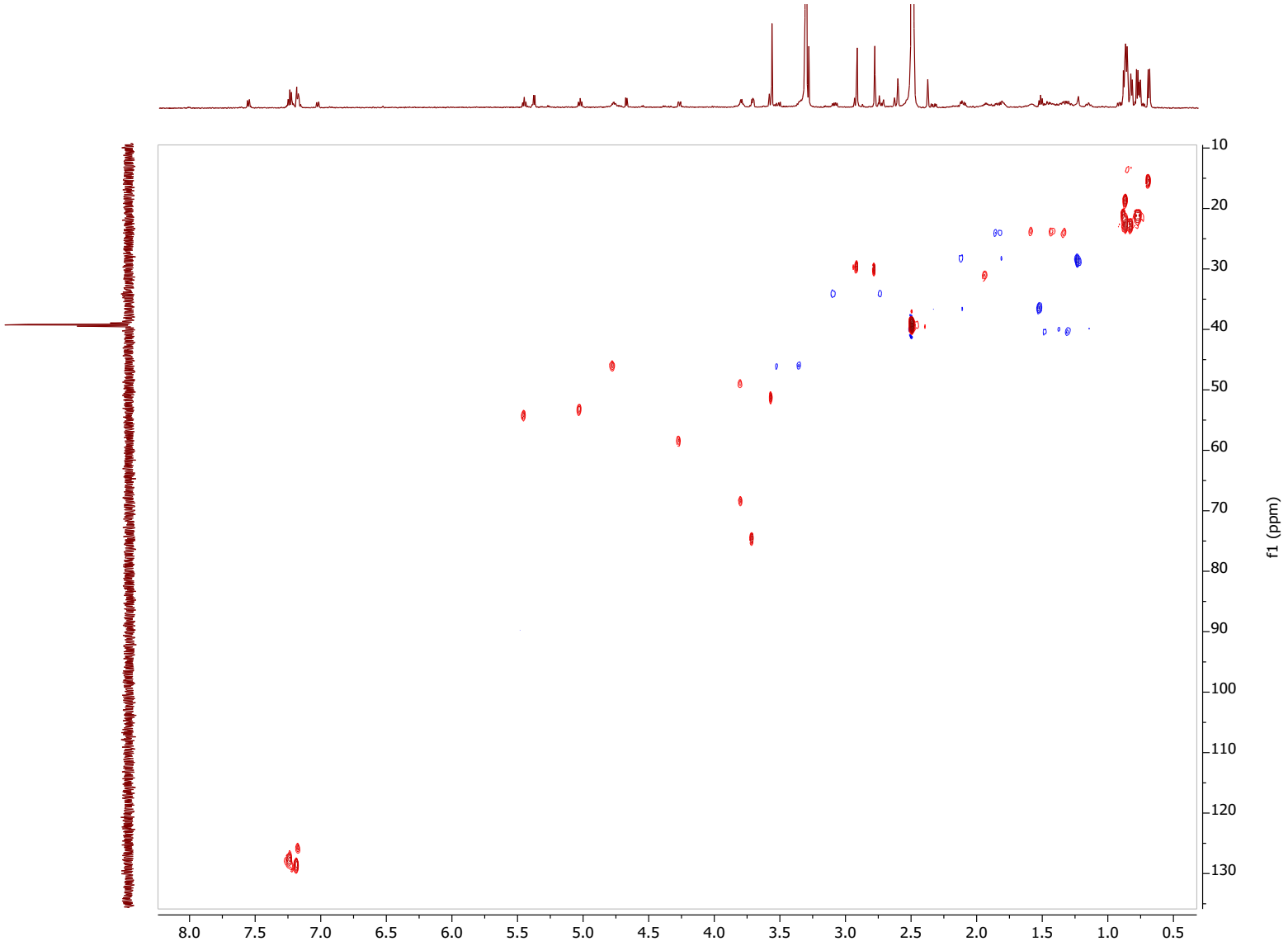

Fig. S8. HSQC spectrum (800 MHz, DMSO-d_6) of natural product grassystatin G (**1**)**.**

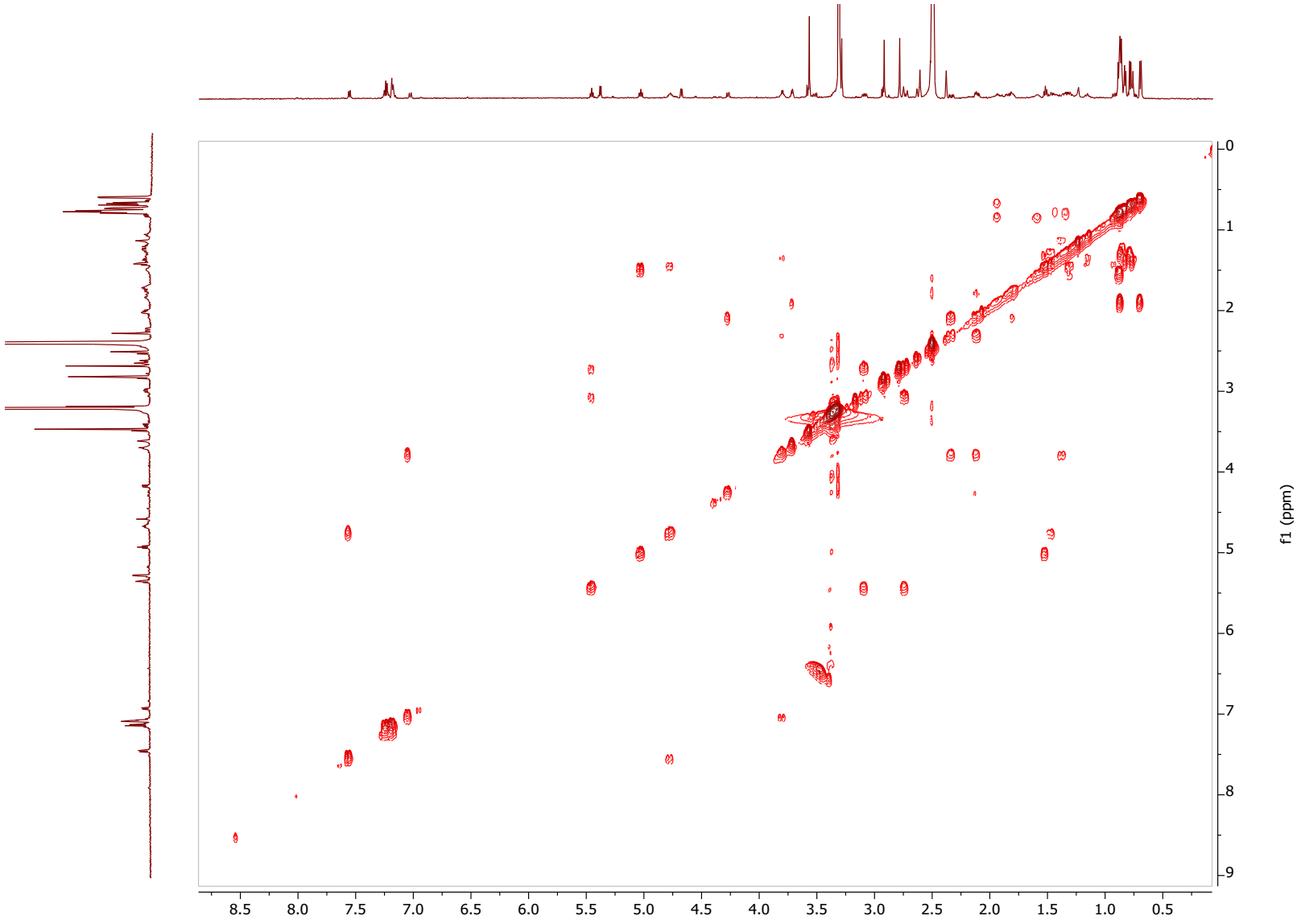

Fig. S9. COSY spectrum (600 MHz, DMSO-d_6) of natural product grassystatin G (**1**).

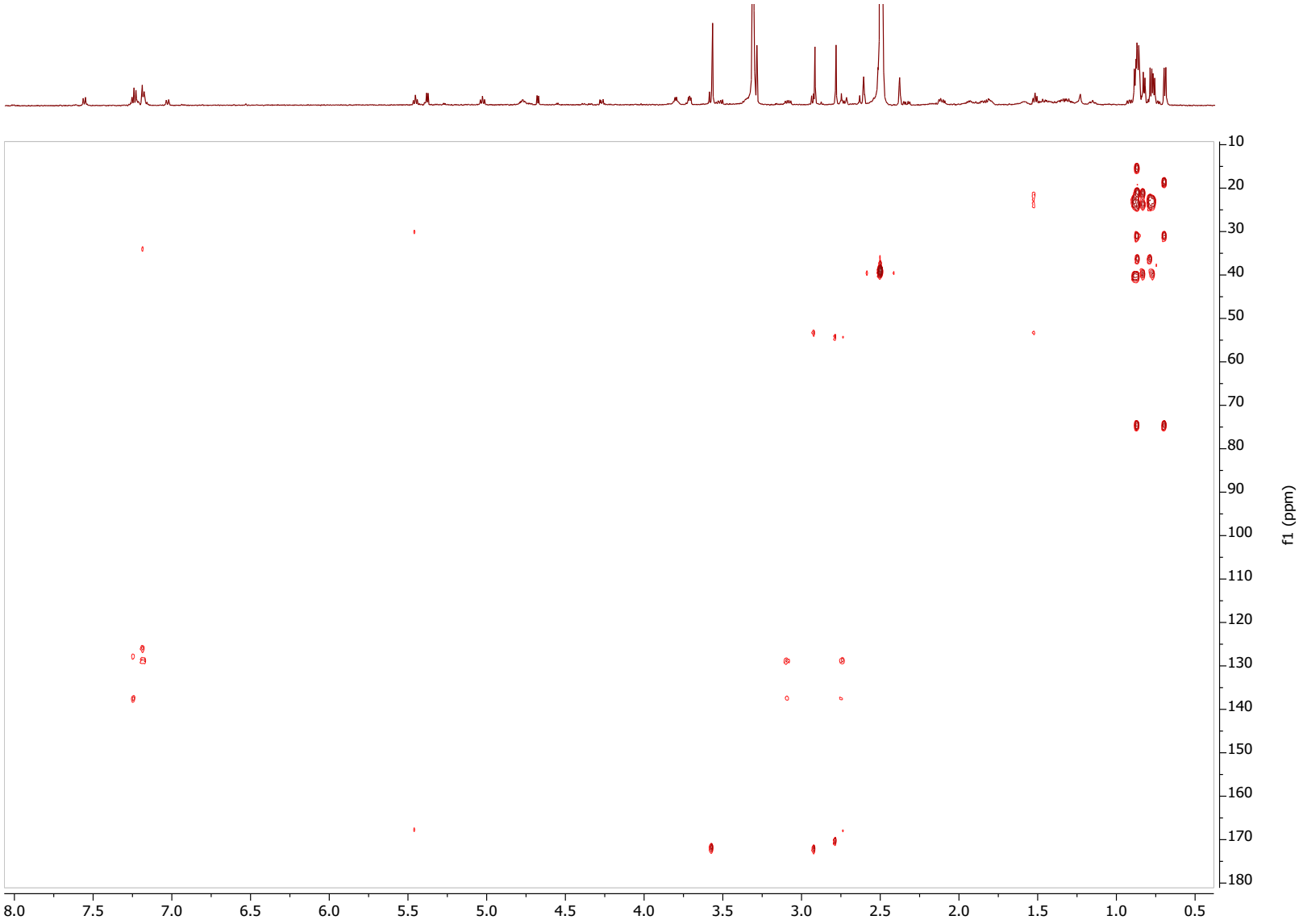

Fig. S10. HMBC spectrum (800 MHz, DMSO-d_6) of natural product grassystatin G (**1**).

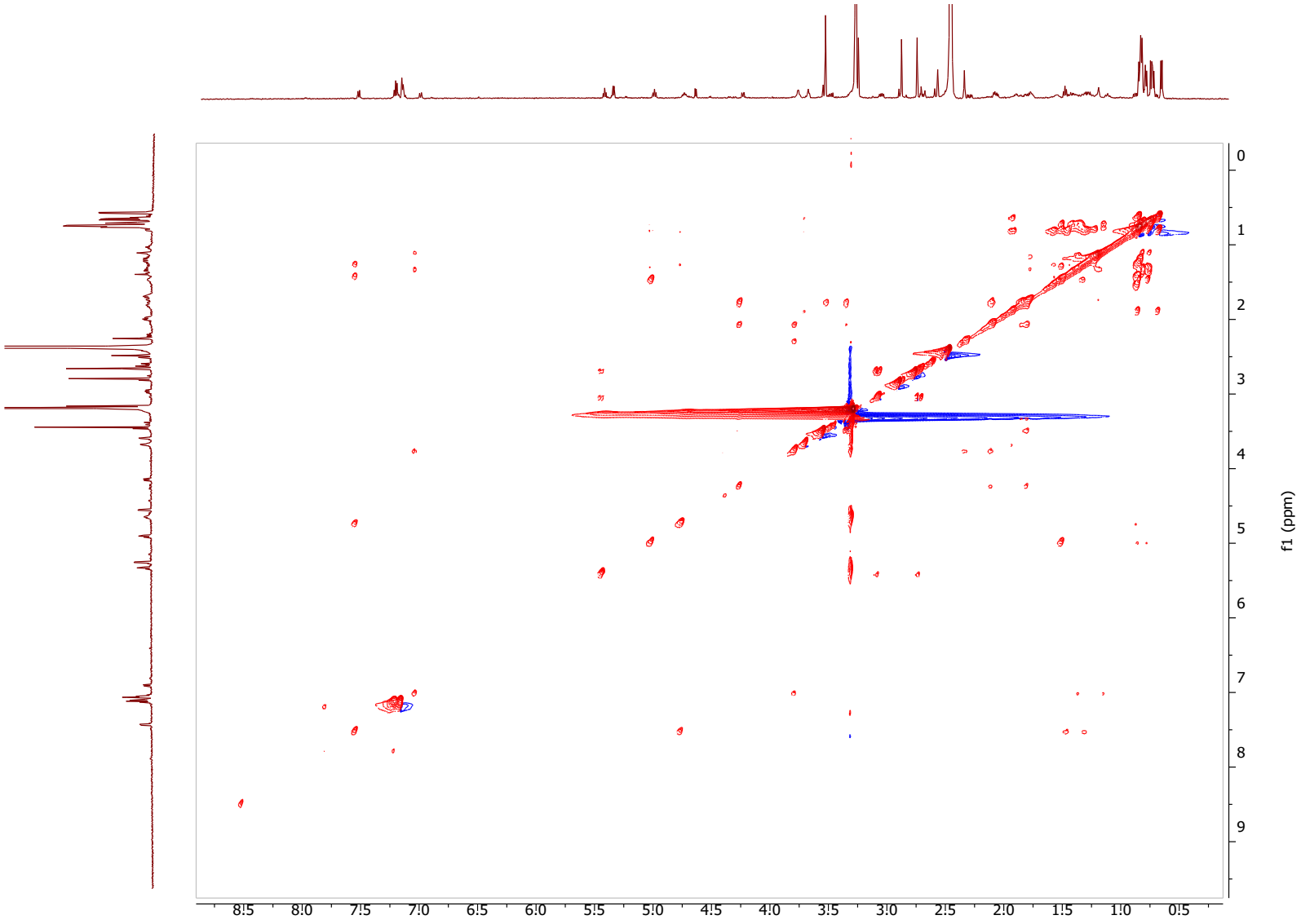

Fig. S11. TOCSY spectrum (600 MHz, DMSO-d_6) of natural product grassystatin G (**1**).

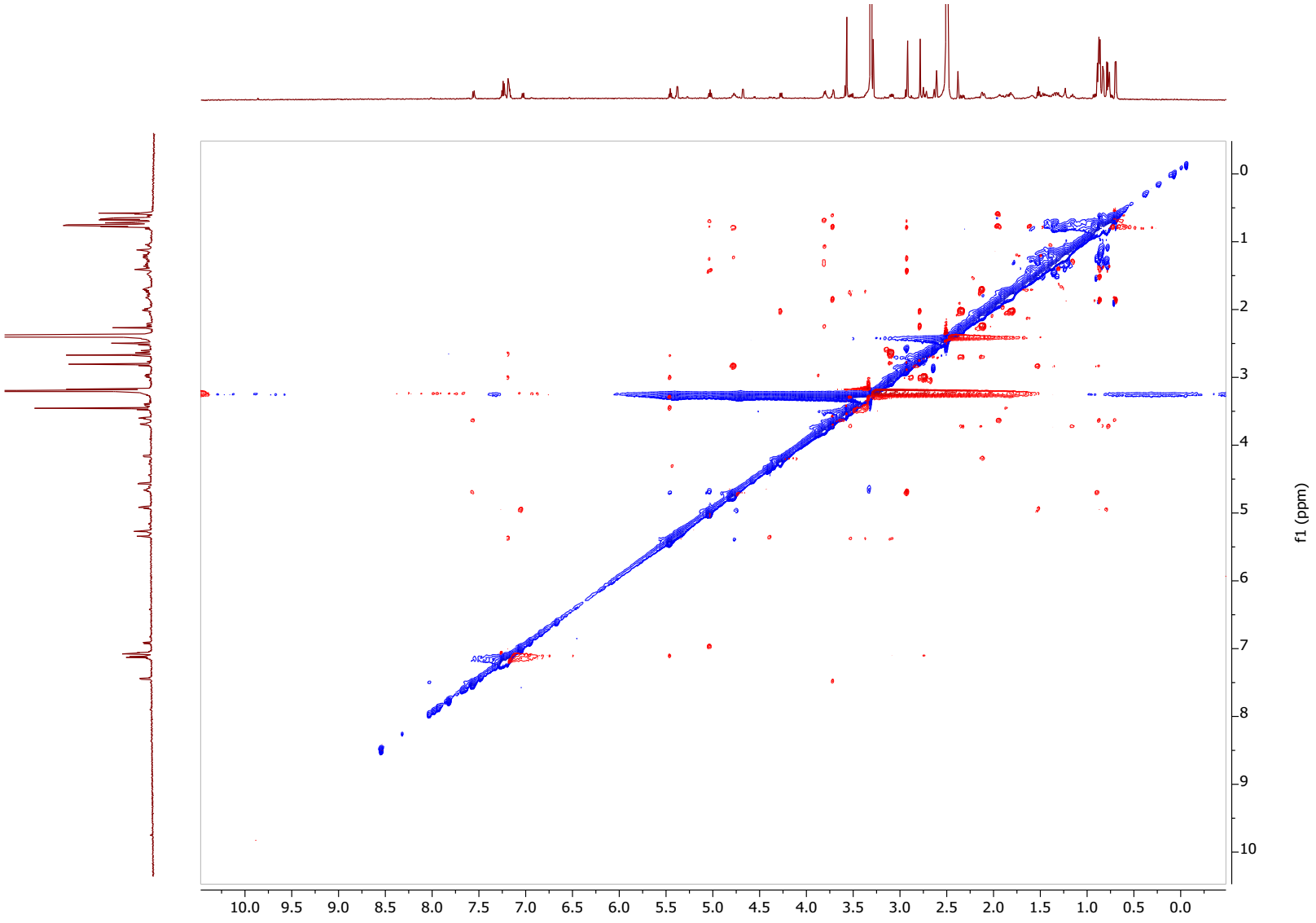

Fig. S12. ROESY spectrum (600 MHz, DMSO-d_6) of natural product grassystatin G (**1**).

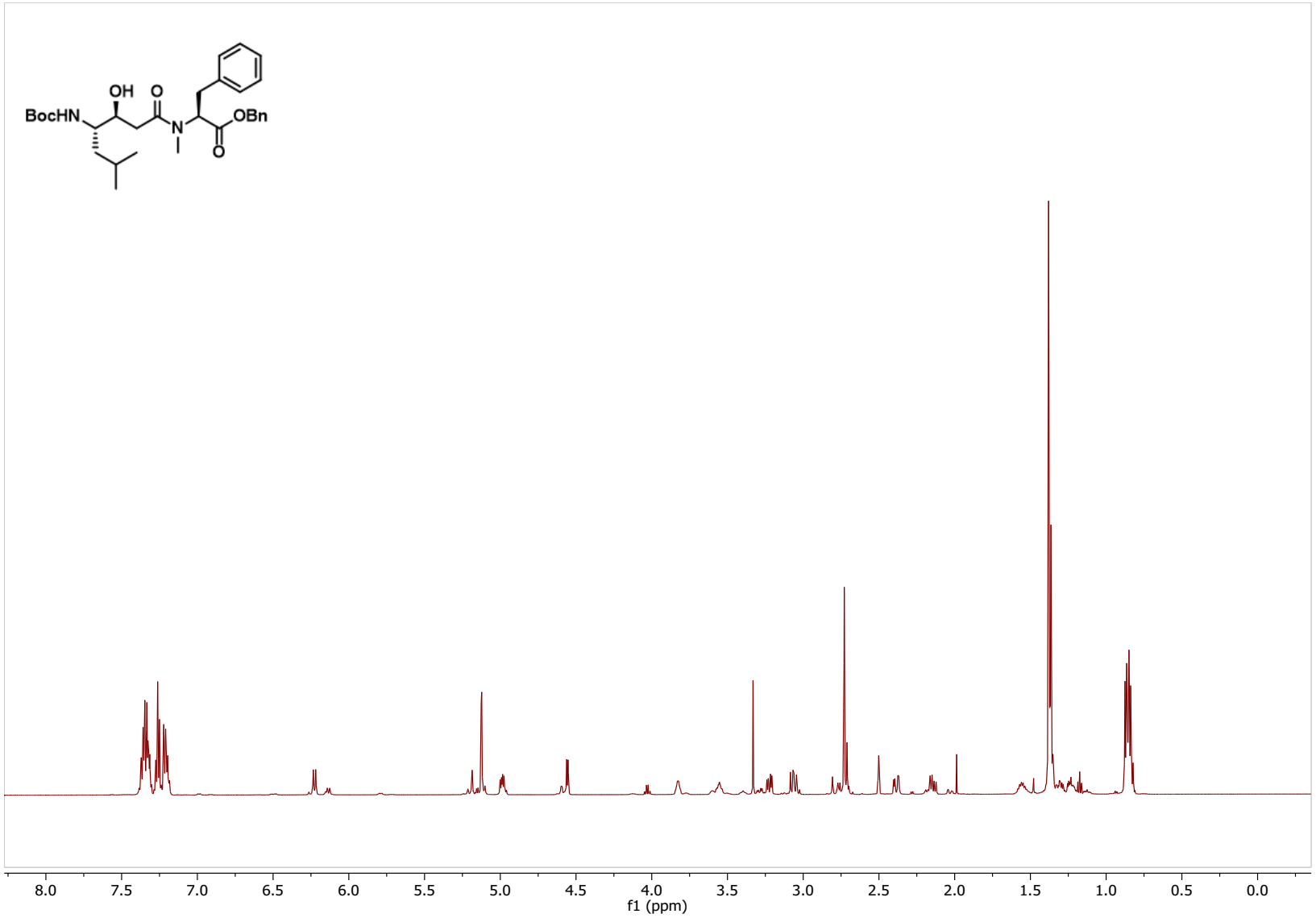

Fig. S13. ^{1}H NMR spectrum (600 MHz, DMSO-d_6) of compound **4.**

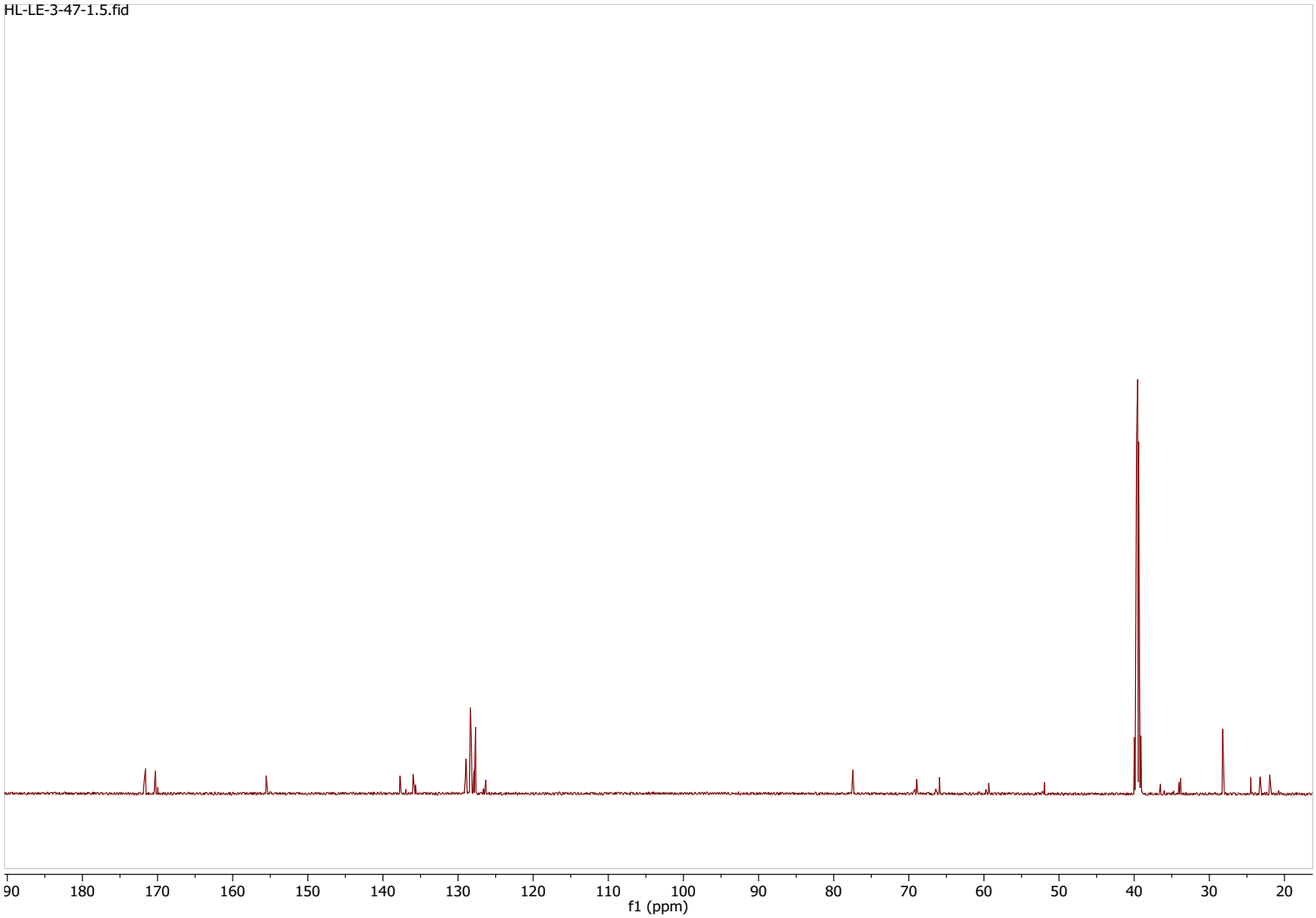

Fig. S14. ^{13}C NMR spectrum (150 MHz, DMSO-d_6) of compound **4.**

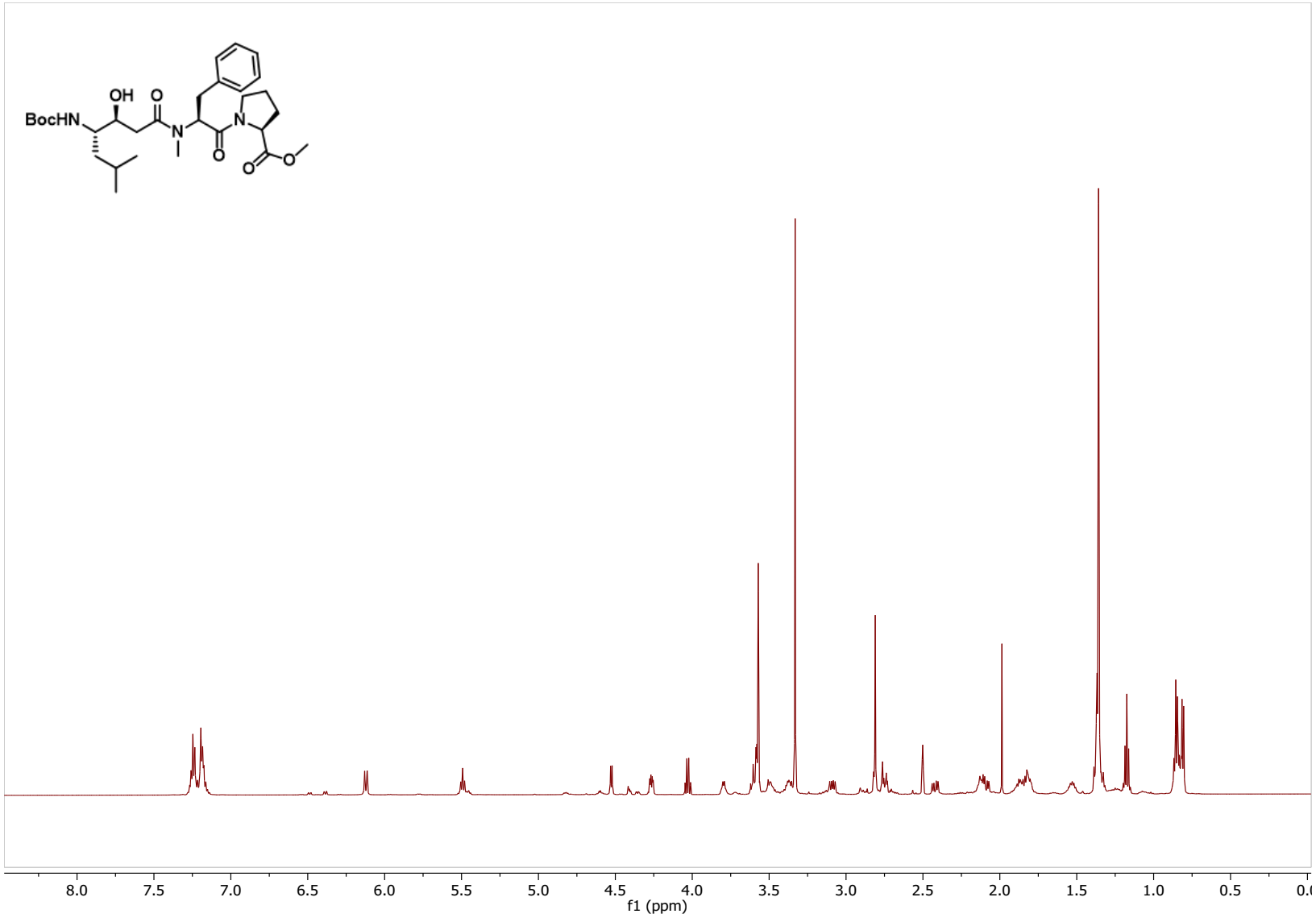

Fig. S15. ^{1}H NMR spectrum (600 MHz, DMSO-d_6) of compound **6.**

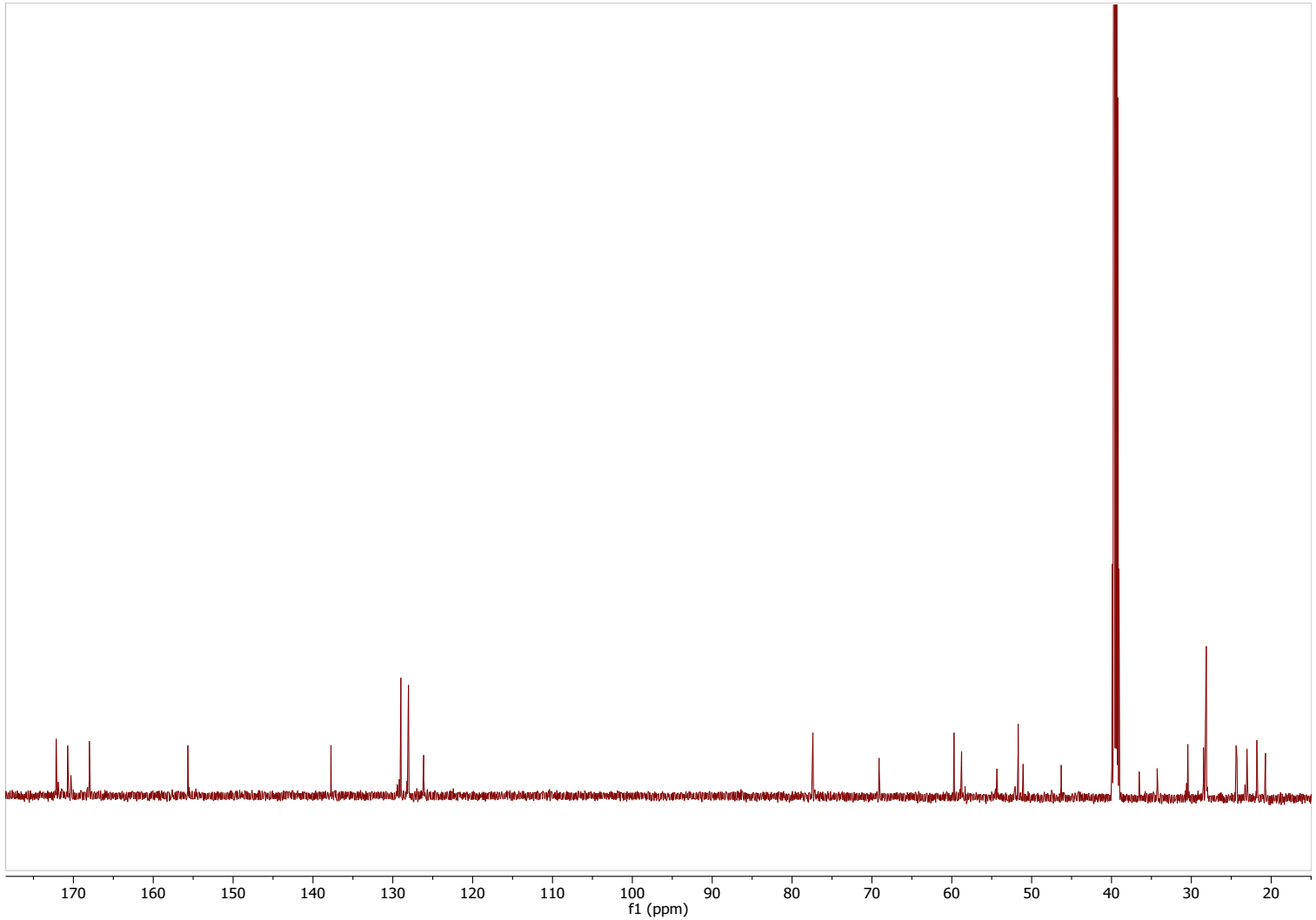

Fig. S16. ^{13}C NMR spectrum (150 MHz, DMSO-d_6) of compound **6.**

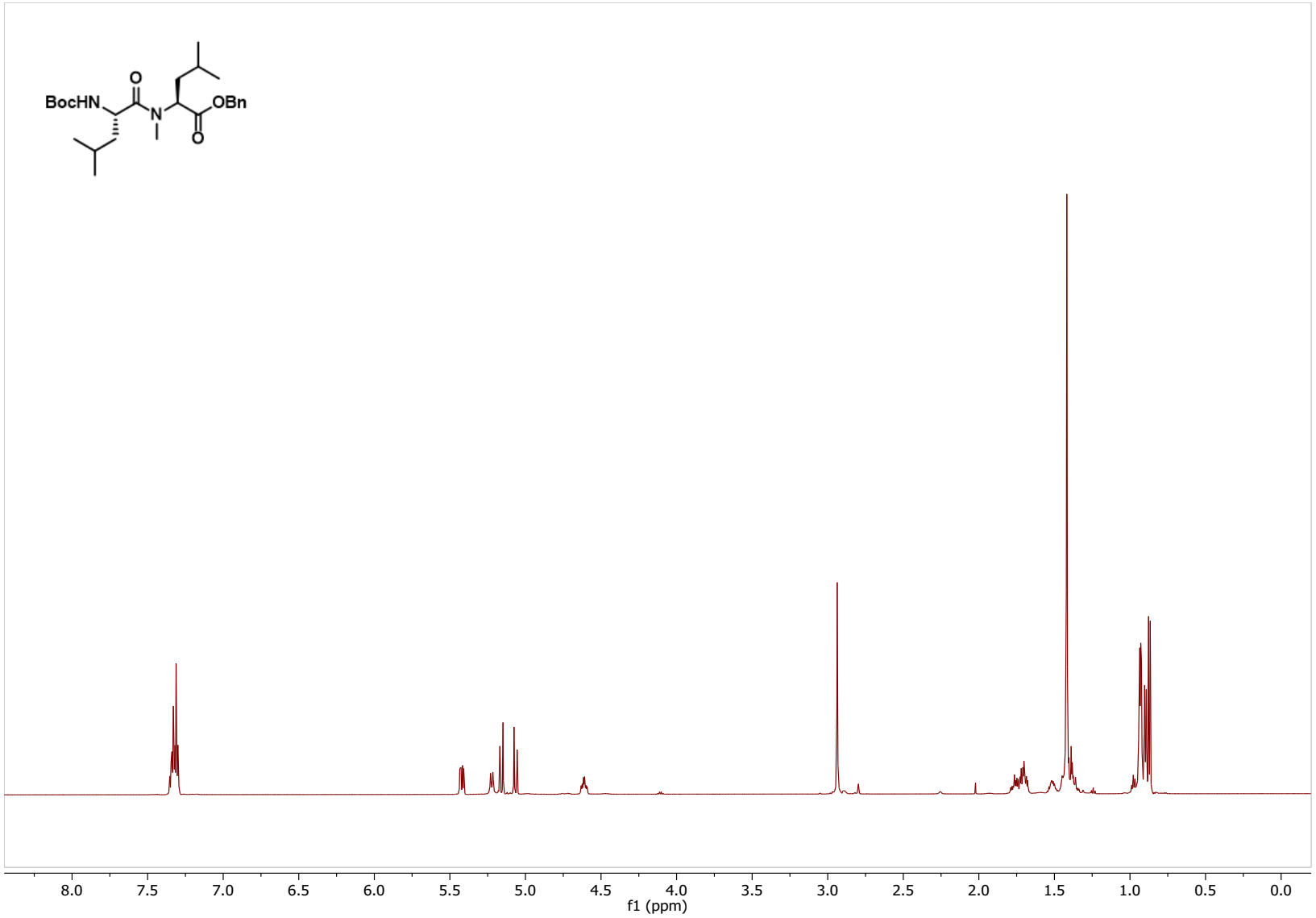

Fig. S17. ^{1}H NMR spectrum (600 MHz, DMSO-d_6) of compound **9.**

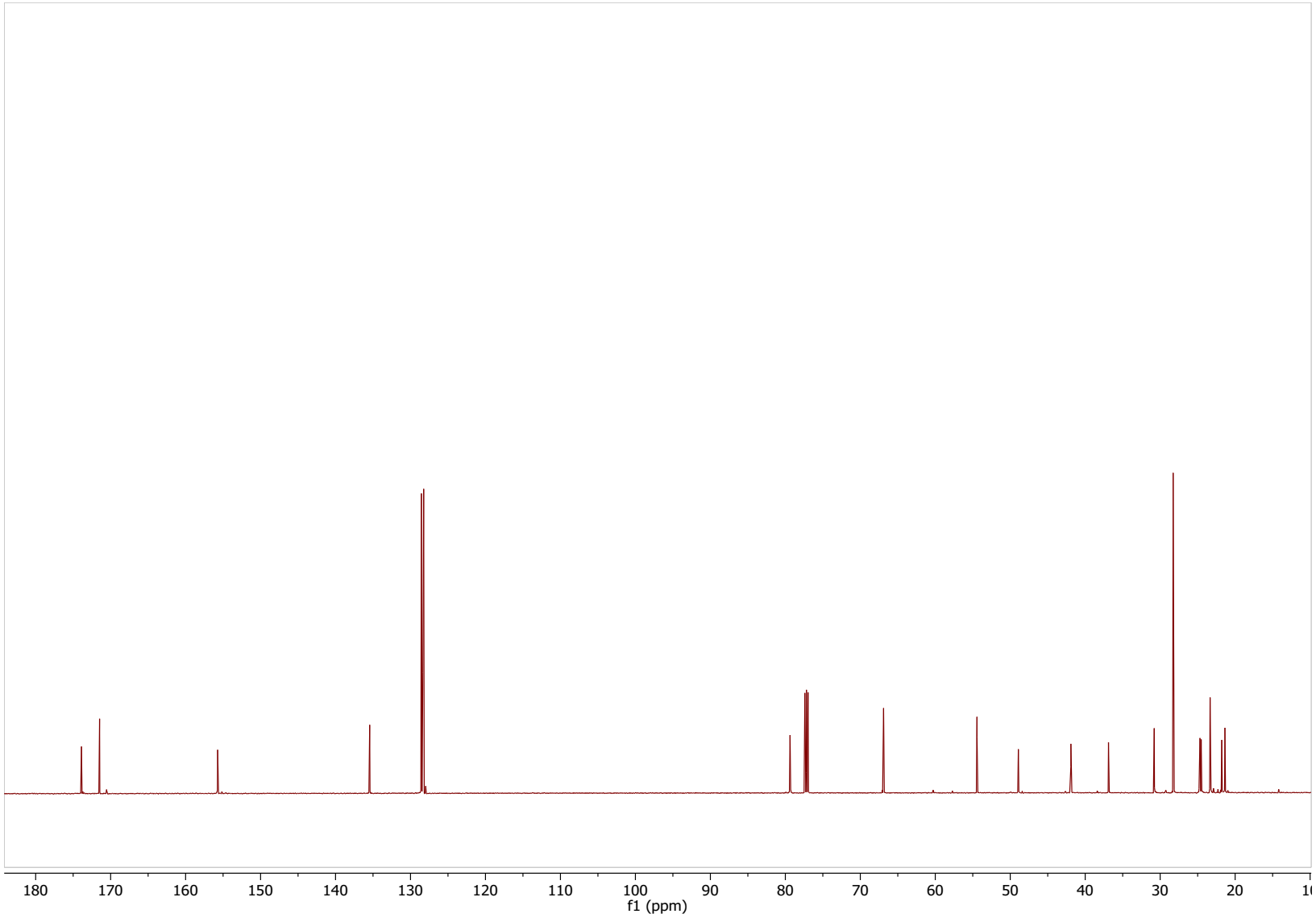

Fig. S18. ^{13}C NMR spectrum (150 MHz, DMSO-d_6) of compound **9.**

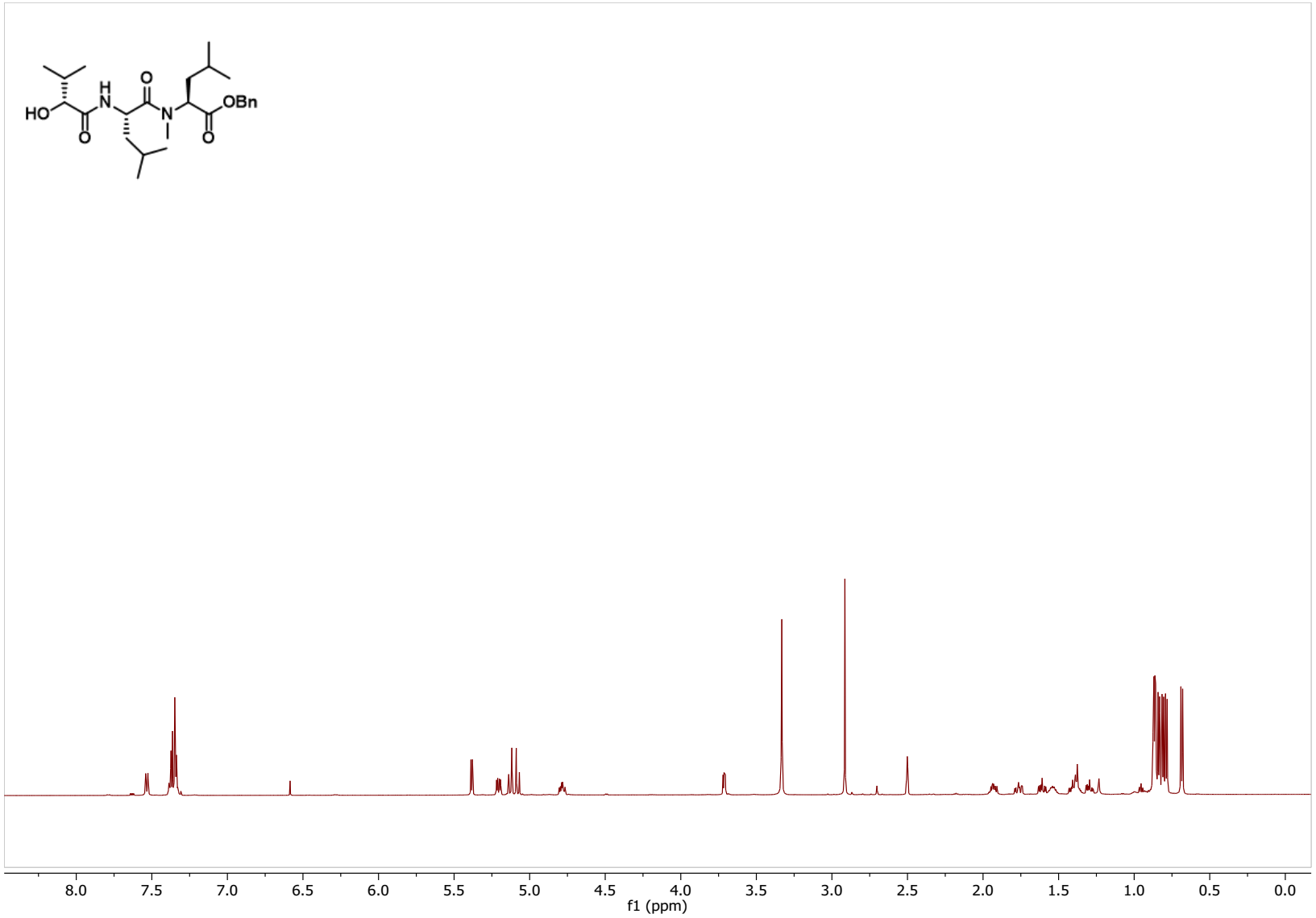

Fig. S19. ^{1}H NMR spectrum (600 MHz, DMSO-d_6) of compound **11.**

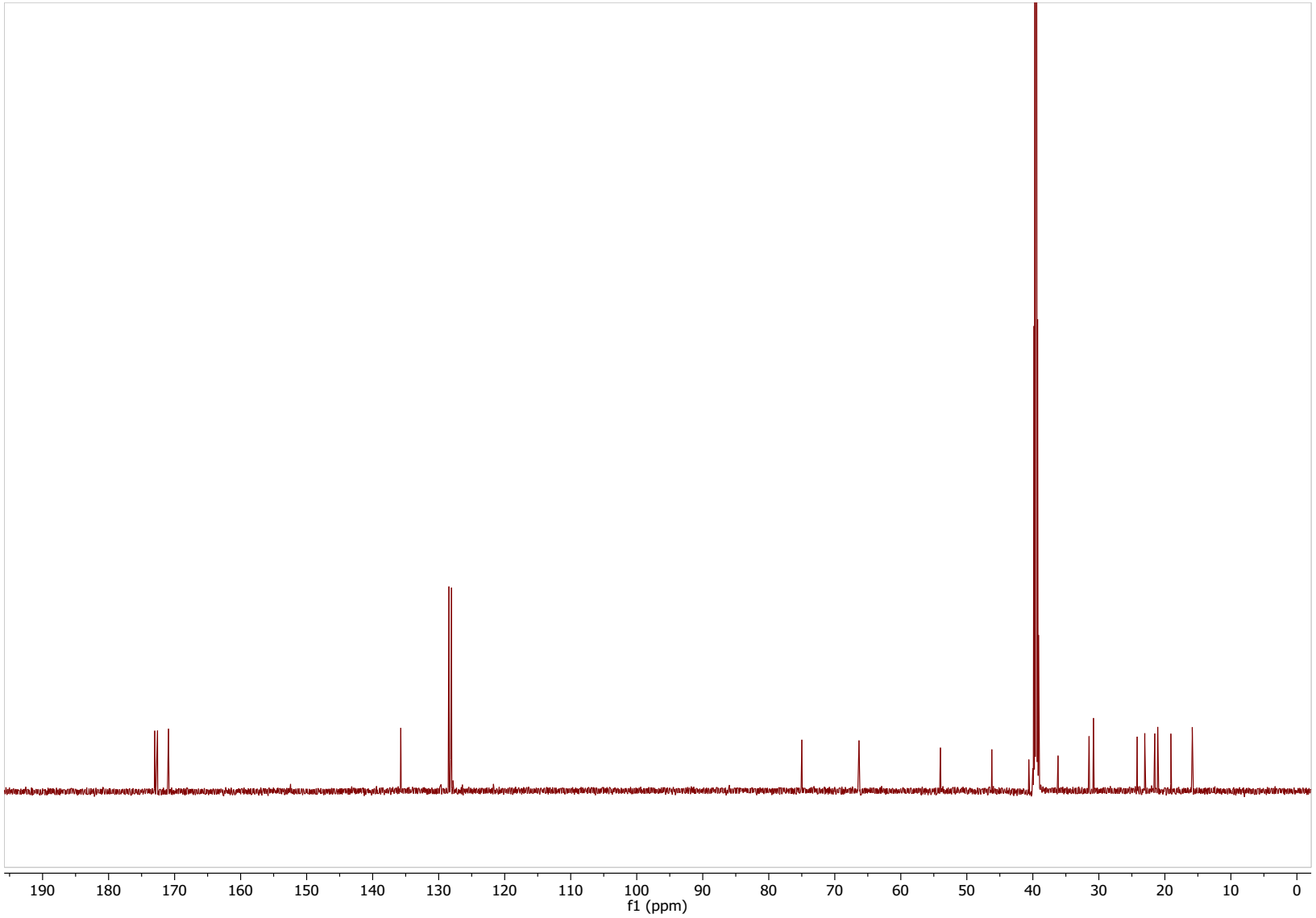

Fig. S20. ^{13}C NMR spectrum (150 MHz, DMSO-d_6) of compound **11.**

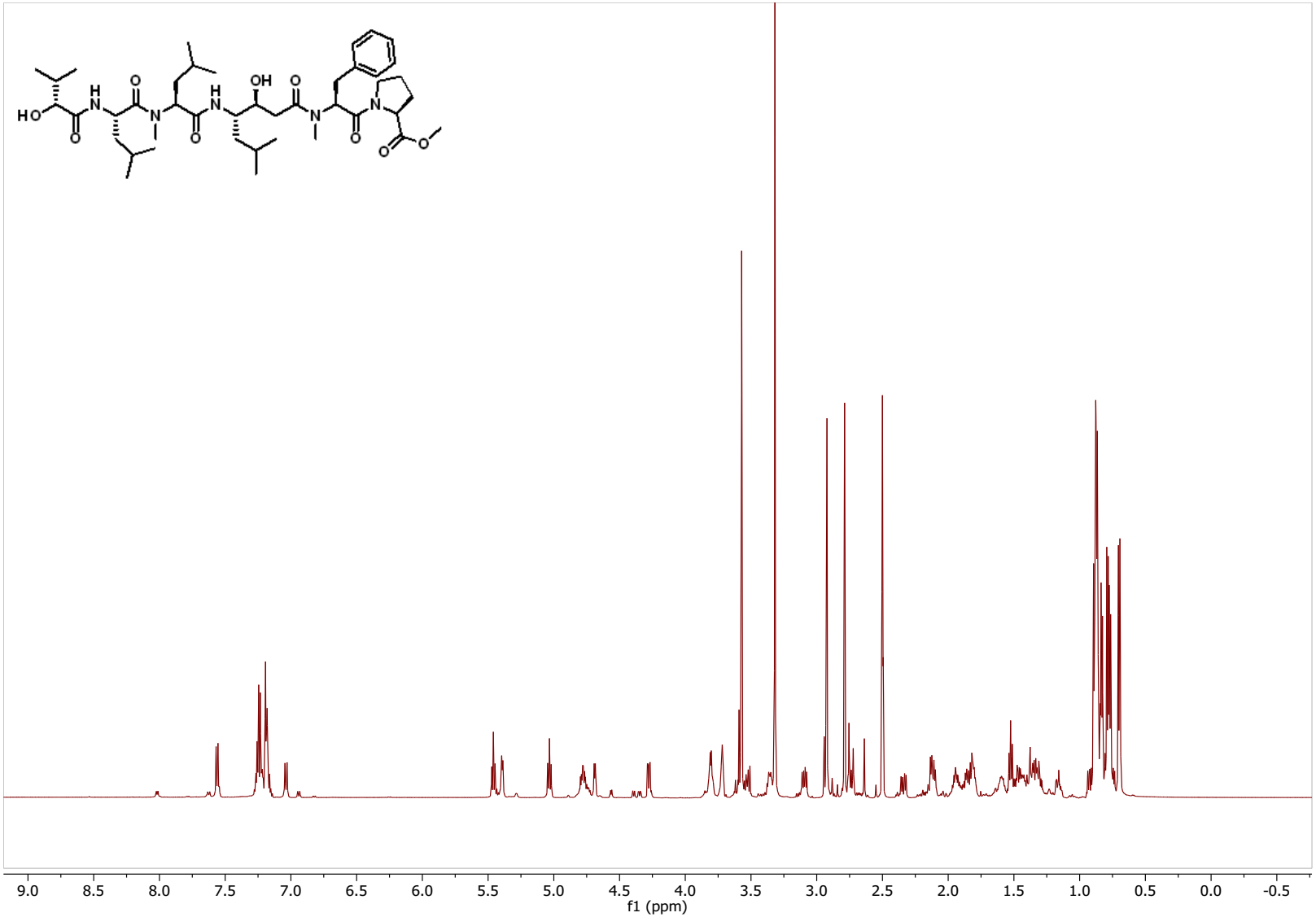

Fig. S21. ^{1}H NMR spectrum (600 MHz, DMSO-d_6) of synthetic grassystatin G.

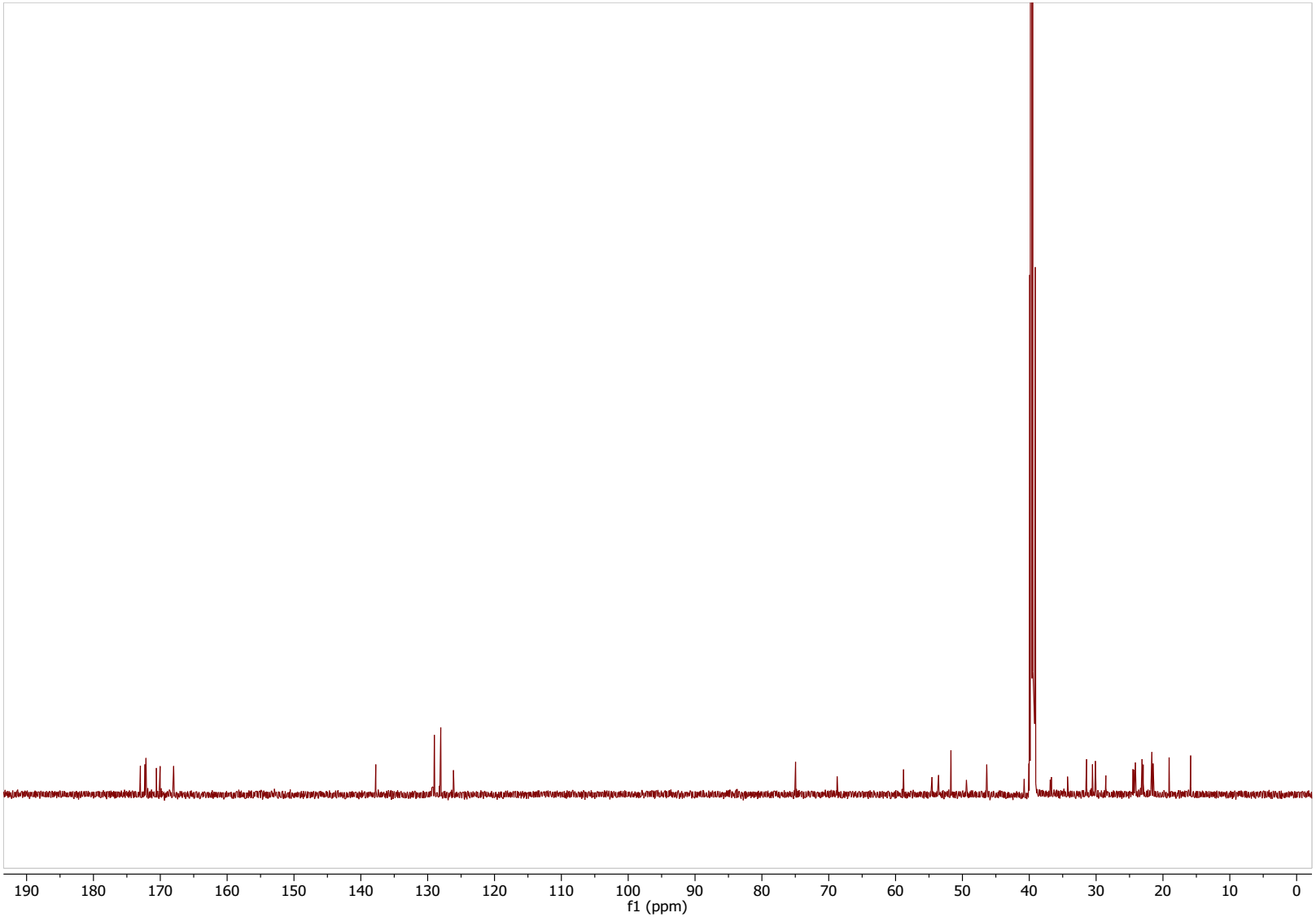

Fig. S22. ^{13}C NMR spectrum 150 MHz, DMSO-d_6) of synthetic grassystatin G.

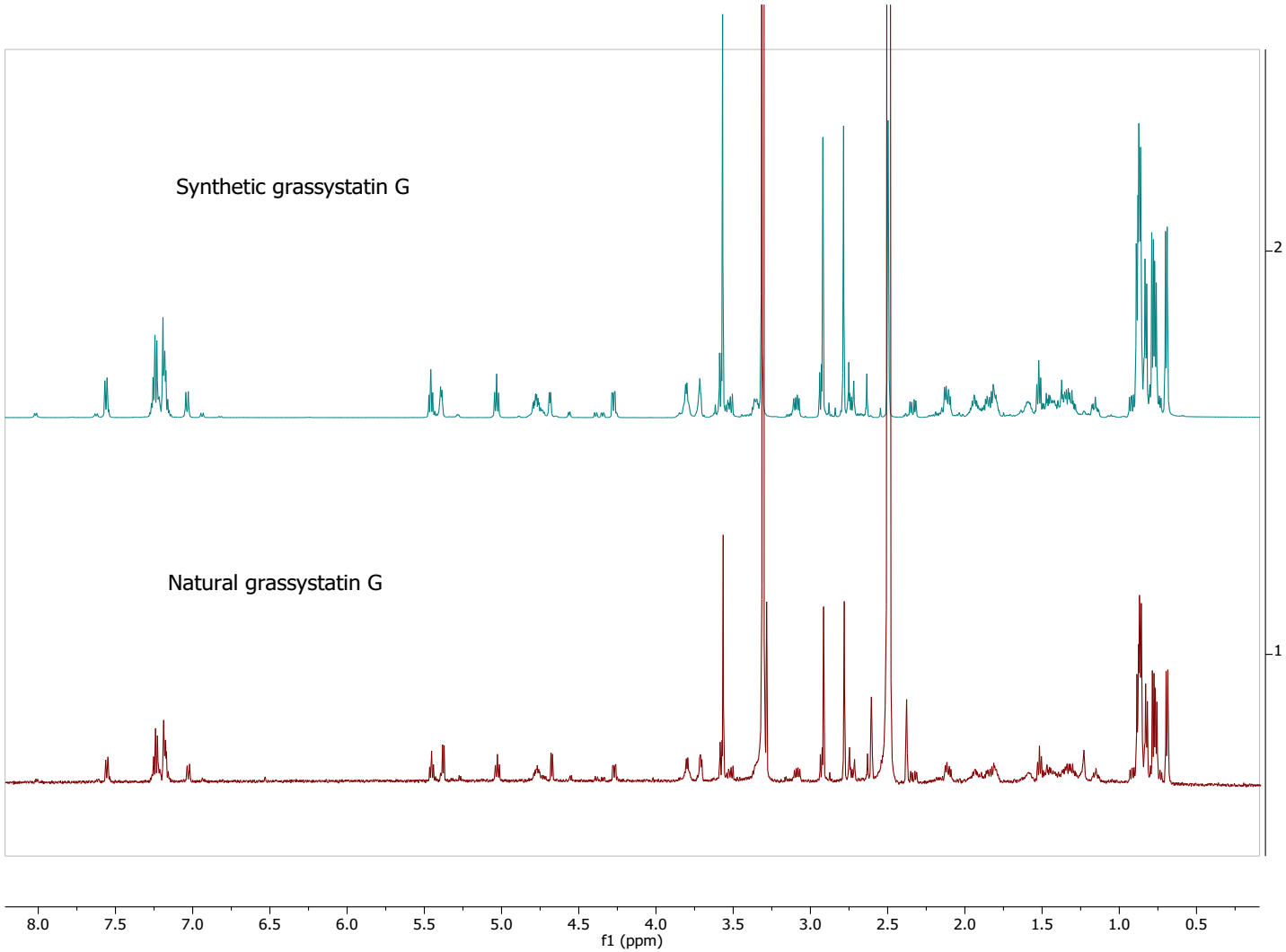

Fig. S23. Comparison of ^{1}H NMR spectrum of synthetic grassystatin G and natural grassystatin G in DMSO-d_6 (600 MHz).

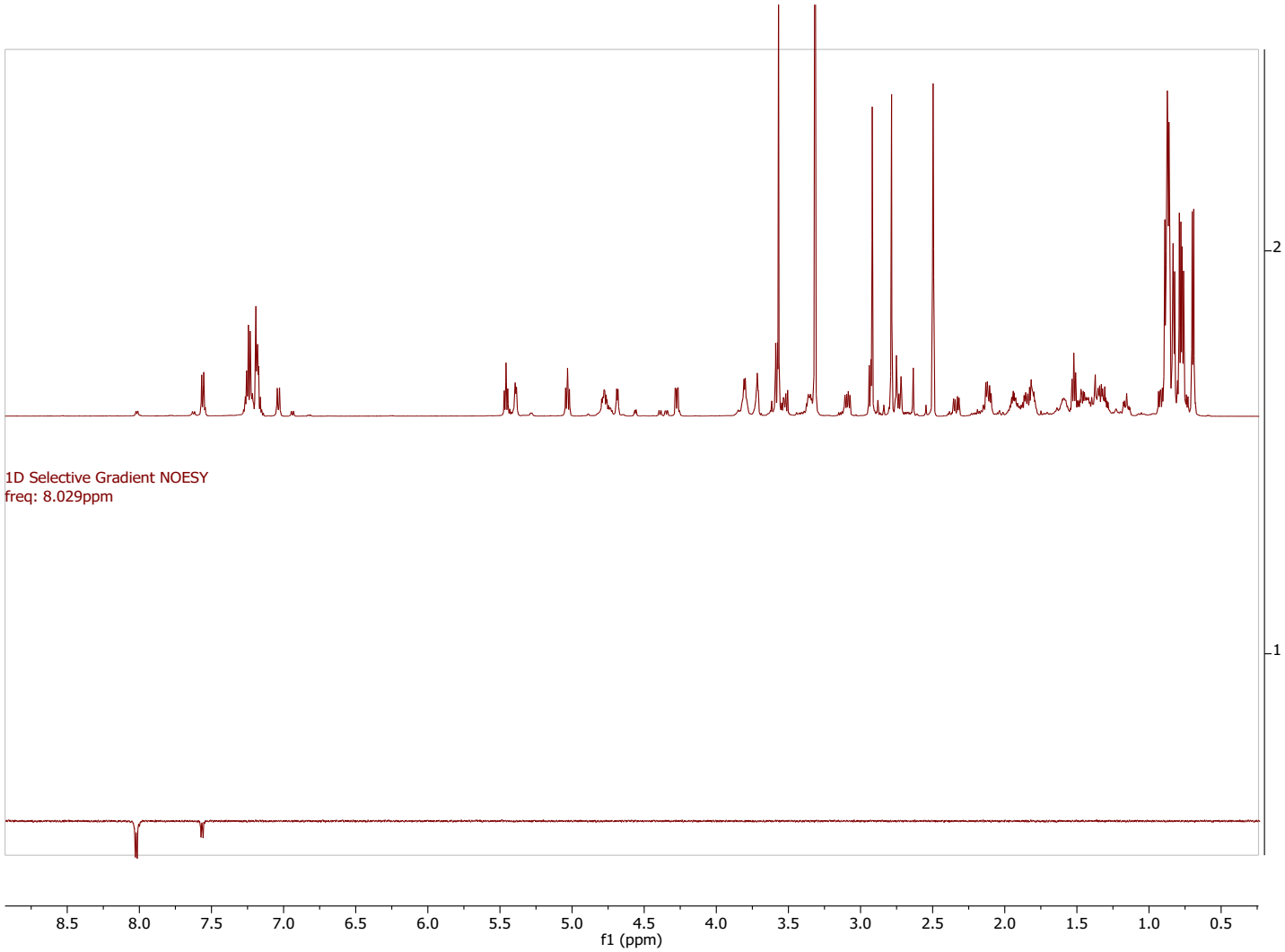

Fig. S24. 1D selective gradient NOE spectrum of synthetic grassystatin G with irradiation of signal at 8.0 ppm creates a new peak of the same phase at 7.56 ppm due to rotameric chemical exchange.

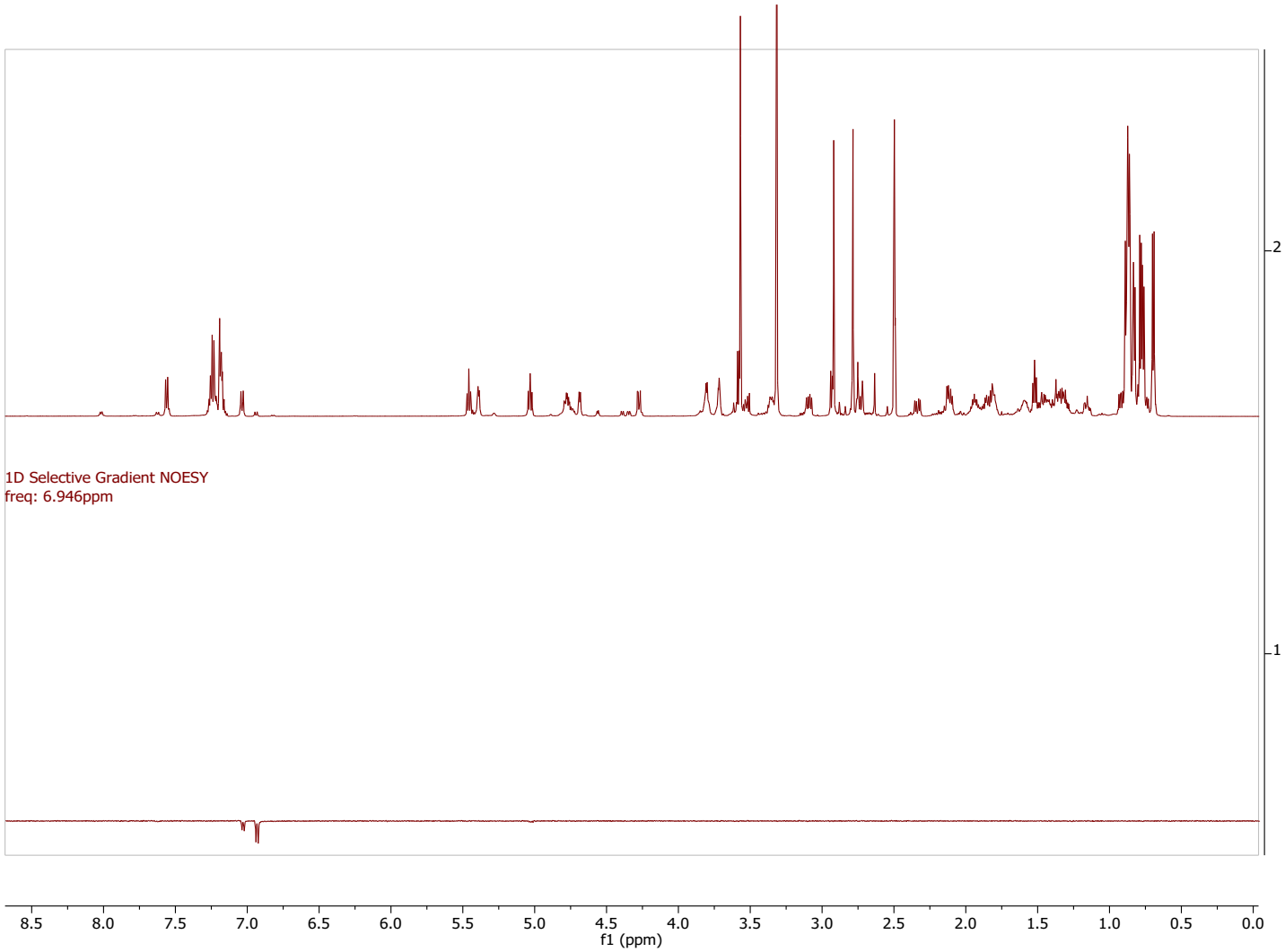

Fig. S25. 1D selective gradient NOE spectrum of synthetic grassystatin G with irradiation of signal at 6.94 ppm creates a new peak of the same phase at 7.05 ppm due to rotameric chemical exchange.

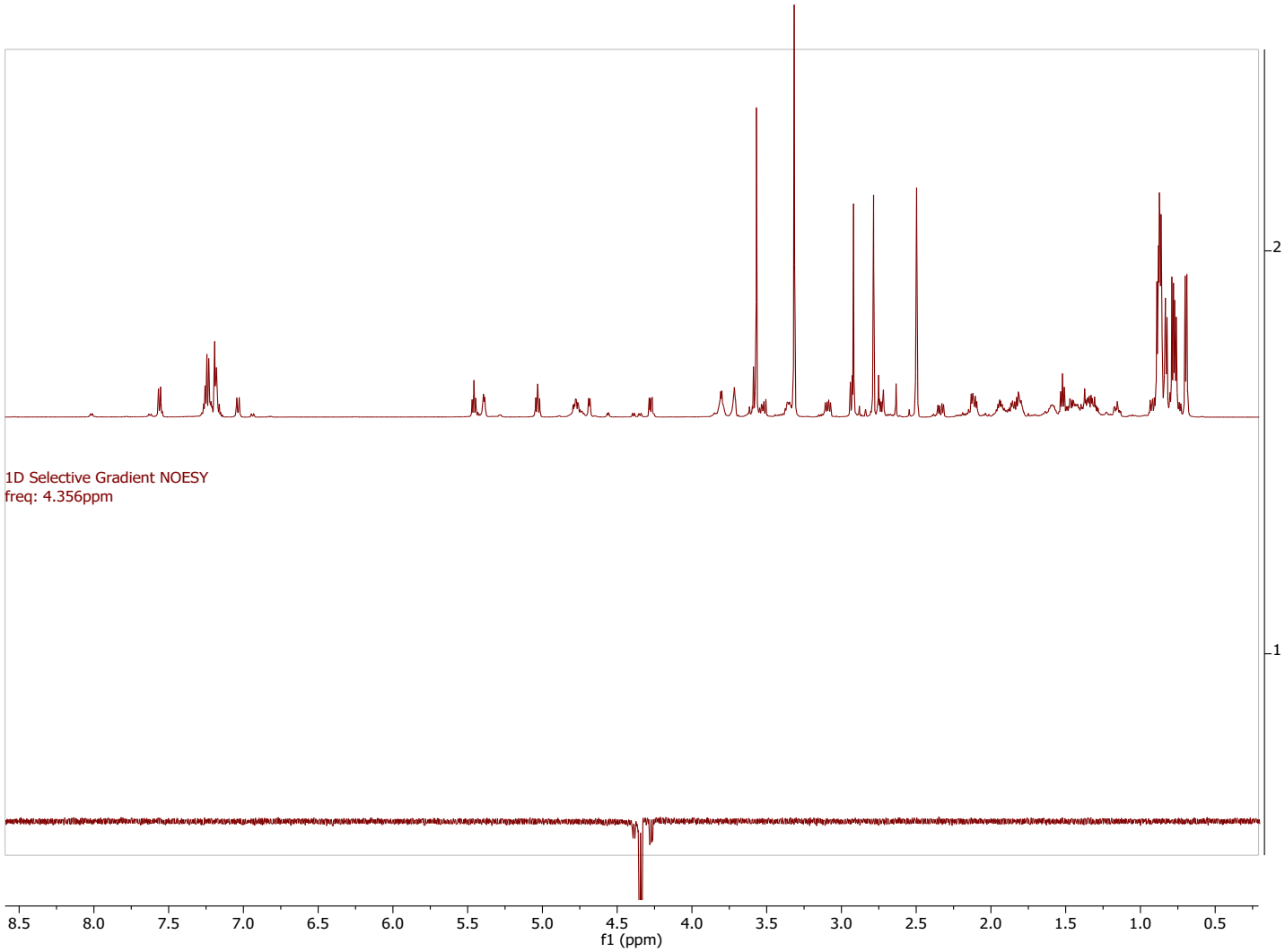

Fig. S26. 1D selective gradient NOE spectrum of synthetic grassystatin G with irradiation of signal at 4.35 ppm creates new peaks of the same phase at 4.28, 4.39 ppm due to rotameric chemical exchange.