Supporting Information

Resistance gene-guided genome mining: Serial promoter exchanges in *Aspergillus nidulans* reveal the biosynthetic pathway for fellutamide B, a proteasome inhibitor

Hsu-Hua Yeh,[1,6,‖] Manmeet Ahuja,[2,7,‖] Yi-Ming Chiang,[1,3] C. Elizabeth Oakley,[2] Shauna Moore,[2] Olivia Yoon,[2] Heather Hajovsky,[2] Jin-Woo Bok,[4] Nancy P. Keller,[4] Clay C. C. Wang,[1,5,*] Berl R. Oakley[2,*]

[1]Department of Pharmacology and Pharmaceutical Sciences, School of Pharmacy, University of Southern California, Los Angeles, California 90089, USA

[2]Department of Molecular Biosciences, University of Kansas, Lawrence, Kansas 66045, USA

[3]Department of Pharmacy, Chia Nan University of Pharmacy and Science, Tainan City 71710, Taiwan

[4]Department of Bacteriology and Department of Medical Microbiology and Immunology, University of Wisconsin, Madison, Wisconsin 53706, USA

[5]Department of Chemistry, University of Southern California, Dornsife College of Letters, Arts, and Sciences, Los Angeles, California 90089, USA

[6]Current address: Drug Discovery and Development Center, Chia Nan University of Pharmacy and Science, Tainan City 71710, Taiwan

[7]Current address: Industrial Biotechnology Division, Reliance Technology Group, Reliance Industries Limited, Reliance Corporate Park, Thane Belapur Road, Ghansoli, Navi Mumbia 400701, India

*Corresponding author: clayw@usc.edu; boakley@ku.edu

[‖]These authors contributed equally to this work.

Table of Contents

Supplemental Methods

Isolation of secondary metabolites

For scale up to isolate compound **1**, strain LO7860 was grown in 30 ml of liquid LMM in 125 ml flasks and induced with 50 mM of methyl ethyl ketone as described in **Materials and Methods**.

Thirty three flasks of LO7860 culture medium were collected and the combined medium (~ 950 ml) after filtration was extracted with 1 litre of EtOAc twice. The crude extract in the EtOAc layer (226 mg) was applied to a reverse phase C18 gel column (COSMOSIL 75C18-OPN, 20 × 30 mm) and eluted with MeOH-H_2O mixtures of decreasing polarity (fraction A, 1:9, 100 ml; fraction B, 3:7, 100 ml; fraction C, 7:3, 100 ml; fraction D, 1:0, 100 ml). Fraction B (84.2 mg), containing compound **1,** was subjected to additional purification by semi-preparative reverse phase HPLC [Phenomenex Luna 5 μm C18 (2), 250 × 10 mm] and monitored by a PDA detector at 254 nm. The gradient system was MeCN (solvent B) in 5 % MeCN/H_2O (solvent A) with the following gradient condition: 5 to 20 % B from 0 to 15 min, 20 to 100 % B from 15 to 17 min, maintained at 100 % B from 17 to 20 min, 100 to 5 % B from 20 to 21 min, and re-equilibration with 5 % B from 21 to 25 min. 12.3 mg of compound **1** eluted at 10.3 min was purified from 22.3 mg of Fraction B.

For compound **2** isolation, 33 flasks of LO8080 culture medium as described in the **Materials and Methods** were collected and filtered through filter paper to separate liquid media and fungal hyphae. The hyphae were then soaked in 400 ml of MeOH overnight for the extraction of compound **2** from the hyphae. It was found that the majority of compound **2** was present in the hyphal extract after LCMS analysis. The MeOH in the hyphal extract was then evaporated to obtain about 690 mg of dried crude material. The dried crude extract was suspended in 10 ml of DCM and the insoluble component was spun down. Compound **2** was found mainly in the DCM insoluble component. A total of 624.3 mg of DCM insoluble component was collected. This was applied to a reverse phase C18 gel column (COSMOSIL 75C18-OPN, 30 × 40 mm) and eluted with MeOH-H_2O mixtures of decreasing polarity (fraction 1, 5:5, 100 ml; fractions 2 – 11, 7:3, 100 ml; fractions 12 – 19, 1:0, 100 ml). All fractions were analyzed by LCMS and fractions 4 – 7 were found to contain compound **2** with > 95% purity, estimated from the total ion current (TIC)

trace of the chromatogram. Fractions 4 – 7 (40 ml) were then combined to obtain 24.2 mg of compound **2** after evaporation of the solvent.

Compound Spectral Data

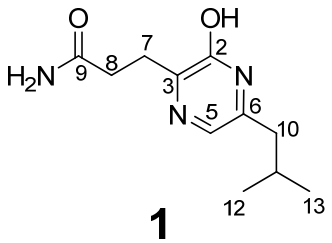
NMR spectra were collected on a Varian Mercury Plus 400 spectrometer. High resolution electrospray ionization mass spectra were obtained on an Agilent 6210 time of flight LC-MS. Optical rotations were measured on a JASCO P-2000 digital polarimeter.

2-Hydroxy-5-isobutyl-3-propanamidylpyrazin (**1**)

Colorless solid; For UV-Vis and ESIMS spectra, see Figure S8A; For ^{1}H and ^{13}C NMR (CD_3OD), see Table S1; HRESIMS, $[M + H]^+$ *m*/*z* found 224.1400, calc. for $C_{11}H_{18}N_3O_2$: 224.1399.

(2*RS*)-Fellutamide B (**2**)

White powder; $[\alpha]_D^{25}$ -30.5° (MeOH *c* 0.5); The solubility of **2** in CD_3OD was estimated to be > 17.0 mg/ml at 25.0 °C, calculated from ^{1}H NMR in CD_3OD based on the fact that 99.8% of deuterium in CD_3OD. For UV-Vis and ESIMS spectra, see Figure S8B. Note that due to the molar absorptivity of **2** and the high UV cut-off of formic acid used in the mobile phase, the UV-Vis spectrum of **2** is similar to its background signal. The UV-Vis spectrum shown in Figure S8B was the absorption of **2** with its background subtracted ; For ^{1}H and ^{13}C NMR (DMSO), see Table S2; HRESIMS, $[M + H]^+$ *m*/*z* found 556.3712, calc. for $C_{27}H_{50}N_5O_7$: 556.3710.

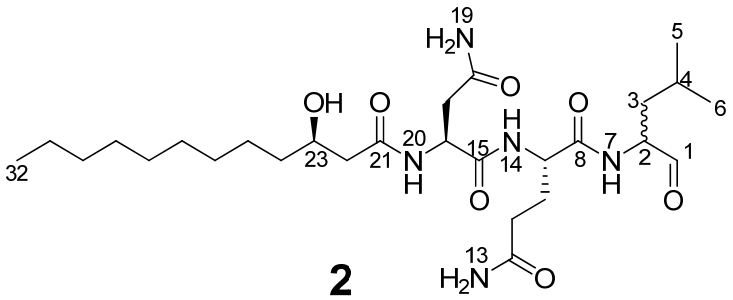


1

Table S1. NMR data for compound **1** in CD_3OD (400 and 100 MHz)[a]

	1	
position	δ_C	δ_H
2	158.9 (C)	—
3	156.3 (C)	—
5	123.9 (CH)	7.13 (1H, s)
6	140.7 (C)	—
7	29.4 (CH_2)	2.99 (2H, d, 7.2)
8	32.8 (CH_2)	2.61 (2H, d, 7.2)
9	178.3 (C)	—
10	40.3 (CH_2)	2.34 (2H, d, 7.6)
11	29.7 (CH)	1.93 (1H, m)
12 & 13	22.6 (CH_3)	0.93 (6H, d, 6.4)

[a]Figures in parentheses are multiplicities and coupling constants (*J*) in Hz.



2

Table S2. NMR data for fellutamide B (**2**) in DMSO-d_6 (400 and 100 MHz)[a]

	(2*S*)-Fellutamide B[b]		(2*RS*)-Fellutamide B (**2**)	
position	δ_C	δ_H	δ_C	δ_H
1	201.6 (CH)	9.35 (1H, s)	201.69 (CH)	9.34 & 9.38 (1H, s)
2	56.6 (CH)	4.00 – 4.08 (1H, m)	56.58 & 56.64 (CH)	4.00 – 4.08 (1H, m)[d]
3	36.3 (CH_2)	1.44 – 1.56 (2H, m)	36.26 (CH_2)	1.44 – 1.52 (2H, m)[d]
4	23.9 (CH)	1.58 – 1.68 (1H, m)	23.91 (CH)	1.56 – 1.66 (1H, m)[d]
5	21.3 (CH_3)	0.84 (3H, d, 6.4)	21.29 & 21.33 (CH_3)	0.84 (3H, m)[d]
6	23.1 (CH_3)	0.89 (3H, d, 6.8)	23.12 (CH_3)	0.89 (3H, m)[d]
7-NH	—	8.27 (1H, d, 7.2)	—	8.25 & 8.29 (1H, d, 7.2)
8	171.9 (C)	—	171.78 & 171.82 (C)	—
9	52.4 (CH)	4.16 – 4.21 (1H, m)	52.44 & 52.54 (CH)	4.08 – 4.22 (1H, m)[d]
10	27.5 (CH_2)	1.68 – 1.82 (1H, m), 1.90 – 2.02 (1H, m)	27.49 (CH_2)	1.68 – 1.82 (1H, m)[d], 1.92 – 2.04 (1H, m)[d]
11	31.4 (CH_2)	2.04 – 2.14 (2H, m)	31.34 (CH_2)	2.04 – 2.14 (2H, m)[d]
12	173.8 (C)	—	173.81 (C)	—
13-NH_2	—	6.76 (1H, br s), 7.21 (1H, br s)	—	6.78 (1H, br s), 7.23 (1H, br s)
14-NH	—	8.10 (1H, d, 7.6)	—	8.12 – 8.20 (1H)[d]
15	171.1 (C)	—	171.11 & 171.17 (C)	—
16	49.8 (CH)	4.50 (1H, q, 7.2)	49.85 (CH)	4.45 – 4.55 (1H, m)
17	37.0 (CH_2)	2.44 (1H, dd, 15.5, 7.2), 2.55 (1H, dd, 15.5, 7.2)	37.06 (CH_2)	2.44 (1H, dd, 15.6, 6.8), 2.51 – 2.59 (1H, m)
18	171.8 (C)	—	171.88 (C)	—
19-NH_2	—	6.93 (1H, br s), 7.42 (1H, br s)	—	6.95 & 6.98 (1H, br s), 7.45 & 7.46 (1H, br s)
20-NH	—	8.09 (1H, d, 7.2)	—	8.02 – 8.20 (1H)[d]
21	171.2 (C)	—	171.21 (C)	—
22	43.5 (CH_2)	2.16 – 2.24 (2H, m)	43.48 (CH_2)	2.15 – 2.25 (2H, m)[d]
23	67.5 (CH)	3.72 – 3.82 (1H, m)	67.47 (CH)	3.72 – 3.82 (1H, m)[d]
24	36.9 (CH_2)	1.20 – 1.40 (2H, m)	36.91 (CH_2)	1.16 – 1.40 (2H, m)[d]
25	25.1 (CH_2)	1.20 – 1.40 (2H, m)	25.17 (CH_2)	1.16 – 1.40 (2H, m)[d]
26	29.1 (CH_2)	1.20 – 1.40 (2H, m)	29.15 (CH_2)[c]	1.16 – 1.30 (2H, m)[d]
27	29.1 (CH_2)	1.20 – 1.40 (2H, m)	29.14 (CH_2)[c]	1.16 – 1.30 (2H, m)[d]
28	29.0 (CH_2)	1.20 – 1.40 (2H, m)	29.03 (CH_2)	1.16 – 1.30 (2H, m)[d]
29	28.7 (CH_2)	1.20 – 1.40 (2H, m)	28.76 (CH_2)	1.16 – 1.30 (2H, m)[d]

30	31.3 (CH_2)	1.20 – 1.40 (2H, m)	31.34 (CH_2)	1.16 – 1.30 (2H, m)[d]
31	22.1 (CH_2)	1.20 – 1.40 (2H, m)	22.13 (CH_2)	1.16 – 1.30 (2H, m)[d]
32	14.0 (CH_3)	0.85 (3H, t, 6.8)	13.99 (CH_3)	0.85 (3H)[d]
OH		4.63 (1H, d, 5.2)		4.66 (1H, br d, 4.8)

[a]Figures in parentheses are multiplicities and coupling constants (*J*) in Hz.

[b]Data obtained from previous studies (Wu et al., *Mar. Drugs* **2014**, *12*, 1815-1838).

[c]Values may be interchanged.

[d]Data obtained from ^{1}H-^{13}C gHMQC correlations.

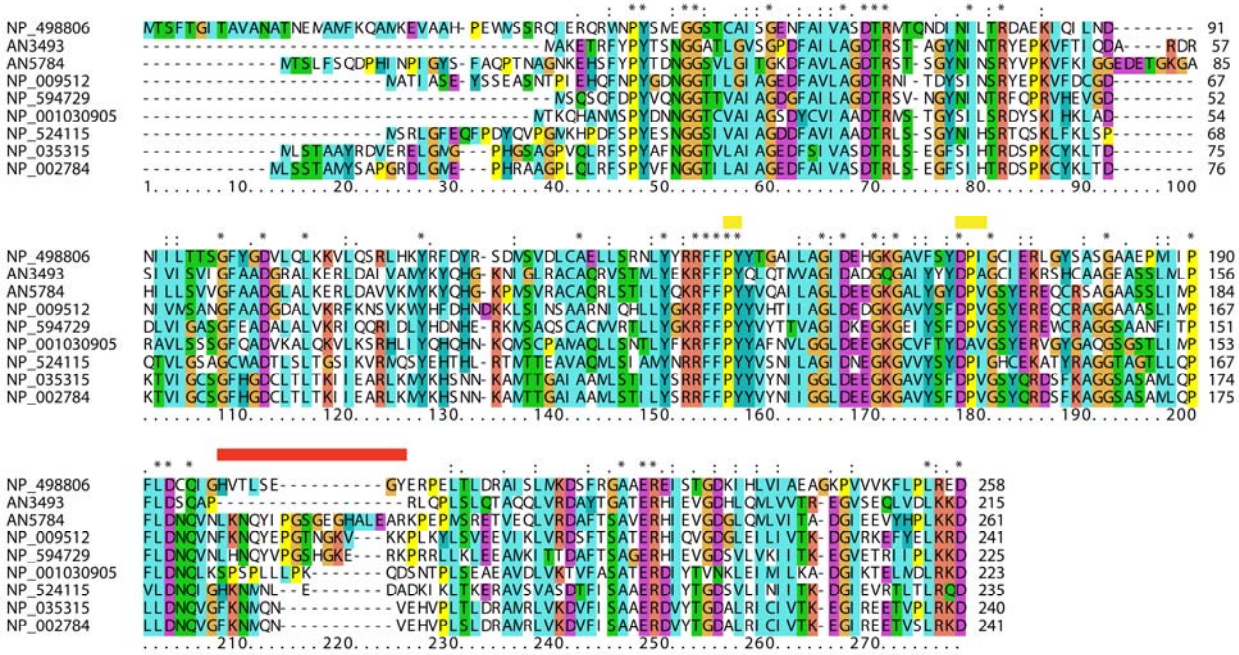
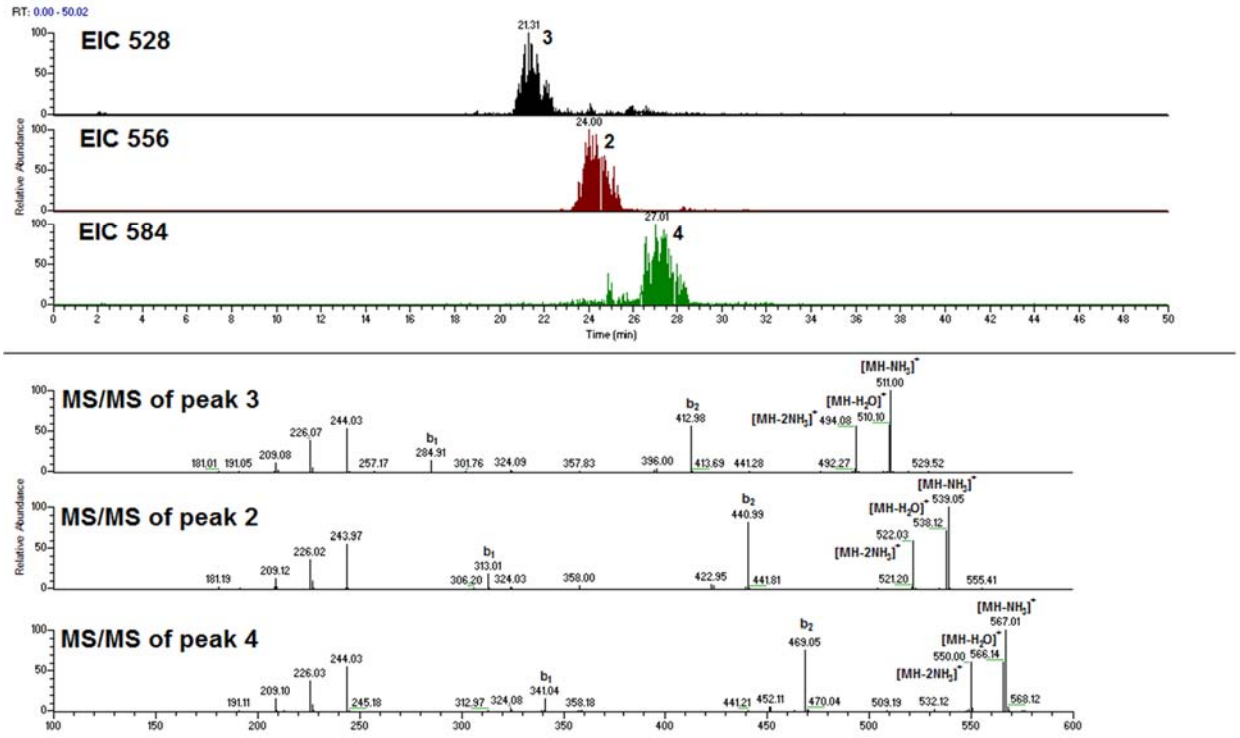
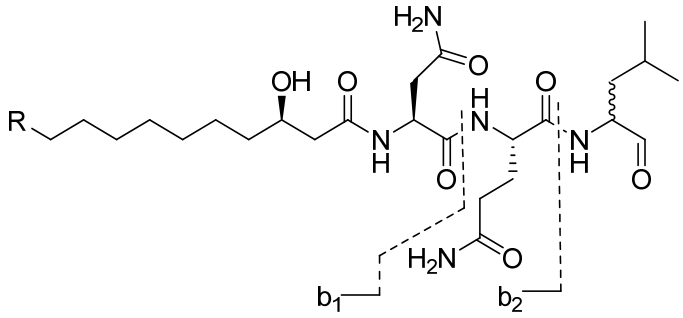


Figure S1. Alignment of the amino acid sequences of 20S proteasome subunit β6 (also known as Proteasome Subunit Beta Type-1, PSMB1). The alignment was done by Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo). This gene is conserved from yeast to humans (HomoloGene: 2087). Species used for this alignment including *Saccharomyces cerevisiae* (NP_009521), *Schizosaccharomyces pombe* (NP_594729), *Caenorhabditis elegans* (NP_498806), *Drosophila melanogaster* (NP_524115), *Arabidopsis thaliana* (NP_001030905), *Mus musculus* (NP_035315), and *Homo sapiens* (NP_002784).





R = H (**3**)	285	413
R = CH_3CH_2 (**2**)	313	441
R = CH_3 CH_2 CH_2 CH_2 (**4**)	341	469

Figure S2. MS/MS Spectra of peaks **2**, **3**, and **4**. The parent ions of peaks **3**, **2**, and **4** are *m*/*z* 528, 556, 584, respectively. The b_1 and b_2 fragments of fellutamides with different lengths of the side chains were labeled in MS/MS spectra.

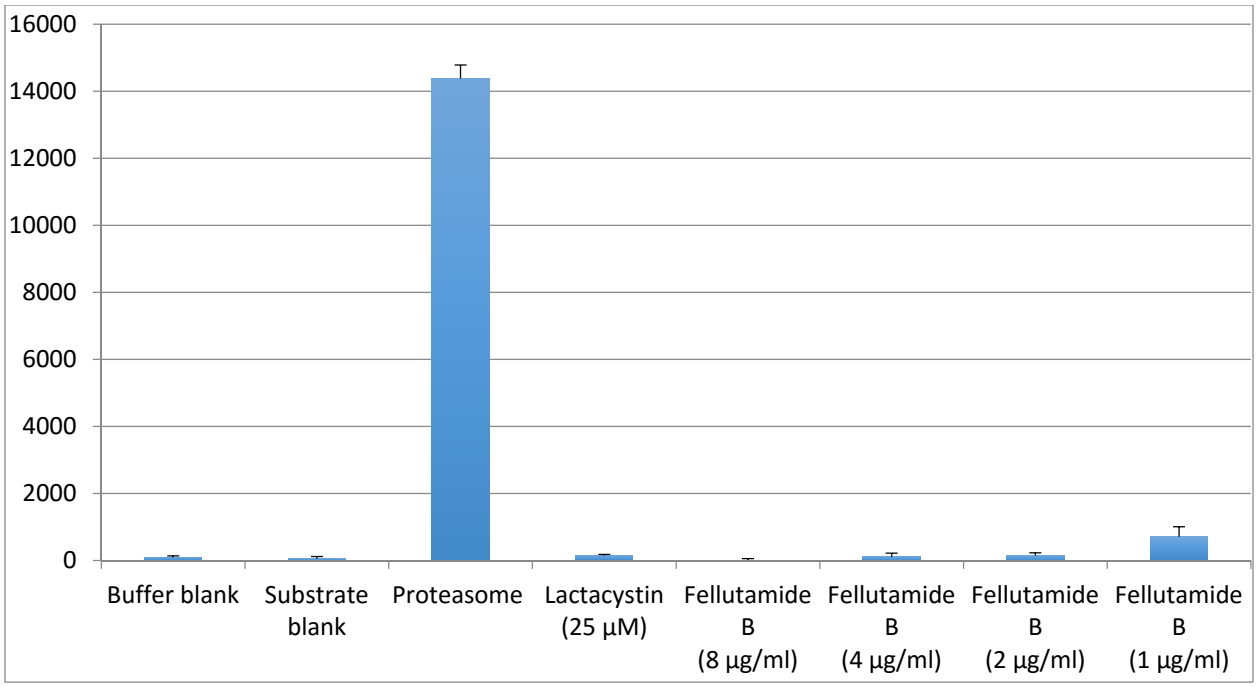


Figure S3. Inhibition of proteasome activity by fellutamide B (**2**) from *A. nidulans*. The activity of a proteasomal preparation from rabbits was assayed using a Millipore 20S Proteasome Activity Kit. The Y axis is relative fluorescence units. Results are means of three assays. Error bars are standard deviations. All values were significantly different from the positive control (Proteasome) ($P < 10^{-6}$ student t test). Lactacystin is a proteasome inhibitor control that is supplied with the kit. Fellutamide B from *A. nidulans* strongly inhibited proteasomal activity with approximately 95% inhibition at 1 µg/ml (1.8 µM). Relative fluorescence units are linear with respect to proteasomal activity over the range tested.

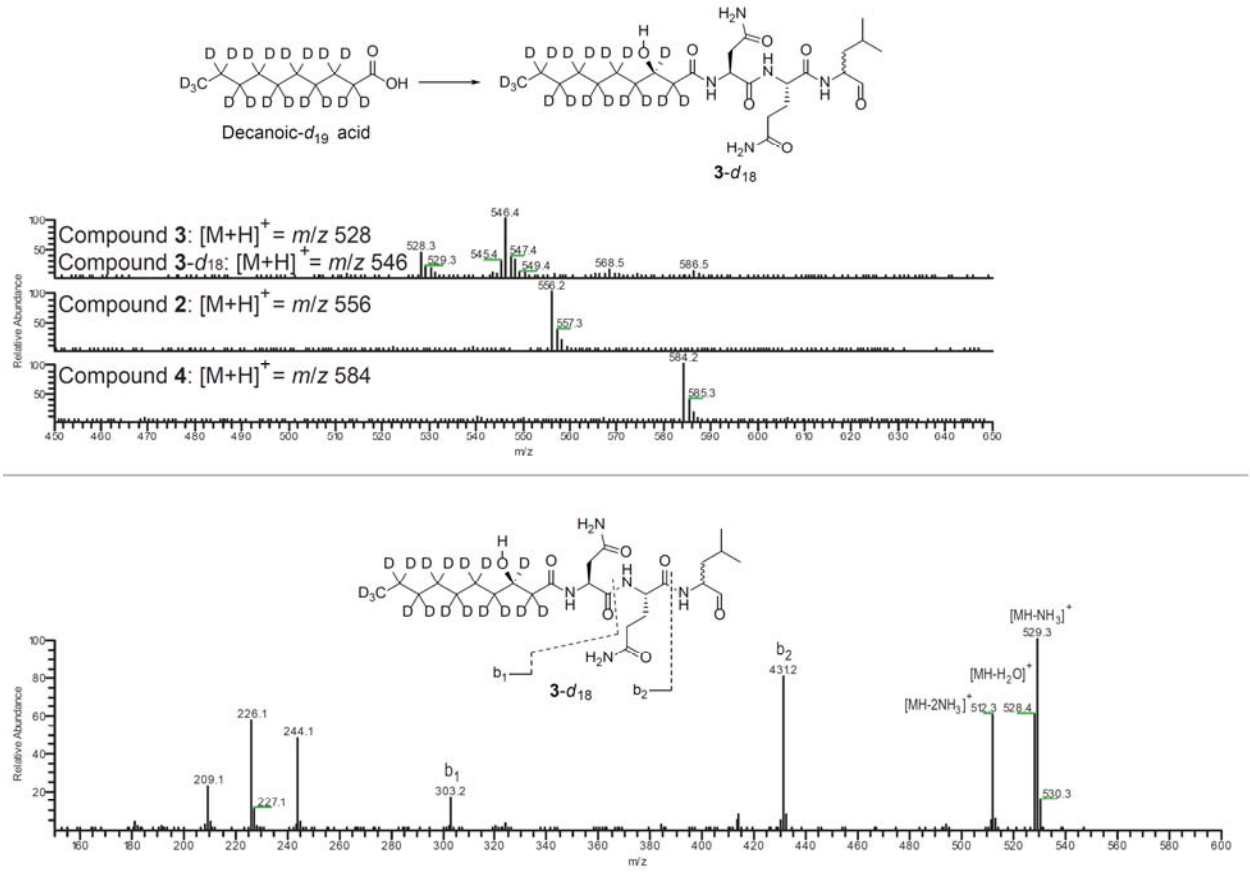


Figure S4. Fatty acids incorporated into the fellutamide biosynthesis pathway. Decanoic-d_{19} acid was fed to the liquid culture of fellutamide producing strain (LO8080). Decanoic-d_{19} acid was incorporated into **3**. This incorporation produced **3**-d_{18} and caused the mass shift (upper panel). The identity of **3**-d_{18} was further confirmed by the MS/MS spectrum of the parental ion (m/z = 546) of **3**-d_{18} (lower panel).

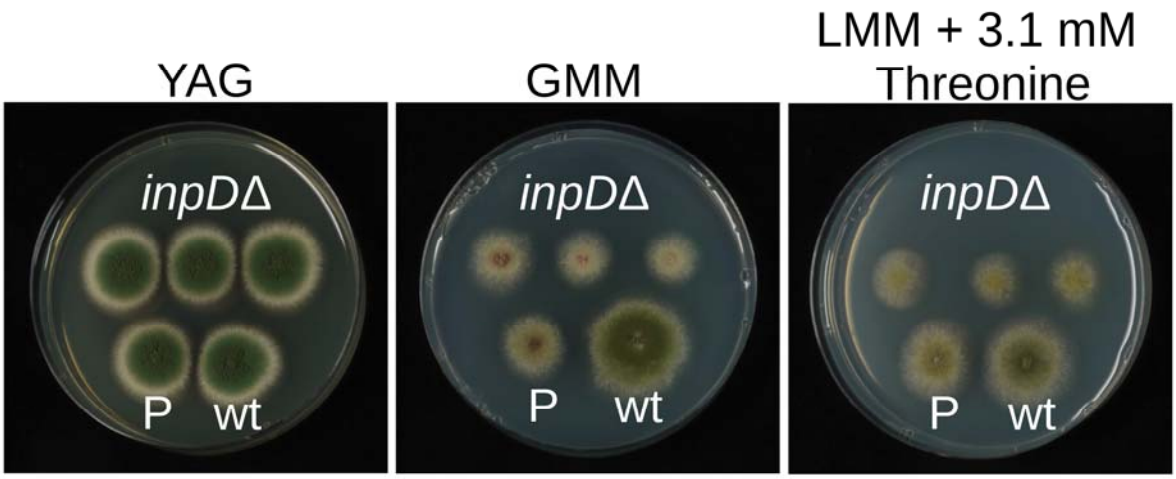


Figure S5. Induction of fellutamide production causes a slight reduction of growth of *inpD* deletion strains relative to a parental control. The top three colonies are three sister transformants (LO10093, LO10096, LO10097) in which the *inpD* gene has been deleted by replacing it with *AtpyrG*. The bottom left strain (P) is LO9222, the parental strain in which *inpD* was deleted. At the bottom right is a wild type strain, FGSC4. (See note on FGSC4 in the Experimental section.) On YAG, a complete medium that strongly represses the *alcA* promoter (and therefore fellutamide production), all strains grow very well and colony size is nearly indistinguishable. On glucose minimal medium (GMM), a medium that represses the *alcA* promoter but not quite as strongly as YAG, growth of the *inpD* deletants is slightly reduced relative to the parental strain. On lactose minimal medium (non-repressing) with the *alcA* inducer 3.1 mM threonine, growth of the *inpD* deletants relative to their parental strain (which grows slightly better on LMM plus threonine than it does on GMM) is reduced slightly more than in GMM, suggesting that *inpD* may make a small contribution to fellutamide resistance. The YAG plate was incubated for two days at 37ºC and the GMM and LMM + threonine plates were incubated for three days at 37ºC. Uridine and uracil were added to all media to supplement the *pyrG* marker that is uncomplemented in LO9222.

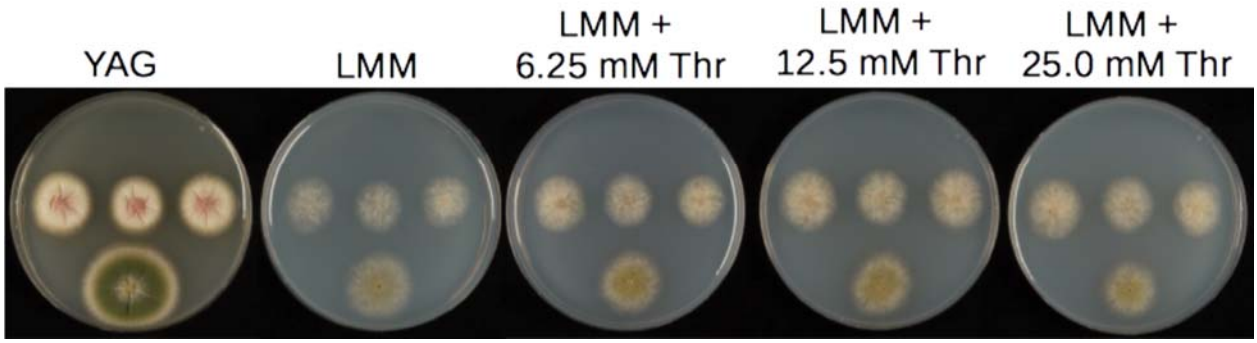


Figure S6. Deletion of AN5784 does not confer fellutamide sensitivity. The colony at the bottom of each plate is LO9222 in with *inpA-D* and *inpF* are under control of the *alcA* promoter. The three colonies above LO9222 are transformants in which AN5784 has been deleted by replacing it with a pyrithiamine resistance gene. AN5784 is distant from the *inp* cluster and it encodes a proteasome β6 subunit. The *alcA* promoter is strongly repressed on YAG (Yeast extract, agar, glucose) medium. The four plates at the right contain the same colonies growing on lactose minimal medium (LMM) with various concentrations of threonine as shown. The *alcA* promoter is derepressed on LMM and induced by threonine such that fellutamide B is produced. Deletion of AN5784 causes reduction of colonial growth and conidiation but does not confer sensitivity to internally produced fellutamide B.

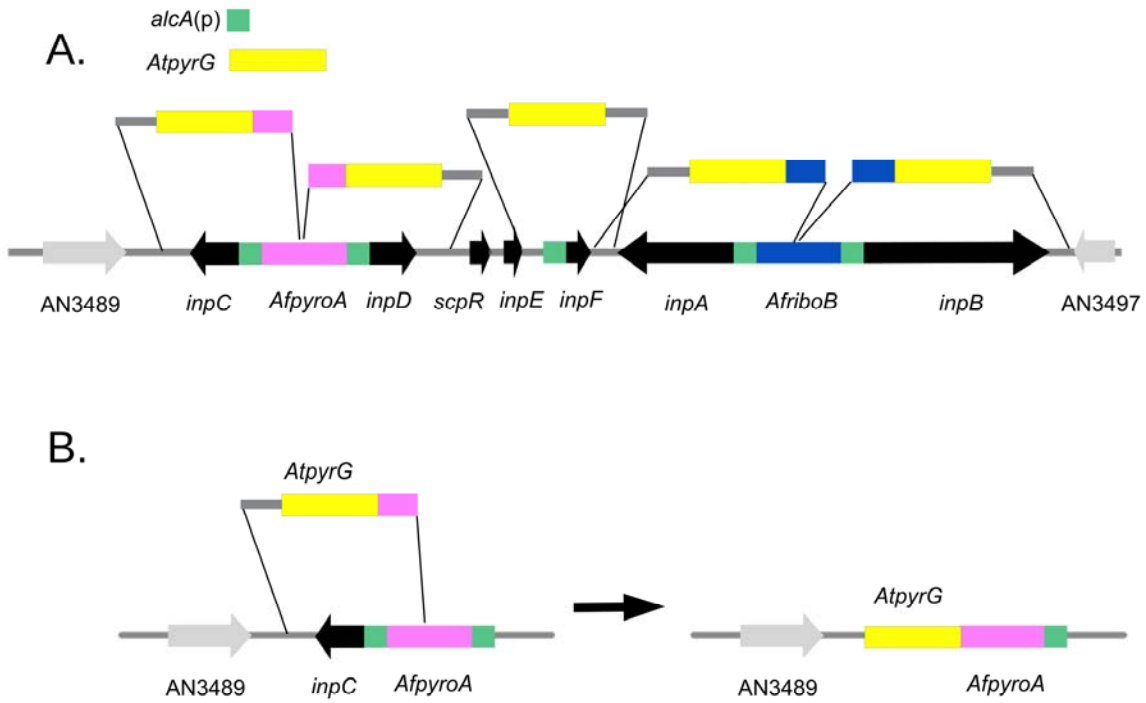


Figure S7. Schematic deletion strategy for *inp* cluster genes. (**A**) Previous transformations had placed *inpA, B, C, D and F* under control of the *alcA* promoter, leaving two selectable markers *AfpyroA and AfriboB*, in the cluster (Figure 2 in the main text). *Inp* cluster genes were individually deleted by replacing them with the *Aspergillus terreus pyrG* gene (*AtpyrG*). (**B**) An example in which the *inpC* gene and the *alcA* promoter driving its expression were transformed with a fragment in which the *AtpyrG* gene was flanked by a genomic sequence adjacent to the *inpC* gene and by a portion of the *AfpyroA* gene. Homologous recombination results in replacement of the *inpC* gene, and the *alcA* promoter driving it, with *AtpyrG*.

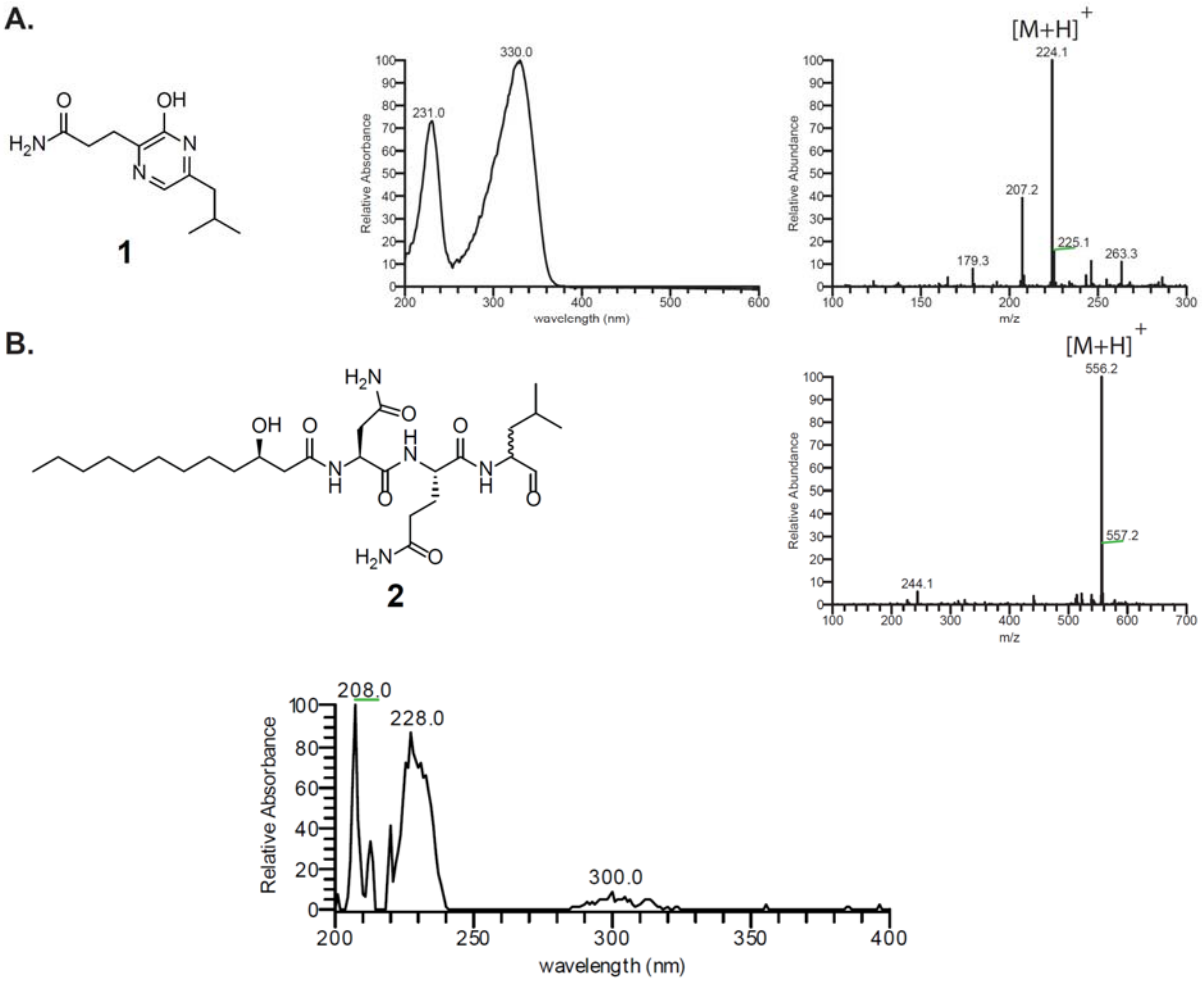


Figure S8. (**A**) UV-Vis and mass spectra of compound **1** and (**B**) **2**.

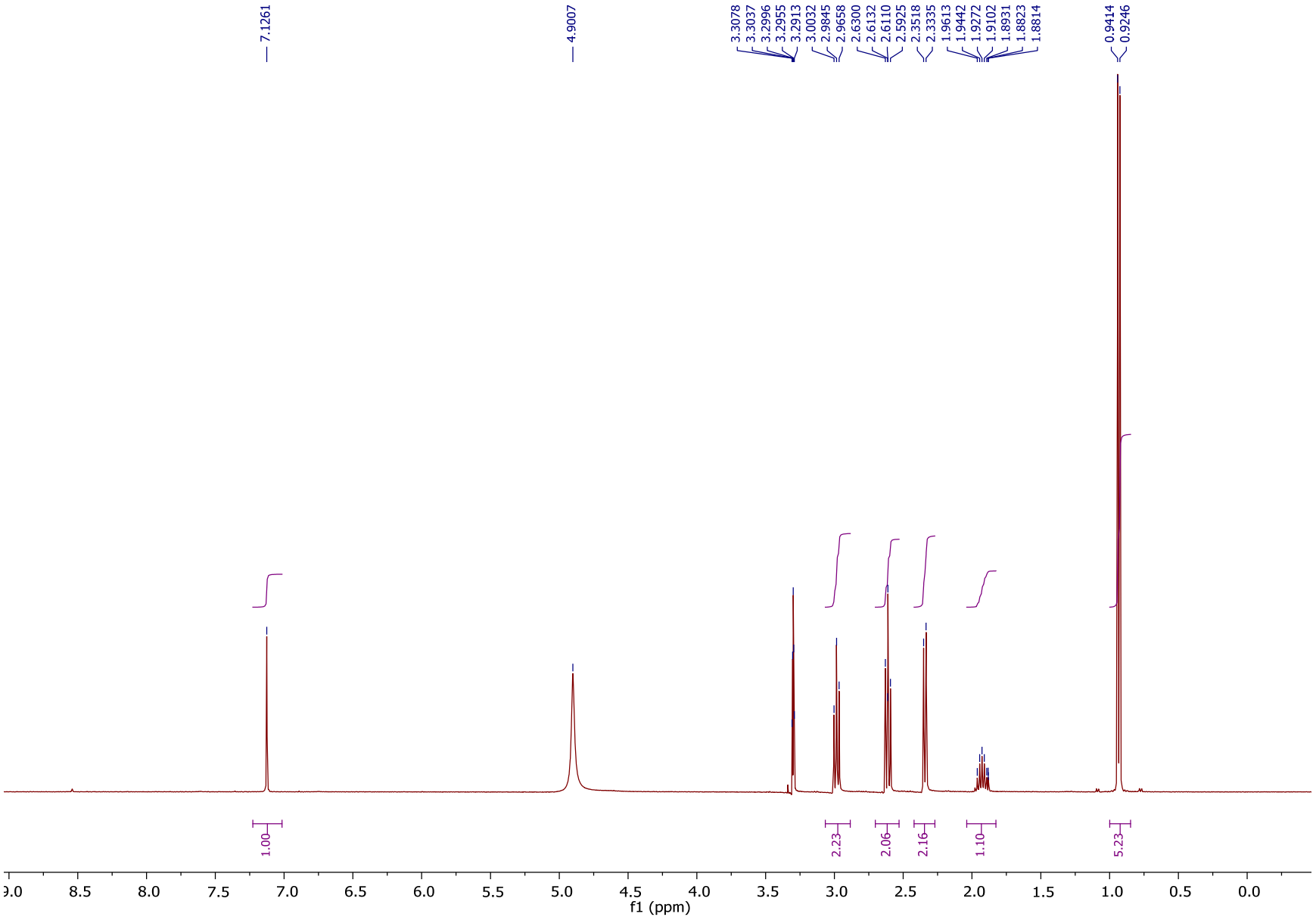


Figure S9. ^{1}H NMR of 2-hydroxy-5-isobutyl-3-propanamidylpyrazin (**1**) in CD_3OD (400 MHz).

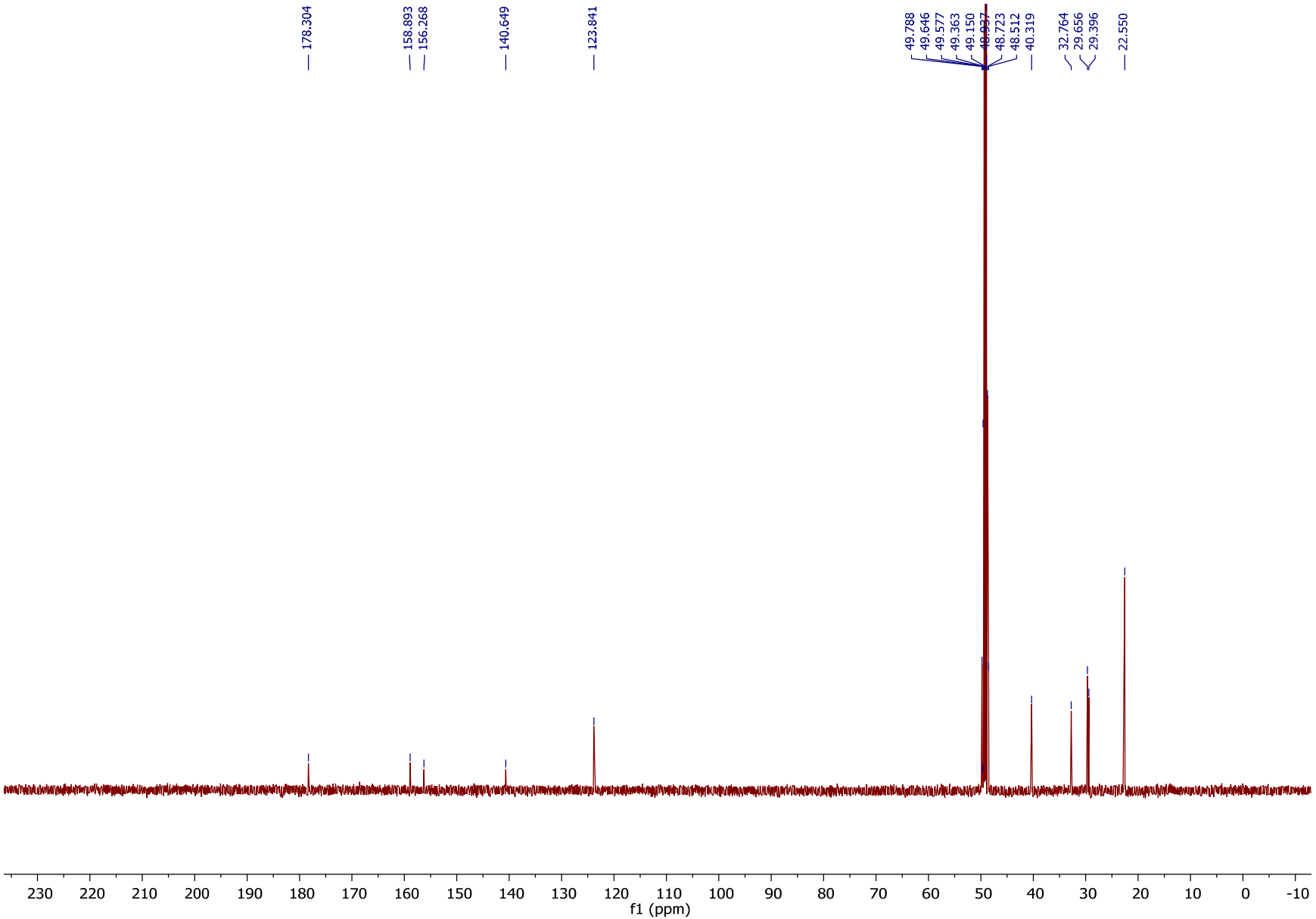


Figure S10. ^{13}C NMR of 2-hydroxy-5-isobutyl-3-propanamidylpyrazin (**1**) in CD_3OD (100 MHz).

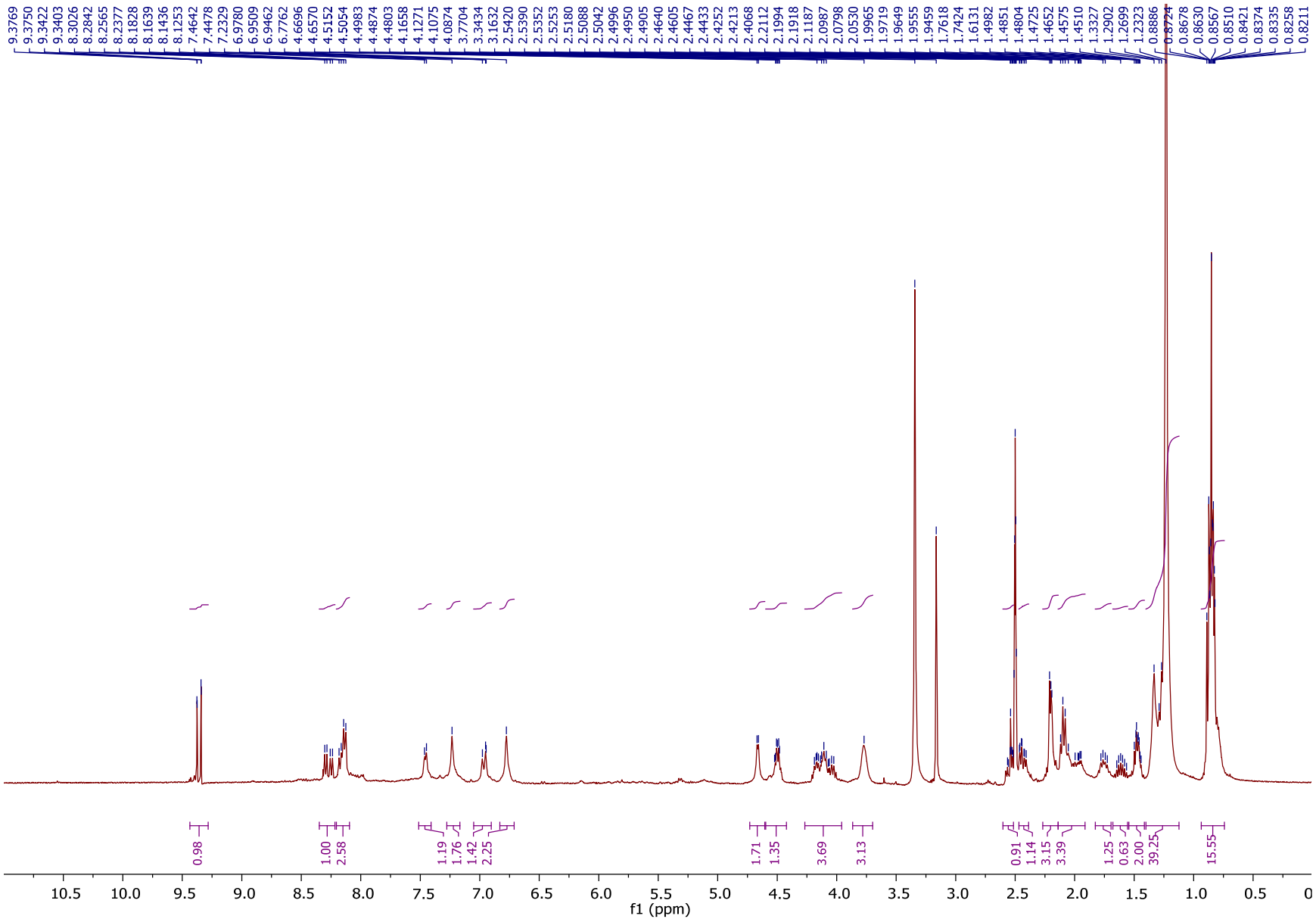


Figure S11. ^{1}H NMR of (2*RS*)-fellutamide B (**2**) in DMSO-*d*6 (400 MHz).

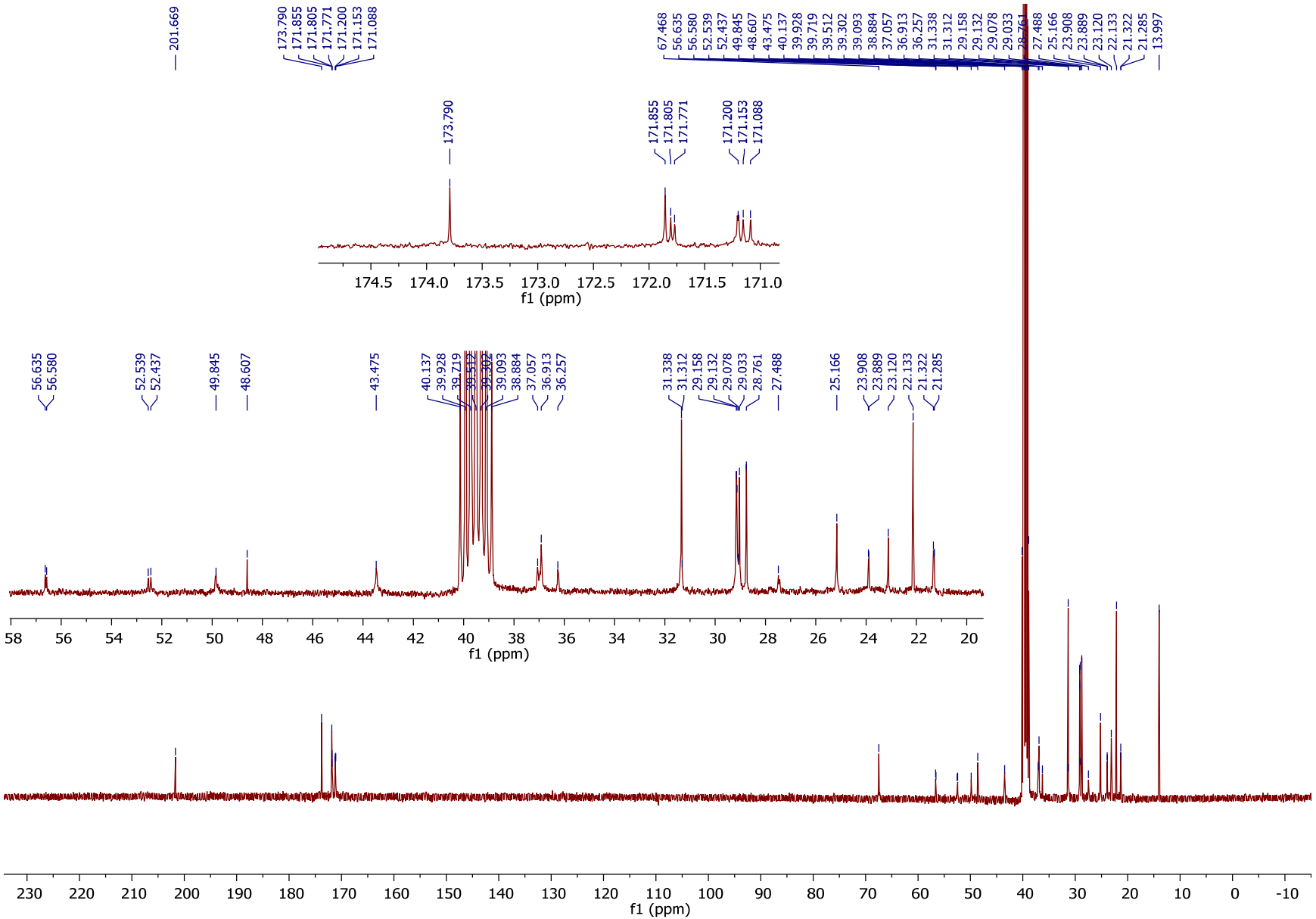


Figure S12. ^{13}C NMR of (2*RS*)-fellutamide B (**2**) in DMSO-*d*6 (100 MHz).