SUPPORTING INFORMATION

Genome Mining for Unknown-Unknown Natural Products

Danielle A. Yee,[1] Kanji Niwa,[1] Bruno Perlatti,[2] Mengbin Chen,[1] Yuqing Li,[1] Yi Tang[1,2]

[1]Department of Chemical and Biomolecular Engineering, [2]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA

Table of Contents

SUPPLEMENTARY TABLES

Table S1. Bioinformatics analysis of the *ank* gene cluster from *A. thermomutatus* using NCBI BlastP search

Protein	Size (aa)	Proposed function	Homologs (ident/coverage)	Strains
AnkA (XP_026617165.1)	518	hypothetical protein	no hits	no hits
AnkB (XP_026617163.1)	524	cytochrome P450	KAG9240506.1 (60%/ 90%)	*Calycina marina*
AnkC (XP_026617167.1)	587	FAD-dependent monooxygenase	P0DUL5.1 (41%/89%)	*Aspergillus hancockii*
AnkD (XP_026617166.1)	563	pyridoxal-phosphate dependent enzyme	KAF7618911.1 (50%/93%)	*Aspergillus flavus* NRRL3357
AnkE (XP_026617162.1)	604	NRPS independent siderophore synthetase (NIS)	CDM31866.1 (48%/93%)	*Penicillium roqueforti* FM164
AnkF (XP_026617168.1)	239	*O*-methyltransferase	XP_015700694.1 (41%/100%)	*Paracoccidioides lutzii* Pb01
AnkG (XP_026617169.1)	480	ATP-grasp protein	KAF7349305.1 (35%/99%)	*Mycena sanguinolenta*

Table S2. Bioinformatics analysis of the *ava* gene cluster from *A. versicolor* dI-29 using NCBI BlastP search

Protein	Size (aa)	Proposed function	Homologs (ident/coverage)	Strains
AvaA	541	hypothetical protein	no hits	no hits
AvaB	709	FAD-dependent monooxygenase	TVY66699.1 (28%/ 98%)	*Fusarium oxysporum* f. sp. cubense
AvaC	299	kynurenine formamidase	A5DNX8.2 (27%/90%)	*Meyerozyma guilliermondii* ATCC 6260
AvaD	290	GNAT family N-acetyltransferase	MBE0409255.1 (38%/85%)	*Anaerolineales bacterium*
AvaE	336	beta-lactamase	KAH6869491.1 (33%/83%)	*Thelonectria olida*
AvaF	537	cytochrome P450	KAH6680460.1 (34%/87%)	*Halenospora varia*
AvaG	319	lysophospholipase catalytic domain family protein	TPR05408.1 (26%/97%)	*Aspergillus niger*
AvaH	492	cytochrome P450	KAF2091838.1 (31%/70%)	*Saccharata proteae* CBS 121410
AvaI	437	cytochrome P450	KAH8801390.1 (28%/84%)	*Xylogone* sp. PMI_703
AvaJ	507	Proton-coupled folate transporter	Q6DCX5.1 (21%/71%)	*Xenopus laevis*
AvaK	450	MFS efflux pump	C8V3Y8.1 (29%/84%)	*Aspergillus nidulans*
AvaL	561	cytochrome P450	KAF8155335.1 (32%/83%)	*Mycena galopus* ATCC 62051
AvaM	255	NAD-dependent epimerase/dehydratase	TQW01081.1 (45%/92%)	*Cordyceps javanica*
AvaN	362	cytochrome P450	XP_033685598.1 (32%/90%)	*Trematosphaeria pertusa*
AvaO	338	dioxygenase	XP_025392866.1 (64%/100%)	*Aspergillus eucalypticola* CBS 122712
AvaP	453	glycosyltransferase	KAE9372740.1 (39%/97%)	*Chalara longipes* BDJ
AvaQ	1542	ABC efflux transporter	P38735.1 (34%/73%)	*Saccharomyces cerevisiae* S288C

Table S3. Plasmids used in this study

Plasmid	Vector	Genes
p3001	pYTU	n/a
p3002	pYTR	n/a
p3003	pYTP	n/a
p3004	pYTR	gpdAp-AnkD; POgpdAp-AnkA
p3005	pYTU	gpdAp-AnkE; POgpdAp-AnkB
p3006	pYTP	gpdAp-AnkC; POgpdAp-AnkF
p3007	pYTR	POgpdAp-AnkA
p3008	pYTU	POgpdAp-AnkB
p3009	pYTP	gpdAp-AnkC
p3010	pYTU	gpdAp-AnkE; POgpdAp-AnkB; PEgpdAp-AnkG
p3011	pYTP	gpdAp-AnkE; POgpdAp-AnkB; PEgpdAp-AnkG(spliced)
p3012	pYTU	gpdAp-PthA
p3013	pYTU	gpdAp-AteA
p3014	pYTU	gpdAp-AmaA
p3015	pYTR	gpdAp-AvaA
p3016	pYTU	gpdAp-AnoA
p3017	pYTU	gpdAp-EshA
p3018	pXW55	n/a
p3019	pXW55	ADH2p-AmaA-ADH2t
p3020	pXW55	ADH2p-*N*-his-AvaA-ADH2t
p3021	pXW55	ADH2p-*N*-his-AvaA(E432A)-ADH2t
p3022	pXW55	ADH2p-*N*-his-AvaA(Y515A)-ADH2t
p3023	pXW55	ADH2p-*N*-his-AvaA(Y515F)-ADH2t
p3024	pXW55	ADH2p-*N*-his-AvaA(C193A)-ADH2t
p3025	pXW55	ADH2p-*N*-his-AvaA(C193S)-ADH2t
p3026	pXW55	ADH2p-*N*-his-AvaA(C193T)-ADH2t
p3027	pXW55	ADH2p-*N*-his-AvaA(D428A)-ADH2t
p3028	pXW55	ADH2p-*N*-his-AvaA(D429A)-ADH2t
p3029	pXW55	ADH2p-*N*-his-AvaA(Y392A)-ADH2t
p3030	pXW55	ADH2p-*N*-his-AvaA(Y392F)-ADH2t
p3031	pET28a	n/a
p3032	pET28a	T7p-*N*-his-AnkD-T7t
p3033	pET28a	T7p-*N*-his-AvaA-T7t
p3034	pET28a	T7p-TrpRS-*C*-his-T7t
p3035	pET28a	T7p-ArgRS-*C*-his-T7t
p3036	pYTU	coxAp-AvaB
p3037	pYTU	coxAp-AvaB; PEgpdAp-AvaC
p3038	pYTP	POgpdA-AvaD

Table S4. cDNA sequences of *ankD* and *ankG* and gDNA sequences of *pthA*, *ateA*, *amaA*, *anoA*, *eshA*, and *avaA-D*.

ankD	ATGGGTGAACTGGGCCCTGCATCAGCTCAACATGGAAGTGACAGTATTTCCTTCTCTGGCTCTTACACGCAGCCGCTCGGAGCTCCTCAACCTCCAAACGAACCTCATGCTATAAGCGTCTCCTTGCCCACCTGGGAGGCTGTGACTGCCGTAATGGCGGGCGCAGACTGGGCTATATGCCAGCTACAAACCAGCTATCCTAGATTTGATATTCATAAATGTGTAAGAGAATTGCACGATGCCGTTCTGGCGAGATTCAAGCACCCGGCGCACACGGTGTGCCGAGCGTTTCCATCTCCTGAAGCTGCTGAGCGTTTTGTCTCAAGACTACAGAGAGAAGATCCCATATTATCTGTGCACACGGCGCGGTTCCATCTGCCTCATGATGCGGTTCCAGAGCTGGCAAAATGGGCAGCGTTCTCAGTCGTATTGTTTAATGAGAGTCTCGAAGAGGTGGCGTTCGAGTTCTGGGAATGGTTTGGGGATGGAATCTCCAGTAGACATGCGGAGTTTTGTCTAACTCAATTCTCCTTCCTGAATACTGGTTCGGATCAGCCTGAATATCAGACTGTGGGCCAAGATTCCCATGATCTGAGTGCGATGAACCTTCCAGAGTGGATTGACTCCTCTACCAAAGACAAAATGGATATCAAGTCTCGGCTTGCCAGCTCAGCAACATCAGCAGATCCAGCCTTGAAGCCCATGGGGAATGATGATGTACTCCTCTATGCCACTGGAATGGCTGCAATCAGCGCCATTGCCCGCGCGCTCGCTCGGACATCAGACGACTCTGGCGCTGTGGTGTACGGATGGCCGTACTCCGGCACCCCTCATTGCGTGCAAGGCTGTGGATTCAAGCGCTACACCATGTACGGTCACGGCTCGAAGGCGGATCTCGATTCGCTCGAGACCTTGCTGGTTTCAGAAACTCGATTCACGGTGCTGTTCTGCGAGATCACGTCCAATCCTCAGCTGTCCACGCCGGATCTGCATCGCATCCGAGATCTCGCCGACCGGTTTGGCTTCATCGTCGTGTGTGACGATACACTCGGCACATCCGTCAACGTGGATATCCTGCCATATGTCGATGTCATCATCACCAGCTTGACCAAGATCTTCAGCGGTGCAGGCAATGTCATGGGAGGAAGCCTCATGATTAACCCCAACTCTGGCCACTACAGTACTCTCCGCGCCCTTCTGACAACAACCTACGAGGACCTTTACTTCCCCCTCGACGCAAAGACAATTGCGCGCAACTCCTCCGACTTCGCAGCCCGCGTCCACAAATGCAACAAGAGCGCTCTGCAGATCGCAAATTTGCTCAACTCGCATGCATCCGTCGAATCCGTCAACTATCCCACCATGGTCCCGACTGCACCTCTATACGAGCGCTACCGCCGCCCCGACGGCGGATATGGATTCCTCCTCAGCGTCATCTTCCGCGAGCCCGAGAGCGCCGTGCTCTTCTACGACAAGCTGGATGTGTGGAAGGGGCCCACTGTCGGCACCAATTTCTCGATTTCCATCCCGTATTCGGCGCTGGCGCATGCGAAGGAGCAGGACTGGGCGGCGTCGCATGGGGTCCCGAAGCATATTGTGCGTCTGAGCGTTGGGTTGGAGGATTATGGCGATTTGTCTGAGCGGGTGAACCGGGCATTGCGGGAGGTGGAACTGAGAGAGAAGATGGGATAG
ankG	ATGTATCAAATTTCTCTAAAGGCGACGAAGAGCGCTGCAGAACCCACCTCTTCAACCGACGCATCTCACGACGACCGTCAAGTAGAGCGCGATTCATACGACACCGCGATGCTGCGCTTCATCAGCGACATGGAACTGTCCTCAATGTACGCGGAGAAGACGTCACCATACCCCATTCTCATGCCGAGAAGCTTCCTTGAGGACCTGAAGAATTTCCAAGATCTGCTATTTGTTGCTGTATCAAACATCCTGGATCGATGGTGGGAGGACAGGGAAGCCGACTTTCCCCGTCGCATGCCATTGGAACCTCACGAGGAGAGCGTATTGAAGCATTACAATGAGCGCCCGCTACACTGGAGGCCGGACATGCTCCTCCCAGCAGCCGGCGATCCAAACACCAACCTACCTTTCAAAATTTGCGAGATCAACGCTCGTTCCCCCTTCAACTCAATGATTAAGAGCATCTGTATGTTTCAGGCAGCGGCCGCAAGCAAAACAGCCCTCCCCGACGGGTTAGAACTCGCTTCGACTGCGGATAGCCTGGTGGATAGTCTGGTCTCGCTGTTCAATCCGGACCTCCCTCTTCATGTTATTTGGCATGAGGGTATCACTGACCCCATCGATGCGTTCAGTTCCTTCTACAAAAAGCGAACAGGCAAGATACCGCGAGTAATTCGCGCGACAGACCTGCGTCTCGCGCCGGATTCCTCATCCCCGACGGGTCGGATTCTGTGCTGCGTTGCCTCCGCCGCGTCCGCAGACCATTCTAACGGCCAGGCAATTGTATCCGAGACAGGCGAACCCCTGGAACGTATCTACCAGGTCGGACTCCAGATGTCCCATCGGGATTACAGTGAGCTCTCTTCGGAAGTGCTGCAGCAGCTCGCCGTGGACGGAATCTGCGACCTACGCAACATCTTCCTGGTCAGCGATAAACGCATGCTCGGCGTAATCTGGCAAGAGCTGGACTCTCTCGTGCACAAGCACCACGTCCTCACGGCTGAGCAGGCCGAGATCCTCCGACAGGGCATCGTACATACCATTTTGCCCGGCTCCGAGGATATGGAACGCCTCTTGCGGCAGACTCGCGAGGGCTCCGTATCCAAGGATAGCTATCTTCTCAAACCTGCCCGTGGACACCGGGGAATGGGAATATTGCTTGGGAAAGACTTGGGGCAGGAGGAGTTTGAGGGTCTCCTCGAGGAGCTGGCGGACCCGTTACTGCCTGCTGACAGGAGATATGTTGTGCAGCCTTTTATTGAGCAGGCACTATTCGGTTTGAGGCTGTATGATGATAGCGAGCCGCAACAATGCCAGATGACGGGGACATA

	TCATGCCATTGGTGGGTGCTTTGCTGGGTTGGGCGTCTGGAGAGCTGATAGCGAGAGAATTTGCTCTCGATTCCACGGTGCCTTTAGTATTCCAGCGATTGTTCCCCGTTAG
pthA	ATGAGTACACTGTATCTGCTTGCGAATCTCCGGAATCGACCCGACGGTCGACACAGCATTGATTTTGGCAGTCTGTCCGCACGCGAGTCGGACCCGGTGAAGATGCCTCATGAAATCACAGGGTTATTCAGTTCAGATTACTGTTCATCAACGTCAACAACCCTTGAAGCAGAAGAGAGTGAAGCGTATTCGAAAGCGCTATCCAGGAATAAGAGAGCGGCTGGGATTCTGGAACATTTCCTGAATGGGGAACAAGCGGCGGCGCATCCGTCGGTCTGTCTGCAAGACAGAGCGCGCTTGCGCCTATGCGAGGAGGCGATCCTAGTTGGGGCACCCCGTGGCACGGCGACAAGGGCCCGAGAGGGCATTGCCTCCTATGACAAAATGGAATTTCTTTTTCCGGGGAGCCGGCAACGGAAGTCAATCCAGATGCTGCATGTGTCCTGTCCAAGCCTGCAGCCGGGCGCAACTCTCCTCGACAGTCACAGCGGCTCCGTTCATGATAGCATCACGATCCTGTCAGGCATGTGTCTGTTTACGCGAAAAATCTTAAATAGTACGACGACCAACGGACGTTTGGATGGCCAGAAGCTACGGCATATGCCACCCACTCTGTATGAAATGGAATGTATTGCTCGACTGTCAGCCATCATTGCTGATGTCACAACTTTAACGCAGTCCAACTTTGGCGAGGATAACCATCCTGCAGTCAACATTGTATTGGATCTCCCAAGTTGGCACTACTATCAATTCATTGAGGACTGTATCAAGAGTGATACCTGCAGTCTAGAGGAGGCTATTGATTGGACGGAGGCTATCAAACTGCGCCGCCAGCAGCTGGCTAGCGTGCTGAAGAAGGCAGTCTGGCACGAGCTTGGACAACGTCAGGTTGCCCACAAATCGACCCTCAAAGCCATCCAGATCTCACCCGAGTCTACGGTGGTGGACGAGCTCATTAAGGAGACCCTGCAACGTGGGCATCAGCCGCGGCTGGATGATATACTGCACGCCCTGAGCGGCACGCCGCATAGTCTCTGGCCGCAATTCTATTCGCTGCTCCGGGACCGCGACAAACCACACTCTATCAGAGATCTCGGTTTTCTCTTCTATGTATTTCAAGTGGTCCGCCCGGCGTTGCTCAAGGCCATCCCATGCCAGGAAGCGCCGAGCTCTAAAGCGCCCGCTTGCGAGCAGCCCAACCTGCATCGCGGGCAGAGGCAATGCCTTCCGCAACGCCCACTCATCATCAGCCTTGATGACCGGGCGGAACGCAAGATTTATTCGCAGGCTCACAGCCTCCTTCTCAGGCTGTCCAGGAGTTCGAACCAGCTTGTCAATCCCACCCTCGTGCAACTGTACATGCCGCGCCGTGTCTACATCGATGGCAACAAGGATGGTCAGCGCCTGTACTGGCATGATCCCAGTCCTGTTCTGCCATTGCTCGAAGGCACGAAAGAGCAGCAGAGGAGGGATGCGCGCAATCACGATCACCATCGGCGAGAACTCCAGCAAACCGACATTCTTACCGAGCTCTACGGTCGGGACTGTTCTGCCAACATCCAGGGTTGGTTCAAAGAAGCTGGCCTTTGTTGA
ateA	ATGGATGTGCAATATGTACTCGCCGATCTCCAAACGGGACCGGAATCTCAATCATCACAGTCAAGTATTCATTTTCATGATATGGCCCCGTTCTATGTGGTGTCCATTGAGCGGCAGCATGGATTGTCTGGACTTGTCAGCATAGTTCCCTGTTCCACGACGCCAACAATTCACTATGATGAAAAGGATAAAAAAGCTTACCAGAAGGCCTGGTCAAGGGGCAGAAGATTTGCTGCATTGTTCGAAAACTTGCTTAATGGGGGGAAAATGGCTACGAAAACATCATGTTCATTGCAGCAGACACCATACTTGCAACTACGCGATGAAGCACTTATTGTTGGGTCACCCCTTGGCACTCCTACACGGACAAGGGAGGGTATTTCTTCCTTCCACAACATGGAATTCAATTTTCCTGGAGCGAAACCAGTTGTTTCATCACAGATGCTGCATATAACATGCCCAGGTCTTCGGTCAGAGGAAAGTCTCTTTGACAAAAATGGCGGAGCACTGTATGATAGCGTCACAGTATTGACCGGCATGTGCTTATTTACTCGCAAAATACTGGACGAAAGTGTTGCTGATGGTTTCCTGGACACCCGGAAGCTGTTGGACATGCAGCTGACTCTATATGAAGTGGAGCGCATCGCCGGATTGTCGTCCGCTATCGCAGATATAGCAGCTCTCGTCTTGTCCAGATACGATGGGAAGTGCCACAAGCCGGTTAATATCATACTGGATCTTCCAACCTGGCATTACTACCAGAGCGTTGAGGATCAACTCCAACATAGAGGCTATAAGCCATGCGAGGTGATGGATTGGATTGAAGCAATTACTCTCCGCAGTCAGCAGTTAGCAGGTTTGCTTAAGAGCGCAGTCCTGCATGAGCTTGATAGGCGTGGTGTCTCAAGCAGTCAACGTCTCTATGATATCCAAATTTCTCCGGGTTCAGCCTTGGTCGATGACGCCTTCCGAGAGACTCTCAAATACGAGAACCTACCATGCCTTGATAATATCCTTGAGGCGCTTAGCGGGAGCGAAGATGGCTCGTGGCAGCAATTCTACTCGCTTCTTCCCGAACGCGATAGGCCTCAGTCTATTAAGGATCTGAGCTATCTCTTCTACATTTTTGAGGTTGTTCGTCCTGCTGTGTTGGCTGGGATTCAATCCGGGAACCAGCAAACACCGACACAATCAGCCGATATGGATTCCACAGTCAGTCACAGGCAACAACAACCTGCTTCCTCCCGCAGCTATACGAGCAAGCAGAATCAAATGCCCCGTCCATTGGTCATCAGCGTTGATGATCCTTCGGAGAGAAAACTATATTCAAAAGCCCATAGCTTCTTCCTTAGGCTTCCTAAGAACCCAATCTACCCAGCTGACCCTACGTTGGTCCAAGTGTACACGACTCGTCGGGTTTTTGTTGACGGGAACAGAAAAGGGGAGCGCCTGTACCGGAATGACCCCAGCCCAGTGATGCCAGAGCTAAGTAATGGTAAGCTCTATGGTGATGGGGGTGATGCTGGGTACAGCAAGTATGCCAGAGAACTCCAGCAAATGGACTTCATAGCAAAGCTGTATGGTGCCGAATGTGCAGCAAATATCCAGCGTTGGTCGAAGGAGGTTGGCTTATGCTGA

amaA	ATGGTTTCAATCTCCAAAAGCCTCGTGGCTGATATATGTGTGGGCTCATCACCGCCAACAGTCCGGTTCGTCACTCTGCTTGAAAATCATGATATCTCCAACACTACCTTCCCATCCTTATCAGGGGTCTTGGCTCAATATGGATCTTCGCTTCAGGGACCGGCTGGCTTGCAACATGTTGGTCTCGCCCAAGACCAAAAGTACCAGCACGCCTTGACCAAAGGGTTGAAAGCTGTAGAAAAACTACAACATTTCCTAGACGGCCCAGCACCAATGTCAGGAGACCCTGTCGTCCTTCGTTGTTCTTCGGCCAGATCCTCAACTCTTTTGGAAATAACCGATGAGACCATTACCAGCGGTCCCCCATGCGCTGGGGCTGCCCGAGAACGAGAGGGGATCACCTCCTTCCATCGAATGAAATTCCACCTCCCCAAACACGTGGGGTGTGTACAAGATGTCCATGAGACCTCCTCGCAGATGCTTCATCTTTCTTGTGACGGATTGAGGTCTGAGGATACCCTCTTTGACCGCTTTGGAAGATGGAGCCGGGATCATGTCACAATCGTTTCCGGGATGTGTCTTTTTTCGGACAAGATTATGAGAGACGCAGTCCAAGGGGATGGAAAACTGGATGTATCAAAGCTGTCTTCTCAGGAACCGACGCTGTACGAGTTGGAAATCATTGCGCGGTCGTCTTCGGCCATCGCAGACGTGGCCAGCCATATCTTGGGGAAGAACGGCGCTGCAAAGCCTATCCCGCTGGAAATTGTTTTGGATGCTCCTAGCTGGCACTACTTCCAGGTGGTTCACGGAAACCTGGCTTCCGGACATTGCACGCCGGCAGAGGCTCTGGACTGGCTGCAGGCCGTGGAGCTTCGCTGCGAACAAGTAACTACGGTTTTTGAGAACTCCGTACGCCATGAGATGGGCCTGAGAGGAGTCCCTGCGGGGTTCTACCACATTCTGGCAGCCCCGGGGACGGCTGGAGTTGGCACCAGCATCAGACAGGCGCTGACCTCTGGCATGGTACCCGACATTGCTGATTGCATGGATGCTATCTGCGAGGTCGAGGGTGAGCGCTGGGCCATGTTCTACTCTCTCATCCCCGAGAAAGATAGACCCTGTGATTTTAGGTCTTTGGGCAATCTGTTTTACATATACGAAGTTGTTCGTCCCGCCTTGGCCGCGAAGAACGGCTCTACAGCTGCAGAGCTGGAAGTCTCTGGTCCGAATTCGGCTCTCGATGACAGCTCATCATCTGGACATTCCGACATCCTCGACAGTGCGTTGCATCTCAAGGCCAACGTGGAACGGCCTCTGATCATGAGCATCGATGACAGGGCCGAGCGCAAAATCTACTCCAAAGCCCAGGCTTTCATCAAGAAGATTCGCAGAGCCCCACAGATCCCTTCACATCCCCTCTTGTTGGAGTTGTACAGCGCACGACGTGTGTTTATCAACGGCAACACGGGTGGCGACAGTCTGTACTGGAATGATTCGAGCCCACACCCACTCCGCATGAACACGACAGCGGGGAGCCCCGAGCTTGAAGCCTTCGGTGTCATCAGGAACCTTTATGGAAATGACTGTGCCCAAAACCTCCAGAGGTGGTTCTCGGAGGCCGGCGTCAATGTTTCAGGACATACGTAG
anoA	ATGATGGGTGCCGTACAACTTCCCGTTGAGACCGAGGCGTGTCTTCACGATGACGCCGCAAGCCGACCGATCTCTGATTCAAACACCGTACGGGTAGCCGTCCTTAGACAACACCAGTATCACTCTAGTGTGACATACCACAGGCTGGATGATAACAGATTAATCCCGGGCGTCGTGGGATATTCCTATCTTCGCGATATTGCAGATGAGGGAAAGAAACAATCCCCTAGCAAACAGCACGCCCGAGTAATACAAGCCTATTCGAAGATTCACTCTCTCCTATCCCCACCAAGAGAACCAGGCCCTATAGACTCAGAAACAAAGACTCGAGAGAAGAAAAGCTCGCCTGTCATCGTTCACGATAAAAGGTGTGCTTTCATAGAGCGACTTGAGCCACCTCCAAACTCCAAGGCGGATATAGTGAACACTGTATTCGCCCAAGTCAATCTGCGGACACCAGTCGGTCCTCCTCTGGATCAATTCGTCAACTGCCGCTCATCGAATATCAAAATAAATGACTTGAAGGACTCCGCGAATGGCCTTTGTCGCCCTATCACAATCAGTACAGGTATATGCCTGTTCTCATCTCGGCTCCTAGGATTCATCCCTACAAATGGACTATCCACTACCGATGGCGCTACAGAGACAGTCATCCCCCCATTGCCTTACAGTGCGGATGCAACTTTCTACGAACTTGAAACATGTGCACGGATGGCAATGACTATTGCAGGCCTTGCCGTTACCGCTGGCACAGGTGGCACAGCTGGTTCTCGGCCAATTGTTGTGAGGTTAGATGTGCCTTGTCTCCAGTATTACTGCTACCCATTGGAGCTTCTTCAGGCAGGACTGGTGAGCTGGAAGTACGTTCAGGAGTGGTTCAGACTGGTCGATCGCCGGCATCGGCAAGTCGCAGGCTTGCTGAAGGATACCATTACTCATGAAGTACTTCGACGGGGCGGTGATATTCAGGTGGAAGTCACCGCTGGCACTATAGCCGCAACTCAGCTTTTGCGTCTCAGTGTTTTCGACCGAATGGAATTACCGAGCGTGGATGATATTCTATTTGTGCTGAAGTGGGTCGGTCCCTATCAGGCGGCATGGCGAGAATTTCTGGATATTCTTGATGATTGTCAAAGGCCCAAAGATCTTCGCAGTGTTGCACTAATGGCTTATGTCTTTGAGGTCATGTATCCCGCATTGCATCAGATAGCTACCAAATCTCTACATGGGAATGGCGAGAAGAGTGGTCGCCCTCTACTGATTCAGGTAGACGATATAGCCGAGTGGAGAATCTTTGACCACGCAGAAAAGCTCCTCAAGCGCTTCAAAGAGCGACAGCATGGTTTTTATCCCCTCCTGGTTGGCGTGTTTCCCTCACCACGAATCTTTACGAGTGAGGATCAAGGGCGCTCTACCCTGTTCCTTCATGACCCTGGTTTAAAGATTTTGCAAACTCGCTCTCCCTCTAGCAATAGTGAAGAAGGCTCATGCGTTGTGCGCCCGTTAGATATCATTGGTCAGATCTATGGAAGAGAGGTTCAGGATACACTTGGGCAACTAACAAGAAAGCACGGTTTATCTGCTGCTGATGGACCCGAATCAGATTAG

eshA	ATGTCTATACCAACCGATATCGCCAACCTACTGATTGGAGAGGATGCAAAACTTTCAATTCACTTTGATAAAATTCCTCTCCACAACGATGCCCATCACAGTGGCTCCAAATTTTATATGCTTTCCACAAATCGTGCGCTGAGAGACCACGAAGAAAAGGAAGACGACCTCAAATACCTGTCTAAGGGCAGACAAGGTGCCATTAGGCTGAACCAGCTACTTAATGCTCCCCATCTCATAAAAGAAAGAGAAAGCACAGCCCAAAATGAGGGAAAAGCTGTGATGCGGCTTGGGGAAACCGCTATGGTTATCGACGAAACATTCGATGCGCAGCCTATCCACAGCCACATCTGGAACTCCTTCATGTCAATTGAATTTCGCTTCCCGCAAGCAGATTTCACATCTCTGCAATATCTACAAACTTCTTGGAAGGGTCTCGAATCCGGCGATACCCTTTACACCAAGGAGGGAACACCAACCTTCGAGGAAGTATCCATCGTTTCTGGAGTTTGCCTCTTCAATGCAAGGCTTTTGGAAATGAGTACAACGTGCCCAAAATCTGGTGGTGTTGATGTTGATCTACTCAAGAGCTCAACCCCTACATACAATGAACTGGACTATATTGCTCGCGCGTCAAGCGCCATAGCTGACGTCGCTGCTATGAACATCTTAAGAGCTTGCGAGCAGGACTCTGGACAGCGACTTACAATAAAGCTTGATATTCCGAGCTGGCACTACTTCCATACAGTAGCCACAAAATTTGTGTCAAAGCAATGTTCGAACACTGAAGTATTGCAATGGATGGACGCGGTTGACCAACGACATGATCAGATCGGCCAAACTTTCGTGGAAGCCATCAGATATGGGCTTGAGCAACGGGGGATTCATGACAGCACCAGCTATGGTATTGGCATGACATCAAGGACAAATACAGCCGCTATTTTAATTCGAACGGCAATTGAACATGAAGAAGTTCCTTCATTGGATGCCATCCTCGCTGCTTTGGACTCTGAAGAAGACGGGTGCTGGAAACGATTCTACGAGATGATCCCGGTCAAGGAACGACCCTCGAATCTGGACCAGCTCGGTTATCTGTACTATGTATATGAGGCAATCAGGCCATCTCTCGCCGAGCGTCCAGCACCAGTACCTCTGGAATCTGACCAAACCAAGAAAGCCAACTTGTCCAAGAAGGCATTGAAGAAGCGAAAACCTCGCCGCTTAATCATCAGTGTTGATGACTCTGCGGAACGACGCATTTACTCTCGAGCTCAAAGGGTACTCAGCAAAATTCGACAGTGTAGCGAAAAGCCCGAGACTTACTTGGTAGAGTCATACGTTTGTAGGAGGTTCTTGGTCAATGGAAACGAGGACAGAGCACGACTAGGTCGGTTAGATCCTCTACCCGATATTCCAGTCCGTACTGGTTTCCACCATGAGCCTATGCTACCTCTCGATGTTGTGCGGCAGCTCTATGGTAATAAATCTGCTCTCAACTTGCAACGATGGCTTCAAGAAGCGGGACTCTCAGTTTAG
avaA	ATGACGGCTACCGTGCAACATACCGTCAAGACCGAGGTGTGTCTTCGCGGCGATGTCGCCAGCCAACCGATCTCTGATTCAAAGCCCCCAACATGGGTAGCGGTCCTTCAGCAGCACGAATATCAGTCCAGTGTTACATACCACAAGCTAGATGACAACAGATTGATCCCAGATGTCTTCGGATATTCCTACCTCCGAGATATCGCAGACCAGGGGAAGAAGCAATCCACCGGCAAACAGCACGCCCGGGTGATACAGGCCTACTCGAAGATTCACTCCCTCCTATCCCCGCCATGGAAATCAAGTTCGATAGACTCCAACACGAAGACCCGAGAGAACAAAGGCTCGCCCGTCATCCTTGAAGATAAACGCTGTGCCTTCATAGAGCGGCTGGAACCACCTCCAAACTCAAAGGCGGATATCGTGAGCACTGTGTTCGCCCAGGTCAATTTGCAGACACCAACCGGTCCTCCGTTAGAGCAATTCCTCAGCTGCCGCGCATCGAATATCAAAGCAAATGACTTAAAGAACGCTGCAGAGGGCGTTCGCCGCCCCATCACAATCAGTACAGGCATCTGCCTGTTCTCATCTCGGCTGCTAGGATTCATCCCTACAAATGGACTATCCGTTTCCCATGATGCTACAGAGACCGTGGTCCCGCCATTGCCTTACACTACGGTTGCGACTTTCTACGAACTCGAAACATGCACGCGGATGGCAATGACTATTGCAGGCCTGGCCGTTACGGCCAGCACAGGGGGTACAGCTGGTTCGCGGCCCATTGTCGTGAGACTAGATGTGCCCAACCTCCAGTACTACTGTTACCCGTTGGAGCTTCTCGAAGCAGGACTTGTGAGCTGGGAGTATGTAGAAGAGTGGTTCCGGCTAGTCGATCGTCGGCATCGACAGGTCGCAACGTTGTTAAAAGACACCATTATCCACGAAGTACGTCGACGCAACTGTGATGTCCAGGTAGATGTTACCTCGGGCACTATAGCCGCAACCCAGCTTCTACGTCTCTGTGTCCTCGGCCGAAGGAAGATACCGAGCGTGAACGATATGCTGTTTGTGCTATCGTGGATCGGTCCGTACCAAGCAGCGTGGCGAGAATTTCTCGCTATTGTCGACGACTGCCAAAGGCCCAAAGATCTTCGCAGCCTTGCACTAATGGCATATGTCTTTGAAGTGATGTATCCGGCATTACAGCAGGCAACTACCAAAACCAAAACTCCACATGGGAAGGGCGAGGAGAGTAGTGGCCGCCCTCTACTGATTCAGGTGGACGATATCGCCGAGTGGAGAATCTTTGACCGTGCAGAAATGCTGCTTAAGCGATTCAAACACCGACAGCATGGTCTTGATCCCCTCCTCGTTGGCGTGTTCCCTTCACCACGGATCTTTACGAGTGAGGATCAAGGGCGCTCGACACTGTTCCTTCATGACCCTGGTCTGAAAATTTCGCAAACTCGCCCTCCCTCTAGTGGTGGTGATAGCGAAGACCACTCATGTGTTGTGGGTCCATTGGATATCATTGGTCAGATCTATGGAGCGGAGGTTCAGGACACACTTATGCGATTGATAATCGAGCGCGGTTTATCTCCTGAGGATGAATTCGAATTAGACTAG
avaB	ATGCCGCGAGTGGTCATTGTTGGTTCTGGTATGCAATATTCATATCGGGTTATCGATCTATTATTGACGTTTTTTTTTTATTCAGGGCTATACGGGCTCATCGCAGCCAAAACCTATCTCCAAGTAACGGACCCTCGTTATCCCAGCAACCAGCAATGCGAAAATACCGACGACGTCCCCCAAT

	CCTTCCAGAGACCTAGAGATACCCGGTCAACACACGATGACACTGACCTTCTTTTACTGGAATCCGGATCAAGCCTTGGGGGAACATGGGCGGAGGAGCGGCTTTACCCAAATCTACTTAGCCAAAACTCCGAGGGTCTGTACGAGTTCAGTGACATGTCTCTGTCGGAGGCCATGGATGGGGTCGATGATGAAGGACAGGCAGATCGAGACACCCCGGTGGAGGATCGCTTCATCCCAGGCTGGAAGCTGAGTCGTTACCTCACAGCGTGGAGTCGCAAATGGAACCTGCCACAGTACATGCGGTTCAATTGGCAGGTTAGTGACGATGGAGAATGCAAGACATATGAATCACGATCTAACCGAAGAATGACAATAGGTGAAGAAGATCAGCCGCCTGCCTACGAAGGAATGGGAGTTGGGAATAGTCATTCACCCGGATCAACCATCGCAGACACGACATACAACTATTCTCTGTGACAAACTTATCCTGTCCCCCGGACTGACATCGGTGCCGAATATATTGAACTTCTCGCCTGCAAGTGGCCGGCCATCTGCCCCGGCCGAGAATGTCATCCACGCCAAAGAAGTCGGTCCGTGGTGCCGTGACAATCTGGGTTACCAGCCGATCCCAGGACAAGATGCCCATCCGAGTCCCACGAAGCCTAGTCGTGTGCCCCGGCGGGTTGCAATCTATGGTGGTGCAAAATCTTCGTTTGATCTTGTCCACATGTTTGCAACTCTGCACTTAAAGGACCCCAGTTTCCACCTAAAAGGATTGGACGCGACTGATTCTATTGAGCCGGTACAGGTGCACTGGATTATACGCGATGGAGGATCTGGCCCGGCCTGGATGTCCCCGCCTCGCTCCAATATGCCCAATGGCCAGAGTATCGCCAGTGACAAAGTGGCCAGTACGCGATTTGTAGGCGTTCTAAGCCCGTGTACGCCATTGGTGCCGAAACGTCTGACACTCCGTCGCTCATCCAATTTCTTGGGCTGGAAACTCGCAGTGGAGGGCTCGTGGCTTGCCCGGCTTCTTCATGGGAACCCTATAGGGCGATACTTTGTACGGAGATTCTGGAAAGCACTTGACAAGTCGTGGGGAGATTTCGCTGGATACGACAGTGCCCAGCACGGCGGAAAGATGGAGCATCTCCGTCCAACAAACAGGTACATTCCCTGGCTTATCTTTGATGAAATTGCCTGACCCATCTAGTATTATTTACTGCGGTGCACCGCTGGGAATTGCCAACCAGCGTGGCTTCTGGGACGCTGTTCGTGCACCGAATGTGCACATTCACCGTTCTGCAATTGAGTCGGTCTCCGGGGATGCATCTTCAAAAGGAGTTATCATCGATCTAGCTGATGATACCACTCTCCCTCTGACTGACCTTCTTATCCAAGCCACTGGATGGAAGCCCAATGTTCCTGTAGAATTCACTCCTCCACCACTAACCCTGCAGCTGGGTCTCTCATGCCGAGTCCCGAGAGCTATCCTTACACCCACAAATAGCGCCCAAGATGTCAAGGTTGACCCGGAGATCAAAAAGCTCATTCATTACTGGGATGGTATTGACTCAGTATCTGCGTCTCGCATCCGCCGCGTCTTCGGGCCCAACAGCAGCCCGCCGAAGGATGTGGTTGAAAGCACCACAGCCCCGACAGATGAATTTGAATTCAGTCCATATCGATTATTCCGCCGCATGGTAGCTCCAGAGCTTGTTGAAGAAGGCGACCGGTCTTTTGTGGCTCTGGGCTTTGTCCTAACTGCCACCACGGCTGTTGTGGCAGAGGTTCAGGCGCTGTGGGCTGCTGCTTTTTTGACGGGTGGTCTAGATGATACCAGGACCCATGACGCCCTAAGCATCAACGCCATGTCGCGAAGTGACGTTGATAGAGATGTCTCGGAGGATGTGGTCTGGGGAGGCTTGACTGGGGTTGGCCCTGGTGTTGATACTCTCAATGTAGGTTCATGCCCTACCTTATTTCAGACATACACTGACACCACCAGTACAATGACATGCTTCTCCGAGATCTGGGTCTCTCTCCGTACCGGATGGGTGGGGGCTTTATCAATGAGCTTACAAGTGTCTATACACCAAAAGCATACCGTGGAATTGTAGAAGAGTGGAAGGCTAAGCACGGTAGACCATAG
avaC	ATGTCTCAACAATATCCACAGACGAACGATGTCCCGTATCTATTCTCTAATACCTCGACTGCCTACCACACTGTCGACATTTGCAAGGTGAACACATCACTCAAAGATTCAGTCCCCCGCTTTTGGTTAATGTAGGTAAACGAGAACTTCCAACGAAGAACAGGCGTTAAAATCATGCCAGCTACATTCACGGAGGCGGTTGGAGCGATCCGAACATCCAAGCCAACACTTTCAGCCAGATACGCGACGAGCTGTTGAAAGAAAAAGATATCCTGGCTCACATTTCTGGCATTGCCGCAGTCAACTACCGTTTATCTGGAGATCCCTCGGAGGGCCGCAACGCGCGATACCCAGACCATCTTGACGATGTGAAGCGCGCTATTGGGTTCCTGGATTCCCGATACGGATTCGAAGACCGGTACATTGTTGTAGGGCACAGTGCGGGCGCTACGCTGGCATTTCAACTTGCCATGGAGCCTGAGACTTCGTCACTTCCGACTCCGAAGAAGCCGCTGGCGATTCTGGGGGCGAGTGGGATTTACGACCTGAGGAGACTTCTGAACTCTATATGGCATGTGGTGAAGTATAGGGAGGAGTATGTGCATATGTTGGAAGGCGCTTTTGGGCCGGGGGGATTTGCTTACGGGGAACAGAGTGAAGGTCGTTGCAGCTCGGAATGGGATCTTGCATCTCCTGCCAAGGCAACGAGTTATGGATCGTCGTGGTGCAATGCGCAATTGGCGATGCTGGTTCATTCTCCCGACGATGAGCTTGTTCCGTTGGAACAGTCTACGCAGTTTGCAACTGTTTTGGAATCTTCTCTTCCCCCTGGAGTGGTGCAGACCCGGTTTGACTTGGCTGGTTCGCATGATGATATTTGGAGAAAGCCCGAAGGTATAGTGCGTCTTATTTCAGAAGCAGTCATGACGCTGTTAGAAAGACATAAATAG
avaD	ATGGCAGGCATTGATCTACAGTCACTGGGGATCAGTCTGCCTCCAGGCTACTCCCTCCAAACCGCAGCAGCTCGGCCAGACATGTACGAGACCTTGTTTGATCCCGAACATCCAATGAGCTCTCTCTGGCCCCAGTTCATTACTTCGACACCTGTGTCCGAAAAGTACTGGTCTCAATTGACAGACATACCGCTCTTCGCATCCTACCAGCTAATGATTCTCCACAGAGATGTGGAGAACAAAC

	ACGAATCAGTGGTGGCTTGTGGAAATTCAGTTCCGGTATATTGTCCACTCAATGACTTGCCCGATGGTGGTTGGGAAGCTATCCTGCAGACGGGCATAGAGAACTATCGTGCCGGCCATAAACAGCACCCCAATTTGCTGTCAGCGCTCGGTGTCACGGTGGCCTCTGCTCACCGGCAGCAGGGGCTGGCTGATATTGTCATTCAAACTCTGAGAGCGCTCGCCAATCAGGCCCACTTTGAAGCCCTGGTCGTACCACTGCGTCCAACCAAGAAGAGCGAACACCCCATGGTTCCTTTGGAGGAATACGTGCATTGGAAACTTGATGAGCACGCACGTCCCGATGATCAGGTTTCATACGATCCGTGGCTGCGGAAACACCTCAGCTATGGAGGCCAAATGGTTCGAATTGCTCCCCGCAGCATGACAATCGCGGCCCACGCGGATCAATGGAAAGACTGGACTGGATGTGACCTGGCTGCCTTGGCTGAGTCCGGCTCGGACACCTCCTCGAAGGGGCTTGTCGAGATGCCTATCCCCCGTGCTCTGGTACCTGTGCAATACGACCCCGTCTCGAAGGTTGCTTCGTATGTGGAGCCCAATGTTTGGGTAGTTCATCCAATCCATTAG

Table S5. Primers used in this study

Primer	Sequence (5’-3’)
A thermomutatus NRPS check R	ACTGGGGAAGTGGCGAAAG
A thermomutatus NRPS F	ATGACTGTGCCACCAATAAGC
A thermomutatus PLP check F	CGTTCTGGCGAGATTCAAGC
A thermomutatus PLP check R	GTGATCTCGCAGAACAGCAC
A therm NRPS POgpdA F	CCTTTTATCAATCGGGGACTTCGTCGTAGTTGACATCTTTGCTCCAGGAATACATGTGAG
POgpdA Atherm P450 R	GGCCAGAAATGGCCCAGGTTGAGTTGAATTGTGCCATTTTTGCGATTGTTTGAAGTGTTC
Atherm P450 pYTU R	CGGGGGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAATCAGTCAGGATGCCACTACTC
gpdA atherm NRPS R	CGAACGAGGCTCGTTCGCTTATTGGTGGCACAGTCATTGTTTAGATGTGTCTATGTGGCG
gpdA atherm PLP R	TTCCATGTTGAGCTGATGCAGGGCCCAGTTCACCCATTGTTTAGATGTGTCTATGTGGCG
atherm PLP POgpdA F	CATCAAAGTCTAACAGAGCCAGCAAAGTAATCCAAACTTTGCTCCAGGAATACATGTGAG
POgpdA atherm HP R	TGCTGGAGGTCGGCAACTGGAATCCTGTCTTCCATTTTTGCGATTGTTTGAAGTGTTCTG
atherm HP F	ATGGAAGACAGGATTCCAGTTGC
atherm HP pYTR R	CGAATTCCTGCAGCCCGGGGGATCCTTAATTAACACAATATGTAACTATCGCTTTGTTGG
gpdA atherm hydroxylase R	TAATCAACACATCAACCGCAGACTCTTTCTTGGTCATTGTTTAGATGTGTCTATGTGGCG
atherm hydroxylase F	ATGACCAAGAAAGAGTCTGCG
atherm hydroxylase POgpdA F	GCTGGCGGTCCATCCTCAAGCAGTGTAAGTAAAATTCTTTGCTCCAGGAATACATGTGAG
atherm hydroxylase R	GAATTTTACTTACACTGCTTGAGGATGG
POgpdA atherm MT R	GTTTCATCGGGGAAGGGAATGATGCTTTCGTACATTTTTGCGATTGTTTGAAGTGTTCTG
atherm MT F	ATGTACGAAAGCATCATTCCCTTC
atherm MT pYTP R	GCTTGATATCGAATTCCTGCAGCCCGGGGGATCCTTAATTAAGGTTCTGCAGCGCTCTTC
atherm PLP pYTR R	TATCGAATTCCTGCAGCCCGGGGGATCCTTAATTAAGTTTGGATTACTTTGCTGGCTCTG
atherm PLP F3	GGCACCAATTTCTCGATTTCCATC
atherm PLP F4	GGAGGTGGAACTGAGAGAGAAG
atherm PLP F5	TGCTGGTTTCAGAAACTCGATTC
atherm PLP R3	CTATCCCATCTTCTCTCTCAGTTCC
atherm HP R2	GTCCAGTTTTCGACGGTGAG
atherm hydroxylase R2	CATCTTGAATCCCAGTTGCTTCC
atherm MT R2	CAAGAGGACGTCGAGCTG
atherm P450 R2	CTCTCGAATGAAGGGGCTAGC

atherm nrps F2	CCCTGTTTAATCGTCCGTCAG
atherm NRPS F3	CATGCGGAGATTGTCTTTGGTC
atherm P450 R3	AGATCCTGGCCGTTGTCG
atherm hydroxylase F2	CTGTGTCTGCCTCACTTGTTATC
atherm hydroxylase R3	GTGCCACAGGTATTGAACCTG
atherm P450 PEgpdA F	CTCGTGACATCTTTTGACTTGAGTAGTGGCATCCTGACTGAGATTCGTCCAGGGCTTCCC
PEgpdA atherm g20 R	CGCTCTTCGTCGCCTTTAGAGAAATTTGATACATGATTGCGGTTTACTAGAAGTTGTTAG
atherm g20 F	ATGTATCAAATTTCTCTAAAGGCGACGAAG
atherm g20 pYTU R	ATCCACTAGTTCTAGAGCGGCCGCCTTAATTAATAAAGATCACTACTCACACCAACTCAC
atherm hydroxylase check F	AAGACTACAGAGACTGCTGAGG
atherm hydroxylase check R	TGCTTTTTGGCCTCGAAAATCG
atherm MT check F	CCTTTGTCACTATTCTCAGGTCAG
atherm hp check F	ACCCTGTGAAGGGCATGC
atherm hp check R	AGATACACCTCGAGTAGATGAGG
gpdA apimo HP_cdps F	CTAACCATTACCCCGCCACATAGACACATCTAAACAATGGTTTCAATCTCCAAAAGCCTC
apimo hp_cdps pYTU R	GGGGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAAGAACGTTAATGTTTCGGGAGAAC
gpdA av hp_cdps F	ACTAACCATTACCCCGCCACATAGACACATCTAAACAATGTCTATACCAACCGATATCGC
av hp_cdps pYTU R	GGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAATGCCAGGTCTGATGTTTGATTTCTC
av HP F	ATGTCTATACCAACCGATATCGCC
av HP F2	CCTTCGTGGAAGCCATCAGAG
av HP R2	CGCAGCTGGATTTGTCCTTG
av HP R	TGCCAGGTCTGATGTTTGATTTCTC
ADH2p apimo HP F	ACAATCAACTATCAACTATTAACTATATCGTAATACATGGTTTCAATCTCCAAAAGCCTC
apimo HP xw55 R	TGAAGGCATCGGTCCGCACAAATTTGTCATTTAAACTACGTATGTCCTGAAACATTGACG
apimo HP seq R	AAGGCAGATGCCACGAAG
gpdAp av in house HP F	AAGCTTGACTAACCATTACCCCGCCACATAGACACATCTAAACAATGACGGCTACCGTGC
av in house HP term pYTU R	ATCCACTAGTTCTAGAGCGGCCGCCTTAATTAAGGAAAGATTTCTGTATCTACTCGGTTC
gpdAp pthy HP F	ACTAACCATTACCCCGCCACATAGACACATCTAAACAATGAGTACACTGTATCTGCTTGC
pthy HP pYTU R	CACTAGTTCTAGAGCGGCCGCCTTAATTAATTTTAGTGGTAGTTGTTTAGAGGATCAACG
gpdAp aspnom HP F	CTTGACTAACCATTACCCCGCCACATAGACACATCTAAACAATGATGGGTGCCGTACAAC
aspnom HP pYTU R	ATCCACTAGTTCTAGAGCGGCCGCCTTAATTAATCCTTGAAAGATTTTTACTACTGGCAG

gpdAp aspte HP F	ACTAACCATTACCCCGCCACATAGACACATCTAAACAATGGATGTGCAATATGTACTCGC
aspte HP pYTU R	GGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAAGGATCTGGCAGCATTGAATATCTCC
pET28a YDR341C F	ATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGCTAGCACAGCAAATATGATTTC
YDR341C pET28a R	GCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGCATTCTTTCTACGGGAGTTAAACCTAAC
YDR341C check F	GGATTATGGTTCTTGCAAATTGGTTG
YDR341C check R	ACAGTTTTGCGGCAGGTTC
av in house HP pYTR R	GCTTGATATCGAATTCCTGCAGCCCGGGGGATCCTTAATTAAGACATTCAGCCCGCTGTG
adh2p av HP in house F	AAAAGCATACAATCAACTATCAACTATTAACTATATCGTAATACATGACGGCTACCGTGC
av HP in house xw55 R	TGAAGGCATCGGTCCGCACAAATTTGTCATTTGACGATCTAGTCTAATTCGAATTCATCC
av in house HP seq R	GGGCCTTGGTCCAAGTAGC
n-his av in house HP F	ATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGACGGCTACCGTGC
POgpdA av HP3 R	GACTGATCCCCAGTGACTGTAGATCAATGCCTGCCATTTTTGCGATTGTTTGAAGTGTTC
av HP3 F	ATGGCAGGCATTGATCTACAG
av HP3 PEgpdA F	TAGCAATATTCAGATTCCCAGACTCCTCTTTTACCCGCCCAGATTCGTCCAGGGCTTCCC
av HP3 R	TGGGCGGGTAAAAGAGGAG
PEgpdA av KFA R	GTCTCTCATCTTCAGAGTGTTCGGGCGAAGTCATGATTGCGGTTTACTAGAAGTTGTTAG
av KFA F	ATGACTTCGCCCGAACACTC
av KFA F2	CTGGTTCGCATGATGATATTTGGAG
av FMO F2	ACATGCTTCTCCGAGATCTGG
av FMO cox4p F	GAATATTGCATACCAGAACCAACAATGACCACTCGCGGCATTGTCCTGGTGGGTGGGTTG
av FMO F	ATGCCGCGAGTGGTCATTG
cox4p pYTR R	TGCTAAAGGGTATCATCGAAAGGGAGTCATCCAATTTAATTAAAATGCCTGATCCAGCCC
atherm hydroxylase pYTP R	GAATTCCTGCAGCCCGGGGGATCCTTAATTAAGAATTTTACTTACACTGCTTGAGGATGG
pYTU coxAp F	GTAGTGAGTCATTTACTCAGCACACTCGCGCTGACGCTCGTCAATGCCTGATCCAGCCCC
av fmo pYTU R	CCTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCACGACCTGAGGAGACTTCTG
av KFA pYTU R	CTGCAGCCCGGGGGATCCACTAGTTCTAGAGCGGCCGCCTTTGTGGCTCTGGGCTTTGTC
av hp3 pYTP R	CTTGATATCGAATTCCTGCAGCCCGGGGGATCCTTAATTAATGGGCGGGTAAAAGAGGAG
n-his atherm plp F	TCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGGTGAACTGGGCCCTGCATC
atherm plp pet28A R	CGGATCTCAGTGGTGGTGGTGGTGGTGCTCGAGCTATCCCATCTTCTCTCTCAGTTCCAC
trp tRNA F	CGGTGGCTCAATGGTAGAG

trp tRNA R	AACGGACAGGAATTGAACCTG
m13 rev trp tRNA F	CAGGAAACAGCTATGACGTCTAGCGCTTTACGGAAGACAATCGGTGGCTCAATGGTAGAG
trp tRNA m13 fwd R	GTAAAACGACGGCCAGTCAAGGCGATTAAGTTGGGTAACAACGGACAGGAATTGAACCTG
m13 rev trp tRNA F	CAGGAAACAGCTATGACGTCTAGCGCTTTACGGAAGACAATCGGTGGCTCAATGGTAGAG
trp tRNA m13 fwd R	GTAAAACGACGGCCAGTCAAGGCGATTAAGTTGGGTAACAACGGACAGGAATTGAACCTG
atherm g20 spliced OH R	GTCCGGCCTCCAGTGTAGCGGGCGCTCATTGTAATGCTTCAATACGCTCTCCTCGTGAGG
atherm g20 spliced F	AGCATTACAATGAGCGCCCGCTAC
gpdA pensub HP F	CTAACCATTACCCCGCCACATAGACACATCTAAACAATGCTTGGTCCAATTTCGTTACAC
pensub HP pYTU R	GGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAACAAGCCTTCAGTGGAAGAAGGAATG
av HP Y515A F	CATTGGTCAGATCGCAGGAGCGGAGGTTC
av HP Y515A R	CTCCGCTCCTGCGATCTGACCAATG
av HP E315A F	CACCATTATCCACGCAGTACGTCGACG
av HP E315A R	CGTCGACGTACTGCGTGGATAATGGTG
av HP E432A	GTGGACGATATCGCCGCATGGAGAATCTTTGAC
av HP E432A R	GTCAAAGATTCTCCATGCGGCGATATCGTCCAC
av HP Y515F F	CATTGGTCAGATCTTCGGAGCGGAGGTTC
av HP Y515F R	CTCCGCTCCGAAGATCTGACCAATG
gpdA eupen HP F	ACCATTACCCCGCCACATAGACACATCTAAACAATGTCTATACCAACCGATATCGCCAAC
eupen HP term pYTU R	GGGATCCACTAGTTCTAGAGCGGCCGCCTTAATTAACAAAGTGCTGTTTGATATGCGCAG
avHP C193A F	GTACAGGCATCGCCCTGTTCTC
avHP C193A R	GAGAACAGGGCGATGCCTGTAC
avHP D428A F	GATTCAGGTGGCCGATATCGC
avHP D428A R	GCGATATCGGCCACCTGAATC
avHP D429A F	CTGATTCAGGTGGACGCTATCGC
avHP D429A R	GCGATAGCGTCCACCTGAATCAG
avHP C193S F	GTACAGGCATCTCCCTGTTCTCATC
avHP C193S R	GATGAGAACAGGGAGATGCCTGTAC
avHP C193T F	GTACAGGCATCACCCTGTTCTCATC
avHP C193T R	GATGAGAACAGGGTGATGCCTGTAC

Table S6. AnkA homologs

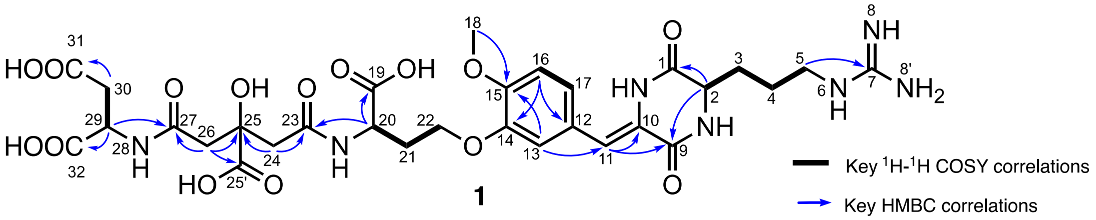
Strain	AA size of homolog	% identity to AnkA	e-value
Aspergillus udagawae CBS 114217	523	83.779	0
Aspergillus siamensis CBS 137452	523	84.895	0
Penicillium variabile	515	73.552	0
Fusarium tricinctum MPI-SDFR-AT-0068	524	42.857	7.40E-137
Astrocystis sublimbata CBS 130006	501	40.076	1.49E-125
Xylaria curta CBS 114988	501	40.571	4.94E-124
Aspergillus appendiculatus CBS 374.75	501	41.078	6.90E-124
Xylaria berteri	501	39.619	6.40E-120
Hypoxylon argillaceum CBS 527.63	500	38.783	6.69E-120
Aspergillus megasporus DTO 048-I3	524	40.49	1.04E-119
Nemania diffusa NC0034 v1.0	500	38.593	1.12E-119
Xylaria venustula FL0490	498	39.313	2.88E-119
Xylogone sp. PMI_703	297	61.074	1.22E-118
Eurotium rubrum	501	40.338	4.31E-117
Aspergillus cumulatus DTO 311-F5	501	40.49	6.64E-117
Aspergillus penicilloides CBS 540.65	490	43.75	1.55E-116
Xylariaceae sp. FL1272	501	38.692	6.82E-116
Aspergillus endophyticus DTO 354-I2	501	41.055	3.75E-114
Aspergillus endophyticus DTO 354-I2	501	37.121	2.67E-112
Poronia punctata CBS 180.79	507	38.015	1.00E-110
Podosordaria jugoyasan CBS 670.77	516	39.074	2.26E-110
Aspergillus teporis DTO058-E5	485	39.279	3.47E-105
Atropellis piniphila CBS 197.64	524	38.095	6.57E-99

Alternaria carthami BMP1963	519	35.019	1.06E-94
Nemania sp. FL0916 v1.0	452	34.63	1.82E-94
Monascus ruber NRRL 1597	584	59.211	4.85E-23
Monascus purpureus	589	60.526	3.50E-23
Aspergillus caperatus DTO 337-E6	410	40.541	3.81E-81
Apiospora montagnei NRRL 25634	543	32.299	4.23E-78
Aspergillus chevalieri CBS 522.65	405	39.394	0.015
Aspergillus multiplicatus CBS 646958	544	32.673	1.19E-70
Aspergillus carlsbadensis CBS 123894	524	32.24	1.60E-69
Pseudogymnoascus sp. VKM F-103	548	28	6.35E-66
Zalerion varium MPI-CAGE-AT-0135	519	31.216	1.82E-65
Leptosphaeria microscopica UNIPAMPA013	327	39.45	4.98E-65
Pseudogymnoascus verrucosus UAMH 10579	548	27.826	8.70E-65
Arthroderma benhamiae CBS 112371	509	33.69	3.54E-57
Trichophyton tonsurans CBS 112818	471	32.8	8.43E-56
Trichophyton interdigitale MR816	509	32.353	2.14E-54
Aspergillus egyptiacus CBS 656.73	508	33.245	6.00E-53
Thelebolus microsporus ATCC 90970	393	32.836	1.2E-52
Thelebolus globosus UNIPAMPA015	393	32.587	1.83E-51
Trichophyton equinum CBS 127.97	330	32.584	2.08E-51
Trichophyton verrucosum HKI 0517	509	33.422	7.53E-51
Aspergillus spiritus DTO 241-G7	305	32.031	1.42E-49
Endocarpon pusillum Z07020	489	29.197	1.04E-48
Aspergillus versicolor	508	32.62	2.65E-48
Nannizzia gypsea CBS 118893	509	31.818	2.89E-48
Aspergillus robustus CBS 428.77	310	36.928	3.78E-48

Aspergillus terreus NIH 2624	431	32.967	1.26E-47
Aspergillus shendaweii IBT 34197	400	32.817	1.96E-47
Aspergillus pseudofelis IBT 34107	334	31.622	1.19E-46
Aspergillus parafelis IBT 34187	402	31.635	5.93E-46
Antarctomyces sp. UNIPAMPA016	387	33.036	2.03E-44
Cucurbitaria berberidis CBS 394.84	387	31.549	6.57E-44
Thelebolus stercoreus	384	30.667	6.50E-43
Penicillium thymicola DAOMC 180753	494	28.571	1.62E-40
Aspergillus insolitus CBS 384.61	528	29.867	2.42E-39
Aspergillus pseudonomius CBS 119388	528	29.459	4.25E-37
Aspergillus nomius NRRL 13137	528	29.459	4.25E-37
Aspergillus versicolor dI-29	532	27.604	8.21E-37
Byssochlamys spectabilis No. 5	528	25.094	9.57E-35
Aspergillus affinis CBS 129190	527	27.968	1.81E-34
Aspergillus sclerotiorum CBS 549.65	436	27.331	2.15E-33
Aspergillus roseoglobulosus CBS112800	427	25.278	1.85E-31
Aspergillus tamarii CBS 117626	449	26.75	3.30E-31
Microcera lavarum	528	26.35	1.24E-30
Aspergillus nutans CBS 121.56	465	23.333	1.39E-30
Rhizina undulata CBS 300.56	497	23.475	3.46E-30
Lomentospora prolificans JHH-5317	546	22.922	1.11E-29
Aspergillus nomius IBT 12657	483	25.145	1.07E-28
Aspergillus pseudonomius CBS 119388	483	25.723	1.14E-28
Neofusicoccum parvum UCRNP2	519	26.05	2.58E-28
Tolypocladium paradoxum NRBC 100945	496	26.269	3.05E-28
Gyromitra infula GyrinfSpk-SM18	518	26.861	5.92E-28

Geopyxis carbonaria CBS 144460	497	27.393	6.47E-28
Tricharina praecox CBS 144465	476	30.345	1.64E-27
Aspergillus undulatus CBS 261.88	501	24.933	5.38E-27
Aspergillus siamensis CBS 137452	492	23.16	1.34E-26
Aspergillus pseudofelis IBT 34107	492	23.116	1.38E-25
Aspergillus appendiculatus CBS 374.75	491	23.497	2.66E-25
Aspergillus neowarcupii IBT 29024	1003	24.149	5.39E-25
Aspergillus brevistipitatus	491	24.923	7.78E-25
Aspergillus avenaceus IBT 18842	480	25.895	2.36E-23
Aspergillus parafelis IBT 34187	492	22.613	6.27E-23
Penicillium thomii	498	24.747	2.42E-22
Aspergillus shendaweii IBT 34197	491	21.833	5.99E-22
Aspergillus waksmanii IBT 31900	480	21.981	3.57E-21
Hirsutella minnesotensis 3608	689	24.541	7.14E-21
Aspergillus transcarpathicus CBS 423.68	490	24.069	7.73E-21
Pochonia chlamydosporia 170	509	20.179	1.44E-20
Scedosporium apiospermum IHEM 14462	499	22.684	2.60E-18
Lobaria pulmonaria	528	22.363	2.27E-17
Drechslerella stenobrocha 248	345	26.038	2.19E-16
Monacrosporium haptotylum CBS 200.50	541	22.892	2.02E-15
Metarhizium robertsii ARSEF 23	379	23.009	1.05E-14
Botryosphaeria dothidea	442	26.432	6.73E-14
Stanjemonium grisellum CBS 655.79	460	22.174	2.14E-13
Penicillium griseofulvum	511	32.35	5E-24

Table S7. Spectroscopic data of compound **1**



1

position	**1**	
	^{13}C	^{1}H (J in Hz)
1	168.9	–
2	55.2	4.27 (1H, t, 5.1)
3	31.2	1.94, 1.84 (each 1H, m)
4	23.2	1.66, 1.59 (each 1H, m)
5	40.8	3.17(2H, m)
7	157.1	–
9	162.7	–
10	123.9	–
11	118.9	6.81 (1H, s)
12	125.7	–
13	114.4	7.01 (1H, s)
14	147.7	–
15	149.6	–
16	112.6	7.03 (1H, d, 8.6)
17	123.3	7.07 (1H, d, 8.6)
18	56.1	3.84 (3H, s)
19	175.3	–
20	50.3	4.60 (1H, dd, 4.4, 8.9)
21	30.3	2.35, 2.16 (each 1H, m)
22	65.7	4.16, 4.09 (each 1H, m)
23	171.4	–
24	44.3	2.79, 2.70 (each 1H, d, 14.4)
25	74.1	–
26	44.0	2.82, 2.72 (each 1H, d, 16.6)
27	171.3	–
29	49.2	2.84 (1H, m)
30	35.9	2.88 (2H, m)
31	174.5	–
32	174.1	–
25'	176.6	–

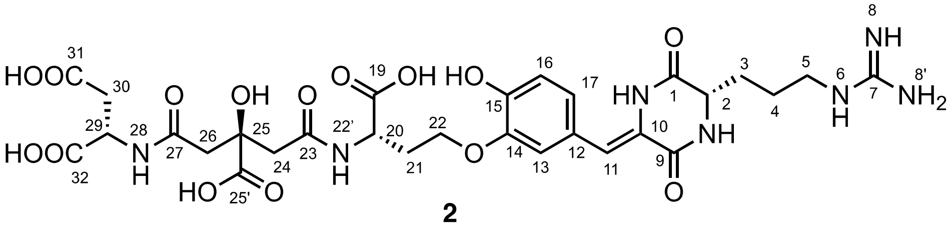
In D_2O 0.5% TFA-*d*, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

HRMS (ESI, $M+H^+$) calculated for $C_{30}H_{40}N_7O_{15}$ 738.2582; found 738.2609

$[\alpha]_D$: -120 (*c* 0.10, DMSO)

Note: The connectivity between homo-Ser and OH-Tyr was indicated by NMR chemical shift similarity of C-13, C-22, and C-15, and H-22 with those of **3b** and **5**. Considering the biosynthetic pathway of **1** and structural similarity with **7**, the geometry of C-11/C-10 is assigned to be *Z*. The *Z* configuration in **7** was determined by NOESY correlations for 1-NH/H-17 and H-11/H-13 as reported in ref [1].

Table S8. Spectroscopic data of compound **2**



2

position	**2**		**Reported 2** (NK13650A)[a]	
	^{13}C	^{1}H (*J* in Hz)	^{13}C	^{1}H (*J* in Hz)
1	166.7	–	166.9	–
2	54.5	3.99 (1H, m)	54.7	4.01 (1H, m)
3	31.1	1.71 (2H, m)	31.3	1.72 (2H, dt, 9.9, 6.1)
4	24.0	1.52 (2H, m)	24.2	1.53 (2H, m)
5	40.5[b]	3.10 (1H, q, 6.7)	40.5	3.11 (1H, q, 6.6)
6		7.51 (1H, t, 4.9)		7.48 (1H, brt, 5.7)
7	156.6	–	156.8	–
8		8.39 (1H, d, 2.1)		8.39 (1H, d, 2.7)
9	160.7	–	160.9	–
10	124.4	–	124.6	–
11	115.3	6.61 (1H, s)	115.5	6.62 (1H, s)
12	124.5	–	124.7	–
13	115.5	7.00 (1H, s)	115.7	7.01 (1H, d, 1.7)
14	146.4	–	146.6	–
15	147.3	–	147.6	–
16	115.8	6.80 (1H, d, 8.1)	116.0	6.81 (1H, d, 8.2)
17	122.9	6.98 (1H, d, 8.4)	123.1	7.01 (1H, dd, 1.7, 8.2)
18		9.82 (1H, s)		9.81 (1H, s)
19	173.3	–	173.4	–
20	49.2	4.40 (1H, m)	49.5	4.41 (1H, dt, 4.8, 8.5)
21	30.7	2.00 (1H, m)	30.9	2.00 (1H, ddt, 8.6, 14.4, 5.8)
		2.16 (1H, b)		2.18 (1H, br-dq, 12.4, 6.5)
22	65.2	3.98-4.06 (2H, m)	65.5	3.94-4.07 (2H, m)
22'		8.26 (1H, d, 7.6)		8.26 (1H, d, 4.7)
23	169.5	–	169.7	–
24	43.1	2.51-2.65 (2H, m)	43.2	2.55-2.67 (2H, m)
25	73.2	–	73.5	–
26	43.1	2.51-2.65 (2H, m)	43.2	2.55-2.67 (2H, m)
27	169.3	–	169.5	–
25'	174.7	–	174.9	–
28		8.26 (1H, d, 8.0)		8.26 (1H, d, 7.9)
29	48.6	4.50 (1H, m)	48.6	4.51 (1H, dt, 7.9, 6.2)
30	36.1	2.51-2.65 (2H, m)	36.3	2.53-2.67 (2H, m)
31	171.6	–	171.8	–
32	172.2	–	172.4	–

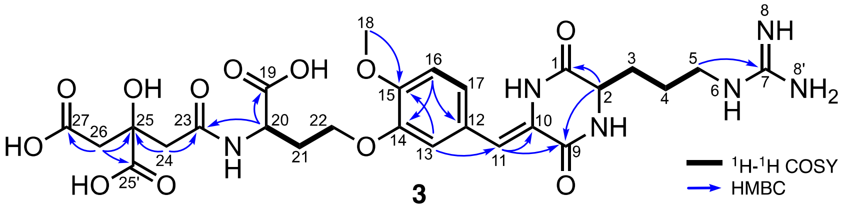
In DMSO-d_6 0.5% TFA-*d*, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm.

[a] Recorded at 600 MHz for ^{1}H and 150 MHz for ^{13}C in DMSO-d_6 0.5% TFA-*d* as reported in ref [2].
[b] The C5 peak could not be detected in the carbon spectra because it overlapped with the DMSO peak, but can be detected in the HSQC spectrum.

HRMS (ESI, M+H$^+$) calculated for $C_{29}H_{38}N_7O_{15}$ 724.2426; found 724.2439.

Measured $[\alpha]_D$ in this study: -115 (*c* 0.10, H_2O); reported $[\alpha]_D$: -157 (*c* 0.15, H_2O).[2]

Table S9. Spectroscopic data of compound **3**



3

position	3	
	^{13}C	^{1}H (*J* in Hz)
1	168.6	–
2	54.8	4.28 (1H, t, 5.1)
3	30.8	1.94, 1.85 (each 1H, m)
4	22.8	1.67, 1.59 (each 1H, m)
5	40.4	3.17 (2H, m)
7	156.7	–
9	162.4	–
10	123.6	–
11	118.6	6.83 (1H, s)
12	125.3	–
13	114.0	7.01 (1H, s)
14	147.3	–
15	149.3	–
16	112.3	7.04 (1H, d, 8.6)
17	123.0	7.08 (1H, d, 8.6)
18	55.7	3.84 (3H, s)
19	174.9	–
20	50.0	4.60 (1H, dd, 4.4, 8.9)
21	30.1	2.36, 2.15 (each 1H, m)
22	65.3	4.17, 4.09 (each 1H, m)
23	170.9	–
24	44.2	2.79, 2.70 (each 1H, d, 14.4)
25	73.4	–
26	42.8	2.93, 2.72 (each 1H, d, 16.6)
27	173.3	–
25'	176.4	–

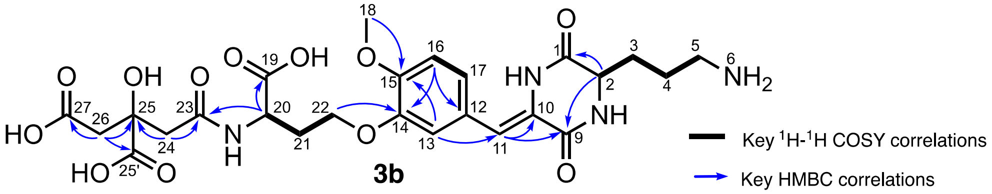
In D_2O 0.5% TFA-*d*, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

HRMS (ESI, M+H$^+$) calculated for $C_{26}H_{35}N_6O_{12}$ 623.2313; found 623.2302.

$[\alpha]_D$: -18 (*c* 0.20, DMSO)

The connectivity between homo-Ser and OH-Tyr was indicated by NMR chemical shift similarity of C-13, C-22, and C-15, and H-22 with those of **3b** and **5**.

Table S10. Spectroscopic data of compound **3b**



3b

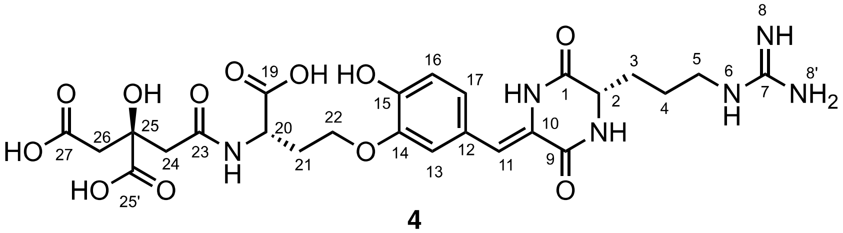
position	**3b**	
	^{13}C	^{1}H (*J* in Hz)
1	168.2	–
2	54.6	4.31 (1H, t, 5.1)
3	30.5	1.98, 1.89 (each 1H, m)
4	21.7	1.76, 1.70 (each 1H, m)
5	38.8	3.00 (2H, m)
9	162.4	–
10	123.5	–
11	118.6	6.83 (1H, s)
12	125.3	–
13	114.0	7.01 (1H, s)
14	147.3	–
15	149.3	–
16	112.3	7.04 (1H, d, 8.6)
17	123.0	7.10 (1H, d, 8.6)
18	55.7	3.84 (3H, s)
19	174.9	–
20	50.0	4.60 (1H, dd, 4.4, 8.9)
21	30.0	2.36, 2.15 (each 1H, m)
22	65.3	4.17, 4.09 (each 1H, m)
23	170.9	–
24	44.9	2.79, 2.70 (each 1H, d, 14.4)
25	73.4	–
26	42.8	2.93, 2.72 (each 1H, d, 16.6)
27	173.3	–
25'	176.4	–

In D_2O 0.5% TFA-*d*, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

HRMS (ESI, M+H$^+$) calculated for $C_{25}H_{33}N_4O_{12}$ 581.2095; found 581.2074.

$[\alpha]_D$: +1 (*c* 0.20, H_2O)

Table S11. Spectroscopic data for compound **4**



4

	4			Reported **4** (NK13650B)[a]	
position	^{13}C	^{1}H (*J* in Hz)		^{13}C	^{1}H (*J* in Hz)
1	166.7	–		166.9	–
2	54.5	4.01 (1H, m)		54.7	4.01 (1H, m)
3	31.1	1.73 (2H, m)		31.3	1.72 (2H, dt, 10.0, 5.8)
4	24.0	1.53 (2H, m)		24.2	1.53 (2H, m)
5	40.5[b]	3.11 (1H, q, 7.0)		40.5	3.10 (1H, q, 6.5)
6		7.52 (1H, t, 5.5)			7.48 (1H, t, 6.0)
7	156.7	–		156.8	–
9	160.7	–		160.9	–
10	124.4	–		124.6	–
11	115.3	6.62 (1H, s)		115.5	6.62 (1H, s)
12	124.5	–		124.7	–
13	115.5	7.02 (1H, s)		115.7	7.01 (1H, d, 1.7)
14	146.4	–		146.6	–
15	147.4	–		147.6	–
16	115.8	6.82 (1H, d, 8.0)		116.0	6.81 (1H, d, 8.0)
17	122.9	7.00 (1H, d, 8.5)		123.1	6.99 (1H, dd, 1.7, 8.0)
18		9.83 (1H, s)			9.80 (1H, s)
19	173.3	–		173.4	–
20	49.3	4.43 (1H, m)		49.5	4.42 (1H, ddd, 8.9, 7.2, 5.2)
21	30.7	2.19, 2.01 (each 1H, m)		30.9	2.00 (1H, ddt, 8.9, 14.1, 5.8) 2.19 (1H, dq, 6.5, 14.1)
22	65.3	4.06, 4.01 (each 1H, m)		65.5	4.05 (1H, dt, 10.9, 5.9) 4.00 (1H, m)
		8.30 (1H, d, 8.0)			8.27 (1H, d, 7.7)
23	169.5	–		169.7	–
24	43.1	2.54-2.76, (2H, m)		43.3	2.55-2.67 (2H, d, 14.8)
25	72.8	–		73.0	–
26	42.9	2.54-2.76, (2H, m)		43.1	2.74, 2.62 (each 1H, d, 15.6)
27	171.3	–		171.5	–
25'	174.5	–		174.8	–

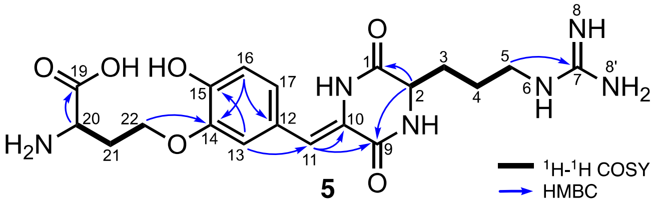
In DMSO-d_6 0.5% TFA-*d*, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm.

[a] Recorded at 600 MHz for ^{1}H and 150 MHz for ^{13}C in DMSO-d_6 0.5% TFA-*d* as reported in ref [2].

[b] The C5 peak could not be detected in the carbon spectra because it overlapped with the DMSO peak, but can be detected in the HSQC spectrum.

HRMS (ESI, $M+H^+$) calculated for $C_{25}H_{33}N_6O_{12}$ 609.2157; found 609.2161.

Table S12. Spectroscopic data of compound **5**



5

position	**5**	
	^{13}C	^{1}H (*J* in Hz)
1	168.6	–
2	54.9	4. 28(1H, t, 6.4)
3	30.8	1.93, 1.85 (each 1H, m)
4	22.9	1.66, 1.58 (each 1H, m)
5	40.5	3.19 (2H, m)
7	156.7	–
9	162.4	–
10	123.4^a	–
11	118.7	6.81 (1H, s)
12	125.0	–
13	114.8	7.02 (1H, s)
14	146.1	–
15	146.4	–
16	116.1	6.94 (1H, d, 8.7)
17	123.3^a	7.03 (1H, d, 8.7)
18	172.3	–
19	51.8	4.23 (1H, dd, 4.4, 8.9)
20	29.4	2.45, 2.38 (each 1H, m)
21	65.8	4.24 (2H, m)

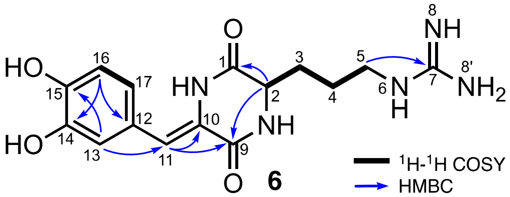
[a] Carbon data for positions 10/17 and 14/15 may be interchangeable.

In D_2O 0.5% TFA-*d*, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

HRMS (ESI, M+H$^+$) calculated for $C_{19}H_{27}N_6O_6$ 435.1992; found 435.2002.

$[\alpha]_D$: -72 (*c* 0.20, H_2O)

Table S13. Spectroscopic data of compound **6**



6

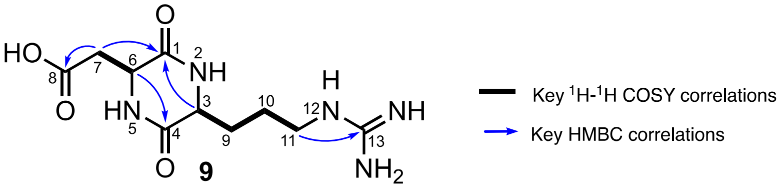
position	**6**	
	^{13}C	^{1}H (*J* in Hz)
1	168.7	–
2	54.9	4.26 (1H, t, 5.0)
3	30.8	1.94, 1.84 (each 1H, m)
4	22.8	1.67, 1.59 (each 1H, m)
5	40.4	3.17 (2H, m)
7	156.7	–
9	162.6	–
10	123.2	–
11	118.9	6.77 (1H, s)
12	125.0	–
13	116.6	6.96 (1H, s)
14	144.2	–
15	145.1	–
16	116.5	6.91 (1H, d, 8.7)
17	122.4	6.92 (1H, d, 8.7)

In D_2O 0.5% TFA-*d*, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

HRMS (ESI, M+H$^+$) calculated for $C_{15}H_{20}N_5O_4$ 334.1515; found 334.1526.

$[\alpha]_D$: -90 (*c* 0.20, H_2O)

Table S14. Spectroscopic data of compound **9**



9

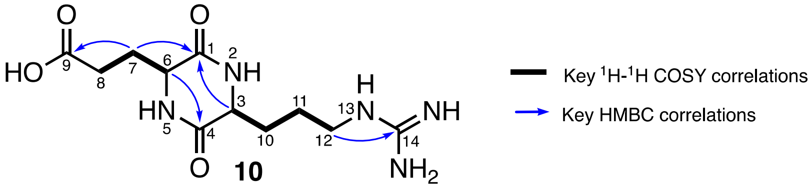
	9	
position	^{13}C	^{1}H (*J* in Hz)
1	169.8	–
3	54.1	4.15 (1H, t, 6.0)
4	169.7	–
6	52.4	4.31 (1H, dd, 7.7, 4.1)
7	40.9	2.71 (1H, dd, 16.3, 4.1) 2.63 (1H, dd, 16.3, 7.7)
8	177.1	–
9	30.3	1.83 (2H, m)
10	23.3	1.55 (2H, m)
11	40.6	3.19 (2H, m)
12	156.7	–

In D_2O, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

HRMS (ESI, M+H^{+}) calculated for $C_{10}H_{18}N_5O_4$ 272.1359; found 272.1381.

$[\alpha]_D$: -18 (*c* 0.10, H_2O)

Table S15. Spectroscopic data of compound **10**



10

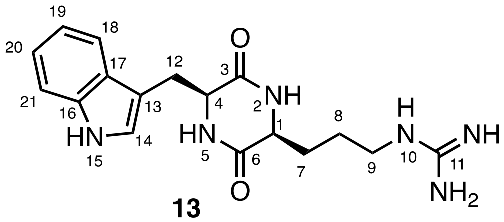
position	**10**	
	^{13}C	^{1}H (*J* in Hz)
1	169.8	–
3	54.0	4.03 (1H, m)
4	169.8	–
6	54.1	4.04 (1H, m)
7	30.3	1.93 (2H, m)
8	32.6	2.11 (2H, m)
9	181.2	–
10	30.2	1.72 (2H, m)
11	23.3	1.46 (2H, m)
12	40.6	3.09 (2H, m)
14	156.7	–

In D_2O, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

HRMS (ESI, M+H$^+$) calculated for $C_{11}H_{20}N_5O_4$ 286.1515; found 286.1516.

$[\alpha]_D$: -14 (*c* 0.20, H_2O)

Table S16. Spectroscopic data of compound **13**



13

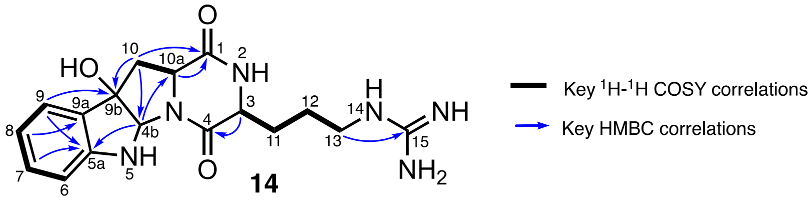
	13		cyclo(L-Arg-L-Trp)^a		cyclo(D-Arg-L-Trp)^a	
position	^{13}C	^{1}H (*J* in Hz)	^{13}C	^{1}H (*J* in Hz)	^{13}C	^{1}H (*J* in Hz)
1	55.3	3.68 (1H, t, 6.3)	55.2	3.68 (1H, ddd, 7.7, 5.2, 1.5)	52.8	2.77 (1H, m)
3	170.0		169.9		169.9	
4	57.6	4.31 (1H, s)	57.5	4.31 (1H, ddd, 4.8, 3.7, 1.3)	56.2	4.23 (1H, ddd, 4.8, 4.0, 1.1)
6	169.7		169.5		168.4	
7	32.2	0.72, 0.50 (each 1H, m)	32.1	0.49 (2H, m)	28.1	1.44 (2H, m)
8	24.7	0.86 (2H, m)	24.5	0.87 (2H, m)	22.8	1.67 (2H, m)
9	41.9	2.62, (2H, t, 6.5)	41.7	2.62 (2H, td, 7.0, 2.0)	40.5	1.44 (2H, m)
11	158.6	–	158.4	–		–
12	30.6	3.15, (1H, dd, 15.0, 3.0) 3.49, (1H, d, 14.5)	30.4	3.14 (1H, dd, 14.7, 4.6), 3.49 (1H, dd, 14.8, 3.6)	29.6	3.15 (1H, ddd, 14.7, 4.5, 0.7), 3.47 (1H, ddd, 14.7, 3.9, 0.7)
13	109.8	–	109.6	–		–
14	126.2	7.08 (1H, s)	126.0	7.08 (1H, s)	124.6	7.06 (1H, s)
16	138.0	–	137.8	–	136.4	–
17	129.6	–	129.3	–	127.3	–
18	120.5	7.63 (1H, d, 8.0)	120.2	7.63 (1H, dd, 8.0, 1.0)	118.3	7.60 (1H, dt, 8.0, 1.0)
19	120.4	7.01 (1H, t, 8.0)	120.2	7.01 (1H, ddd, 8.0, 7.0, 1.1)	118.7	7.00 (1H, ddd, 8.0, 7.0, 1.0)
20	122.7	7.10 (1H, d, 8.0)	122.5	7.09 (1H, ddd, 8.0, 7.0, 1.1)	121.1	7.08 (1H, ddd, 8.2, 7.0, 1.2)
21	112.3	7.34 (1H, d, 8.5)	112.1	7.34 (1H, d, 8.1)	110.7	7.33 (1H, dt, 8.1, 0.9)

^{a}Recorded at 600 MHz for ^{1}H and 150 MHz for ^{13}C in CD_3OD.

HRMS (ESI, M+H$^+$) calculated for $C_{17}H_{23}N_6O_2$ 343.1883; found 343.1869

Note: Marfey's analysis was used to determine the Trp moiety in **13** is in L-configuration (Fig. S20). The proton NMR spectrum of **13** corresponded to that of cyclo(L-Arg-L-Trp) rather than cyclo(D-Arg-L-Trp),[3] with a similar upfield shift in the Arg side chain resonances, supporting L-configuration of Arg in **13**.

Table S17. Spectroscopic data of compound **14**



14

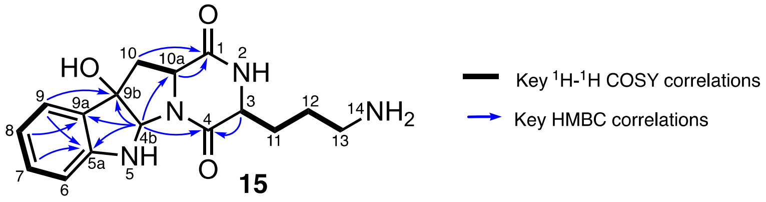
	14	
position	^{13}C	^{1}H (*J* in Hz)
1	170.5	–
3	54.5	4.31 (1H, m)
4	167.2	–
4b	80.1	5.46 (1H, s)
5a	149.4	–
6	111.8	6.83 (1H, d, 7.3)
7	131.1	7.31 (1H, t, 7.3)
8	121.2	7.00 (1H, t, 7.3)
9	124.3	7.45 (1H, d, 7.3)
9a	128.7	–
9b	85.6	–
10	40.3	2.92 (1H, dd, 11.8, 6.3)
		2.50 (1H, dd, 12.0, 11.8)
10a	57.9	4.19 (1H, dd, 12.0, 6.3)
11	26.5	1.98, 1.86 (each 1H, m)
12	22.6	1.69, 1.56 (each 1H, m)
13	40.6	3.22 (2H, m)
15	156.7	–

In D_2O, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

HRMS (ESI, M+H$^+$) calculated for $C_{17}H_{23}N_6O_3$ 359.1832; found 359.1828.

$[\alpha]_D$: -86 (*c* 0.20, H_2O)

Table S18. Spectroscopic data of compound **15**



15

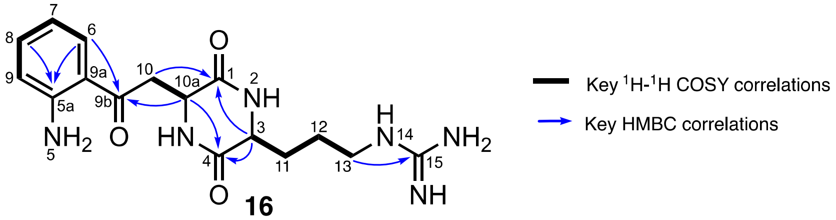
	15	
position	^{13}C	^1H (J in Hz)
1	170.7	–
3	54.2	4.30 (1H, m)
4	167.0	–
4b	80.0	5.45 (1H, s)
5a	149.4	–
6	111.8	6.82 (1H, d, 7.3)
7	131.1	7.29 (1H, ddd, 7.6, 7.4, 1.6)
8	121.2	6.98 (1H, ddd, 7.4, 7.1, 2.2)
9	124.3	7.43 (1H, dd, 7.1, 1.6)
9a	128.7	–
9b	85.6	–
10	40.1	2.91 (1H, dd, 11.8, 6.2)
		2.52 (1H, brt, 11.8)
10a	58.0	4.19 (1H, dd, 12.3, 6.2)
11	26.2	2.01, 1.87 (each 1H, m)
12	21.5	1.77, 1.65 (each 1H, m)
13	39.0	3.00 (2H, m)

In D_2O, 500 MHz for ^1H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^1H-^1H COSY, HMBC and HSQC correlation.

HRMS (ESI, M+H$^+$) calculated for $C_{16}H_{21}N_4O_3$ 317.1614; found 317.1601.

$[\alpha]_D$: -63 (c 0.20, H_2O)

Table S19. Spectroscopic data of compound **16**



16

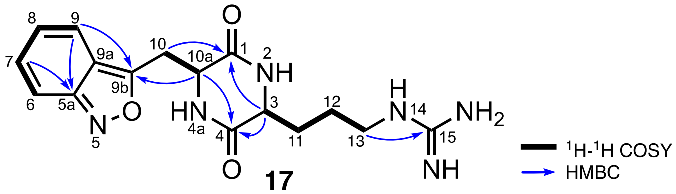
position	**16**	
	^{13}C	^{1}H (J in Hz)
1	169.6	–
3	54.0	4.25 (1H, t, 4.8)
4	169.9	–
5a	132.9	–
6	131.8	8.10 (1H, d, 7.8)
7	127.9	7.53 (1H, t, 7.8)
8	135.4	7.71 (1H, t, 7.8)
9	124.2	7.41 (1H, d, 7.8)
9a	126.5	–
9b	199.6	–
10	42.4	3.80 (1H, dd, 18.2, 3.8) 3.63 (1H, dd, 18.2, 6.8)
10a	50.6	4.64 (1H, m)
11	29.6	1.95, 1.86 (1H, m)
12	23.3	1.76, 1.64 (1H, m)
13	40.6	3.22 (2H, m)
15	156.7	–

In D_2O 0.5% TFA-*d*, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

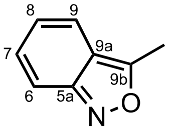
HRMS (ESI, M+H^{+}) calculated for $C_{16}H_{23}N_6O_3$ 347.1832; found 347.1831.

$[\alpha]_D$: -3 (c 0.20, H_2O)

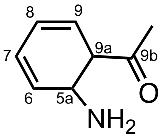
Table S20. Spectroscopic data of compound **17**



17



3-Methylbenzo[*c*]isoxazole



(2-Aminophenyl)ethanone

	17		3-methylbenzo[*c*]isoxazole[a]		(2-aminophenyl)ethanone[a]	
position	^{13}C	^{1}H (*J* in Hz)	^{13}C	^{1}H (*J* in Hz)	^{13}C	^{1}H (*J* in Hz)
1	168.1	–				
3	53.9	3.99 (1H, dd, 5.2, 2.1)				
4	169.2	–				
5a	157.3	–	157.2	–	150.3	–
6	113.8	7.44 (1H, d, 8.9)	115.0	7.43-7.41 (1H, m)	117.3	6.70-6.64 (1H, m)
7	133.0	7.36 (1H, t, 8.9)	130.9	7.25 (1H, ddd, 9.1, 6.4, 0.9)	134.5	7.75-7.71 (1H, m)
8	125.1	7.03 (1H, t, 8.9)	122.9	6.91 (1H, dd, 8.8, 6.4)	118.4	6.70-6.64 (1H, m)
9	120.2	7.50 (1H, d, 8.9)	120.0	7.51-7.48 (1H, m)	132.1	7.31-7.25 (1H, m)
9a	117.7	–	115.8	–	115.9	–
9b	164.8	–	165.8	–	200.9	–
10	30.3	3.92 (1H, dd, 15.5, 3.3)				
		3.58 (1H, dd, 15.5, 5.1)				
10a	54.1	4.62 (1H, m)				
11	29.4	1.14, 1.01 (1H, m)				
12	22.3	0.48, 0.31 (1H, m)				
13	40.7	2.68 (2H, m)				
15	156.7	–				

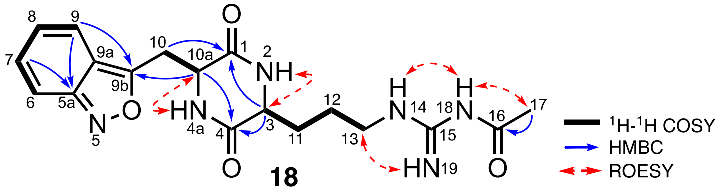
In D_2O 0.5% TFA-*d*, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC and HSQC correlation.

[a] Recorded at 400 MHz for ^{1}H and 100 MHz for ^{13}C in $CDCl_3$. The existence of anthranil moiety was supported by comparison of the experimental NMR data of **17** with the reported NMR data for 3-methylbenzo[*c*]isoxazole in ref [4] and (2-aminophenyl)ethanone in ref. [5].

HRMS (ESI, M+H^{+}) calculated for $C_{16}H_{21}N_6O_3$ 345.1675; found 345.1666.

$[\alpha]_D$: -44 (*c* 0.10, MeOH)

Table S21. Spectroscopic data of compound **18**



18

position	**18**		3-methylbenzo[*c*]isoxazole[a]		(2-aminophenyl)ethanone[a]	
	^{13}C	^{1}H (*J* in Hz)	^{13}C	^{1}H (*J* in Hz)	^{13}C	^{1}H (*J* in Hz)
1	166.0	–				
3	53.2	3.78 (1H, dd, 6.3, 2.0)				
4	166.8	–				
5a	156.4	–	157.2	–	150.3	–
6	114.1	7.49 (1H, d, 8.2)	115.0	7.43-7.41 (1H, m)	117.3	6.70-6.64 (1H, m)
7	131.3	7.31 (1H, t, 8.2)	130.9	7.25 (1H, ddd, 9.1, 6.4, 0.9)	134.5	7.75-7.71 (1H, m)
8	123.3	6.98 (1H, t, 8.2)	122.9	6.91 (1H, dd, 8.8, 6.4)	118.4	6.70-6.64 (1H, m)
9	120.9	7.63 (1H, d, 8.2)	120.0	7.51-7.48 (1H, m)	132.1	7.31-7.25 (1H, m)
9a	116.6	–	115.8	–	115.9	–
9b	165.3	–	165.8	–	200.9	–
10	30.2	3.73 (1H, dd, 15.7, 3.6) 3.53 (1H, dd, 15.7, 5.1)				
10a	53.3	4.42 (1H, m)				
11	29.7	1.28, 1.10 (each 1H, m)				
12	22.5	1.03, 1.00 (each 1H, m)				
13	40.5	3.26 (2H, m)				
15	152.9	–				
16	172.6					
17	24.2	2.15 (3H, s)				
2-NH	–	8.19 (1H, brs)				
4a-NH	–	8.42 (1H, brs)				
14-NH	–	8.98 (1H, brs)				
18-NH	–	11.84 (1H, brs)				
19-NH	–	8.66 (1H, brs)				

In DMSO-d_6, 500 MHz for ^{1}H and 125 MHz for ^{13}C NMR; Chemical shifts are reported in ppm. All signals are determined by ^{1}H-^{1}H COSY, HMBC, ROESY, and HSQC correlation.

[a] Recorded at 400 MHz for ^{1}H and 100 MHz for ^{13}C in $CDCl_3$. The existence of anthranil moiety was supported by comparison of the experimental NMR data of **17** with the reported NMR data for 3-methylbenzo[*c*]isoxazole in ref [4] and (2-aminophenyl)ethanone in ref. [5].

HRMS (ESI, M+H$^+$) calculated for $C_{18}H_{23}N_6O_4$ 387.1781; found 387.1777.

$[\alpha]_D$: -24 (*c* 0.20, H_2O)

SUPPLEMENTARY FIGURES

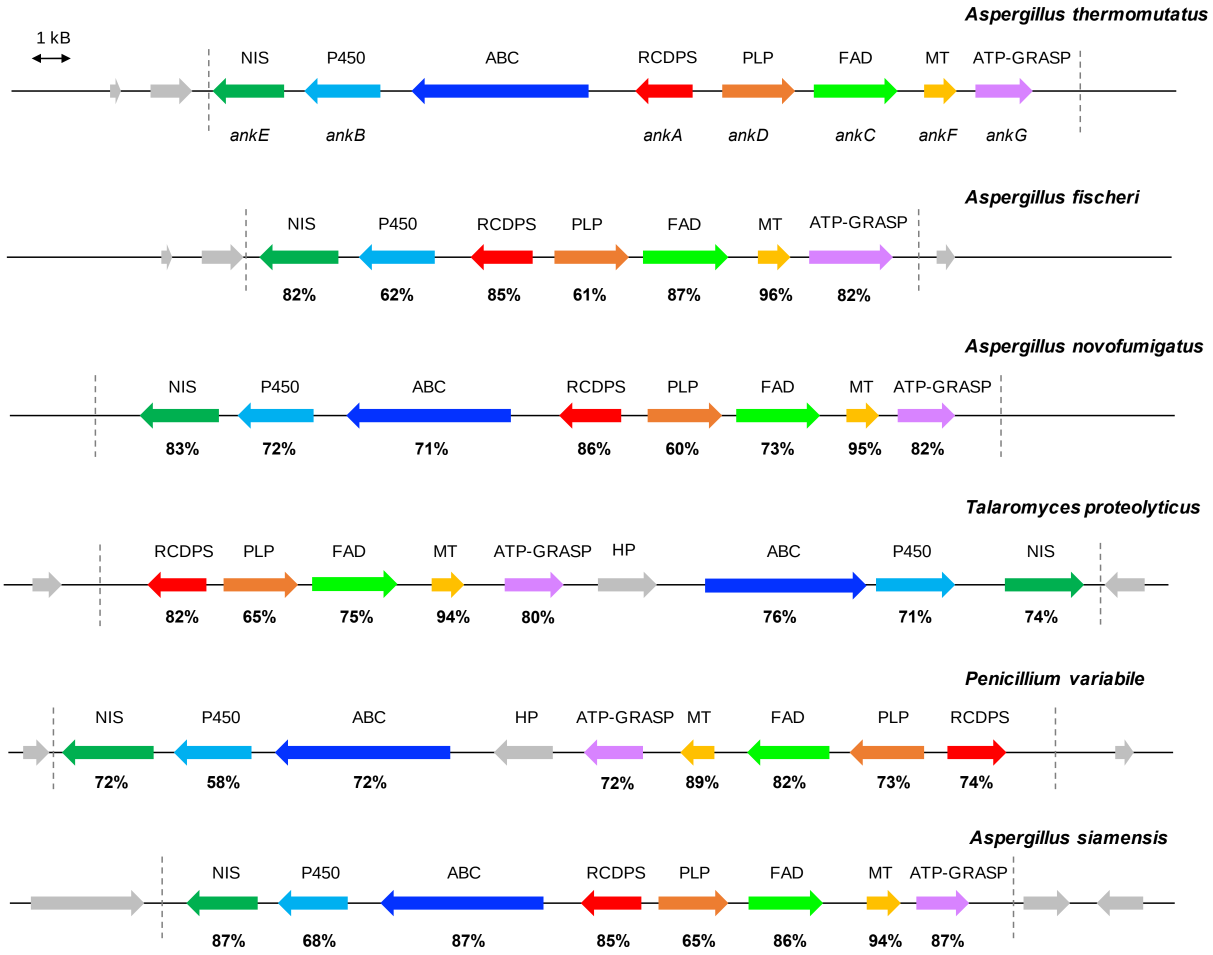


Figure S1. The *ank* cluster and homologous clusters from *Aspergillus thermomutatus*, *Aspergillus fischeri*, *Aspergillus novofumigatus*, *Talaromyces proteolyticus*, *Penicillium variable*, and *Aspergillus siamensis*. Seven biosynthetic genes from *ankA* to *ankG* are conserved in all of the clusters.

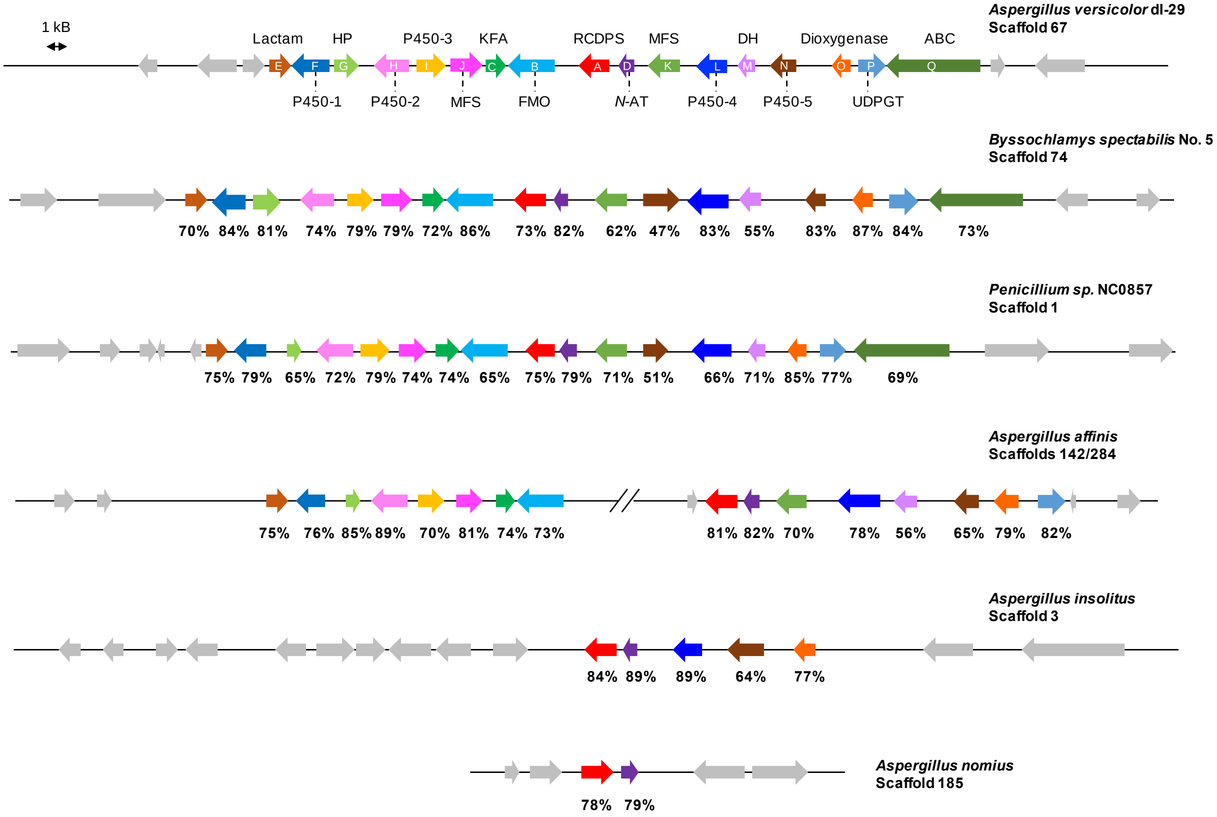


Figure S2. The *ava* cluster and homologous clusters in *Aspergillus versicolor* dI-29, *Byssochlamys spectabilis* No. 5, *Penicillium sp.* NC0857, *Aspergillus affinis*, *Aspergillus insolitus*, and *Aspergillus nomius*. Cluster genes are located on two scaffolds in *A. affinis*. Nearly all biosynthetic cluster genes are conserved in *A. versicolor* dI-29, *B. spectabilis* No. 5, *P. sp.* NC0857, and *A. affinis*.

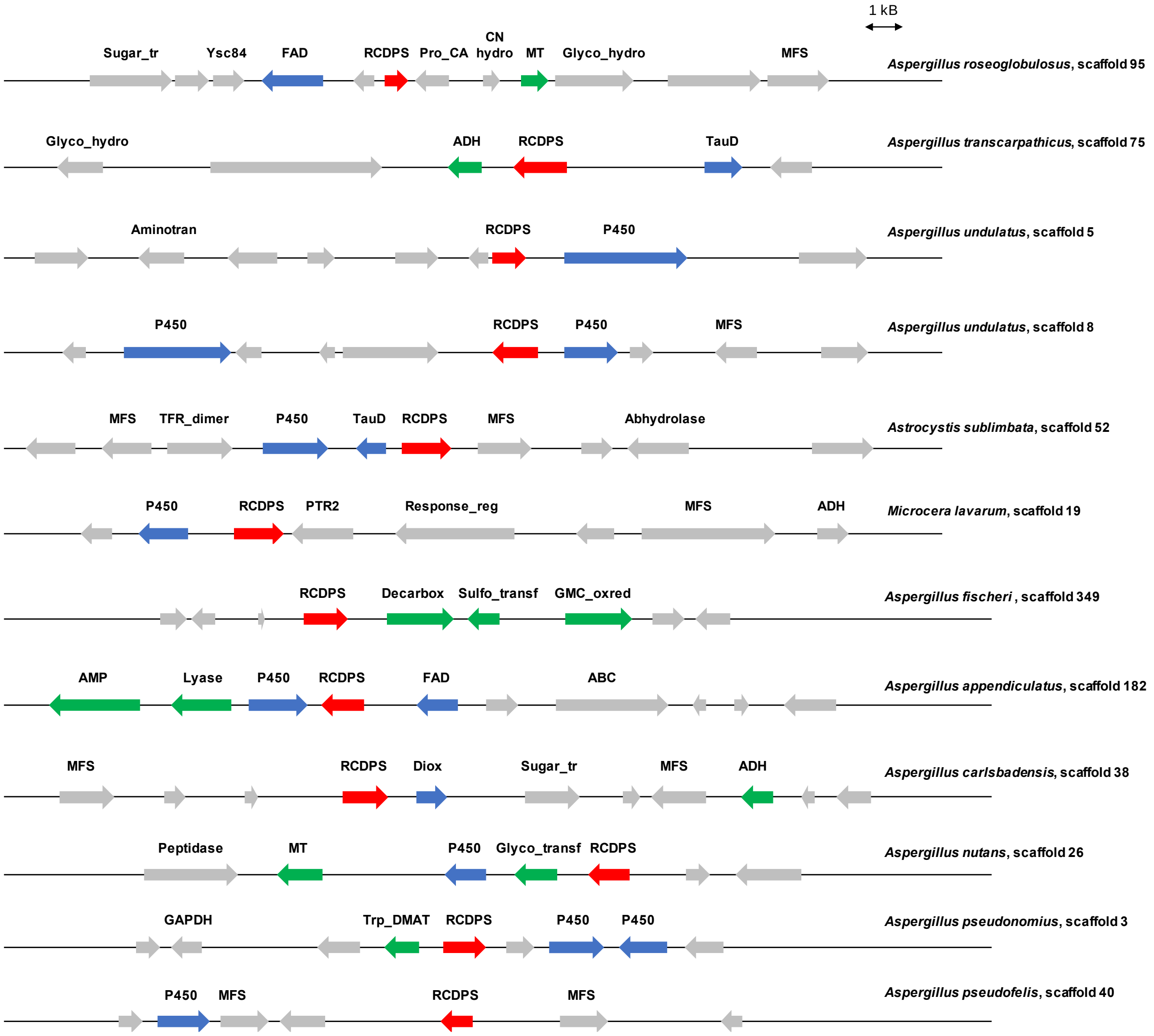
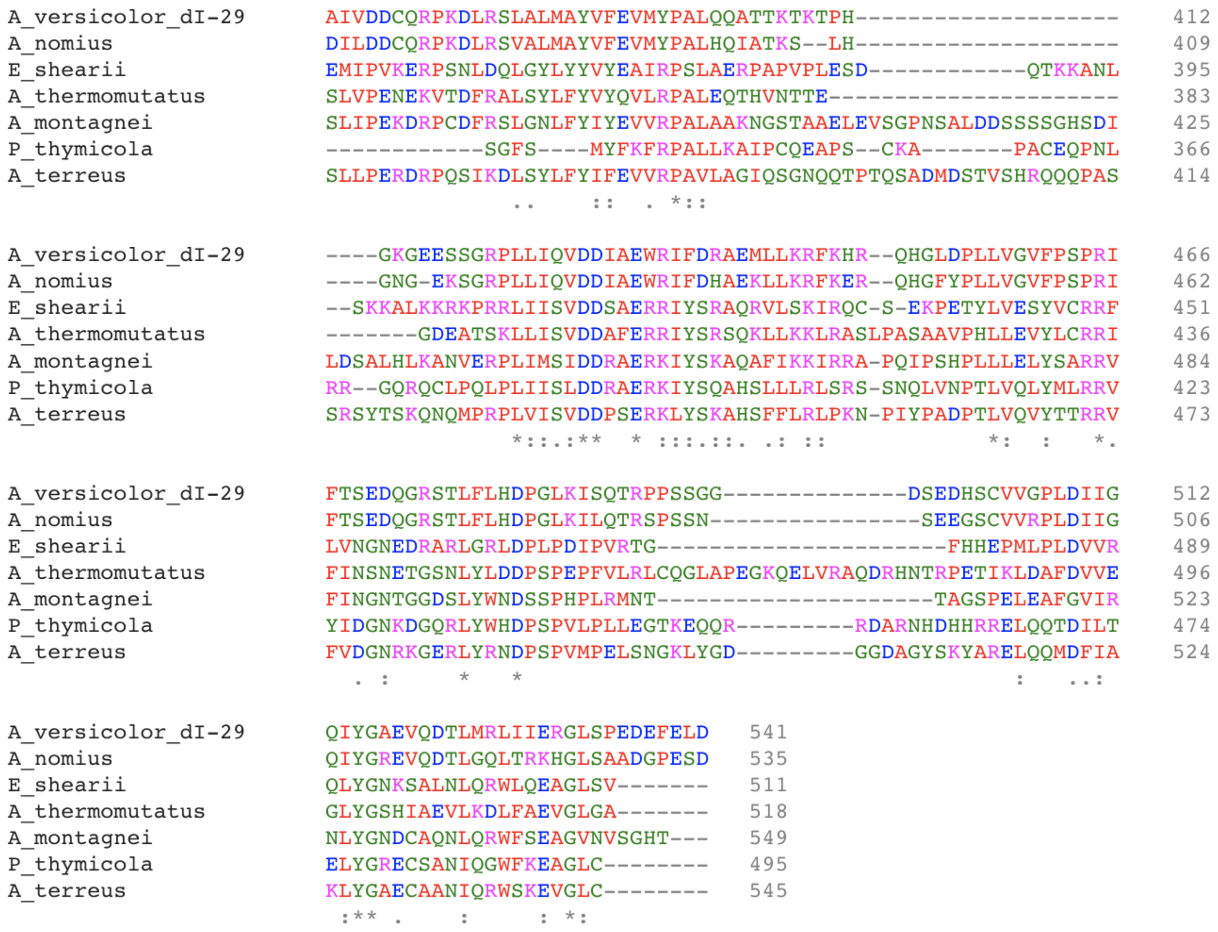


Figure S3. RCDPS homologs clustered with biosynthetic tailoring genes. RCDPS homologs are colored red, potential oxidative enzymes are colored blue, and other potential tailoring enzymes are colored green.

```
CLUSTAL O(1.2.4) multiple sequence alignment


A_versicolor_dI-29      MTATVQHTVKTEVCLRGDVASQPISDSKPPTWVAVLQQHE----YQSSVTYHKLDDNRLI    56
A_nomius                MMGAVQLPVETEACLHDDAASRPISDSNT-VRVAVLRQHQ----YHSSVTYHRLDDNRLI    55
E_shearii               --MSIP------------------------TDIANLLIGEDA---KLSIHFDKIPLHNDA    31
A_thermomutatus         --MEDR------------------------IPVADLQHSAG---SLAAIAFH------NI    25
A_montagnei             MVSISK------------------------SLVADICVGS----SPPTVRFVTLLENHDI    32
P_thymicola             --MSTQ------------------------YLLANLRNRPDGRPS---IDFGSLSARESD    31
A_terreus               --MDVQ------------------------YVLADLQTGPESQSSQSSIHFHDMAPFYVV    34
                                                        :* :            : :

A_versicolor_dI-29      --------PDVFGYSYLRDIAD-------QGKKQSTGKQHARVIQAYSKIHSLLSPPWKS    101
A_nomius                --------PGVVGYSYLRDIAD-------EGKKQSPSKQHARVIQAYSKIHSLLSPPREP    100
E_shearii               ----HHSGSKF-------YMLSTNRALRDHEEKEDDLKYLSKGRQGAIRLNQLLNAPHLI    80
A_thermomutatus         EHCTQKTSSAILGYYHLPHIGSEHSSSDNADEAAKYTKAFIKGRKVAVKVANLINQGKTI    85
A_montagnei             SNTTFPSLSGVLAQYGSSLQGPAGLQHVGLAQDQKYQHALTKGLKAVEKLQHFLDGPAPM    92
P_thymicola             PVKMPHEITRLF---SSDYCTSTST-TLEAEESEVYWKALSRNKRAAGILEHFLNGEQAA    87
A_terreus               SIERQHGLSGLV---SIVPCSTTPTIHYDEKDKKAYQKAWSRGRRFAALFENLLNGGKMA    91
                                  .                    .     :   :  :    .  ::.

A_versicolor_dI-29      SSIDSNTKTRENKGSP-------VILEDKRCAFIERLEPPPNSKADIVSTVFAQVNLQTP    154
A_nomius                GPIDSETKTREKKSSP-------VIVHDKRCAFIERLEPPPNSKADIVNTVFAQVNLRTP    153
E_shearii               KERESTA---QNEGKAVMRLGETAMVIDETFDA------QPIHSHIW--NSFMSIEFRFP    129
A_thermomutatus         HPV-------------PLRVTGSGLVGGPPCEG------HARLIEGI--TWFEKTQFHHQ    124
A_montagnei             SGDPVVLRCSSARSSTLLEITDETITSGPPCAG------AAREREGI--TSFHRMKFHLP    144
P_thymicola             AHPSVY-----LQDRARLRLCEEAILVGAPRGT------ATRTREGI--ASYDKMEFLFP    134
A_terreus               TKTSCS-----LQQTPYLQLRDEALIVGSPLGT------PTRTREGI--SSFHNMEFNFP    138
                                                :  .                       :   ::

A_versicolor_dI-29      TGPP--------LEQF--LSCRASNIKANDLK-NAAEGVRRPITISTGICLFSSRLLGFI    203
A_nomius                VGPP--------LDQF--VNCRSSNIKINDLK-DSANGLCRPITISTGICLFSSRLLGFI    202
E_shearii               QAD-------FTSLQYLQTSWKGLESGDTLYT-KEGTPTFEEVSIVSGVCLFNARLLEMS    181
A_thermomutatus         SLRP--EGNHSTASHYLHFSTPLAQIIPGAEGRTQPMDVVQDITLISGMCLWASRIIQSS    182
A_montagnei             KHVGCVQDVHETSSQMLHLSCDGLRSEDTLFD-RFGRWSRDHVTIVSGMCLFSDKIMRDA    203
P_thymicola             ESR------QRKSIQMLHVSCPSLQPGATLLD-SHSGSVHDSITILSGMCLFTRKILNST    187
A_terreus               GAK------PVVSSQMLHITCPGLRSEESLFD-KNGGALYDSVTVLTGMCLFTRKILDES    191
                                      :    .    .                 ::: :*:**:  :::

A_versicolor_dI-29      ---PTNGLSVSHDATETVVPPLPYTTVATFYELETCTRMAMTIAGLAVTA--STGGTAGS    258
A_nomius                ---PTNGLSTTDGATETVIPPLPYSADATFYELETCARMAMTIAGLAVTA--GTGGTAGS    257
E_shearii               TTCPKSGGVDV---------DLLKSSTPTYNELDYIARASSAIADVAAMNILRACEQDSG    232
A_thermomutatus         --MPRGQRLDV---------SRLLSSKITLYEVEYLARAVPVIADLAAL--SANNPYGP    229
A_montagnei             --VQGDGKLDV---------SKLSSQEPTLYELEIIARSSSAIADVASHIL-GKNGAAKP    251
P_thymicola             ---TTNGCLDG---------QKLRHMSPTLYEMECIARLSAIIADVTTLTQ-SNFGEDNH    234
A_terreus               ---VADGFLDT---------RKLLDMQLTLYEVERIAGLSSAIADIAALVL-SRYDGKCH    238
                             .                      *  *::  :     **.::

A_versicolor_dI-29      RPIVVRLDVPNLQYYCYPLELLEAGLVSWEYVEEWFRLVDRRHRQVATLLKDTIIHEVRR    318
A_nomius                RPIVVRLDVPCLQYYCYPLELLQAGLVSWKYVQEWFRLVDRRHRQVAGLLKDTITHEVLR    317
E_shearii               QRLTIKLDIPSWHYFHTVATKFVSKQCSNTEVLQWMDAVDQRHDQIGQTFVEAIRYGLEQ    292
A_thermomutatus         TPLNIRLDIPSFHYYHSLEERLRDGCCTFPEALQWMHAVEKRHHQLSRVFCRLIDHELSR    289
A_montagnei             IPLEIVLDAPSWHYFQVVHGNLASGHCTPAEALDWLQAVELRCEQVTTVFENSVRHEMGL    311
P_thymicola             PAVNIDC--------------IKSDACSLEEAIDWTEAIKLRRQQLASVLKKAVWHELGQ    280
A_terreus               KPVNIILDLPTWHYYQSVEDQLQHRGYKPCEVMDWIEAITLRSQQLAGLLKSAVLHELDR    298
                          : :                :     .   . :*   :  *  *:   :   : : :

A_versicolor_dI-29      RNCDVQ------VDVTSGTIAATQLLRLCVLGRRKIPSVNDMLFVLSWIGPYQAAWREFL    372
A_nomius                RGGDIQ------VEVTAGTIAATQLLRLSVFDRMELPSVDDILFVLKWVGPYQAAWREFL    371
E_shearii               RGIHDS--TSYGIGMTSRTNTAAILIRT-AIEHEEVPSLDAILAALD--SEEDGCWKRFY    347
A_thermomutatus         RWVGTPHRRKLDVQVSPLADLVFQLICD-SLANSVLPDVDDILQI-V--QTEDTTWVRFY    345
A_montagnei             RGVPA---GFYHILAAPGTAGVGTSIRQ-ALTSGMVPDIADCMDAIC--EVEGERWAMFY    365
P_thymicola             RQVAHK-STLKAIQISPESTVVDELIKE-TLQRGHQPRLDDILHALS--GTPHKI-----    331
A_terreus               RGVSSS-QRLYDIQISPGSALVDDAFRE-TLKYENLPCLDNILEALS--GSEDGSWQQFY    354
                        *           :  :  :  .   :    :     * :   :
```



```
A_versicolor_dI-29      AIVDDCQRPKDLRSLALMAYVFEVMYPALQQATTKTKTPH--------------------    412
A_nomius                DILDDCQRPKDLRSVALMAYVFEVMYPALHQIATKS--LH--------------------    409
E_shearii               EMIPVKERPSNLDQLGYLYYVYEAIRPSLAERPAPVPLESD------------QTKKANL    395
A_thermomutatus         SLVPENEKVTDFRALSYLFYVYQVLRPALEQTHVNTTE----------------------    383
A_montagnei             SLIPEKDRPCDFRSLGNLFYIYEVVRPALAAKNGSTAAELEVSGPNSALDDSSSSGHSDI    425
P_thymicola             ------------SGFS----MYFKFRPALLKAIPCQEAPS--CKA-------PACEQPNL    366
A_terreus               SLLPERDRPQSIKDLSYLFYIFEVVRPAVLAGIQSGNQQTPTQSADMDSTVSHRQQQPAS    414
                                      ..    ::  . *::

A_versicolor_dI-29      ----GKGEESSGRPLLIQVDDIAEWRIFDRAEMLLKRFKHR--QHGLDPLLVGVFPSPRI    466
A_nomius                ----GNG-EKSGRPLLIQVDDIAEWRIFDHAEKLLKRFKER--QHGFYPLLVGVFPSPRI    462
E_shearii               --SKKALKKRKPRRLIISVDDSAERRIYSRAQRVLSKIRQC-S-EKPETYLVESYVCRRF    451
A_thermomutatus         -------GDEATSKLLISVDDAFERRIYSRSQKLLKKLRASLPASAAVPHLLEVYLCRRI    436
A_montagnei             LDSALHLKANVERPLIMSIDDRAERKIYSKAQAFIKKIRRA-PQIPSHPLLLELYSARRV    484
P_thymicola             RR--GQRQCLPQLPLIISLDDRAERKIYSQAHSLLLRLSRS-SNQLVNPTLVQLYMLRRV    423
A_terreus               SRSYTSKQNQMPRPLVISVDDPSERKLYSKAHSFFLRLPKN-PIYPADPTLVQVYTTRRV    473
                                      *::.:**  * :::.::. .: ::            *:  :   *.

A_versicolor_dI-29      FTSEDQGRSTLFLHDPGLKISQTRPPSSGG--------------DSEDHSCVVGPLDIIG    512
A_nomius                FTSEDQGRSTLFLHDPGLKILQTRSPSSN----------------SEEGSCVVRPLDIIG    506
E_shearii               LVNGNEDRARLGRLDPLPDIPVRTG----------------------FHHEPMLPLDVVR    489
A_thermomutatus         FINSNETGSNLYLDDPSPEPFVLRLCQGLAPEGKQELVRAQDRHNTRPETIKLDAFDVVE    496
A_montagnei             FINGNTGGDSLYWNDSSPHPLRMNT---------------------TAGSPELEAFGVIR    523
P_thymicola             YIDGNKDGQRLYWHDPSPVLPLLEGTKEQQR---------RDARNHDHHRRELQQTDILT    474
A_terreus               FVDGNRKGERLYRNDPSPVMPELSNGKLYGD---------GGDAGYSKYARELQQMDFIA    524
                          . :     *   *                                     :   ..:

A_versicolor_dI-29      QIYGAEVQDTLMRLIIERGLSPEDEFELD   541
A_nomius                QIYGREVQDTLGQLTRKHGLSAADGPESD   535
E_shearii               QLYGNKSALNLQRWLQEAGLSV-------   511
A_thermomutatus         GLYGSHIAEVLKDLFAEVGLGA-------   518
A_montagnei             NLYGNDCAQNLQRWFSEAGVNVSGHT---   549
P_thymicola             ELYGRECSANIQGWFKEAGLC--------   495
A_terreus               KLYGAECAANIQRWSKEVGLC--------   545
                         :** .    :     : *:
```

Figure S4. Clustal omega (1.2.4) multiple sequence alignment of RCDPS homologs with activity verified *in vivo*. Conserved residues are starred.

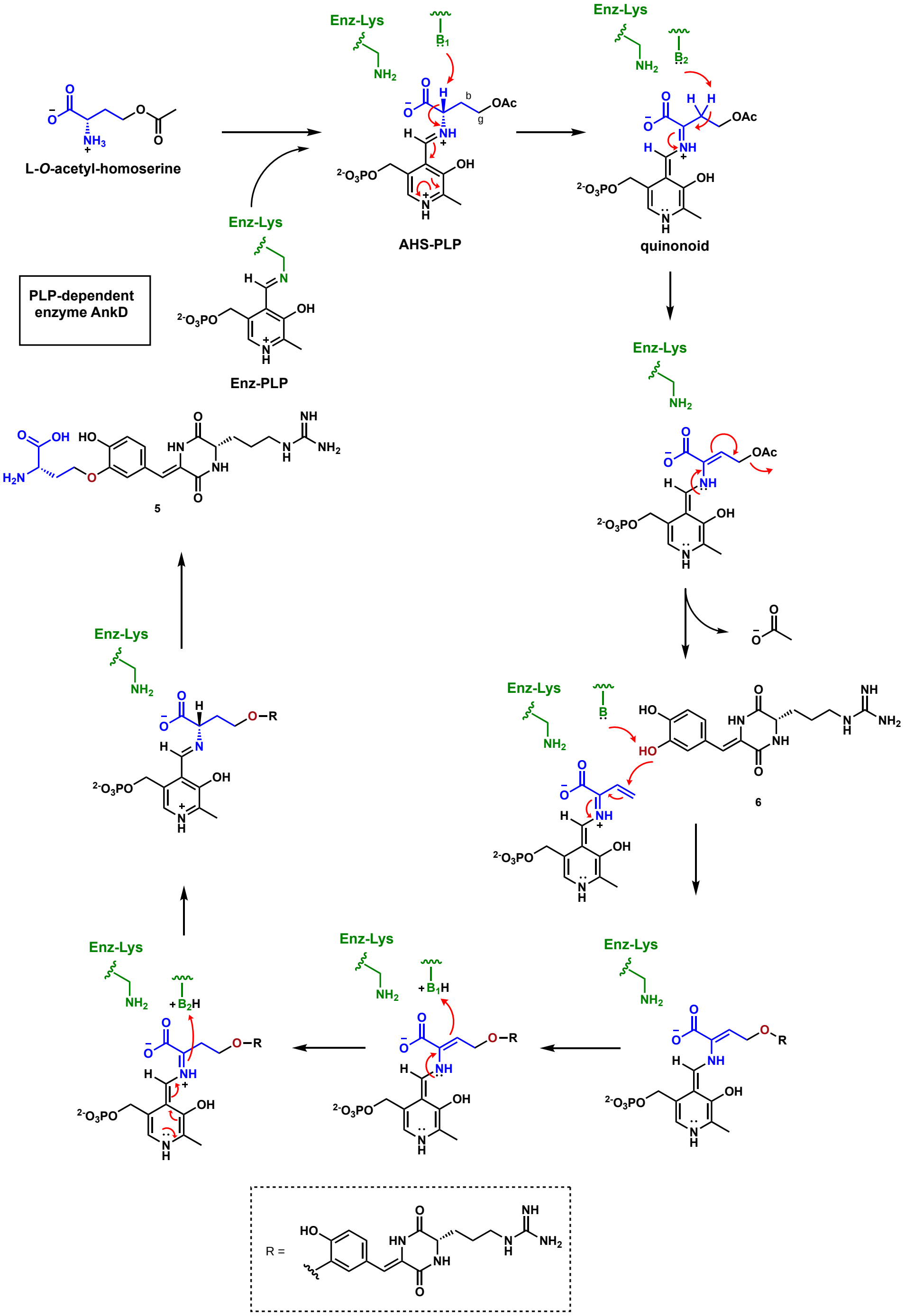


Figure S5. Proposed mechanism of PLP-dependent enzyme AnkD from L-*O*-acetyl-homoserine and **6** to afford **5**.

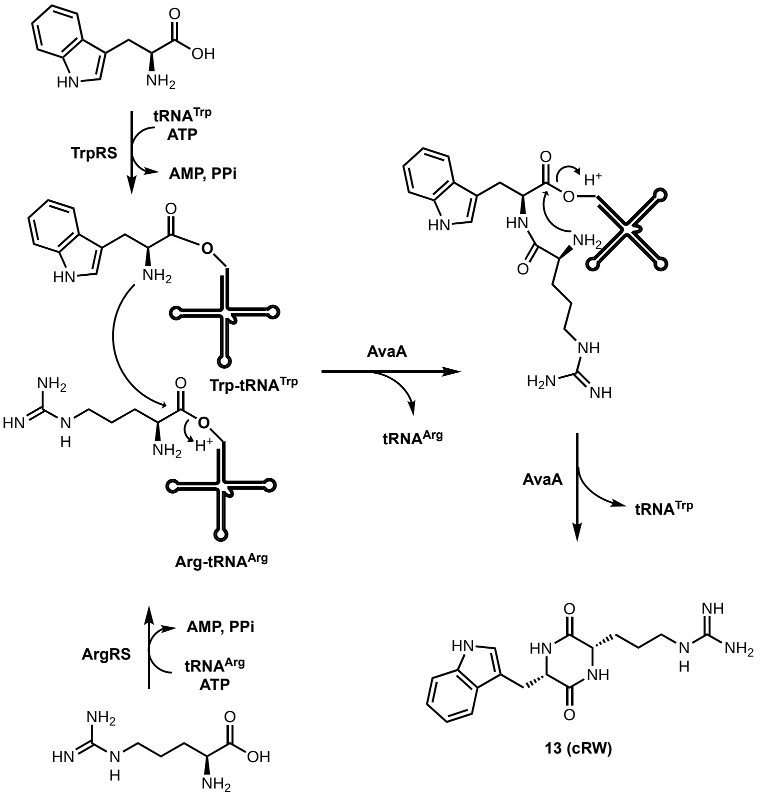


Figure S6. Proposed tRNA-dependent mechanism of AvaA to synthesize **13** from L-Trp and L-Arg. Since the order of amide bond formation is unknown, it is also possible the first amide bond is formed between the amino group of Arg and carboxylate group of Trp.

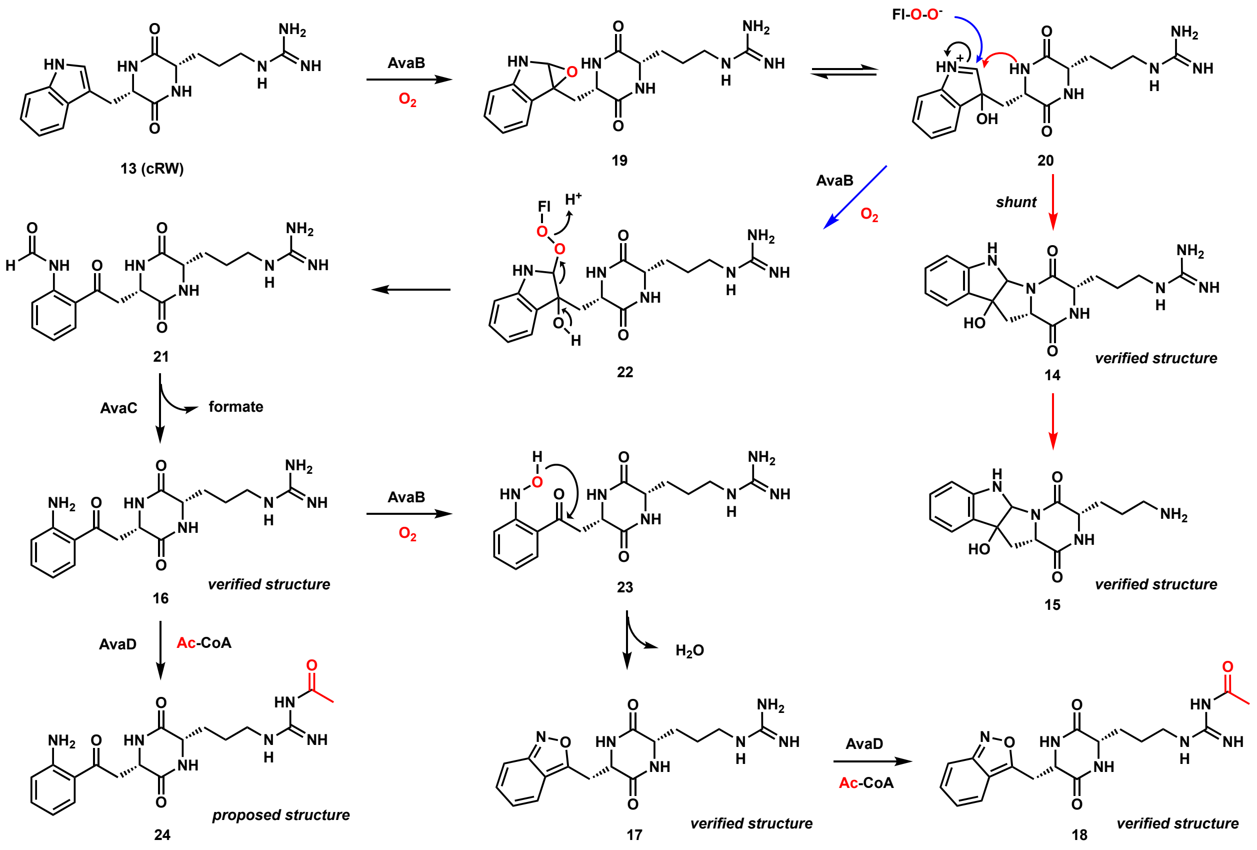


Figure S7. The proposed mechanism for the FMO AvaB. We propose a multifunctional role of AvaB. A Fl-C4a-OO$^-$ species in AvaB can attack the iminium intermediate **20** to give the adduct **22**, which can then undergo C-C bond cleavage with ejection of Fl-C4a-OH to give the cyclo-Arg-formylkynurenine DKP **21**. Hydrolysis of the formyl group can be catalyzed by AvaC, as well as endogenous KFA, to yield **16**. Oxidation of the aniline ring to *N*-phenylhydroxylamine **23** by AvaB can lead to an isoxazoline intermediate that can dehydrate to form **16**. This pathway therefore requires AvaB to perform three different oxidation reactions, which has precedence in fungal flavin-dependent enzymes.[6] The formation of **14** is therefore a shunt product as a result of facile intramolecular cyclization. At this point, we cannot exclude endogenous IDO-like enzyme catalyzing the indole cleavage directly from **13** to **21**. **16** and **17** can then be acetylated by AvaD to afford **24** (proposed based on HRMS (ESI, M+H$^+$) calculated for $C_{18}H_{25}N_6O_3$ 389.1937; found 389.1924) and **18**, respectively (Fig. S21).

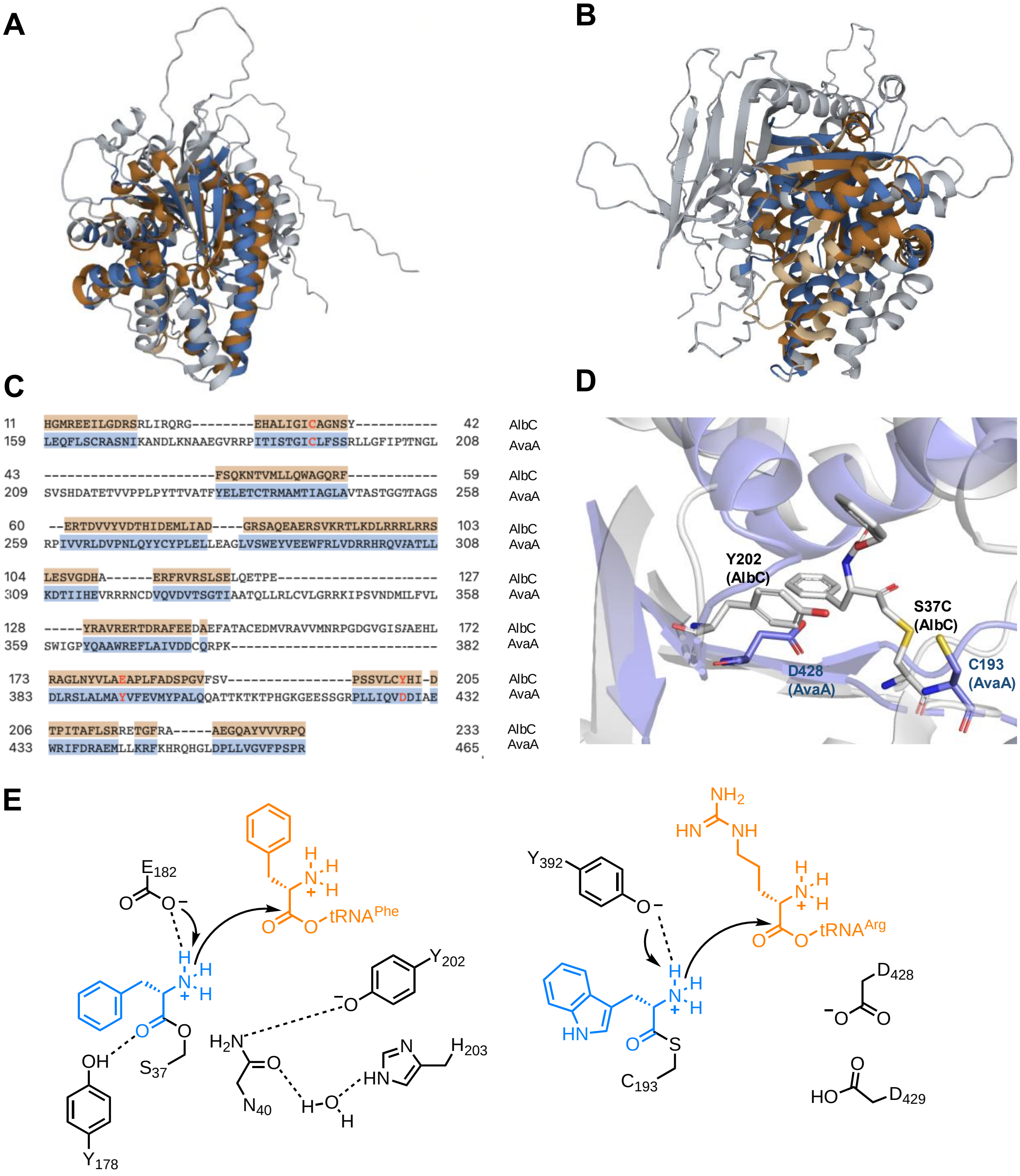


Figure S8. RCSB PDB pairwise structural alignment of the predicted structure of AvaA to the bacterial CDPS AlbC S37C mutant. The AlphaFold[7] structural prediction of AvaA is colored blue and AlbC S37C mutant (4Q24)[8] is colored brown. The dark blue and dark brown portions correspond to regions with similar structure, whereas the gray (AvaA) and beige (AlbC) correspond to regions with low similarity. (**A**) Orientation that highlights the similar core Rossmann fold. (**B**) Orientation that that highlights the additional folds in AvaA not present in AlbC. (**C**) Sequence alignment from the RCSB PDB pairwise structural alignment between AlbC and AvaA. Aligned active site residues are colored red. (D) Overlay of AlbC S37C mutant (gray) and AvaA model (purple) generated by AlphaFold. C193 in AvaA aligned to S37C in AlbC, while D428 aligned to Y202 in AlbC. S37 and Y202 are the catalytic residues in AlbC. In the structure of AlbC S37C (4Q24), the cysteine is covalently modified with a substrate analogue which can be seen in the enzyme active site. (**E**) Comparison of active site architecture of AlbC and AvaA. Using cysteine rather than serine for acyl intermediate formation in AvaA is likely to provide advantages such as kinetic enhancement, as well as reducing competing solvolysis by bulk solvent.[9,10] This has been demonstrated in thioesterases of polyketide synthases and NRPSs,[9,10] exemplifying a strategy of convergent evolution across different enzymes. Y392 in AvaA, equivalent to E182 in AlbC, may deprotonate the positively charged amino group of the acyl acceptor in preparation for formation of the first amide bond. Given the aligned position of D428 and Y202, it is likely that once the dipeptidyl intermediate forms, D428 is responsible for deprotonating the positively charged amino group of the acyl donor to attack the C193-thioester intermediate and form the diketopiperazine ring. The exact mechanism of RCDPS requires X-ray crystal structure validation.

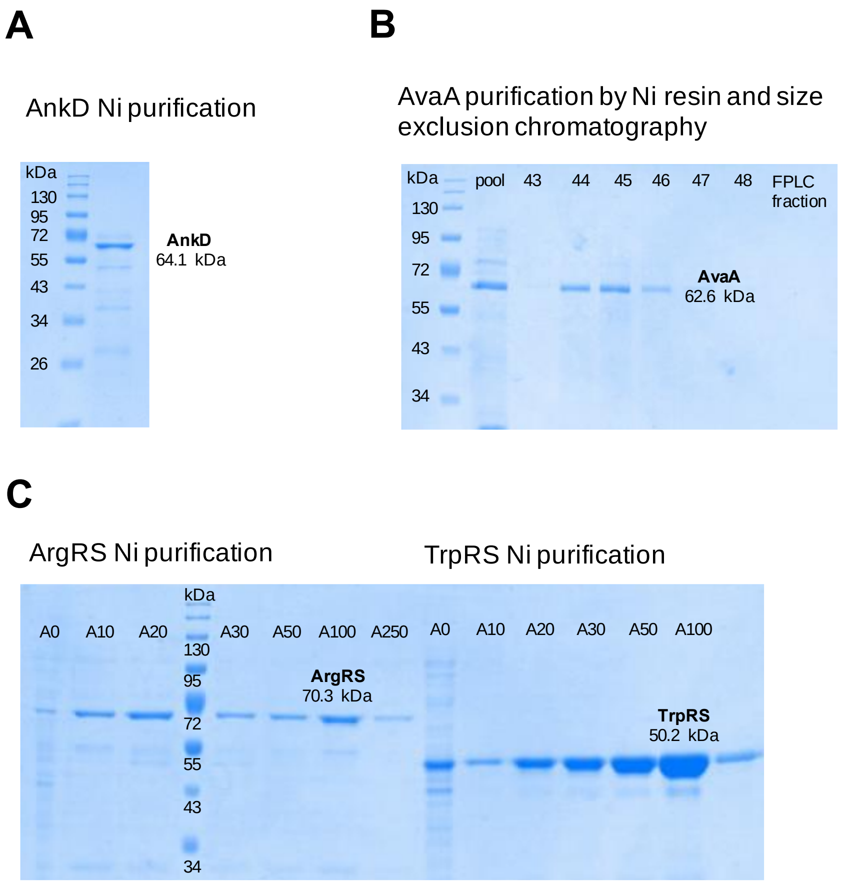


Figure S9. SDS-PAGE (12%) analysis of (**A**) AnkD, (**B**) AvaA, (**C**) Yeast ArgRS and TrpRS expressed and purified from *E. coli* BL21(DE3).

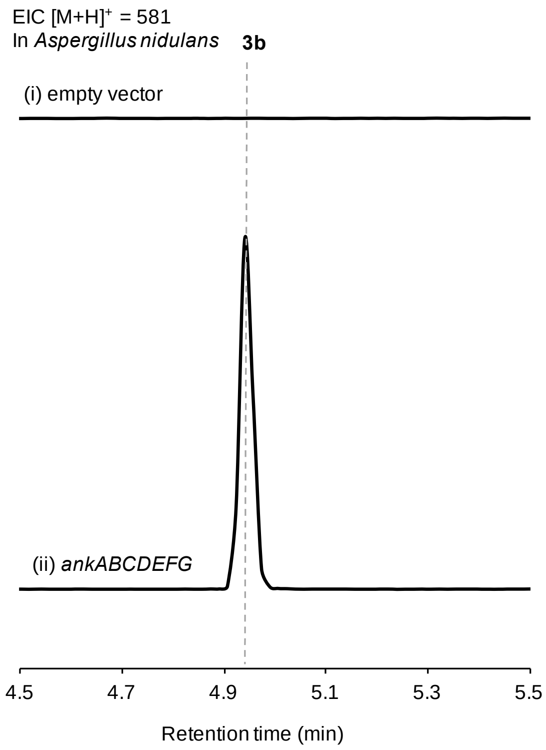


Figure S10. QTOF analysis of extracts from expression of *ankA-G* in *A. nidulans* showing production of compound **3b** (MWT = 580) compared to the empty vector control, retention time 4.5-5.5 min. Ion-extracted QTOF traces ($[M+H]^+ = 581$) are shown.

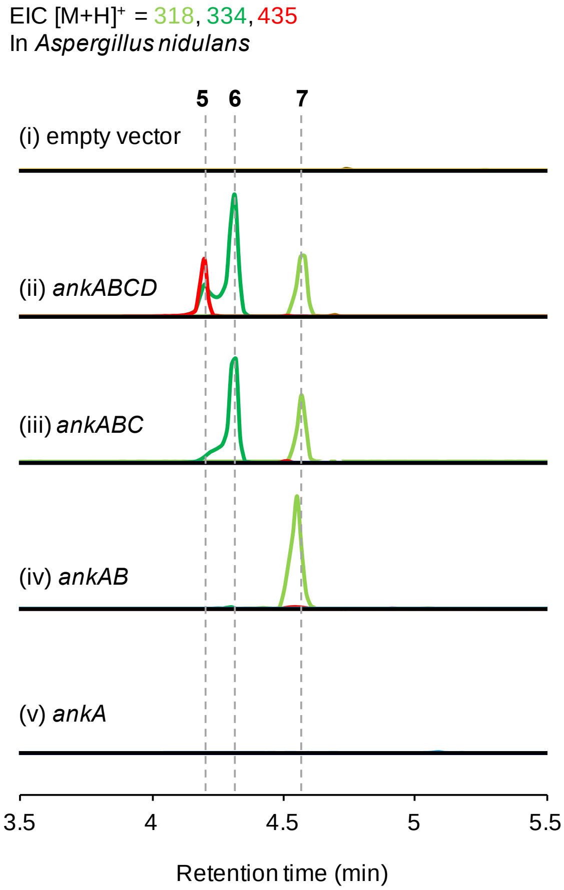


Figure S11. QTOF analysis of extracts from the expression of *ankA-D* in *A. nidulans*. Traces include *ankABCD* (ii), *ankABC* (iii), *ankAB* (iv), *ankA* (v), and the empty vector control (i), retention time 4.5-5.5 min. Ion-extracted traces correspond to the $[M+H]^+$ for **5** ($[M+H]^+ = 435$), **6** ($[M+H]^+ = 334$), and **7** ($[M+H]^+ = 318$).

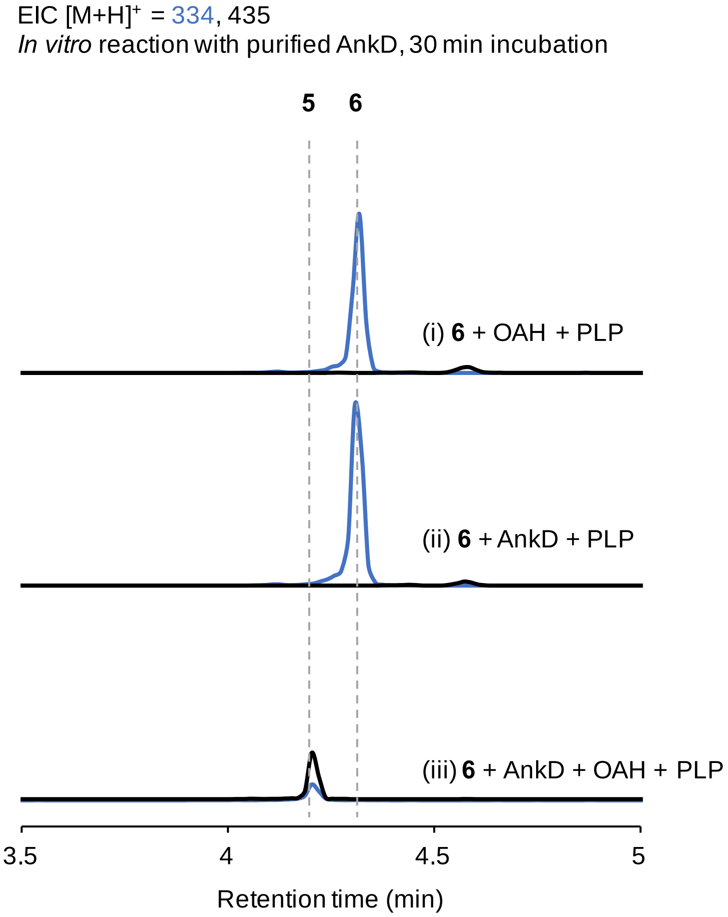


Figure S12. QTOF analysis of the extracts from AnkD *in vitro* reactions. Reactions include **6** and (i) O-acetyl-homoserine (OAH) and PLP, (ii) purified AnkD and PLP, and (iii) purified AnkD, OAH, and PLP. Ion-extracted traces correspond to the $[M+H]^+$ for **6** ($[M+H]^+ = 334$) and **5** ($[M+H]^+ = 435$). The reactions were incubated at room temperature for 30 minutes.

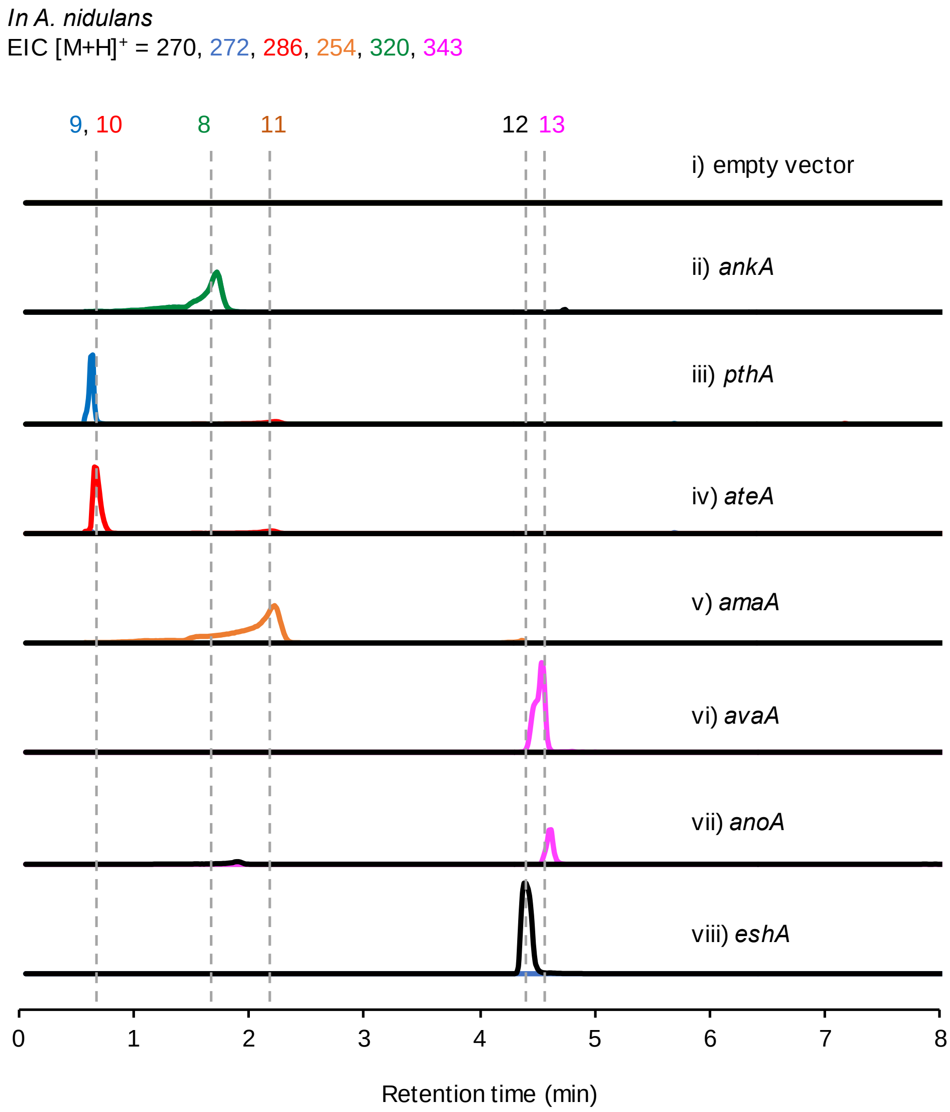


Figure S13. QTOF analysis of extracts from expression of *ankA*, *pthA*, *ateA*, *amaA*, *avaA*, *anoA*, and *eshA* in *A. nidulans* (ii-viii) compared to the empty vector control (i). Ion-extracted traces correspond to the $[M+H]^+$ for **8** ($[M+H]^+$ = 320), **9** ($[M+H]^+$ = 272), **10** ($[M+H]^+$ = 286), **11** ($[M+H]^+$ = 254), **12** ($[M+H]^+$ = 270), and **13** ($[M+H]^+$ = 343).

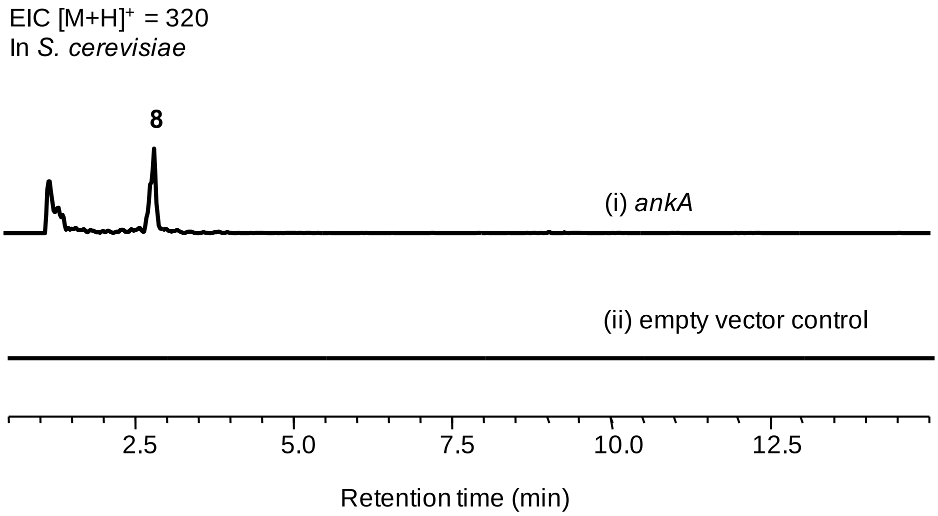


Figure S14. LC-MS analysis of extracts from expression of AnkA in yeast (i) compared to the empty vector control (ii). Ion-extracted traces correspond to the $[M+H]^+$ for **8** ($[M+H]^+$ = 320).

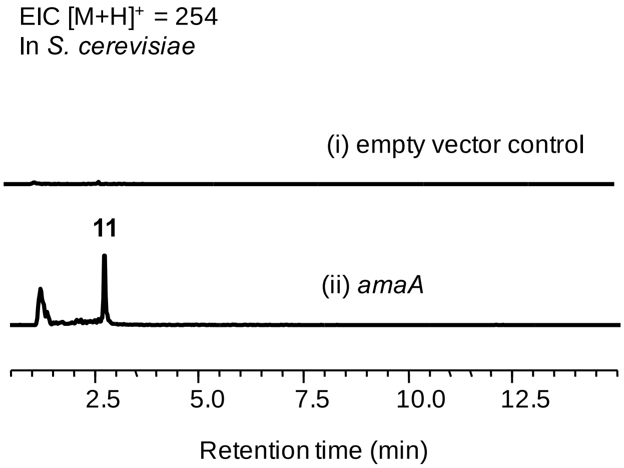


Figure S15. LC-MS analysis of extracts from expression of AmaA in yeast (ii) compared to the empty vector control (i). Ion-extracted traces correspond to the $[M+H]^+$ for **11** ($[M+H]^+$ = 254).

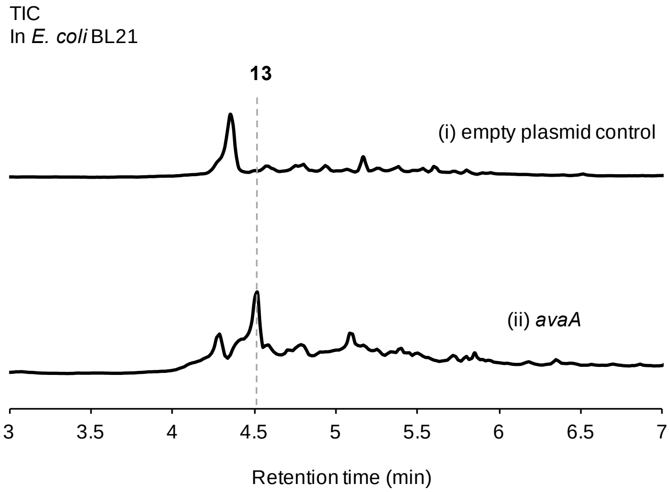


Figure S16. QTOF analysis of the extracts from expression of *avaA* in *E. coli* BL21 (ii) compared to the empty plasmid control (i). The total ion count (TIC) trace is shown.

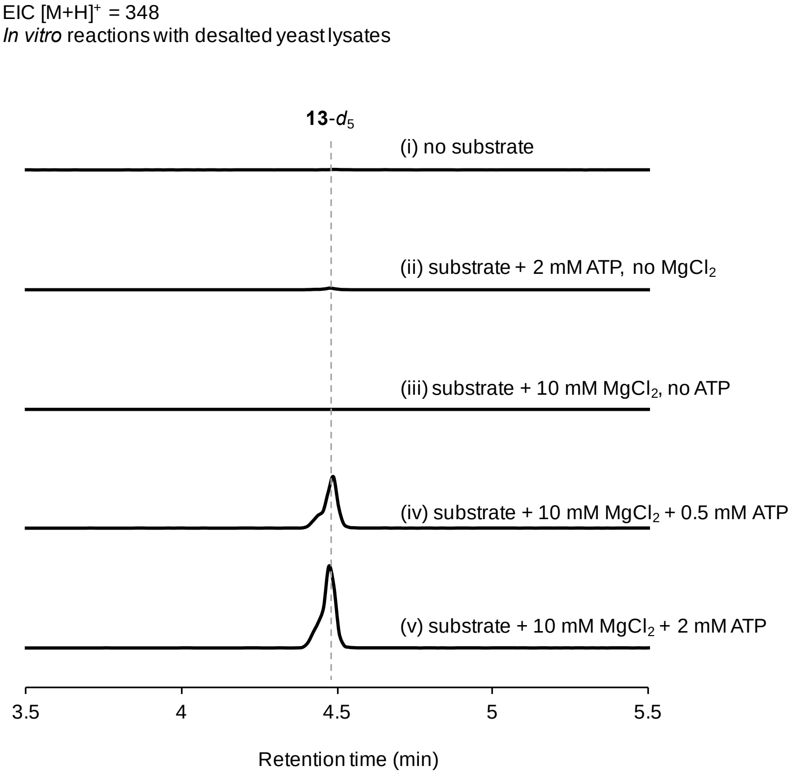


Figure S17. QTOF analysis of the extracts from in vitro reactions with desalted yeast lysates with (i) no substrate or cofactors, (ii) 5 mM L-Arg, 5 mM Trp-d_5, 2 mM ATP, no $MgCl_2$, (iii) 5 mM L-Arg, 5 mM Trp-d_5, 10 mM $MgCl_2$, no ATP, (iv) 5 mM L-Arg, 5 mM Trp-d_5, 0.5 mM ATP, 10 mM $MgCl_2$, and (v) 5 mM L-Arg, 5 mM Trp-d_5, 2 mM ATP, 10 mM $MgCl_2$. Ion-extracted traces correspond to the $[M+H]^+$ for **13** with incorporation of Trp-d_5 ($[M+H]^+ = 348$).

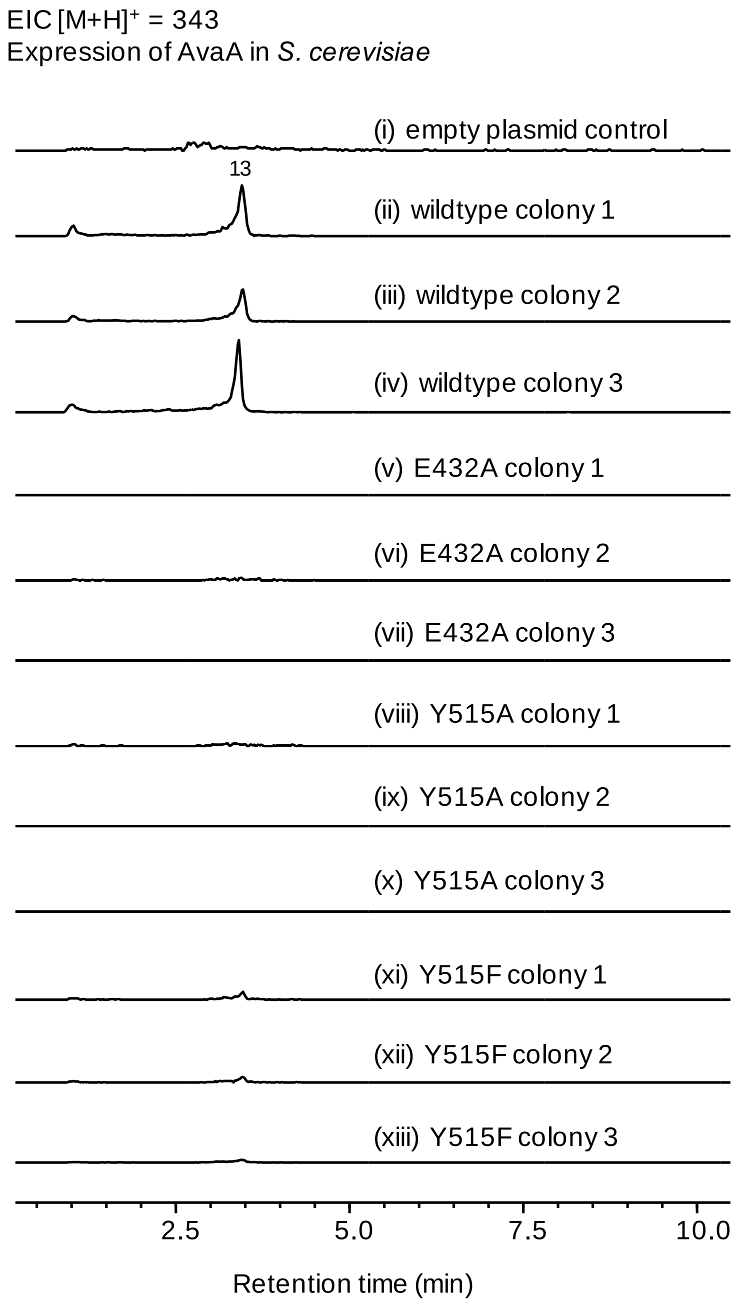
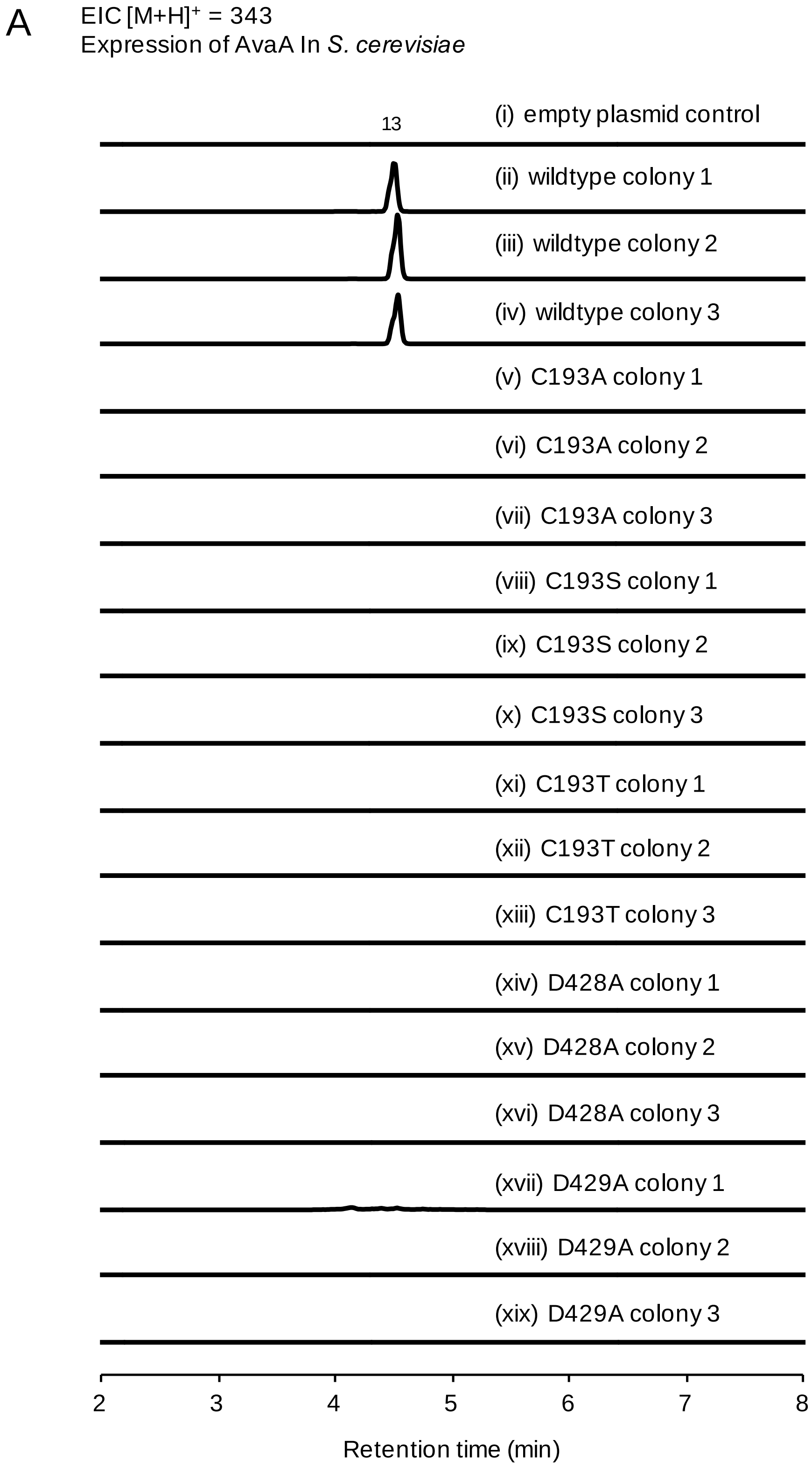


Figure S18. LC-MS analysis of the extracts from yeast expression of AvaA mutants E432A, Y515A, and Y515F (v-xiii) compared to the empty vector control (i) and wildtype AvaA (ii-iv) in biological triplicate. Ion-extracted traces correspond to the $[M+H]^+$ for **13** ($[M+H]^+ = 343$).


A
EIC [M+H]+ = 343
Expression of AvaA In S. cerevisiae
13
(i) empty plasmid control
(ii) wildtype colony 1
(iii) wildtype colony 2
(iv) wildtype colony 3
(v) C193A colony 1
(vi) C193A colony 2
(vii) C193A colony 3
(viii) C193S colony 1
(ix) C193S colony 2
(x) C193S colony 3
(xi) C193T colony 1
(xii) C193T colony 2
(xiii) C193T colony 3
(xiv) D428A colony 1
(xv) D428A colony 2
(xvi) D428A colony 3
(xvii) D429A colony 1
(xviii) D429A colony 2
(xix) D429A colony 3
2
3
4
5
6
7
8
Retention time (min)

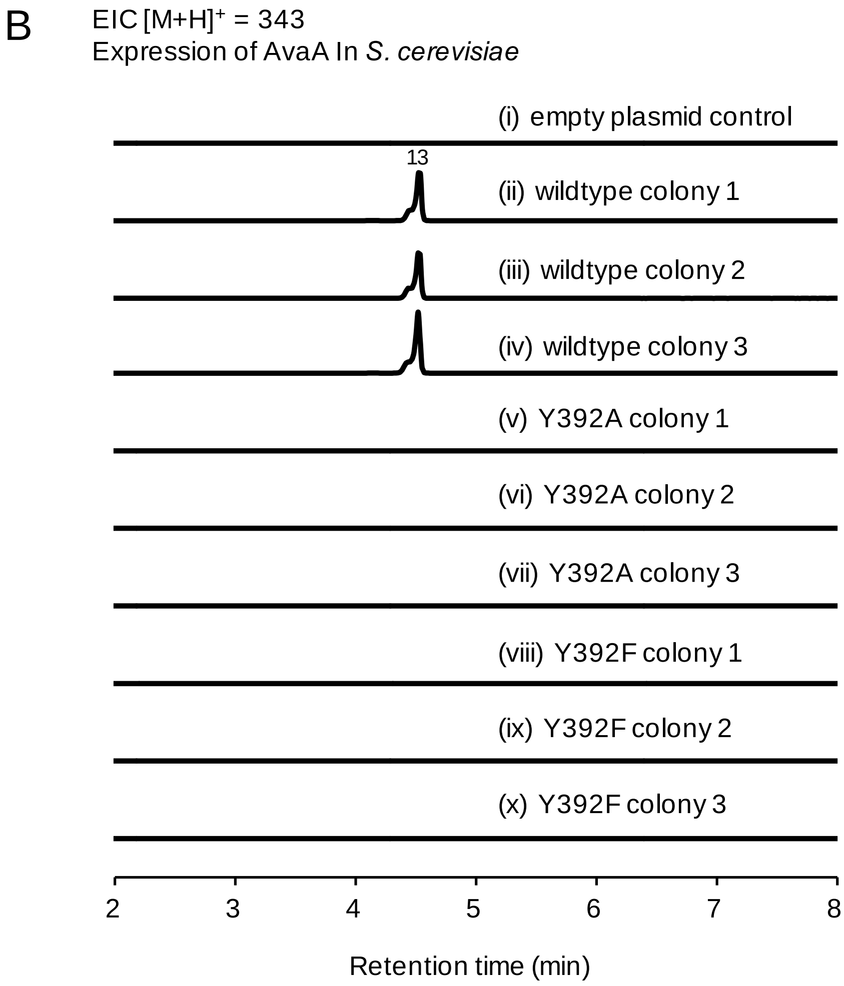


Figure S19. QTOF analysis of the extracts from yeast expression of AvaA mutants C193A, C193S, C193T, D428A, D429A, Y392A, and Y392F (**A**: v-xvi, **B**: v-x) compared to the empty vector control (i) and wild type AvaA (ii-iv) in biological triplicate. Ion-extracted traces correspond to the $[M+H]^+$ for **13** ($[M+H]^+ = 343$).

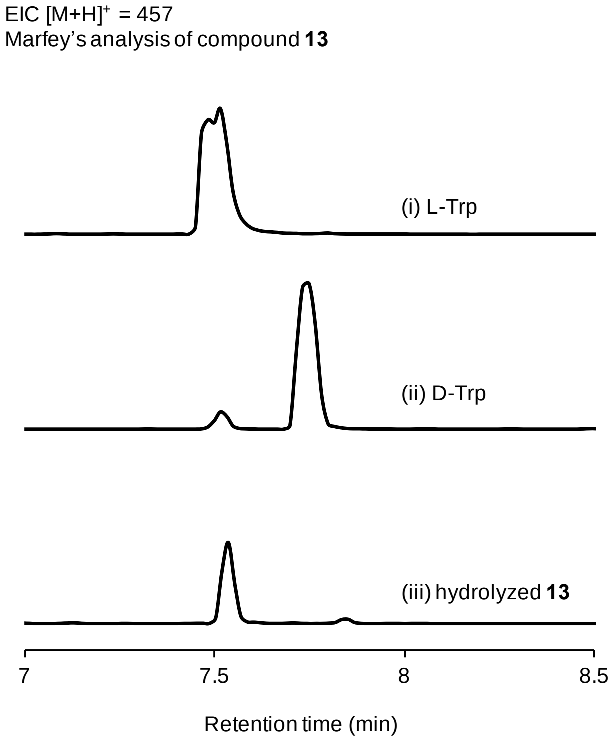


Figure S20. Marfey's analysis to determine the stereochemistry of Trp in compound **13**. Marfey's reagent was used to derivatize (i) L-Trp, (ii) D-Trp, and (iii) hydrolyzed **13**, and the reaction products were analyzed by QTOF, indicating **13** contains L-Trp. Ion-extracted traces correspond to the $[M+H]^+$ for derivatized Trp ($[M+H]^+ = 457$) as reported in ref [11].

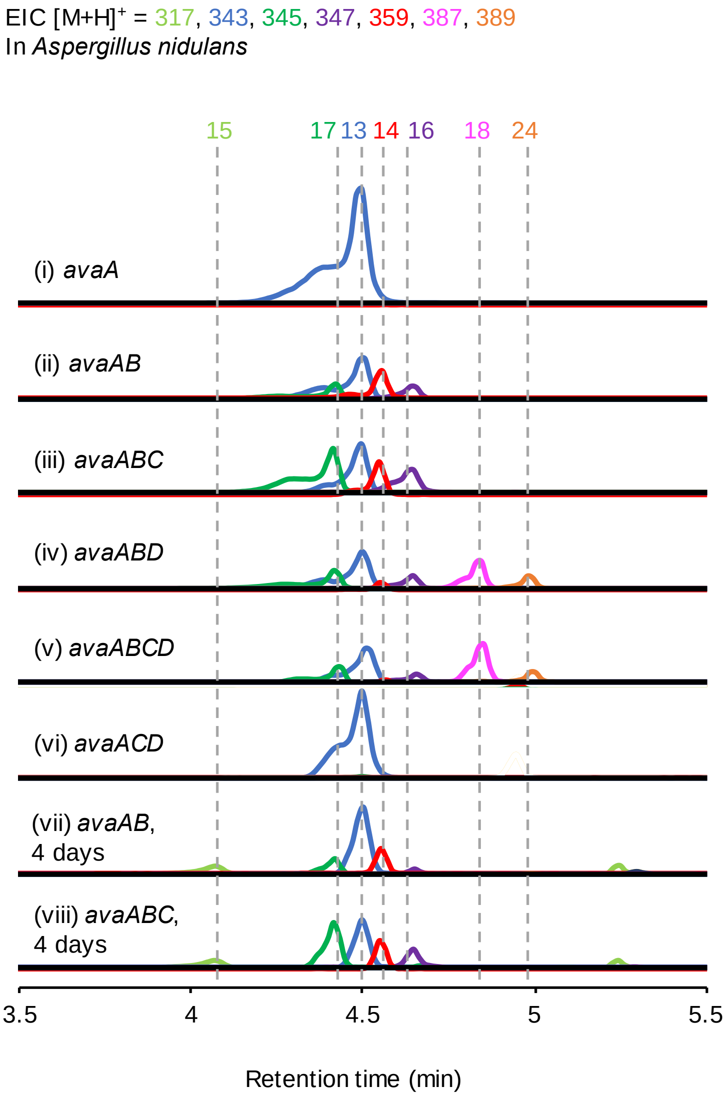


Figure S21. QTOF analysis of extracts from expression of *ava* pathway genes in *A. nidulans*. Ion-extracted traces correspond to the $[M+H]^+$ for **13** ($[M+H]^+ = 343$), **14** ($[M+H]^+ = 359$), **15** ($[M+H]^+ = 317$), **16** ($[M+H]^+ = 347$), **17** ($[M+H]^+ = 345$), **18** ($[M+H]^+ = 387$), and **24** ($[M+H]^+ = 389$). Structures are shown in Fig. S7.

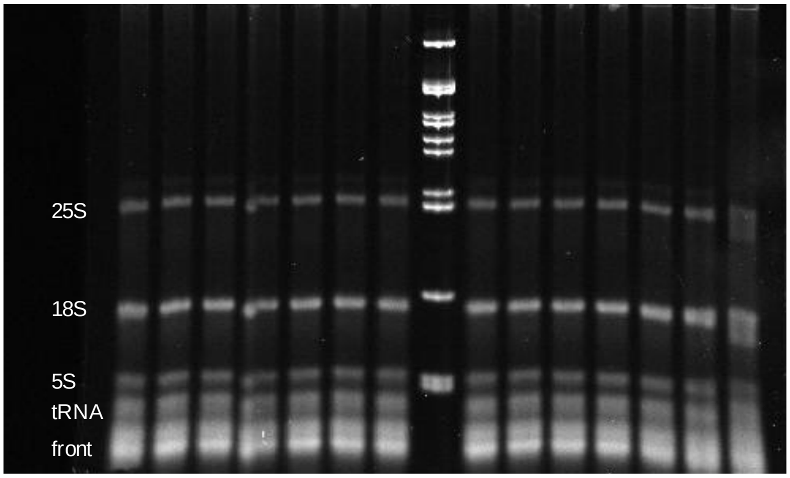


Figure S22. RNA-PAGE purification of total yeast RNA for isolation of tRNA.

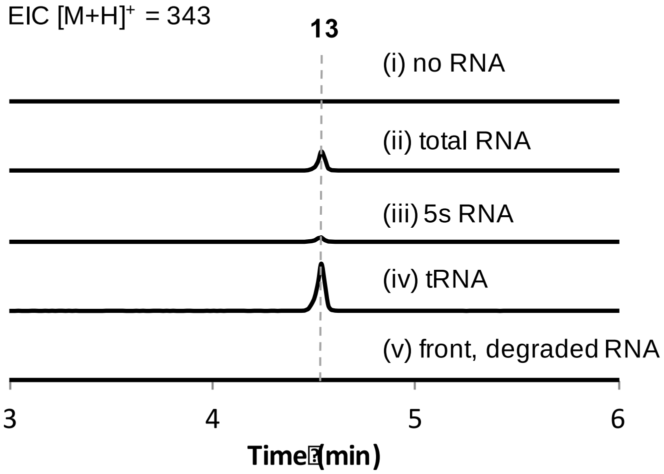
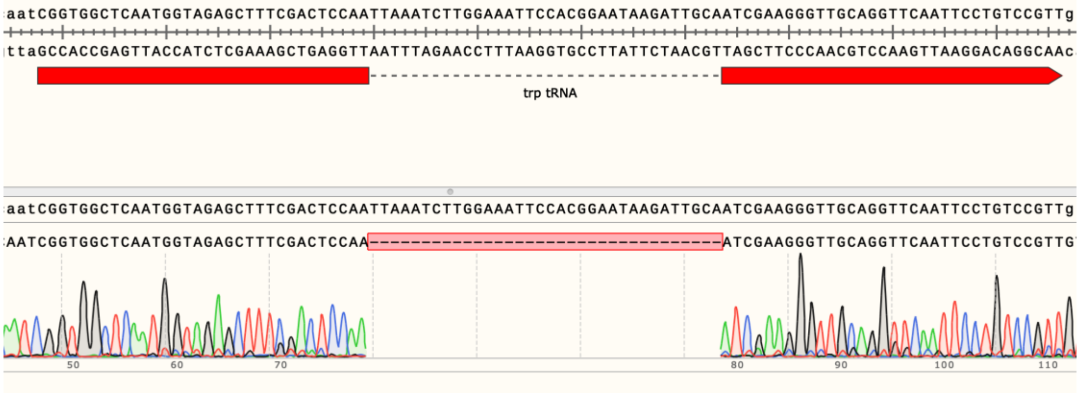


Figure S23. *In vitro* assays with purified AvaA, ArgRS, TrpRS, unlabeled substrates, and cofactors with various RNA fractions, including (i) no RNA, (ii) total yeast RNA, (iii) purified 5s RNA, (iv) purified tRNA, and (v) the degraded RNA front.



```
aatCGGTGGCTCAATGGTAGAGCTTTCGACTCCAATTAAATCTTGGAAATTCCACGGAATAAGATTGCAATCGAAGGGTTGCAGGTTCAATTCCTGTCCGTTg
ttaGCCACCGAGTTACCATCTCGAAAGCTGAGGTTAATTTAGAACCTTTAAGGTGCCTTATTCTAACGTTAGCTTCCCAACGTCCAAGTTAAGGACAGGCAAc
```

```
aatCGGTGGCTCAATGGTAGAGCTTTCGACTCCAATTAAATCTTGGAAATTCCACGGAATAAGATTGCAATCGAAGGGTTGCAGGTTCAATTCCTGTCCGTTg
AATCGGTGGCTCAATGGTAGAGCTTTCGACTCCAA----------------------------------ATCGAAGGGTTGCAGGTTCAATTCCTGTCCGTTG
```

Figure S24. Sanger sequencing of RT-PCR product of tRNA$^{\mathrm{Trp}}$ from the isolated tRNA band. The sequencing results showed the intron was removed, indicating the transcript was not amplified from genomic DNA contamination.

Figure S25. ^{1}H NMR spectrum of compound **1** in D_2O 0.5% TFA-*d* (500 MHz).

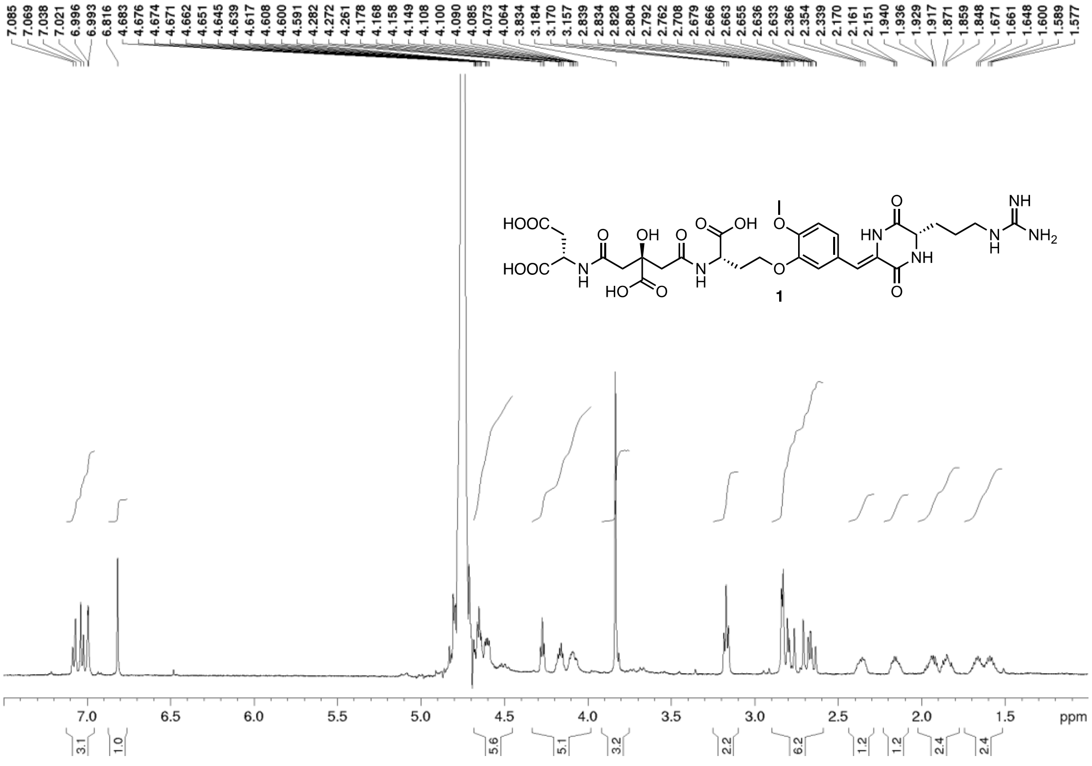


Figure S26. ^{13}C NMR spectrum of compound **1** in D_2O 0.5% TFA-*d* (125 MHz).

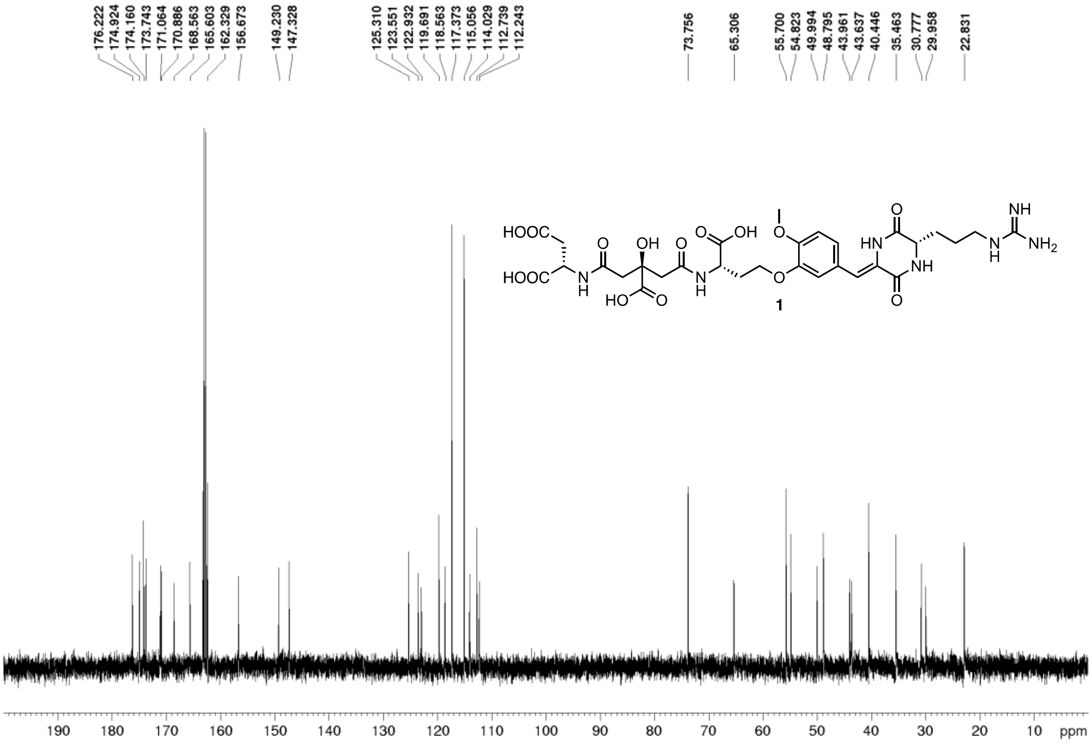


Figure S27. ^{1}H-^{1}H COSY spectrum of compound **1** in D_2O 0.5% TFA-*d* (500 MHz).

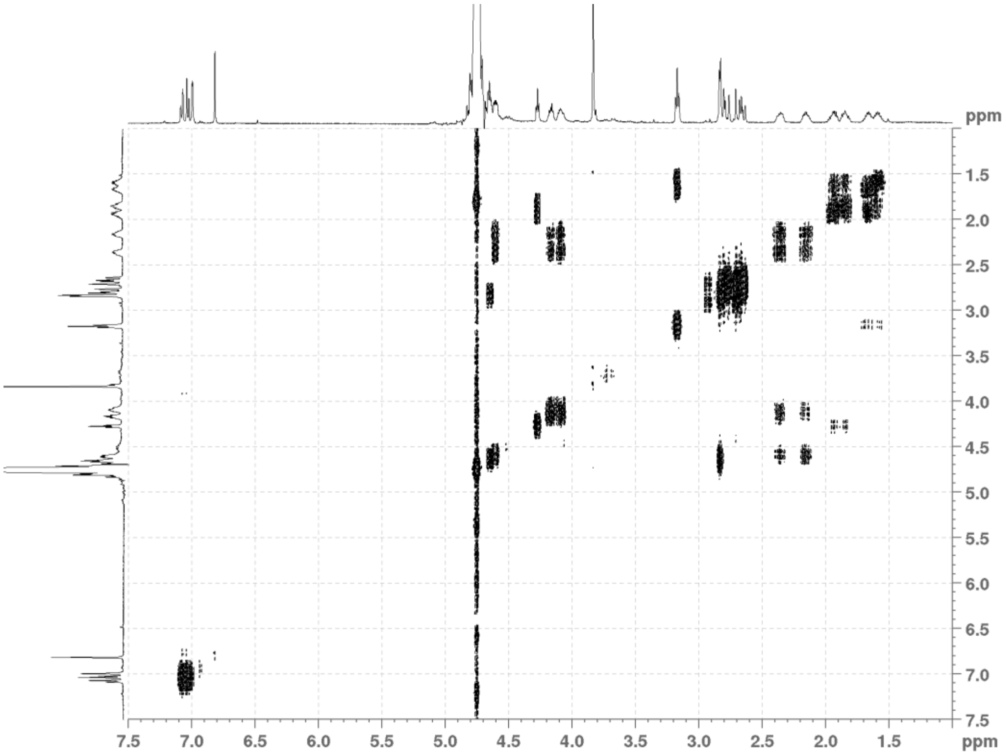


Figure S28. HSQC spectrum of compound **1** in D_2O 0.5% TFA-*d* (500 MHz).

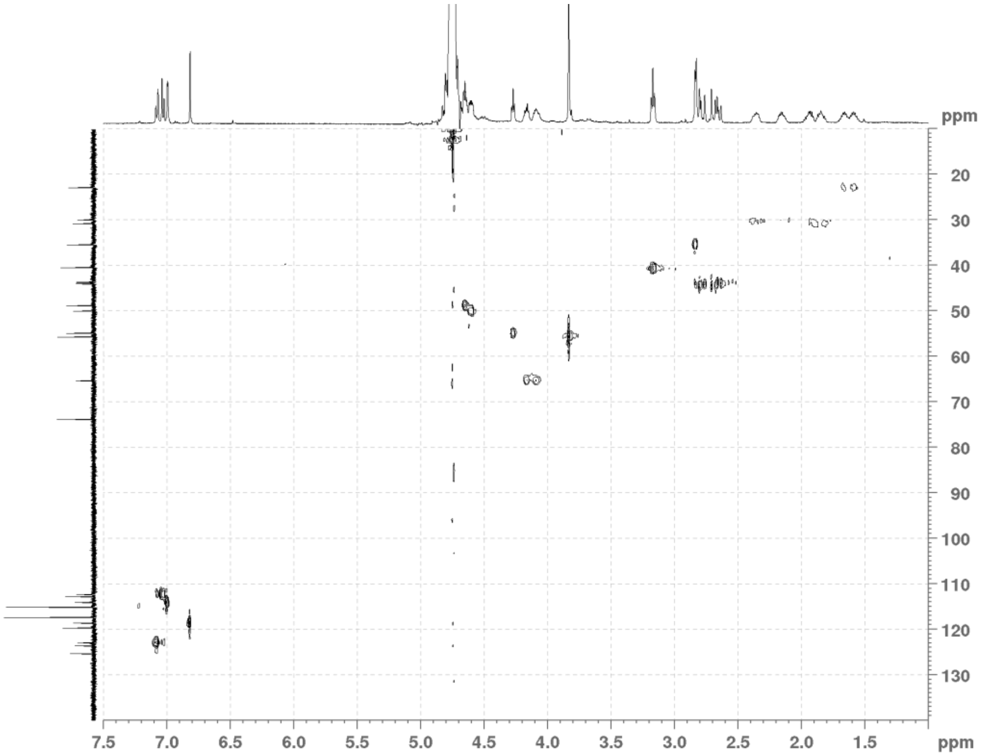


Figure S29. HMBC spectrum of compound **1** in D_2O 0.5% TFA-*d* (500 MHz).

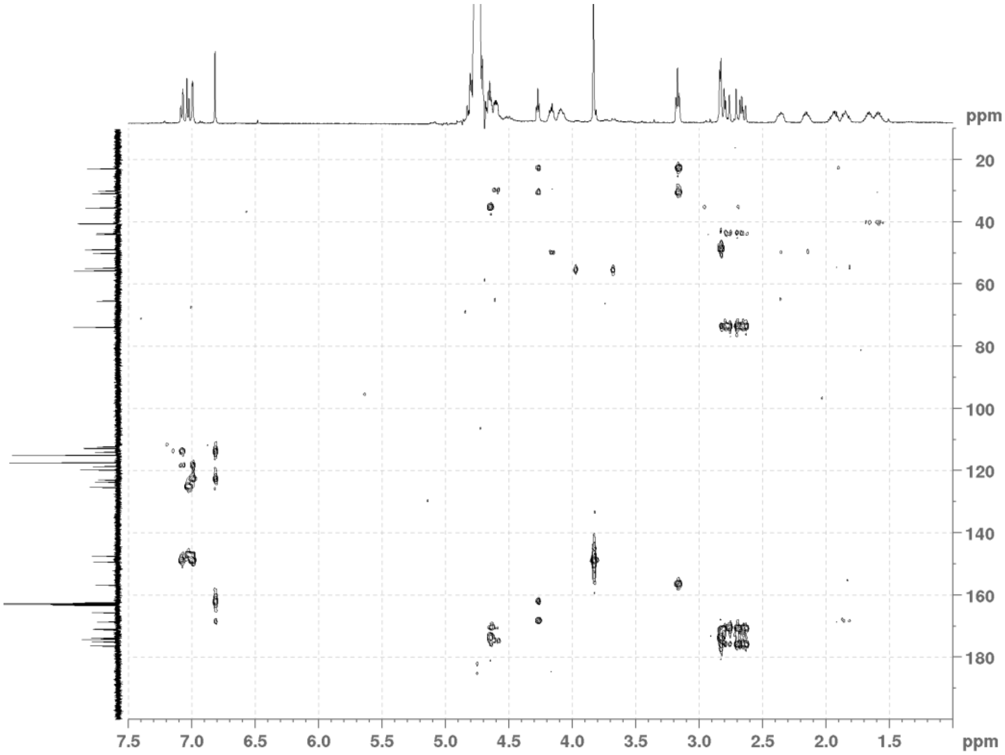


Figure S30. ^{1}H NMR spectrum of compound **2** in DMSO-d_6 0.5% TFA-d (500 MHz).

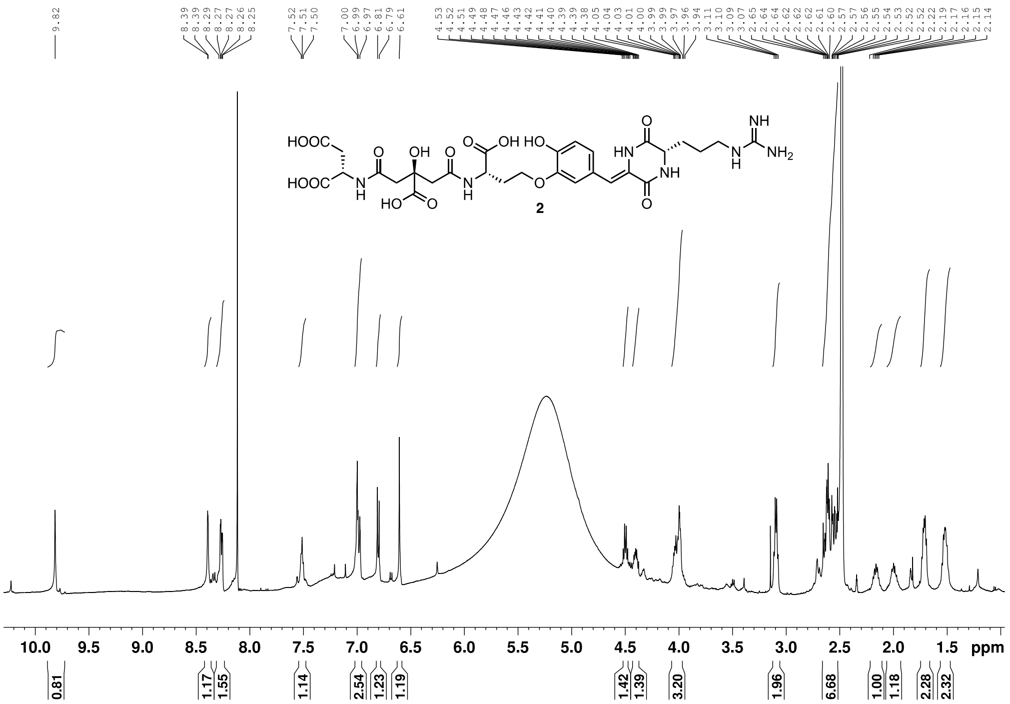


Figure S31. ^{13}C NMR spectrum of compound **2** in DMSO-d_6 0.5% TFA-d (125 MHz).

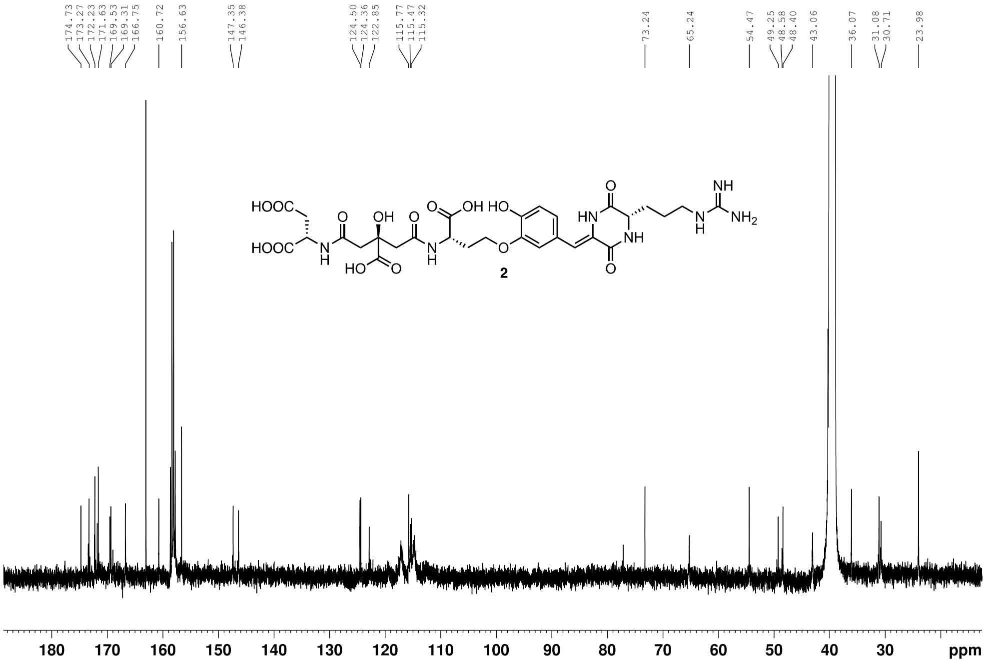


Figure S32. HSQC spectrum of compound **2** in DMSO-d_6 0.5% TFA-d (500 MHz).

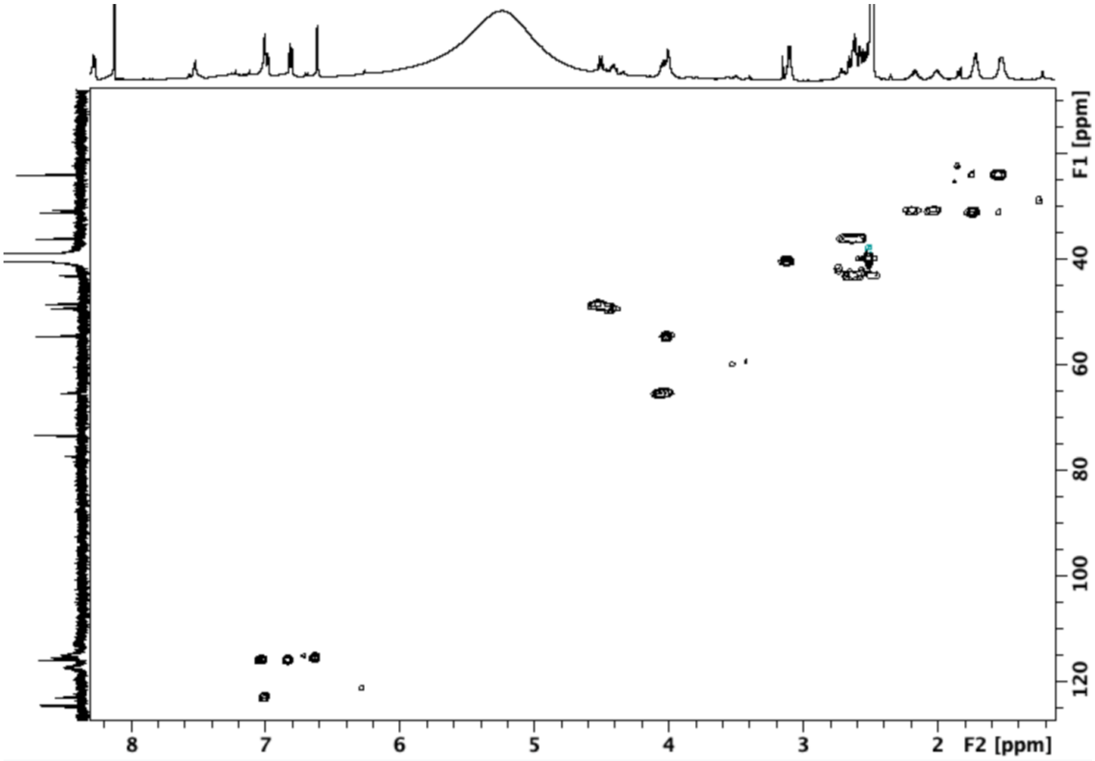


Figure S33. ^{1}H NMR spectrum of compound **3** in D_2O 0.5% TFA-*d* (500 MHz).

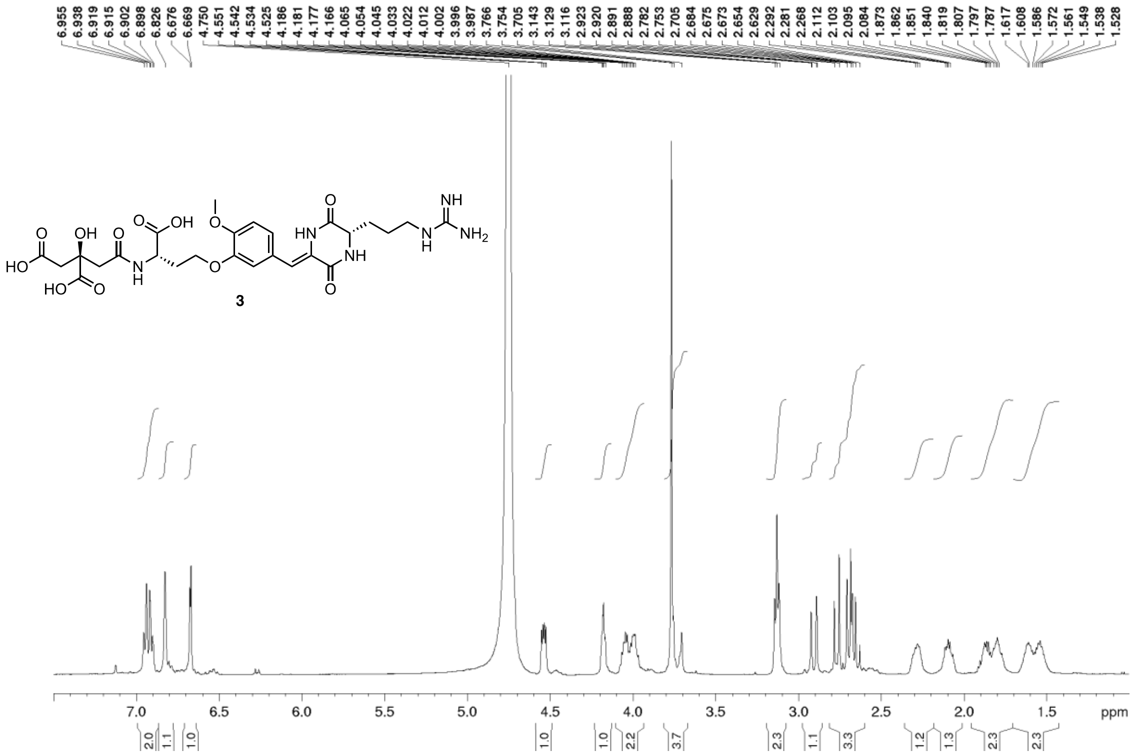


Figure S34. ^{13}C NMR spectrum of compound **3** in D_2O 0.5% TFA-*d* (125 MHz).

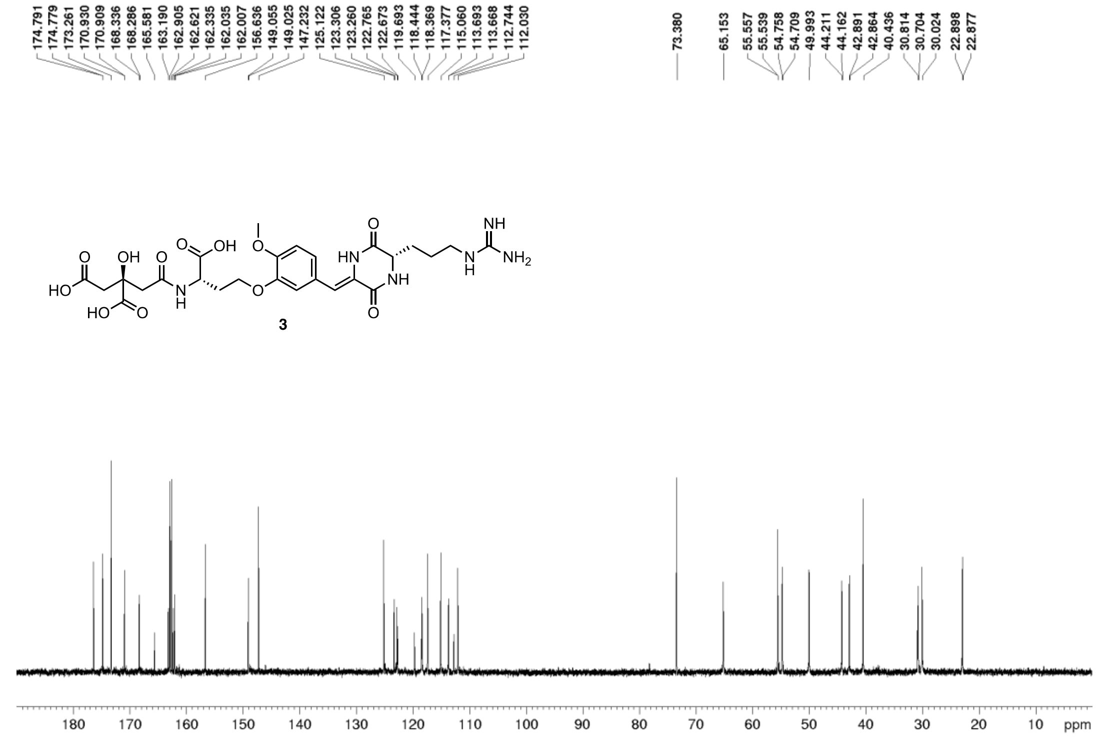


Figure S35. ^1H-^1H COSY spectrum of compound **3** in D_2O 0.5% TFA-*d* (500 MHz).

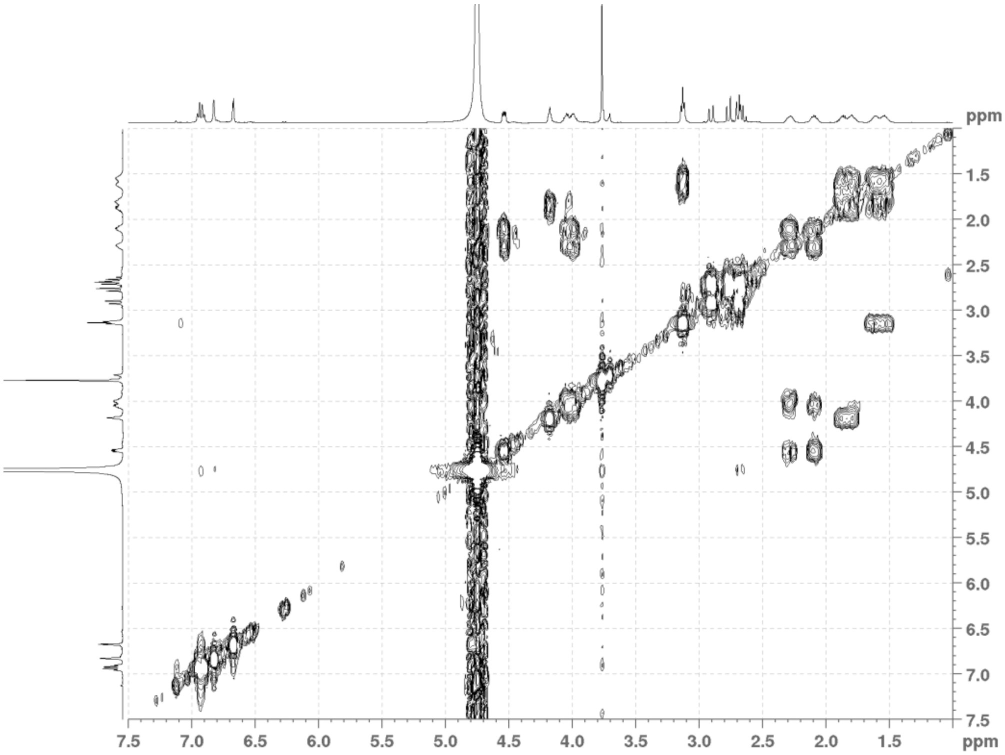


Figure S36. HSQC spectrum of compound **3** in D_2O 0.5% TFA-*d* (500 MHz).

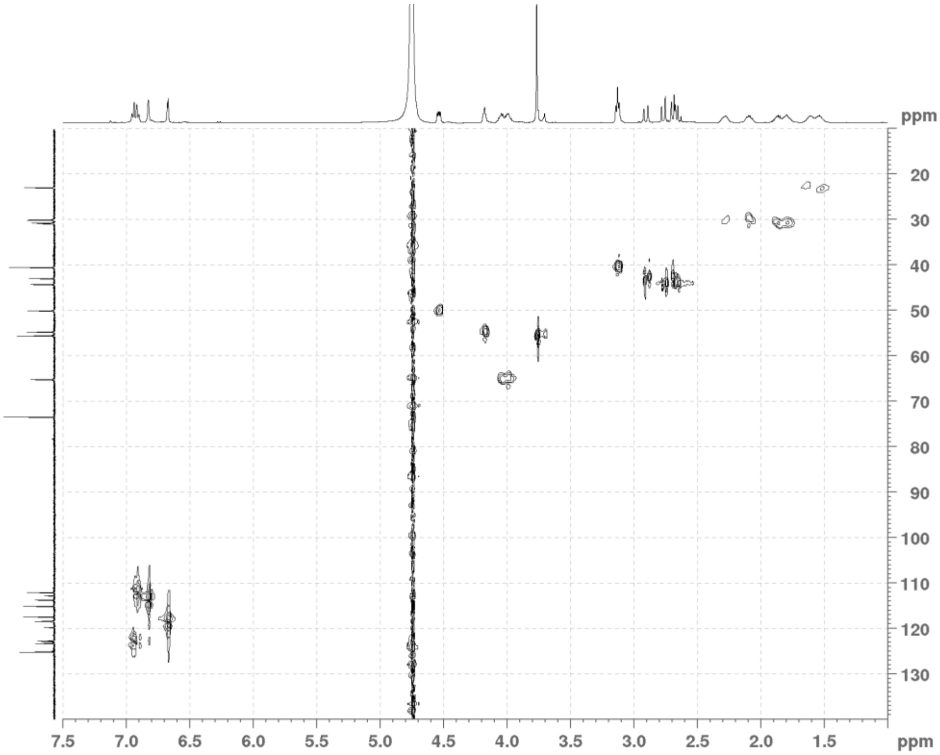


Figure S37. HMBC spectrum of compound **3** in D_2O0.5% TFA-*d* (500 MHz).

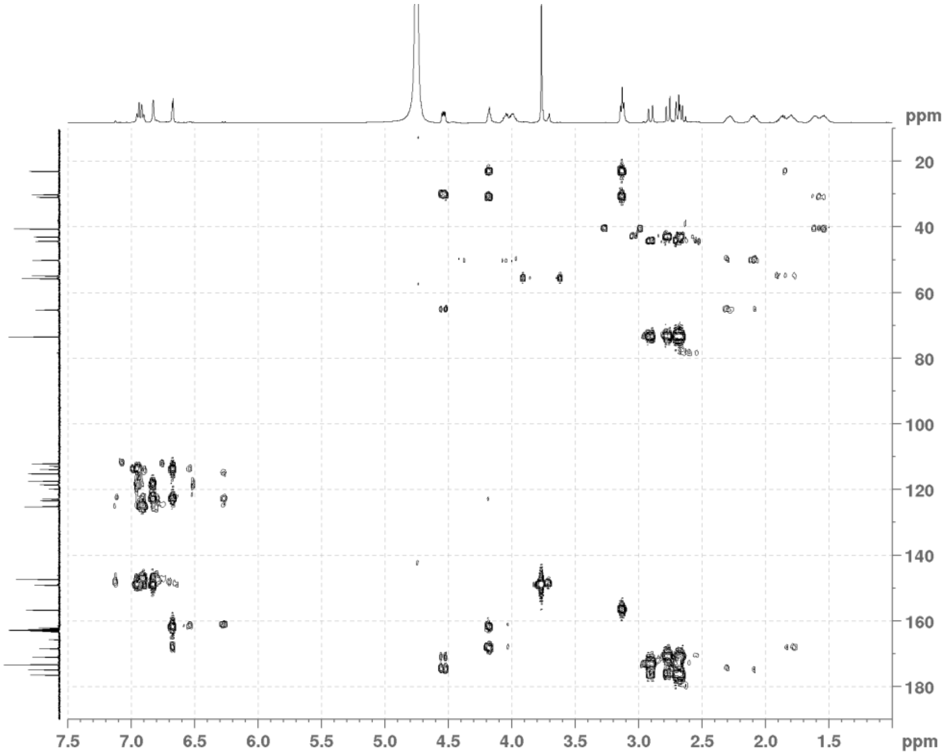


Figure S38. ^{1}H NMR spectrum of compound **3b** in D_2O 0.5% TFA-*d* (500 MHz).

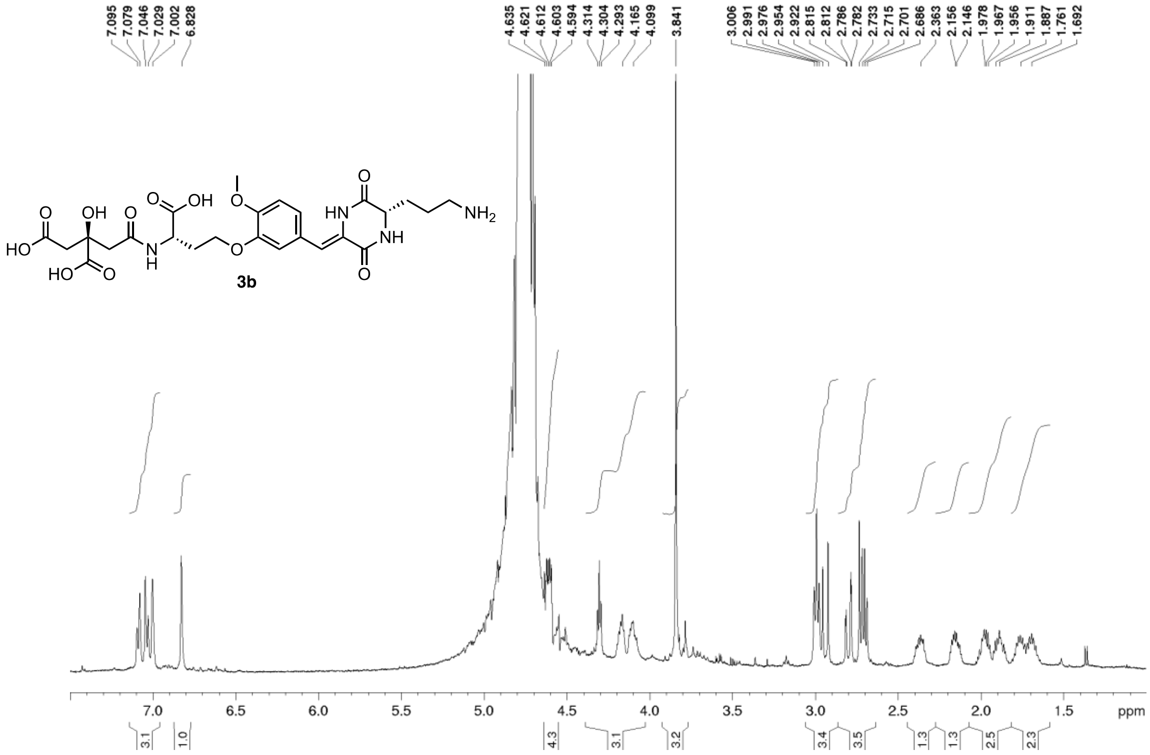


Figure S39. ^{13}C NMR spectrum of compound **3b** in D_2O 0.5% TFA-*d* (125 MHz).

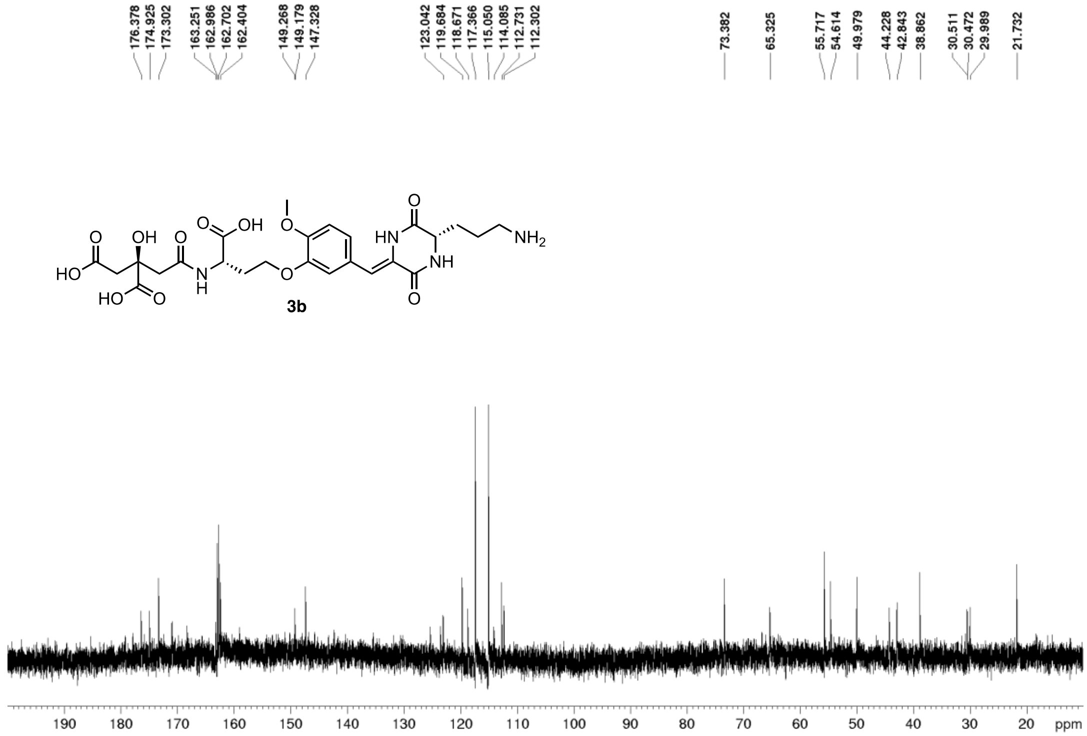


Figure S40. ^{1}H-^{1}H COSY spectrum of compound **3b** in D_2O 0.5% TFA-*d* (500 MHz).

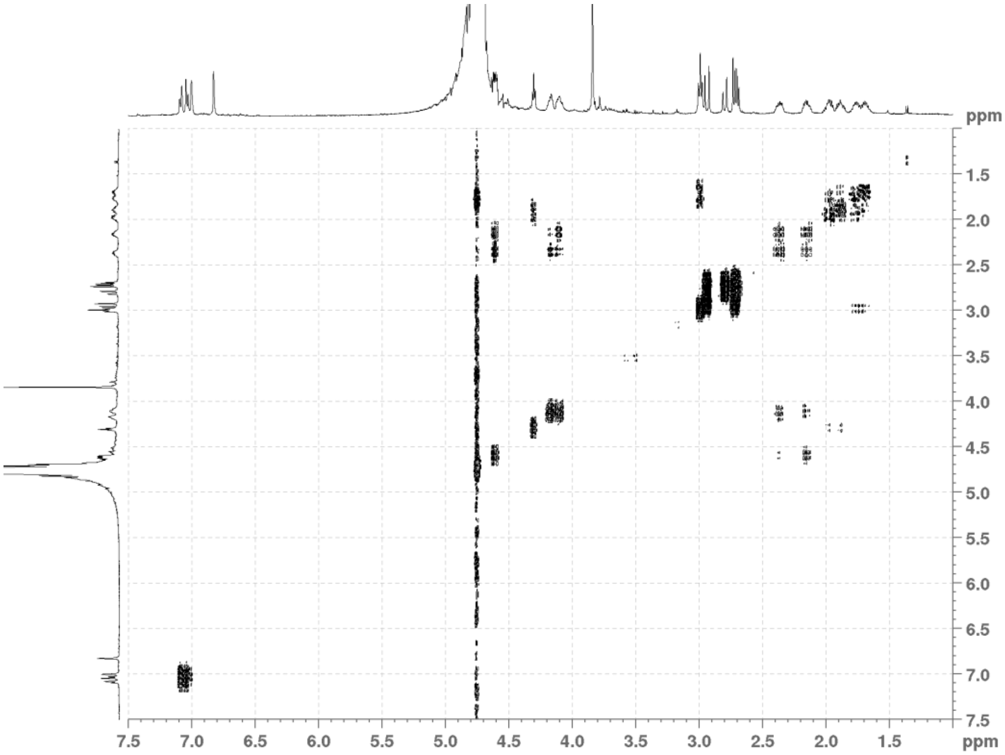


Figure S41. HSQC spectrum of compound **3b** in D_2O 0.5% TFA-*d* (500 MHz).

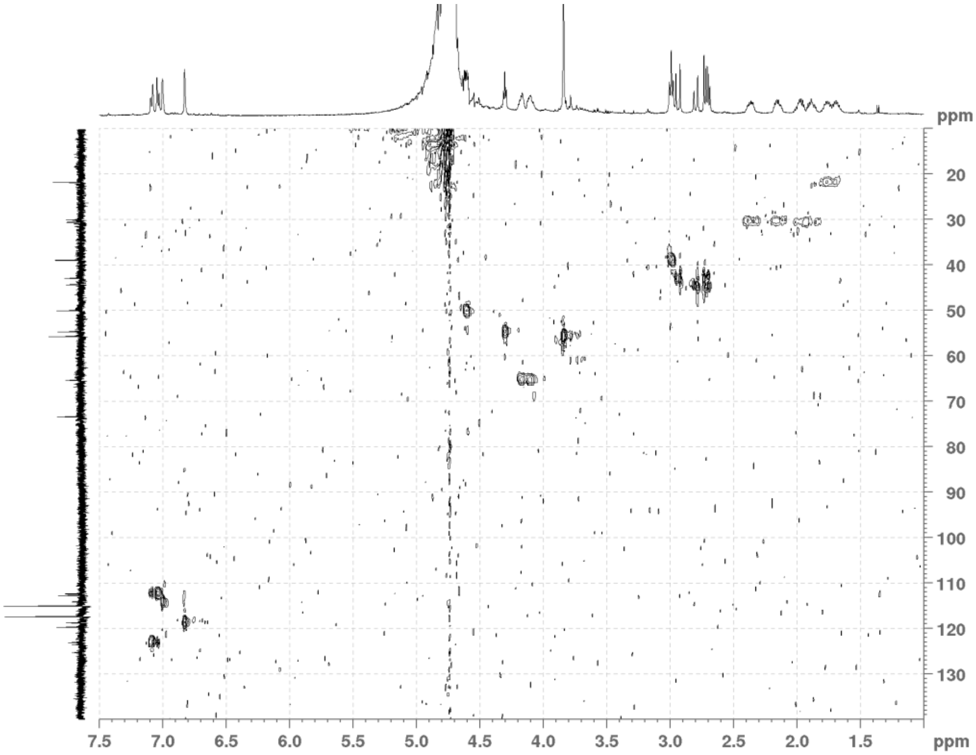


Figure S42. HMBC spectrum of compound **3b** in D_2O 0.5% TFA-*d* (500 MHz).

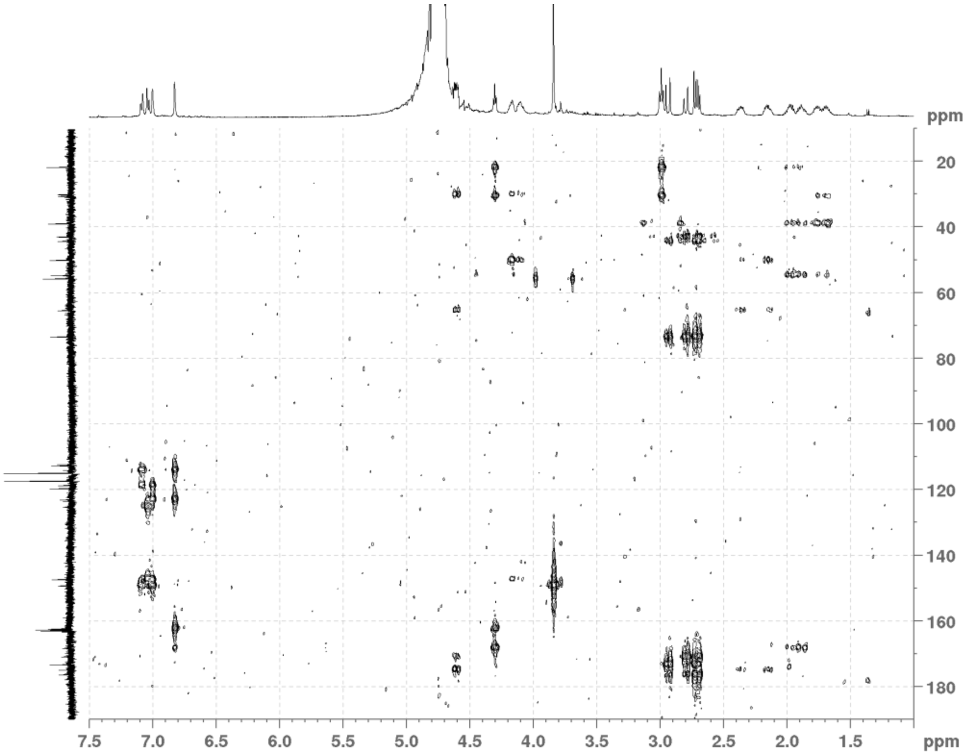


Figure S43. ^{1}H NMR spectrum of compound **4** in DMSO-d_6 0.5% TFA-d (500 MHz).

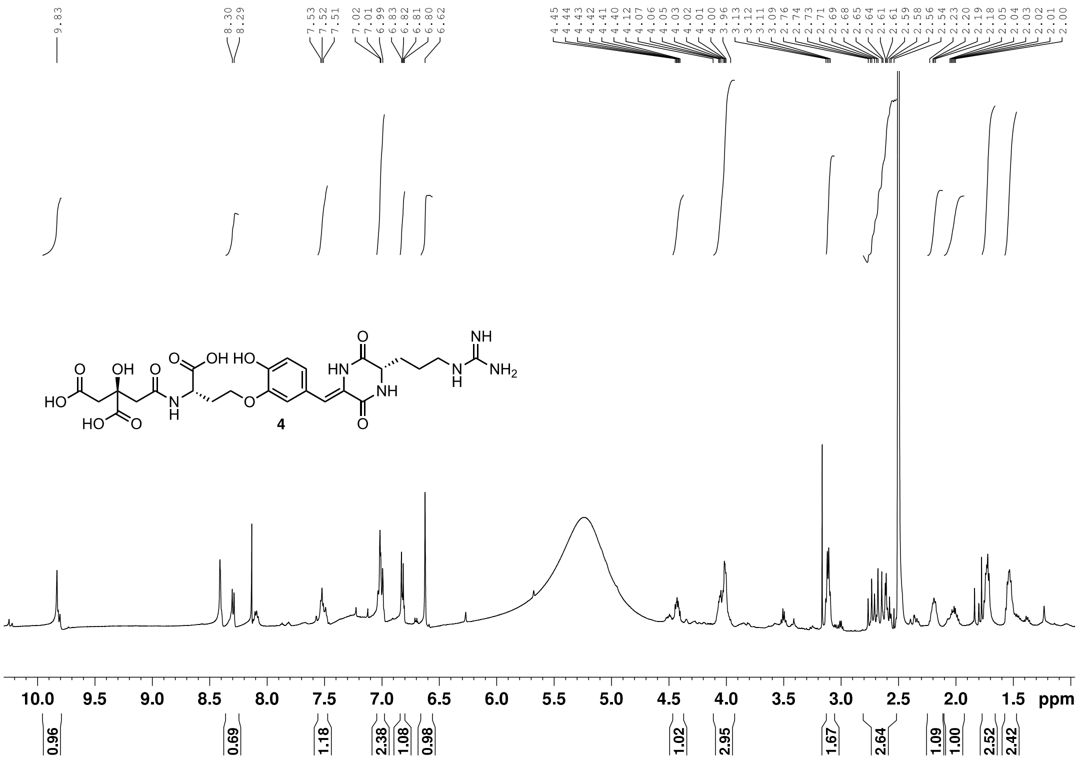


Figure S44. ^{13}C NMR spectrum of compound **4** in DMSO-d_6 0.5% TFA-d (125 MHz).

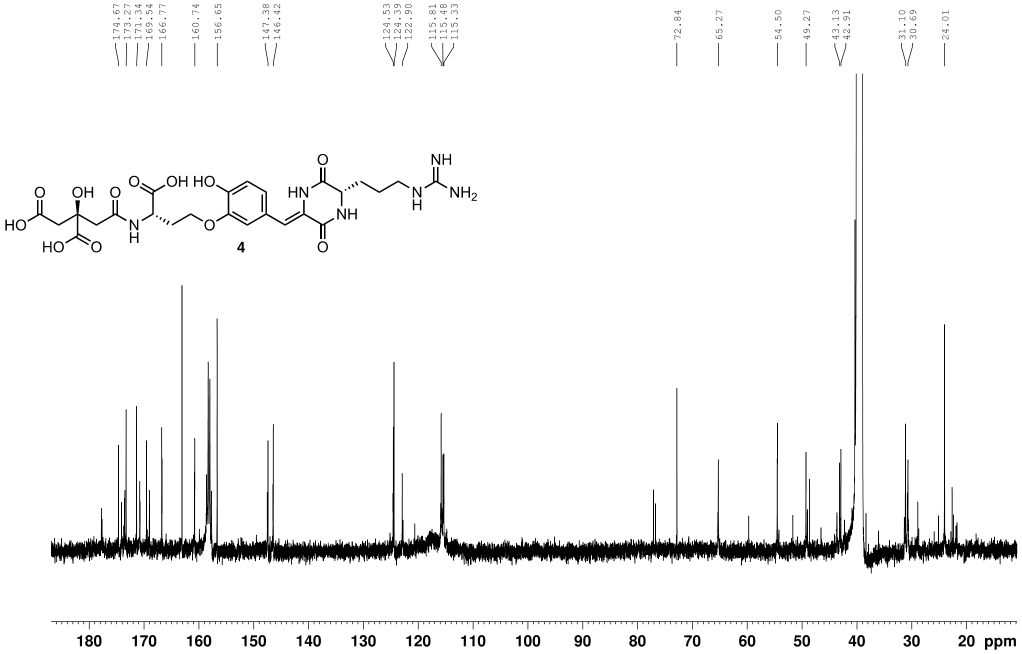


Figure S45. HSQC spectrum of compound **4** in DMSO-d_6 0.5% TFA-d (500 MHz).

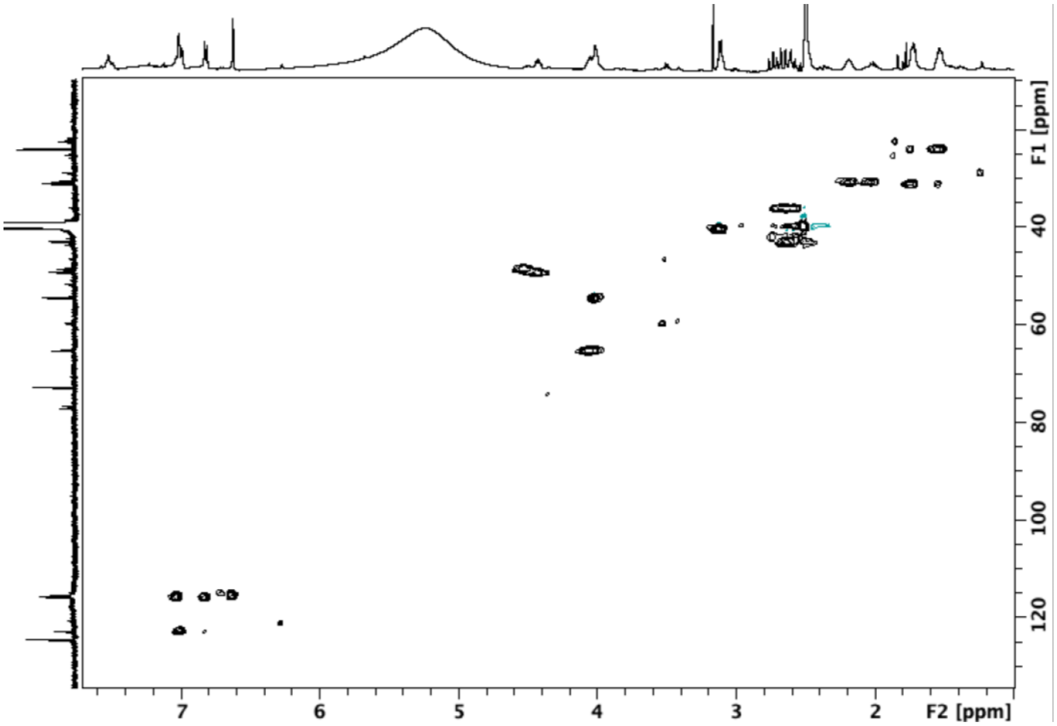


Figure S46. ^{1}H NMR spectrum of compound **5** in D_2O 0.5% TFA-*d* (500 MHz).

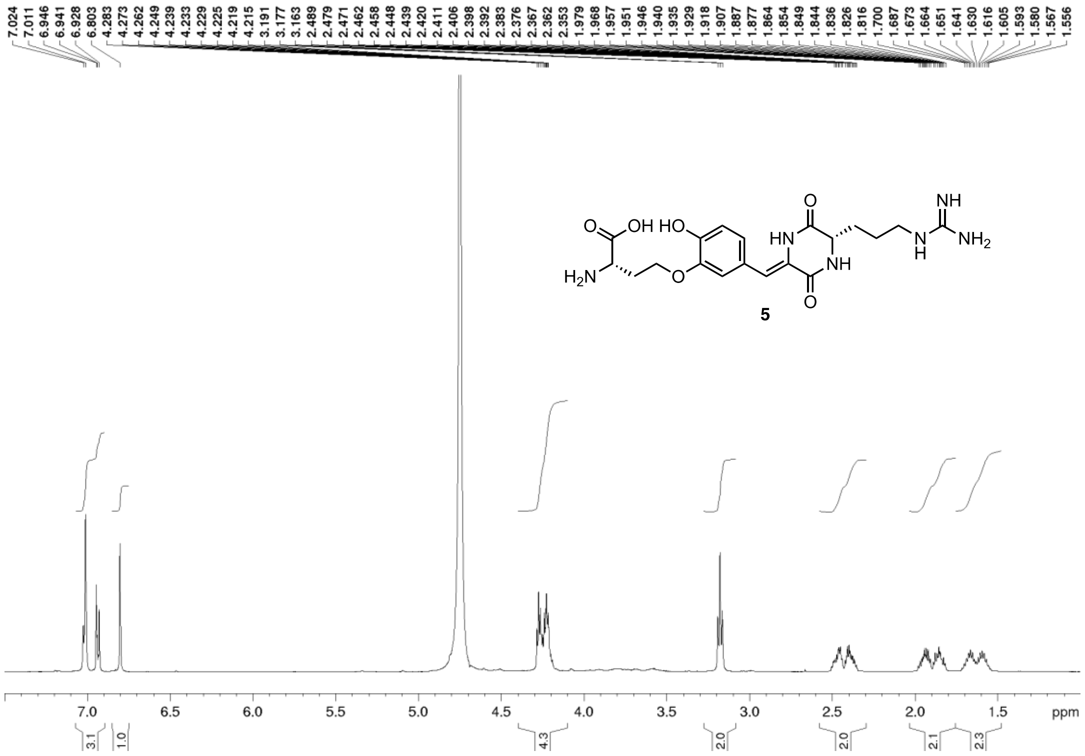


Figure S47. ^{1}H NMR spectrum of compound **5** in D_2O 0.5% TFA-*d* (125 MHz).

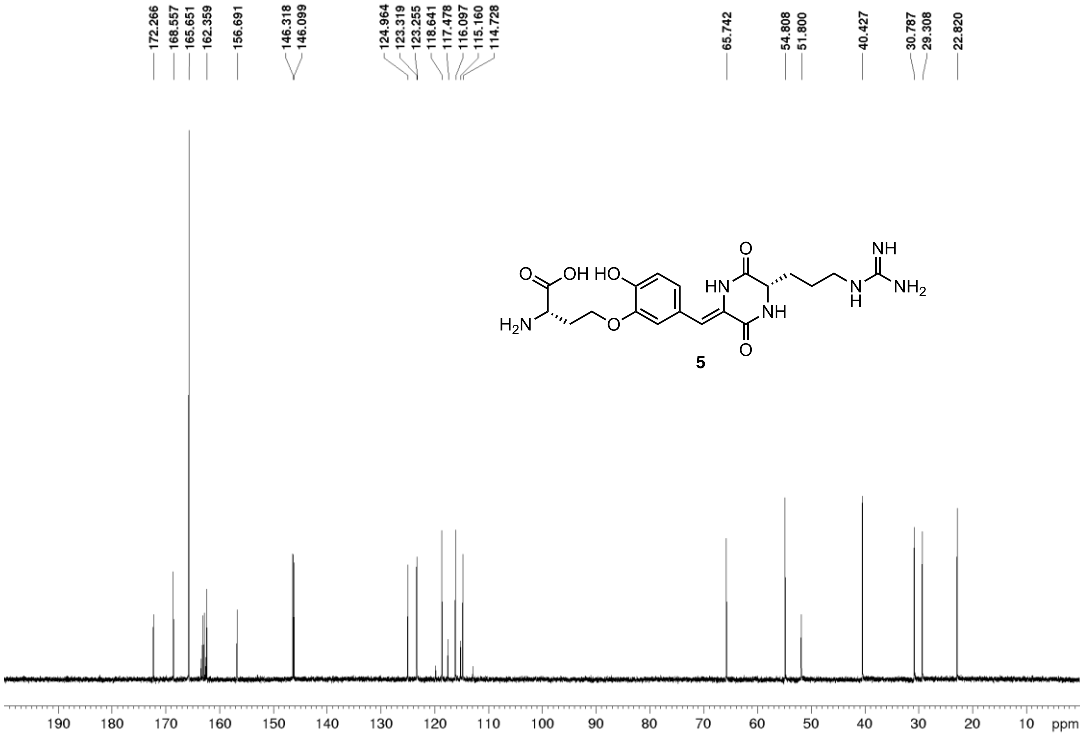


Figure S48. ^{1}H-^{1}H COSY spectrum of compound **5** in D_2O 0.5% TFA-*d* (500 MHz).

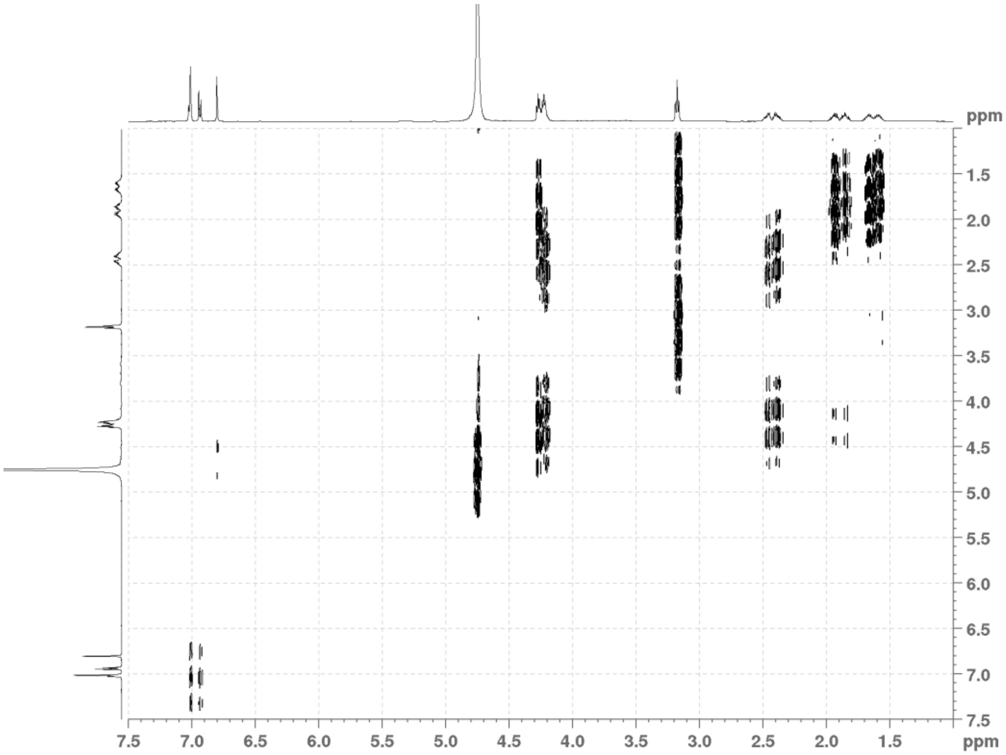


Figure S49. HSQC spectrum of compound **5** in D_2O 0.5% TFA-*d* (500 MHz).

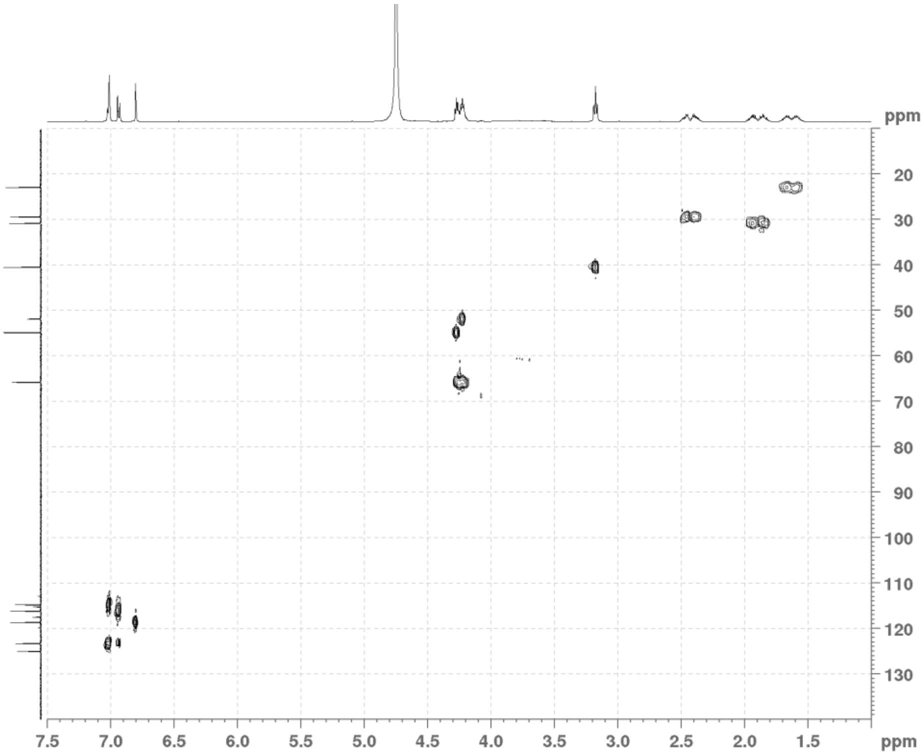


Figure S50. HMBC spectrum of compound **5** in D_2O 0.5% TFA-*d* (500 MHz).

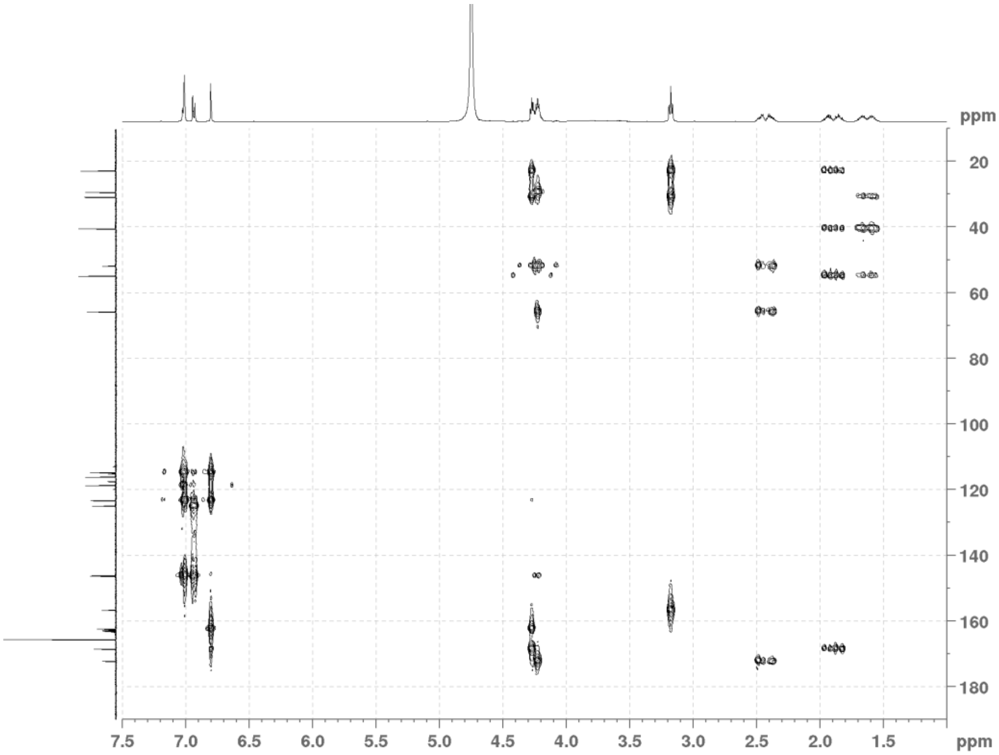


Figure S51. ^{1}H NMR spectrum of compound **6** in D_2O 0.5% TFA-*d* (500 MHz).

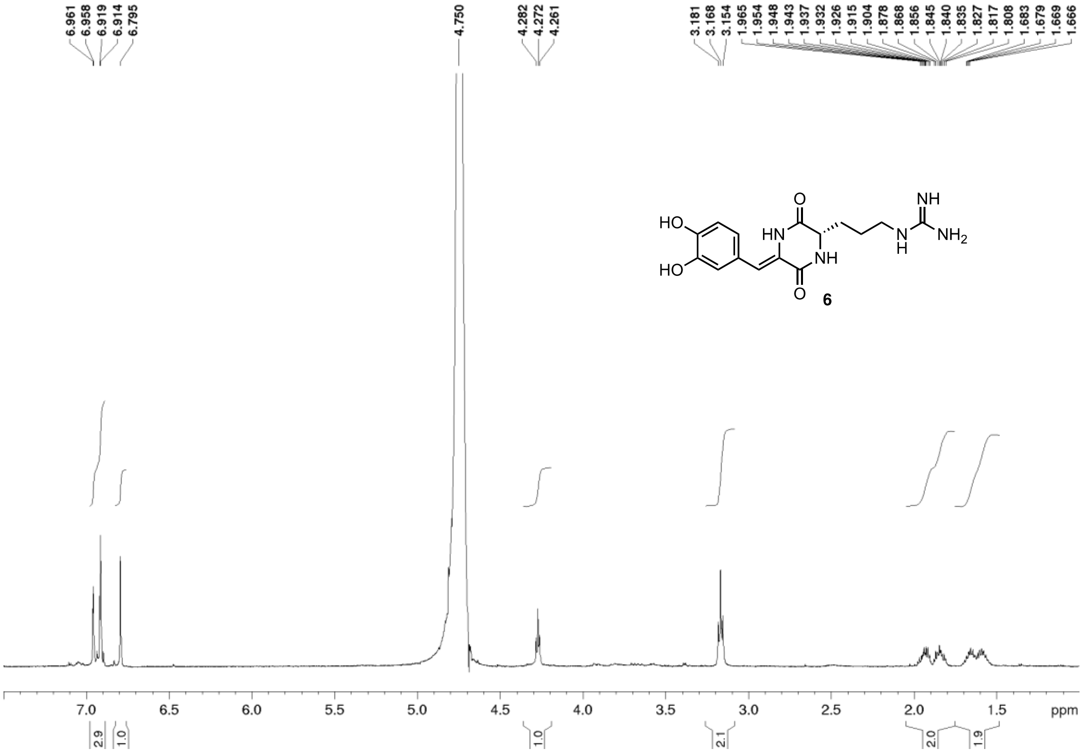


Figure S52. ^{13}C NMR spectrum of compound **6** in D_2O 0.5% TFA-*d* (125 MHz).

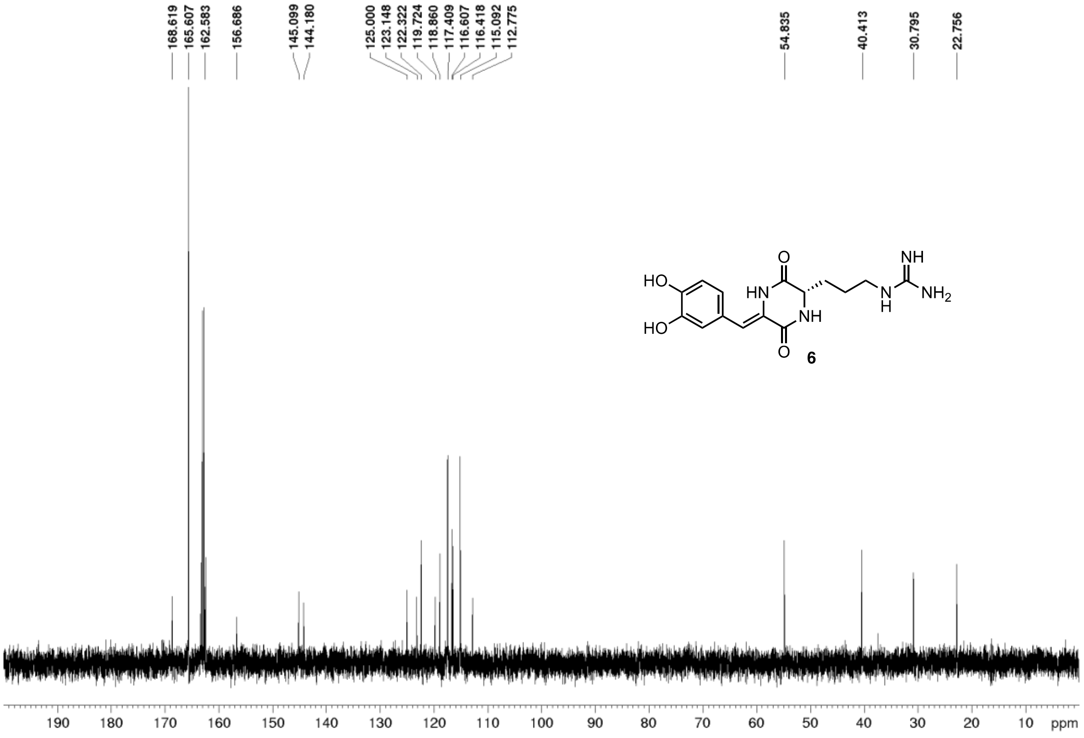


Figure S53. ^{1}H-^{1}H COSY spectrum of compound **6** in D_2O 0.5% TFA-*d* (500 MHz).

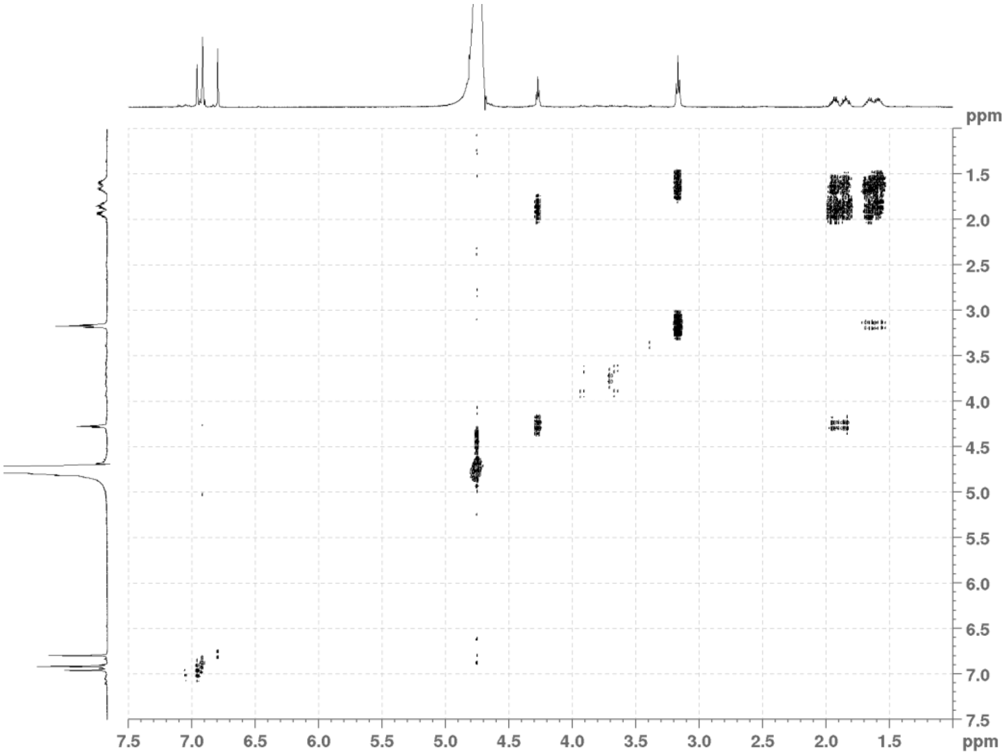


Figure S54. HSQC spectrum of compound **6** in D_2O 0.5% TFA-*d* (500 MHz).

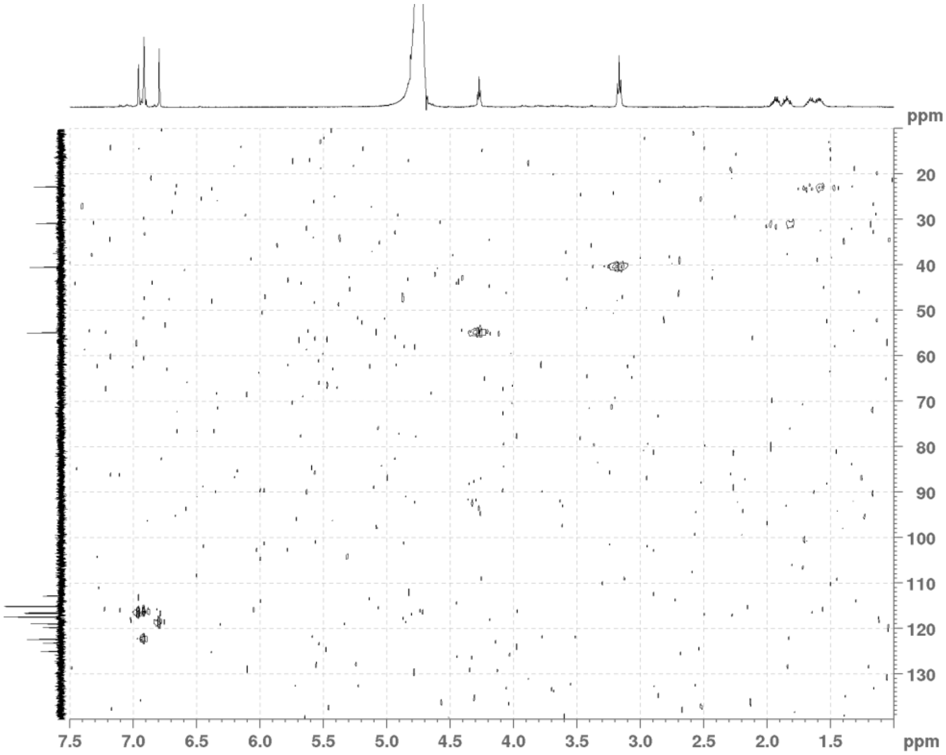


Figure S55. HMBC spectrum of compound **6** in D_2O 0.5% TFA-*d* (500 MHz).

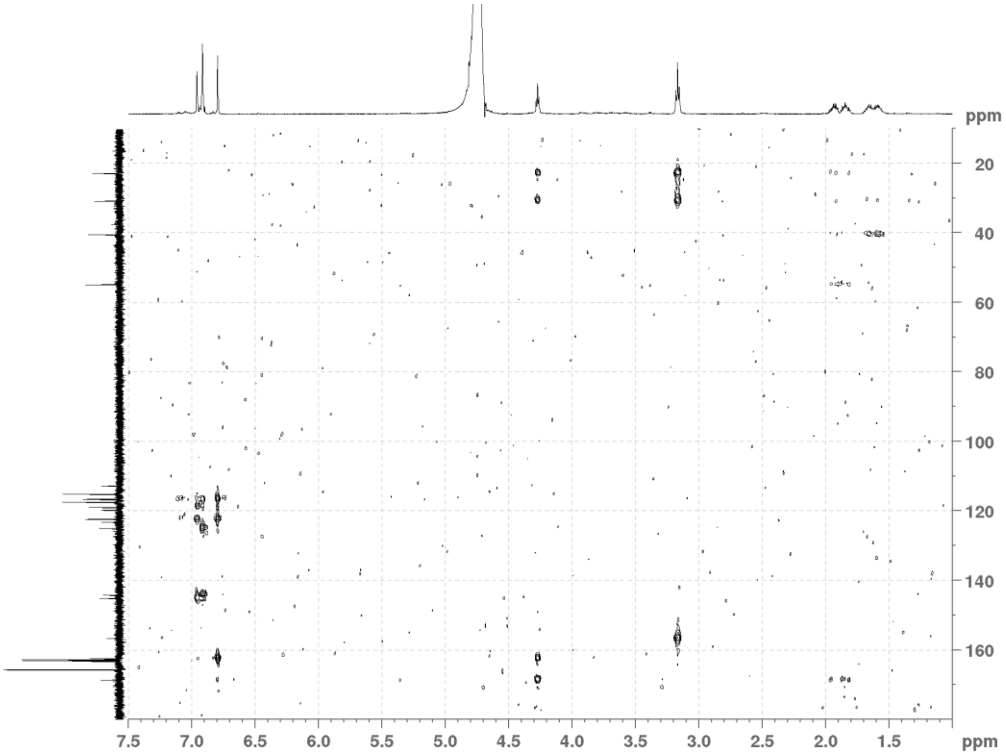


Figure S56. ^{1}H NMR spectrum of compound **9** in D_2O (500 MHz).

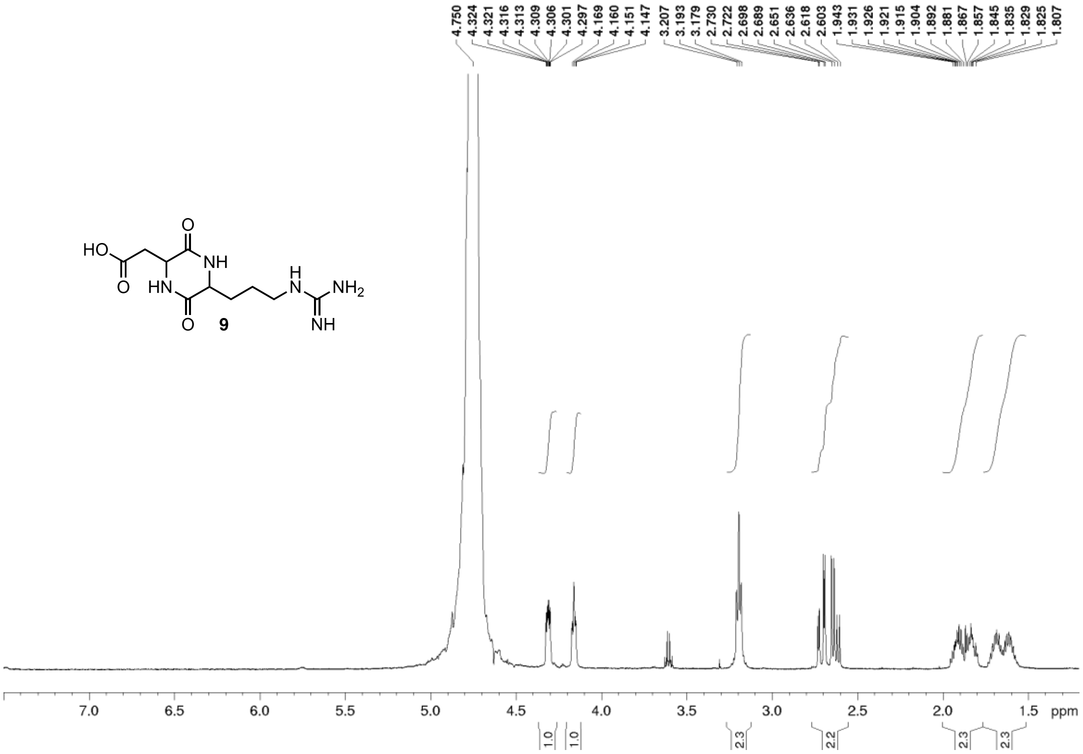


Figure S57. ^{13}C NMR spectrum of compound **9** in D_2O (125 MHz).

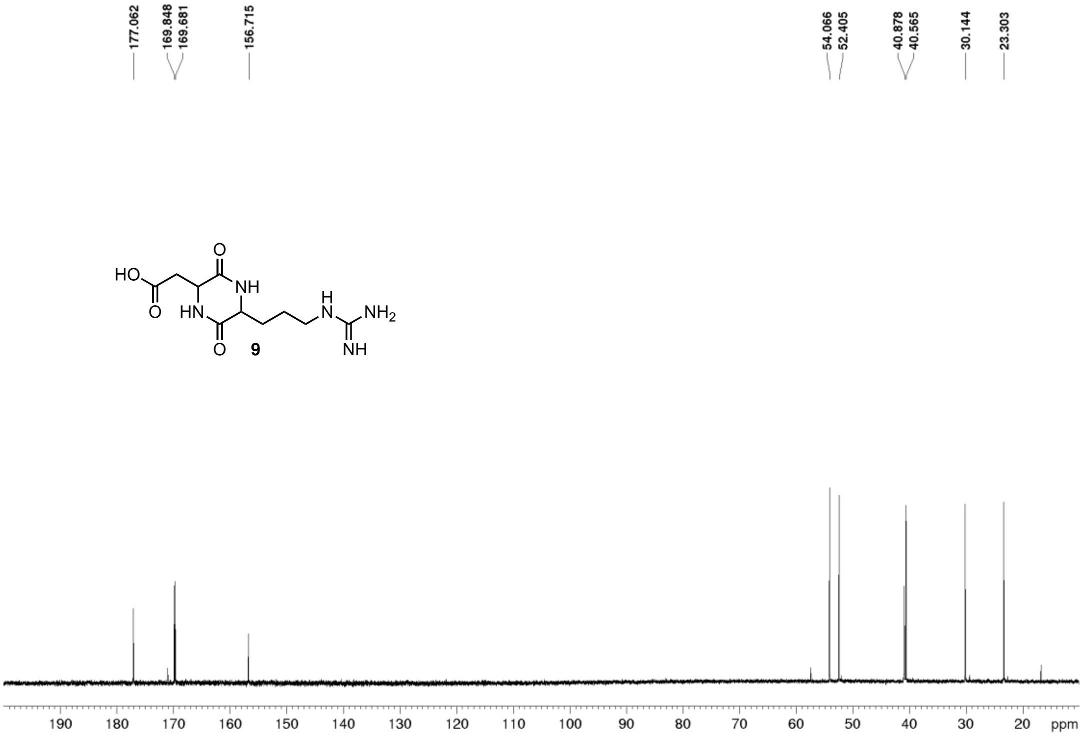


Figure S58. ^{1}H-^{1}H COSY spectrum of compound **9** in D_2O (500 MHz).

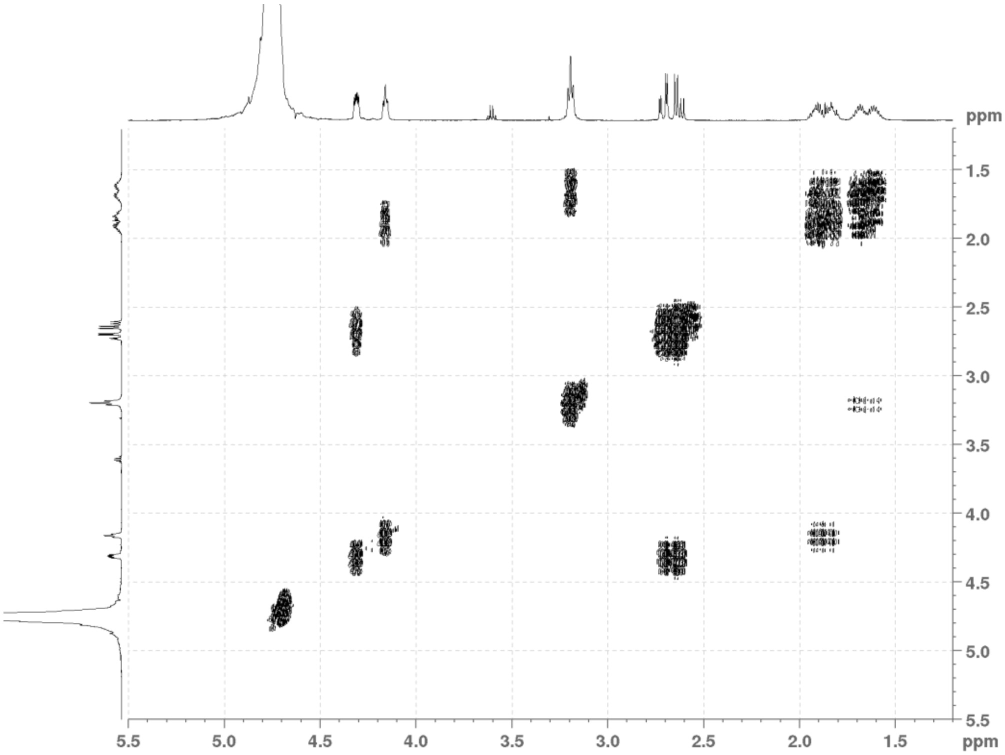


Figure S59. HSQC spectrum of compound **9** in D_2O (500 MHz).

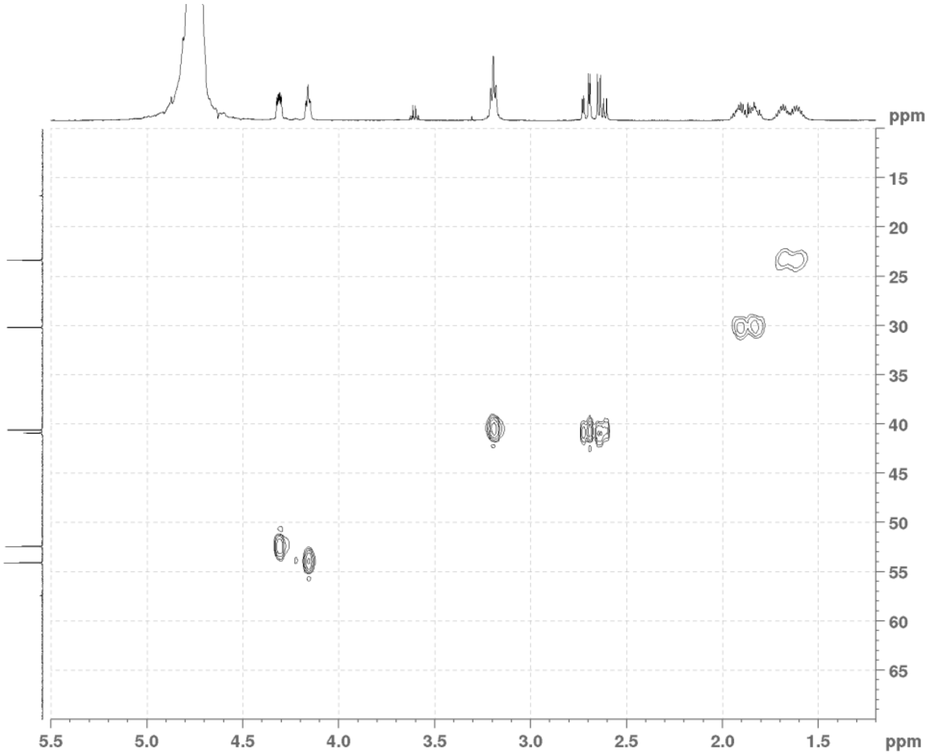


Figure S60. HMBC spectrum of compound **9** in D_2O (500 MHz).

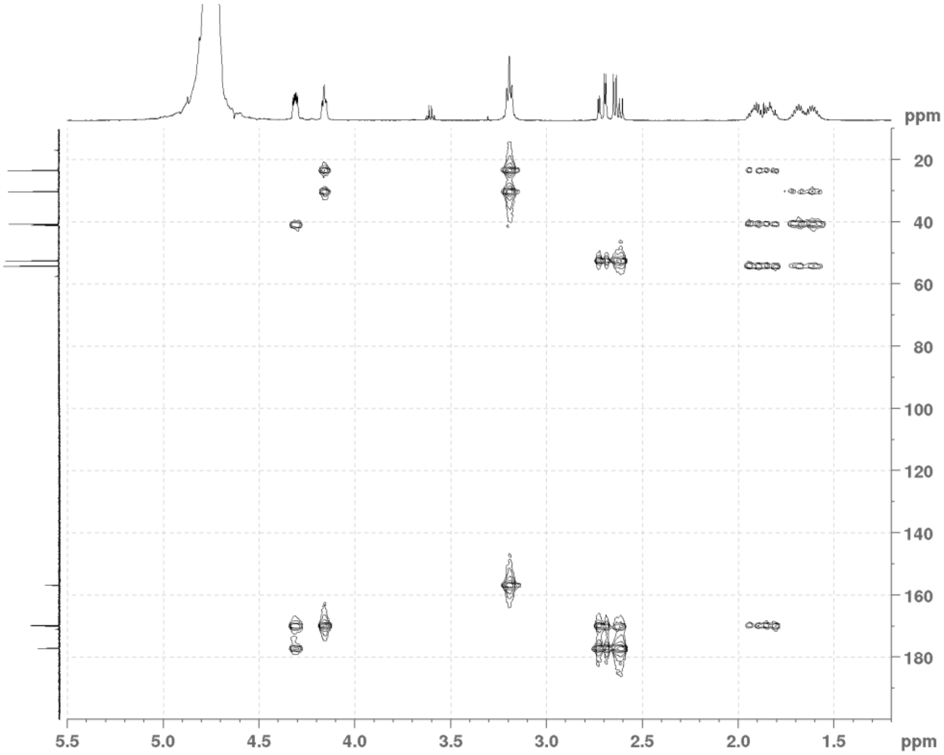


Figure S61. ^{1}H NMR spectrum of compound **10** in D_2O (500 MHz).

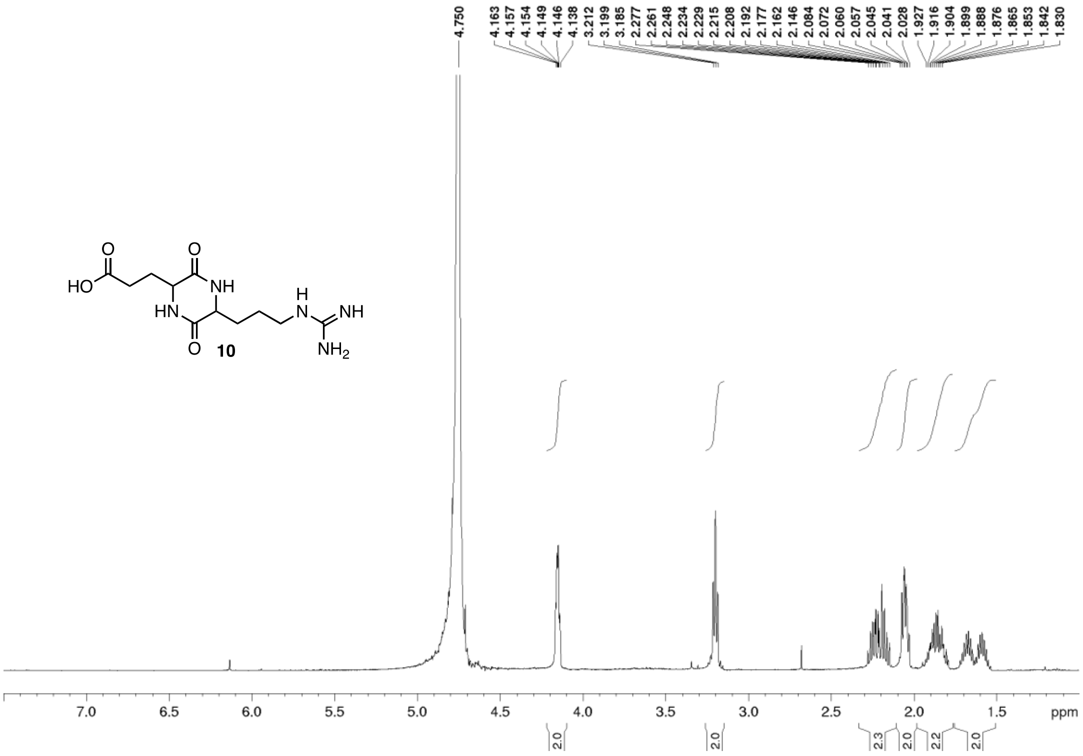


Figure S62. ^{13}C NMR spectrum of compound **10** in D_2O (125 MHz).

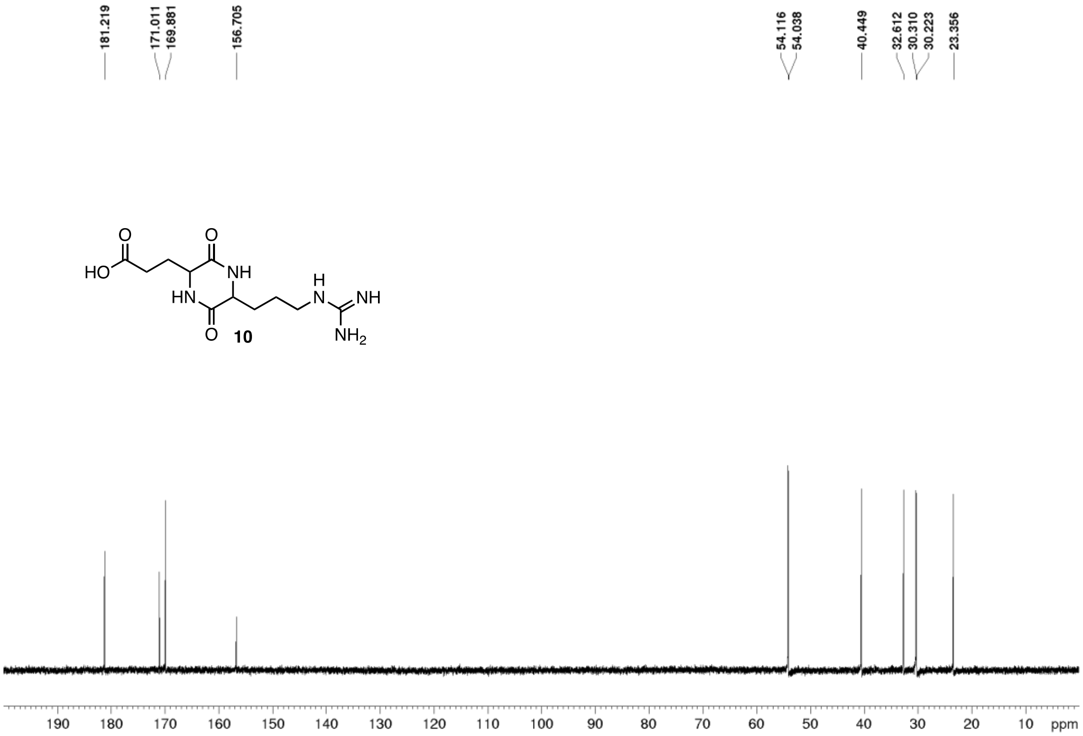


Figure S63. ^1H-^1H COSY spectrum of compound **10** in D_2O (500 MHz).

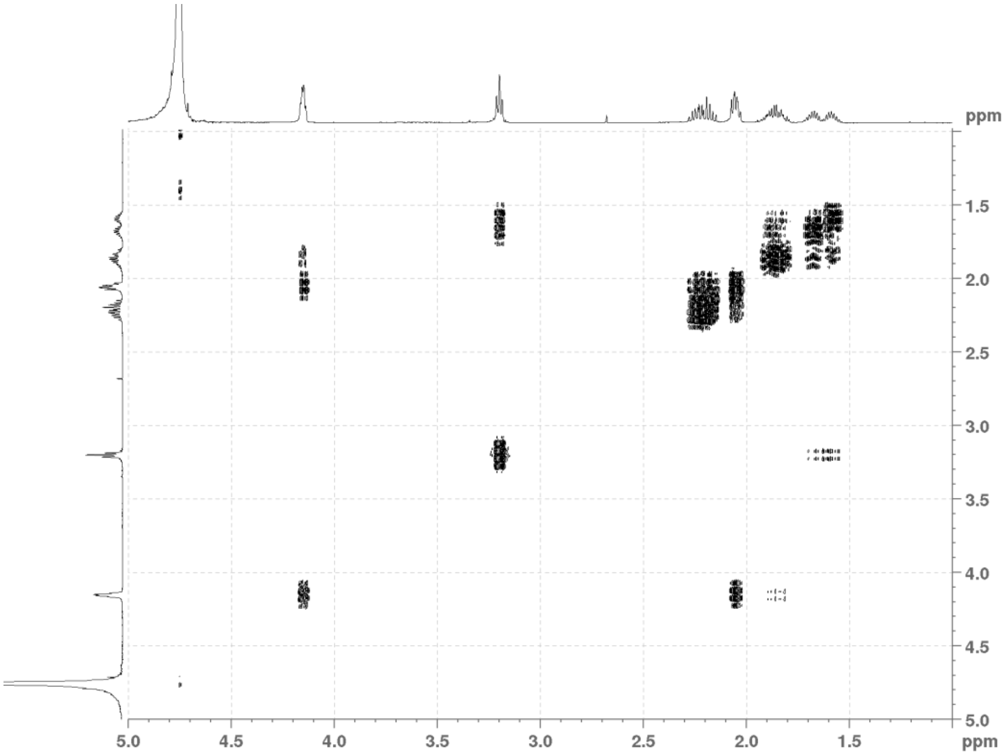


Figure S64. HSQC spectrum of compound **10** in D_2O (500 MHz).

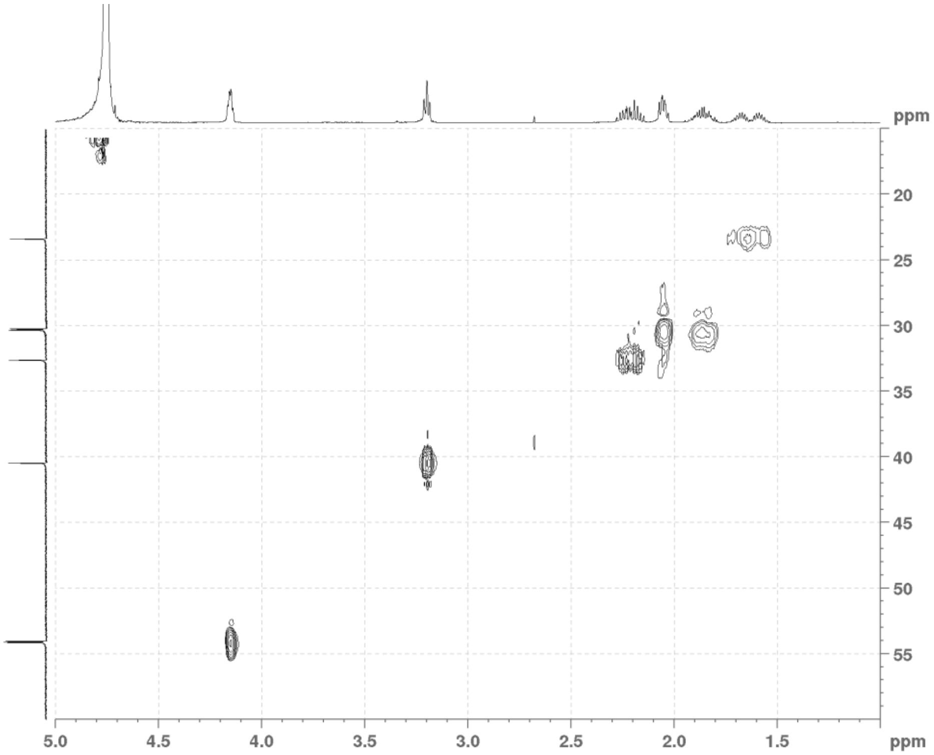


Figure S65. HMBC spectrum of compound **10** in D_2O (500 MHz).

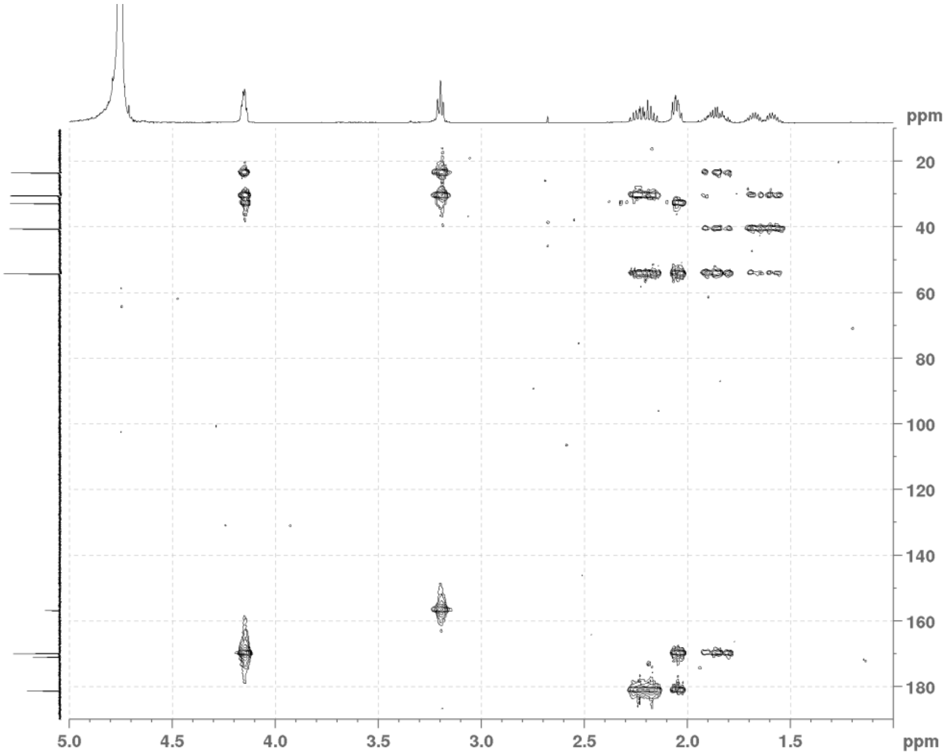


Figure S66. ^{1}H NMR spectrum of compound **13** in CD_3OD (500 MHz).

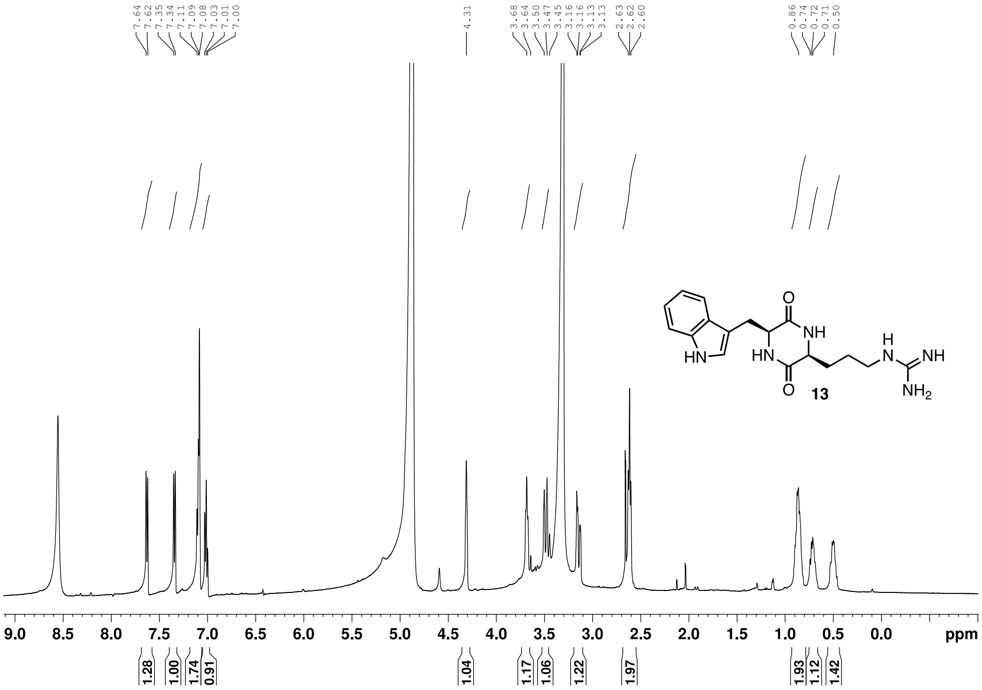


Figure S67. ^{13}C NMR spectrum of compound **13** in CD_3OD (125 MHz).

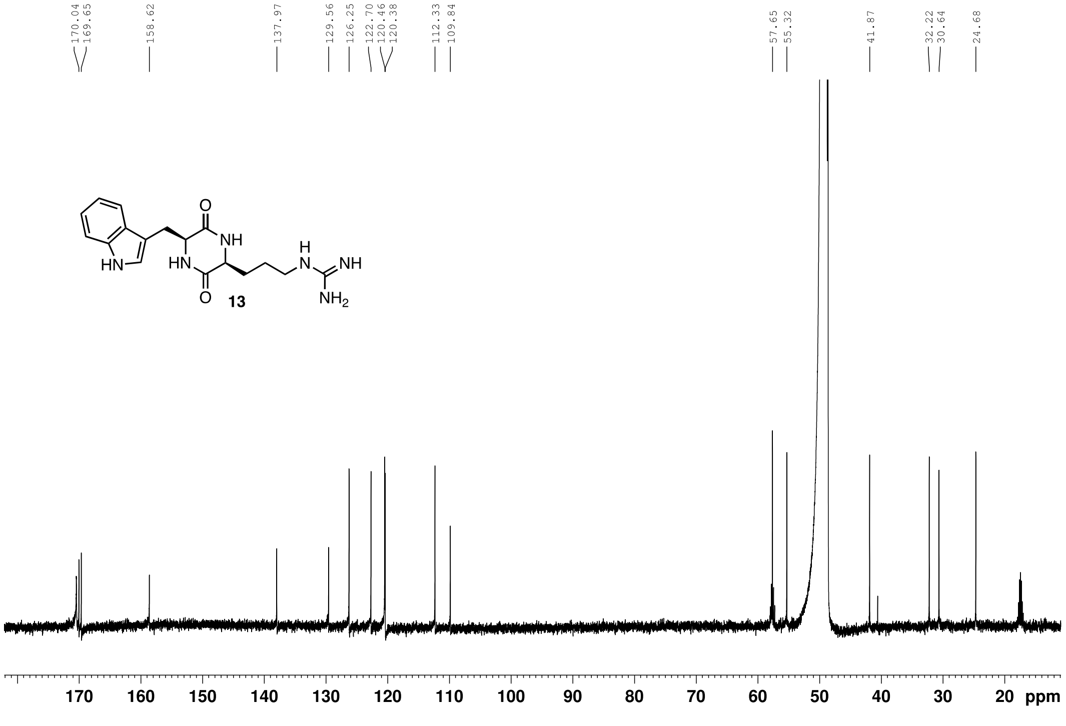


Figure S68. ^{1}H NMR spectrum of compound **14** in D_2O (500 MHz).

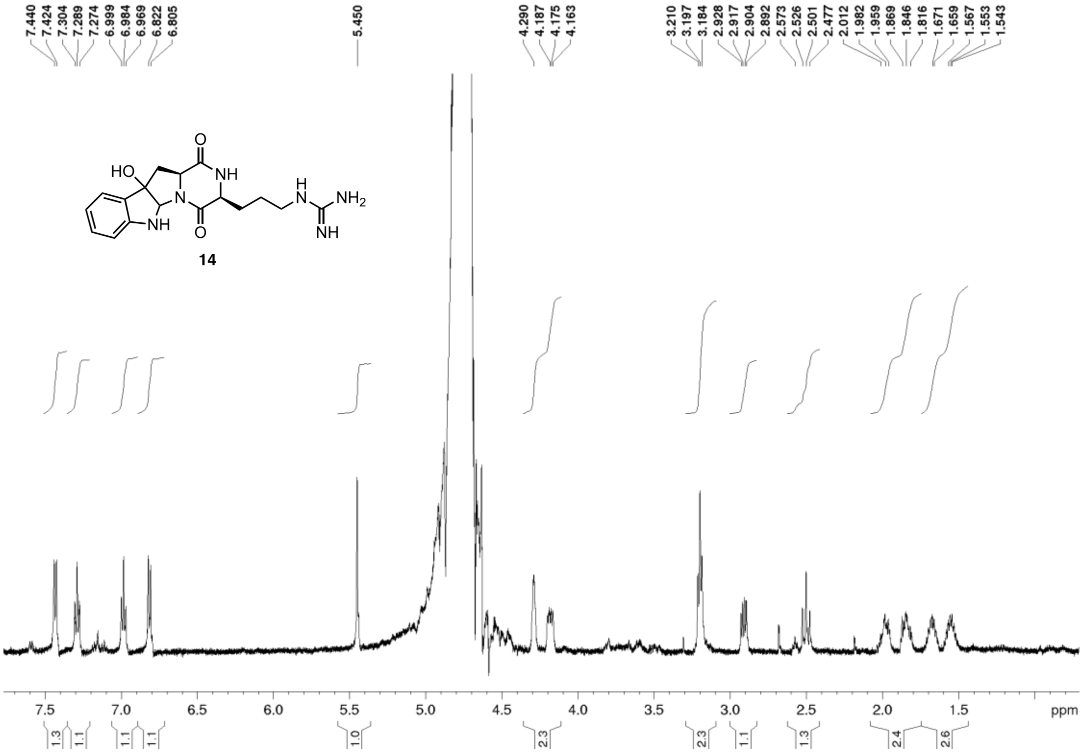


Figure S69. ^{13}C NMR spectrum of compound **14** in D_2O (125 MHz).

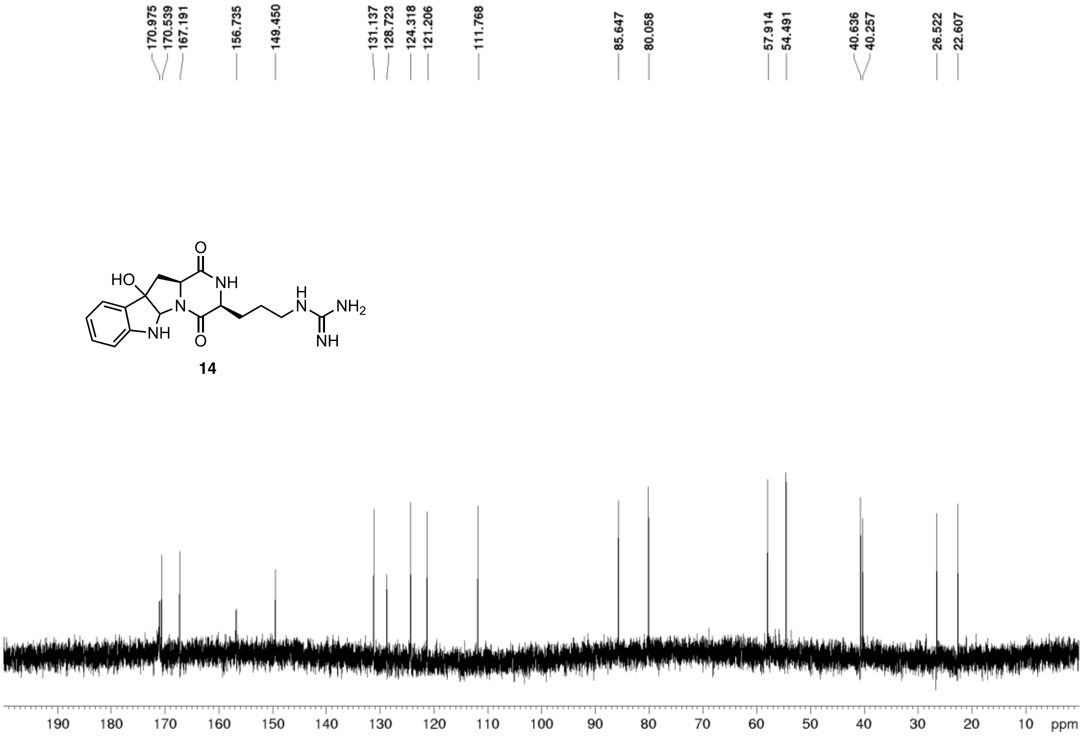


Figure S70. ^{1}H-^{1}H NMR spectrum of compound **14** in D_2O (500 MHz).

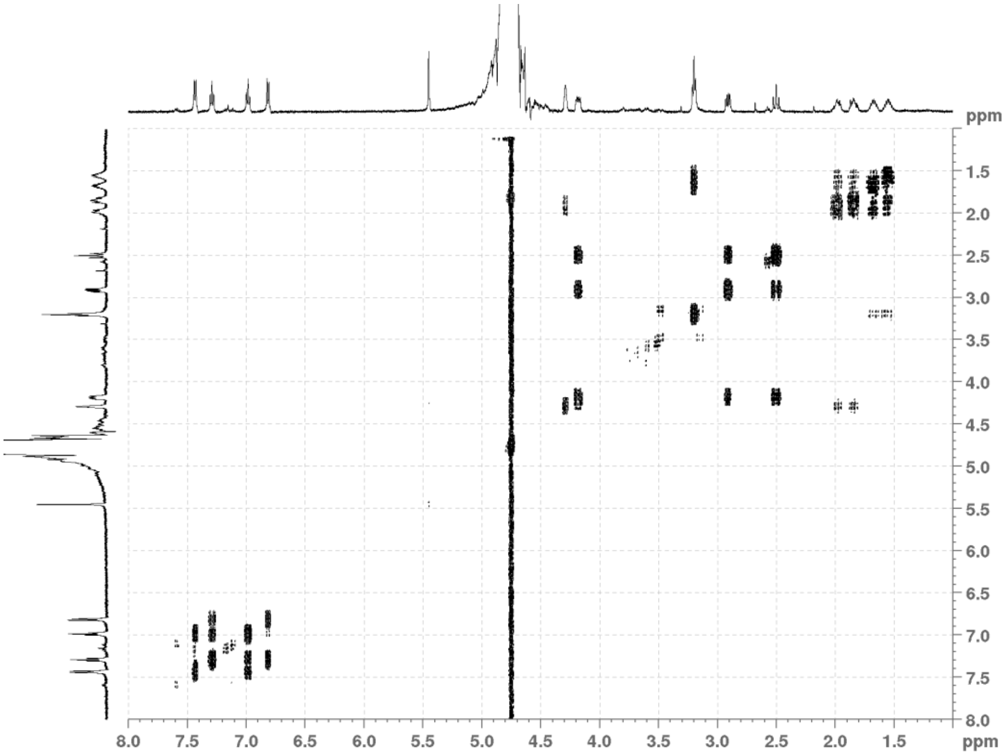


Figure S71. HSQC spectrum of compound **14** in D_2O (500 MHz).

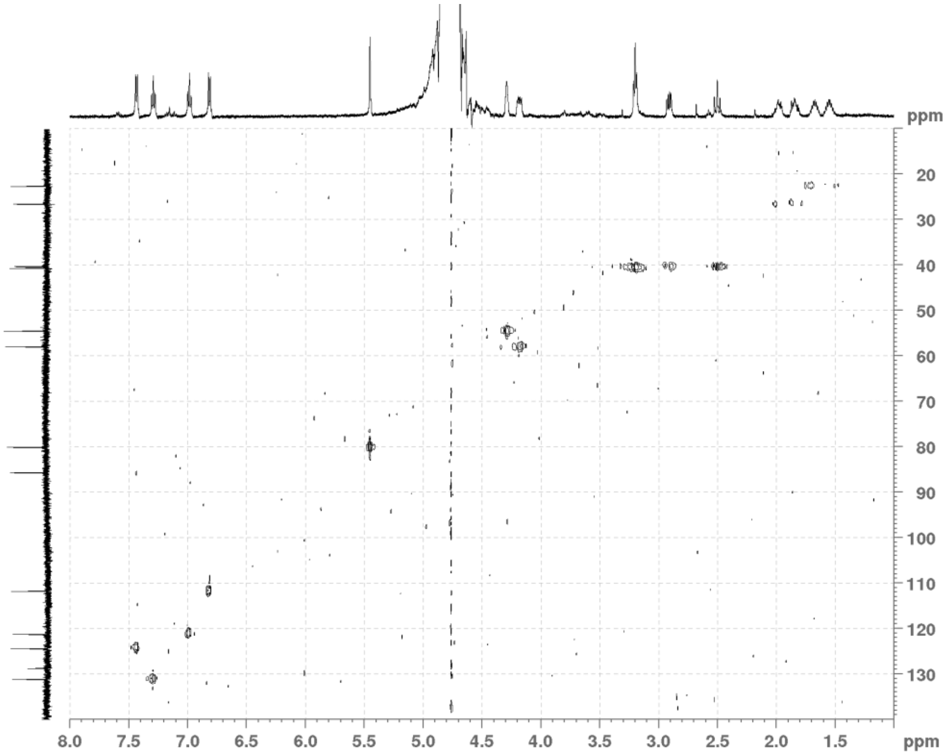


Figure S72. HMBC spectrum of compound **14** in D_2O (500 MHz).

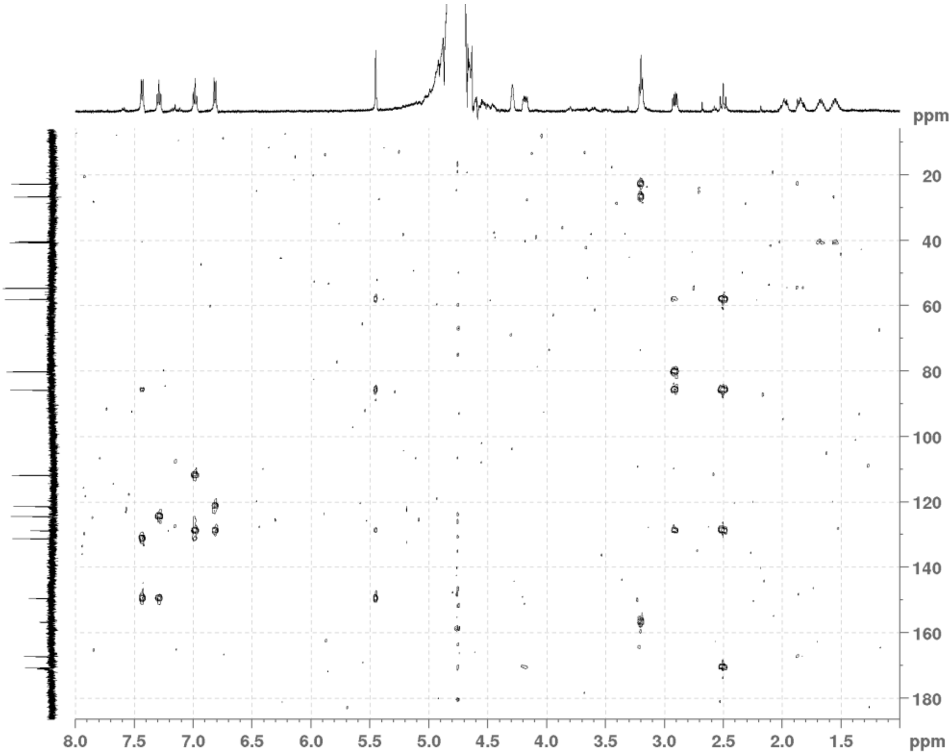


Figure S73. ^{1}H NMR spectrum of compound **15** in D_2O (500 MHz).

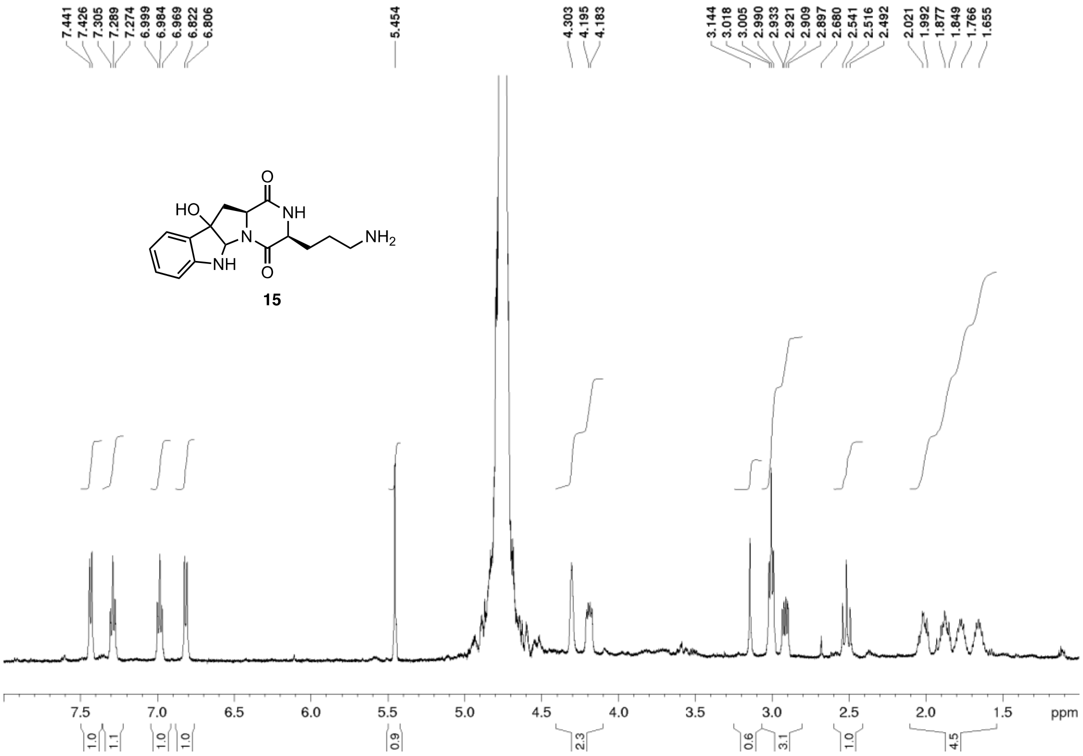


Figure S74. ^{13}C NMR spectrum of compound **15** in D_2O (125 MHz).

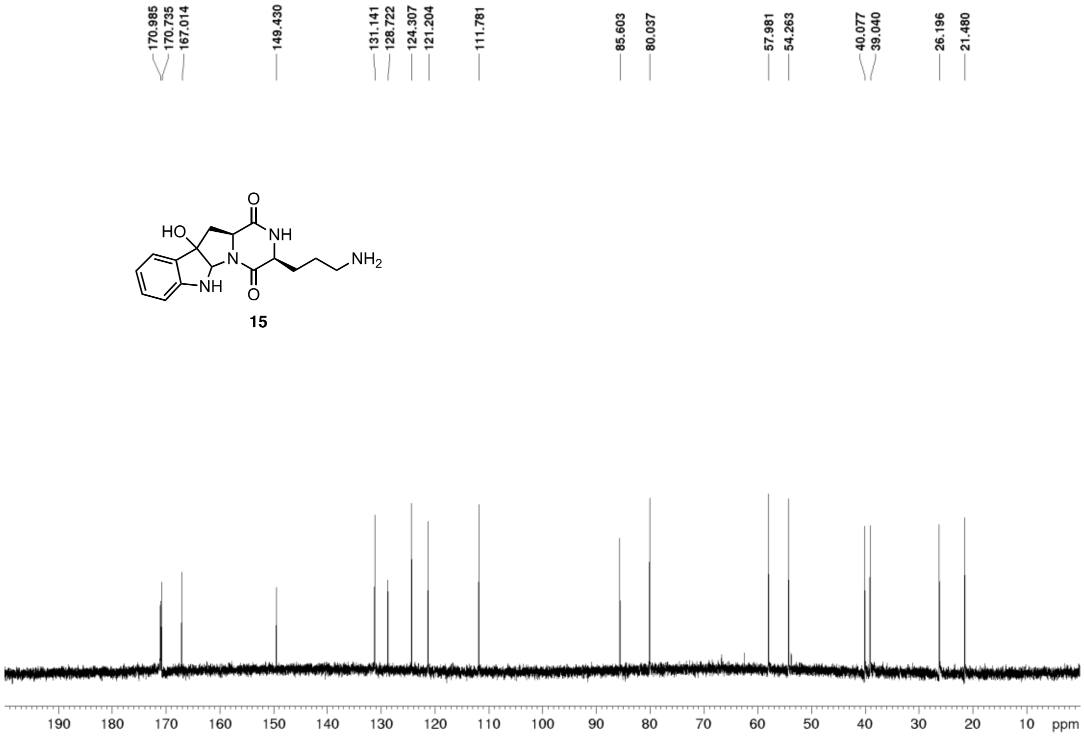


Figure S75. ^{1}H -^{1}H COSY spectrum of compound **15** in D_2O (500 MHz).

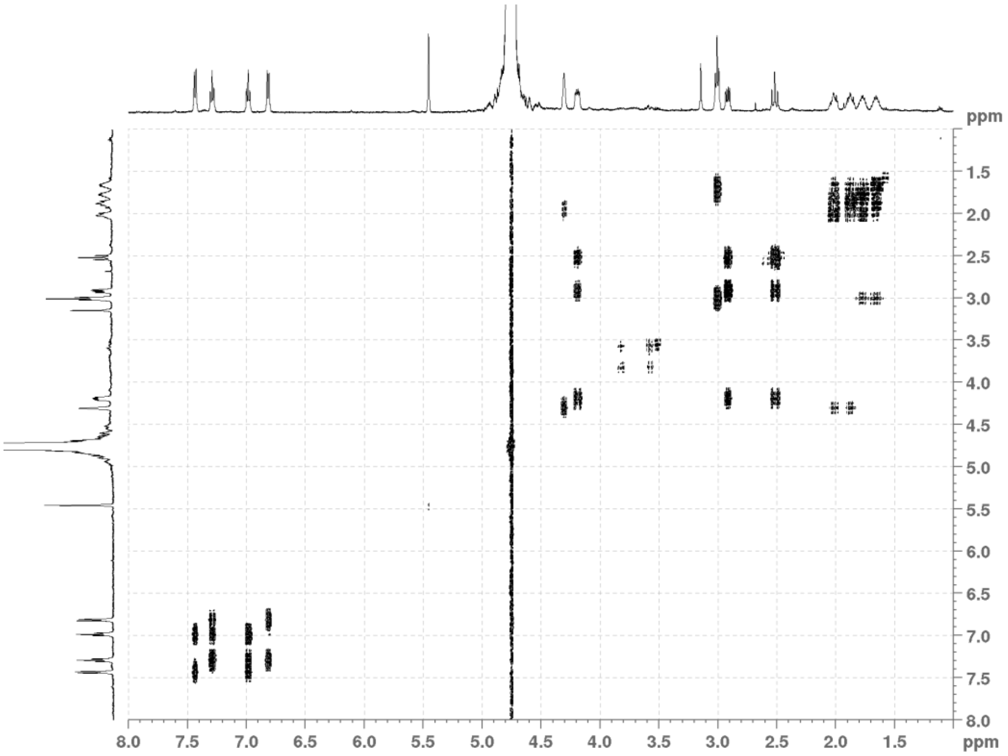


Figure S76. HSQC spectrum of compound **15** in D_2O (500 MHz).

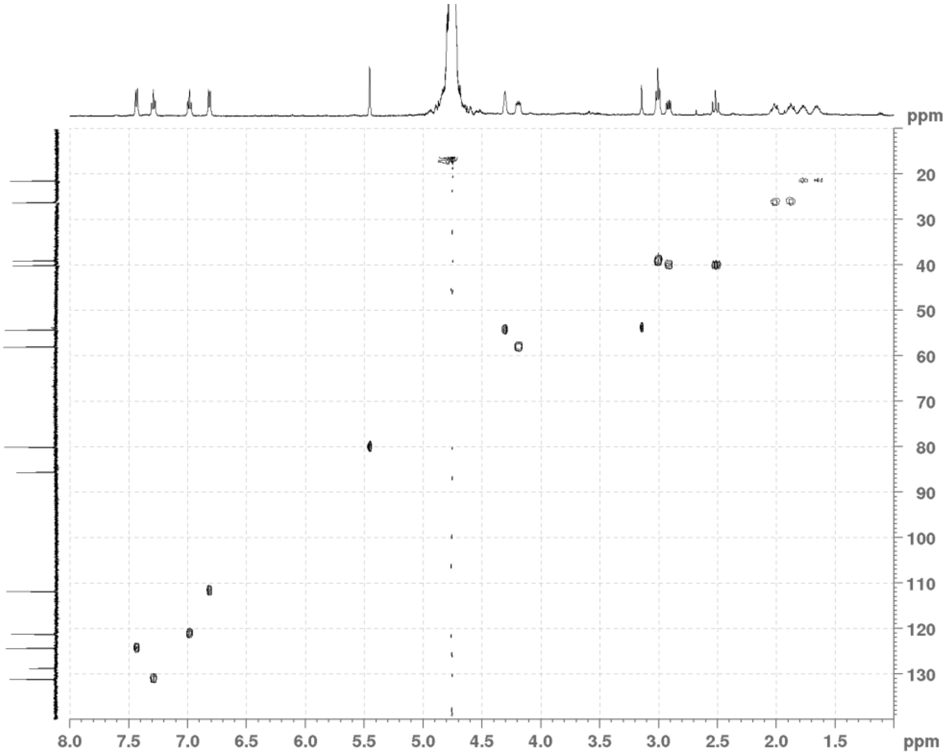


Figure S77. HMBC spectrum of compound **15** in D_2O (500 MHz).

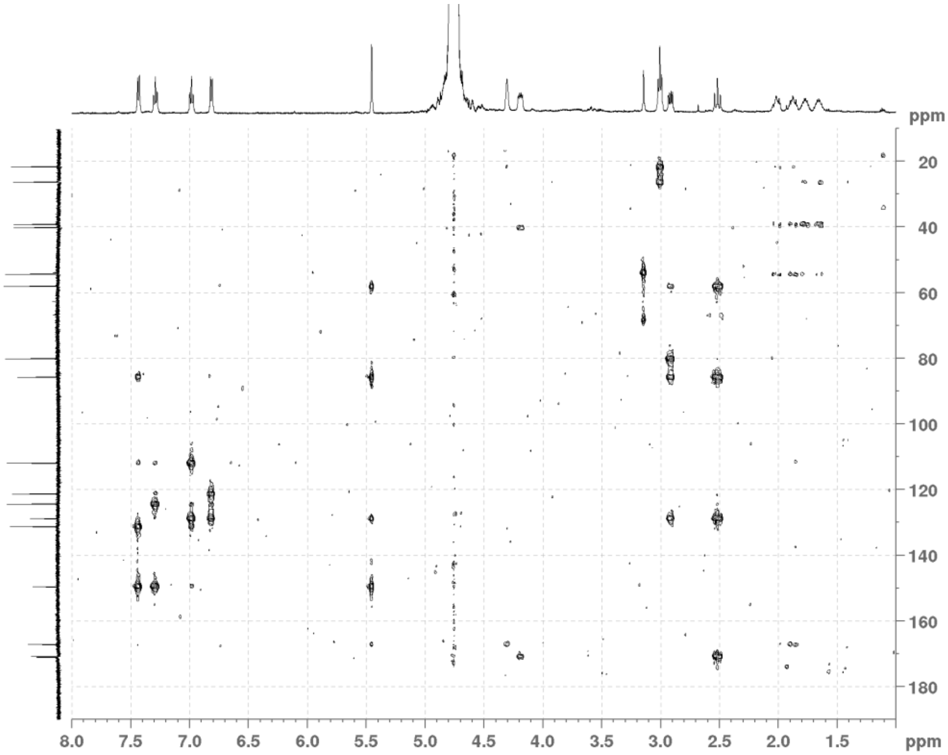


Figure S78. ^{1}H NMR spectrum of compound **16** in D_2O 0.5% TFA-*d* (500 MHz).

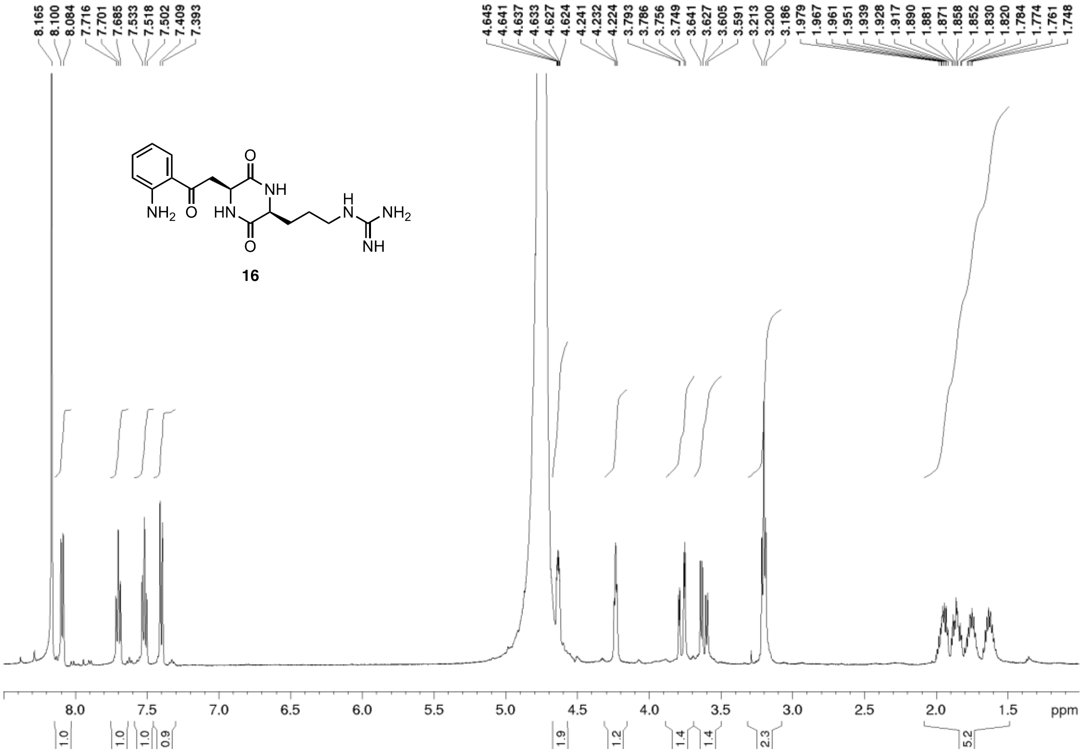


Figure S79. ^{13}C NMR spectrum of compound **16** in D_2O 0.5% TFA-*d* (125 MHz).

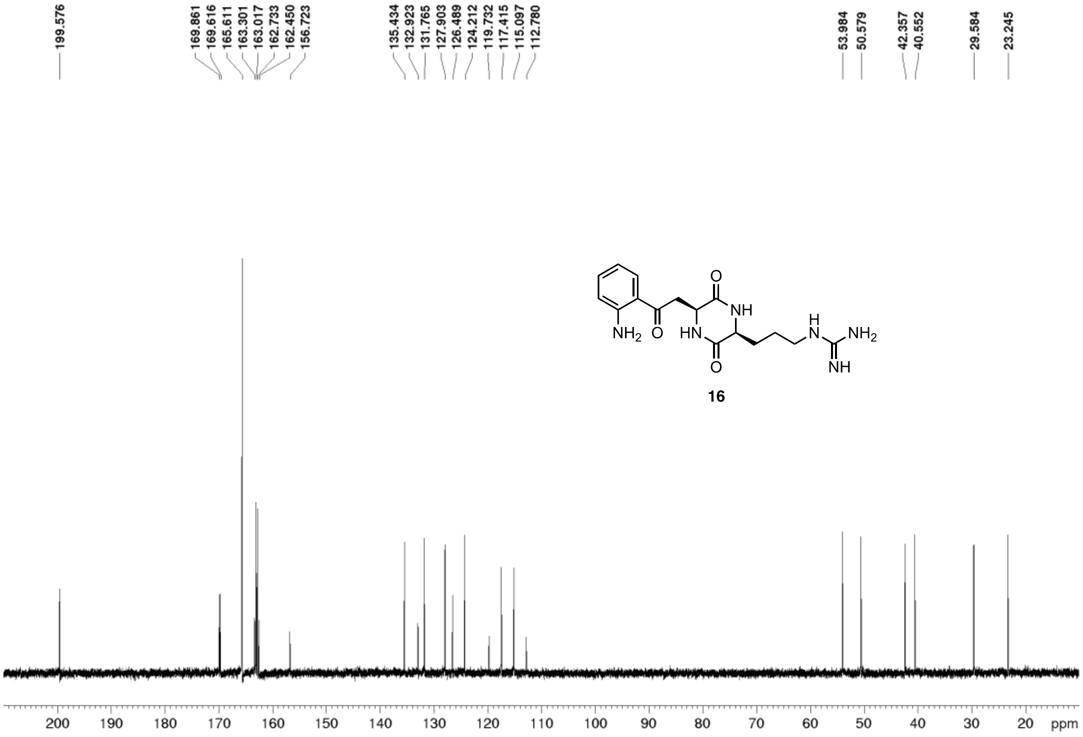


Figure S80. ^{1}H-^{1}H COSY spectrum of compound **16** in D_2O 0.5% TFA-*d* (500 MHz).

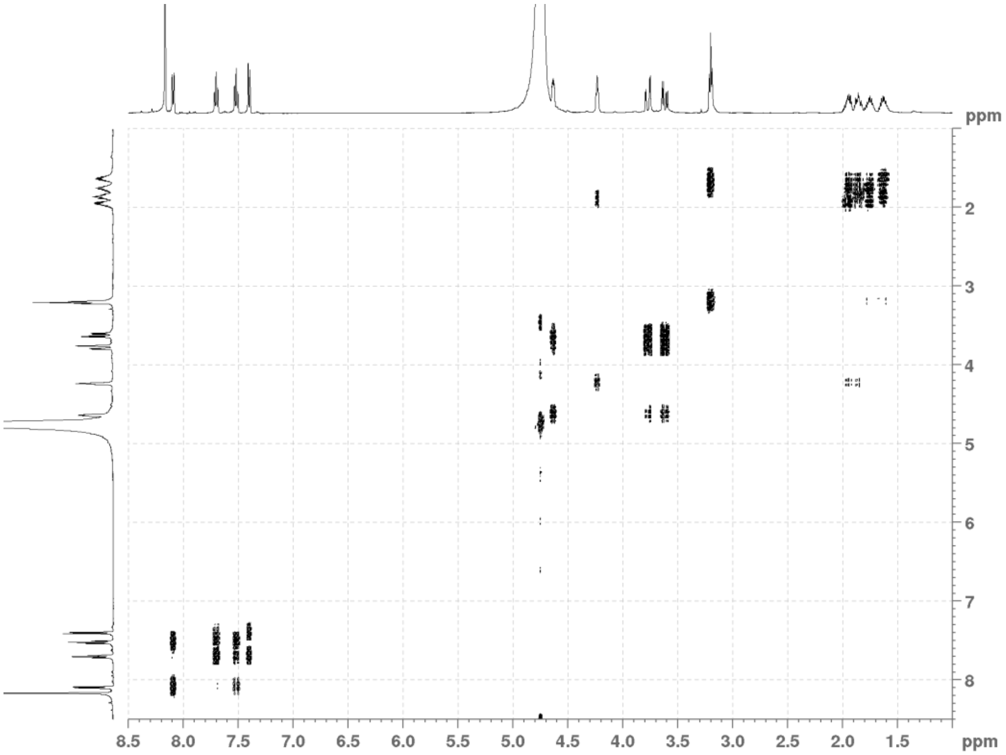


Figure S81. HSQC spectrum of compound **16** in D_2O 0.5% TFA-*d* (500 MHz).

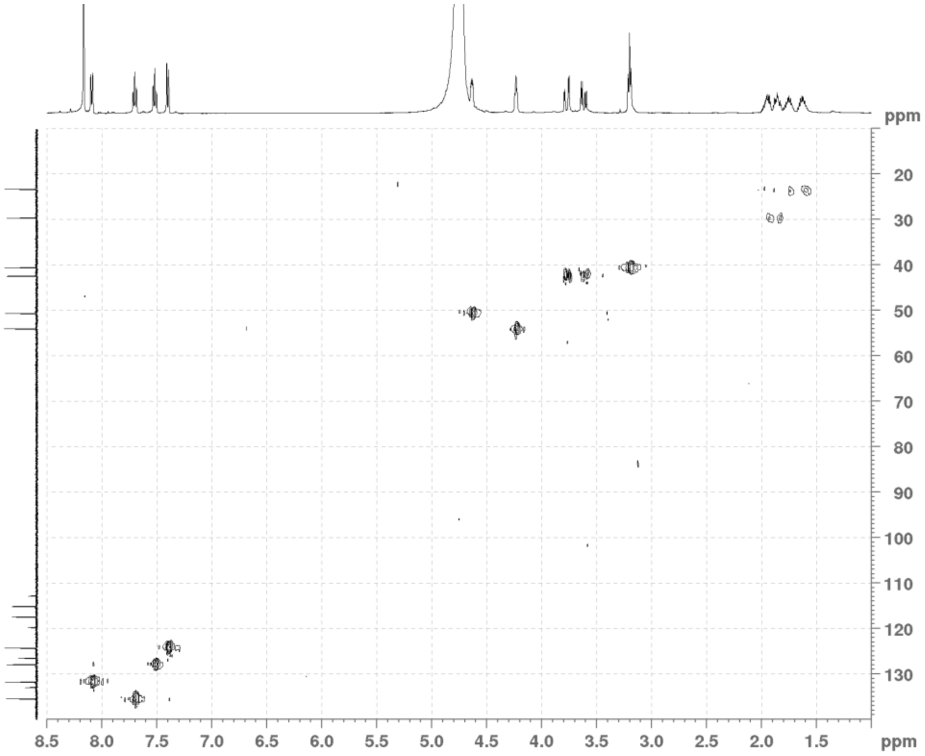


Figure S82. HMBC spectrum of compound **16** in D_2O 0.5% TFA-*d* (500 MHz).

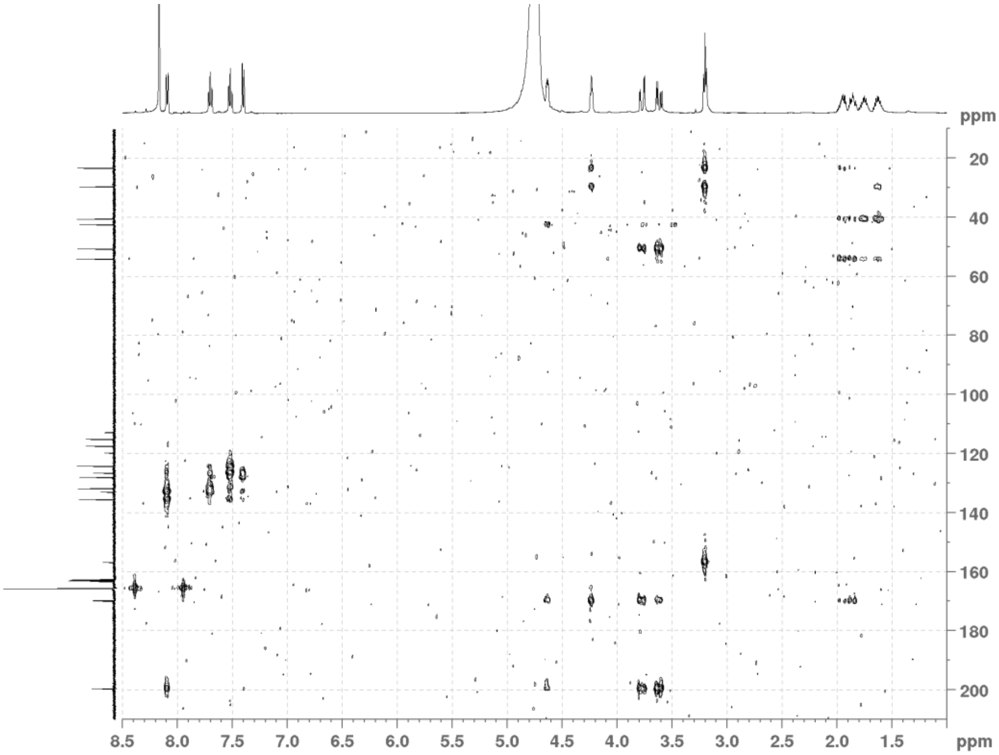


Figure S83. ^{1}H NMR spectrum of compound **17** in D_2O 0.5% TFA-*d* (500 MHz).

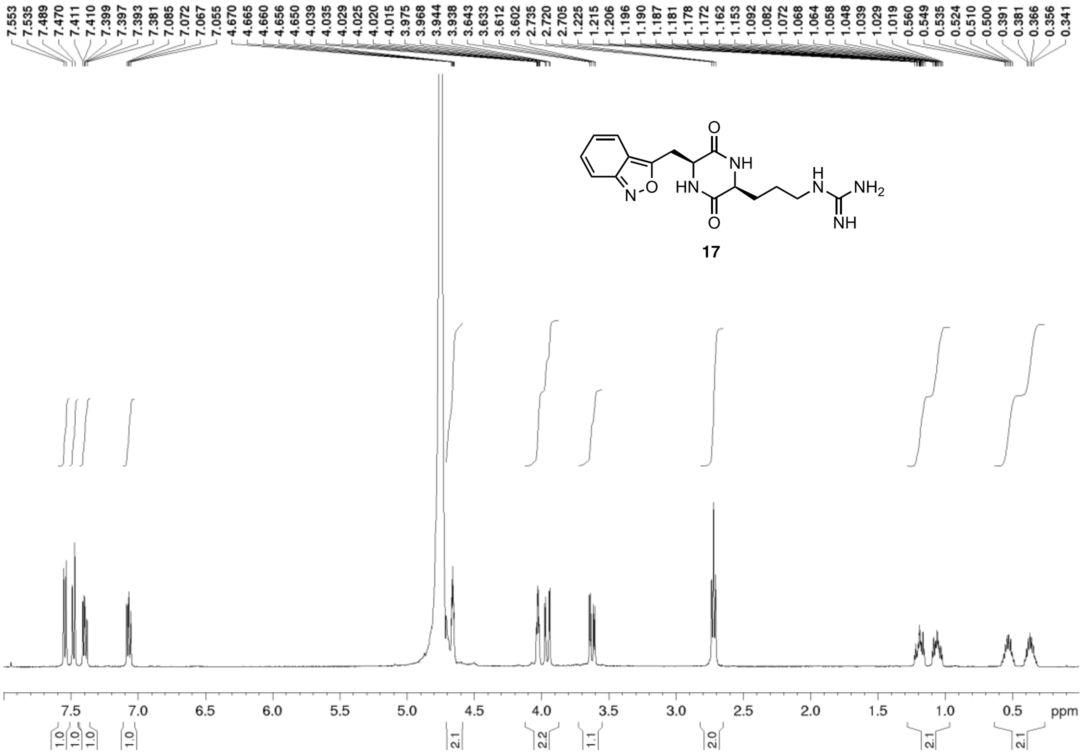


Figure S84. ^{13}C NMR spectrum of compound **17** in D_2O 0.5% TFA-*d* (125 MHz).

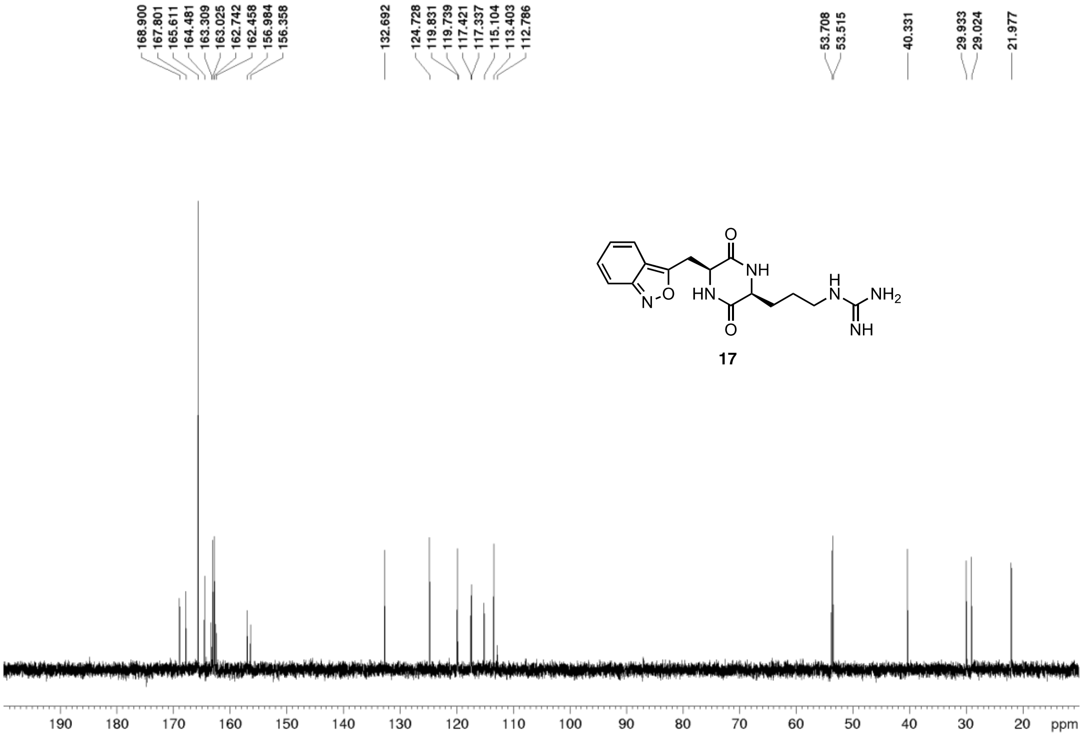


Figure S85. ^{1}H-^{1}H COSY spectrum of compound **17** in D_2O 0.5% TFA-*d* (500 MHz).

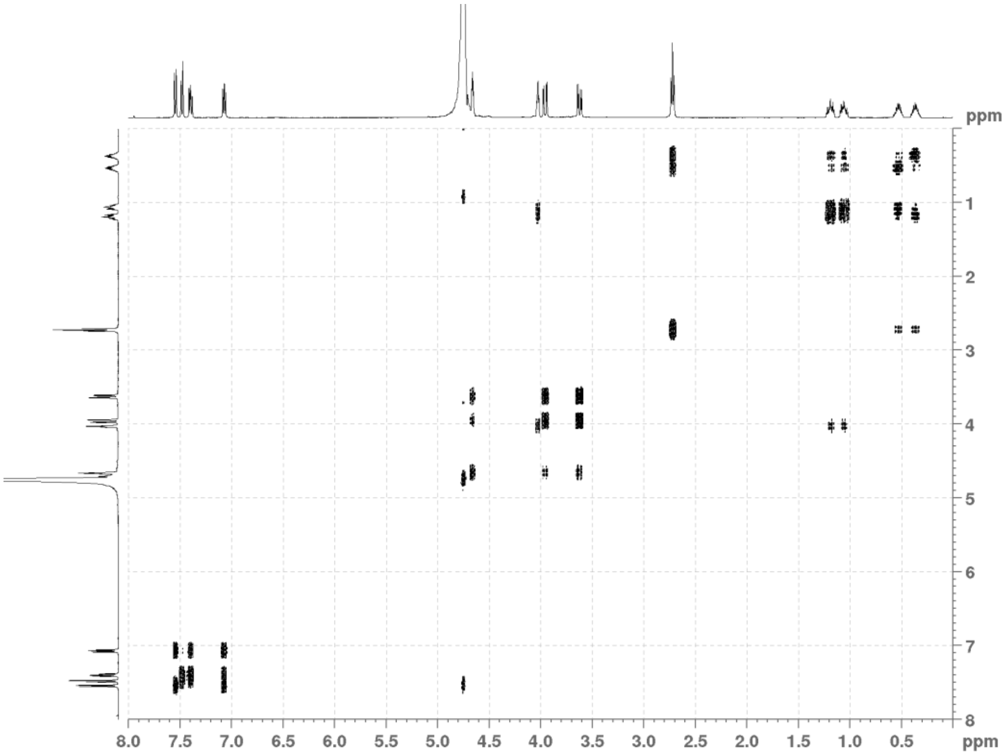


Figure S86. HSQC spectrum of compound **17** in D_2O 0.5% TFA-*d* (500 MHz).

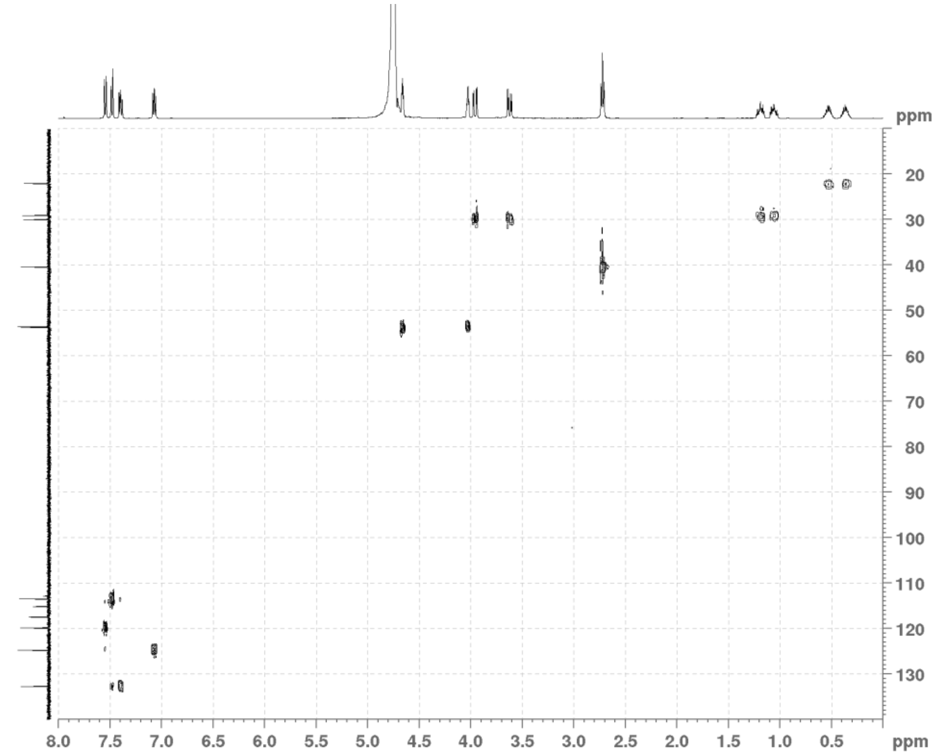


Figure S87. HMBC spectrum of compound **17** in D_2O 0.5% TFA-*d* (500 MHz).

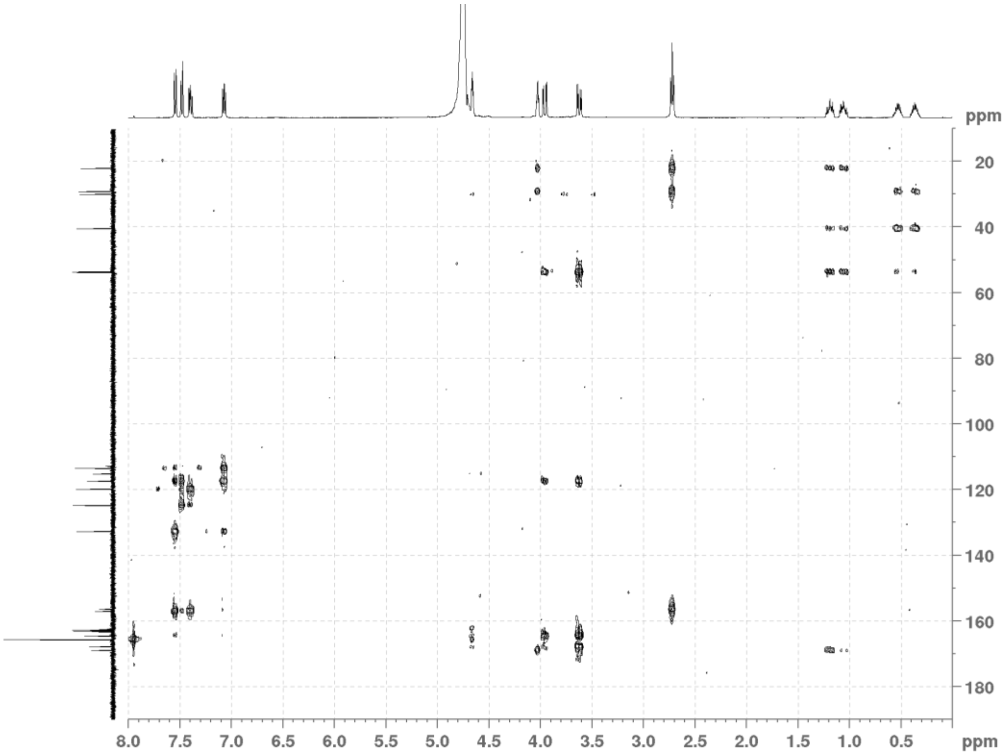


Figure S88. ^{1}H NMR spectrum of compound **18** in DMSO-d_6 (500 MHz).

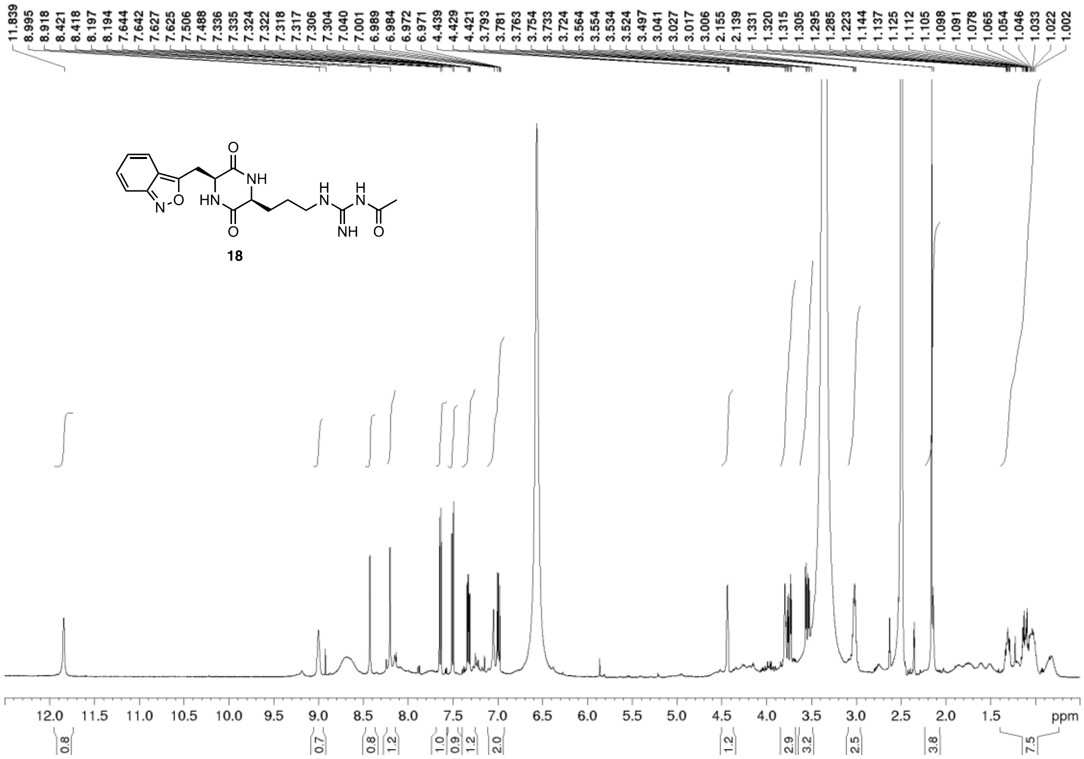


Figure S89. ^{13}C NMR spectrum of compound **18** in DMSO-d_6 (125 MHz).

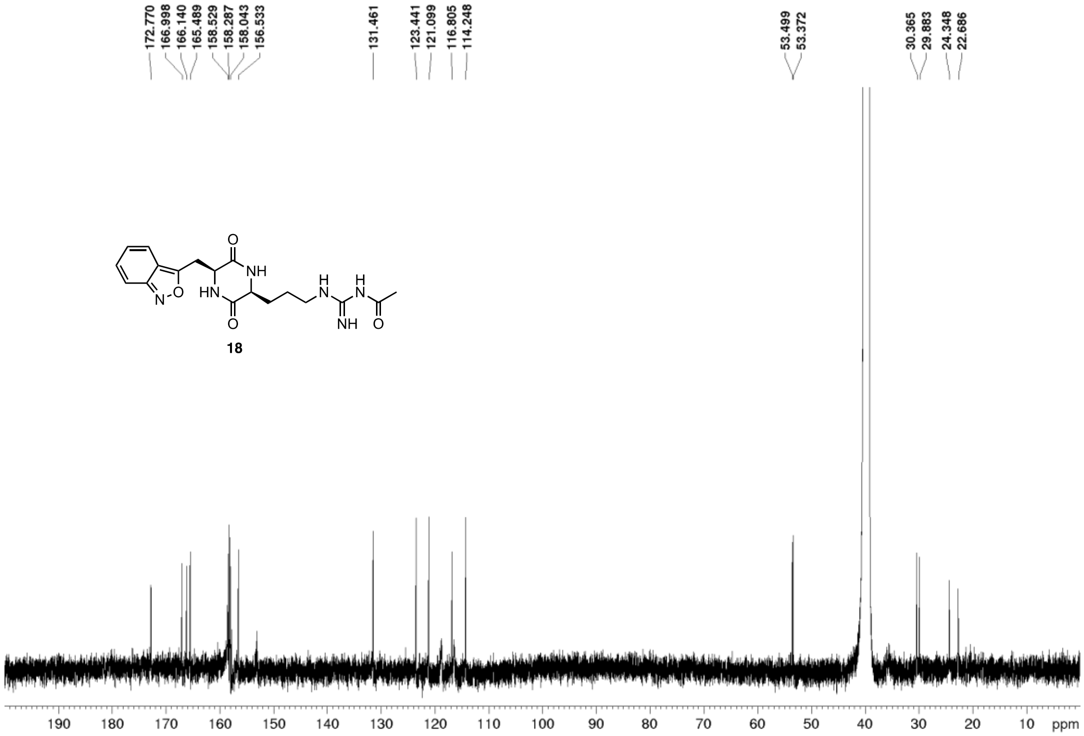


Figure S90. ^{1}H-^{1}H COSY spectrum of compound **18** in DMSO-d_6 (500 MHz).

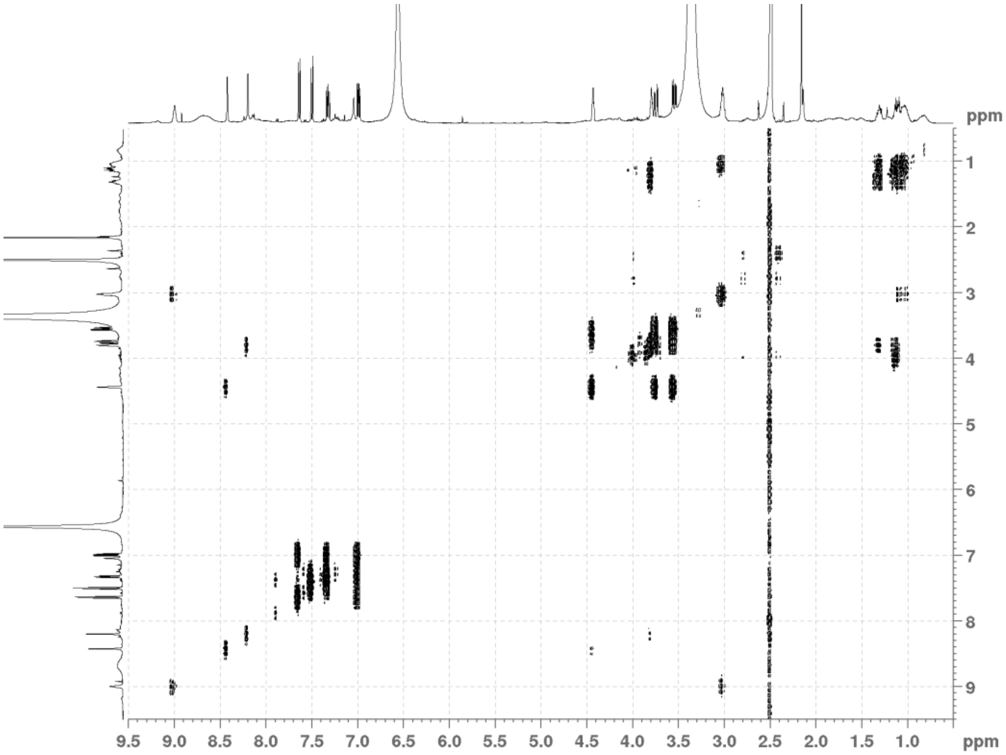


Figure S91. HSQC spectrum of compound **18** in DMSO-d_6 (500 MHz).

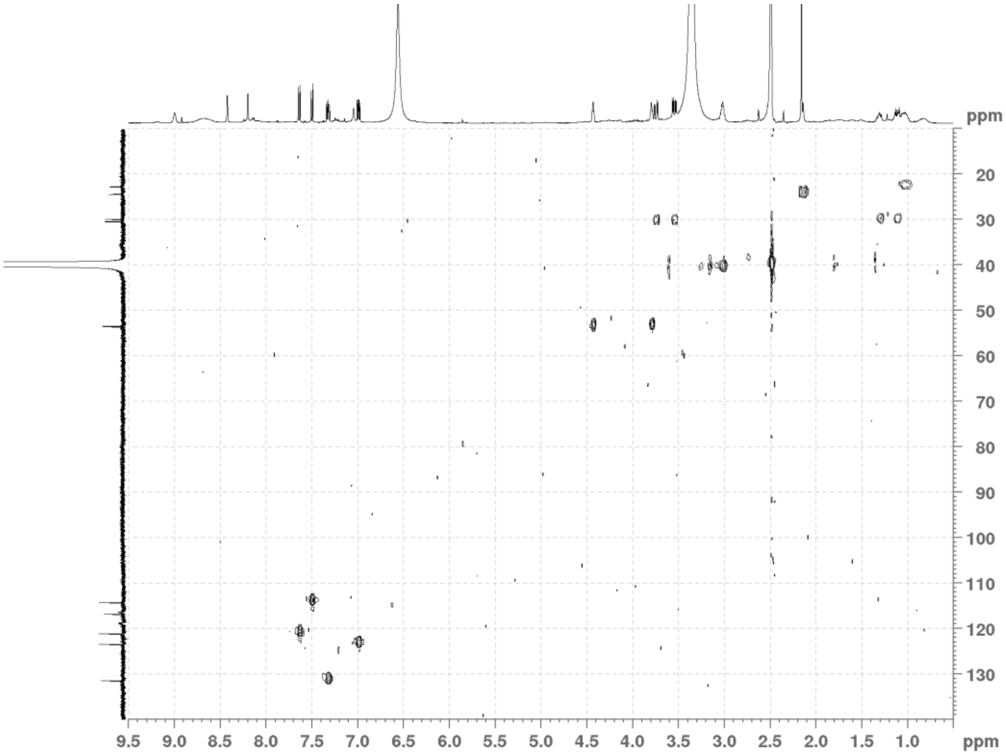


Figure S92. HMBC spectrum of compound **18** in DMSO-d_6 (500 MHz).

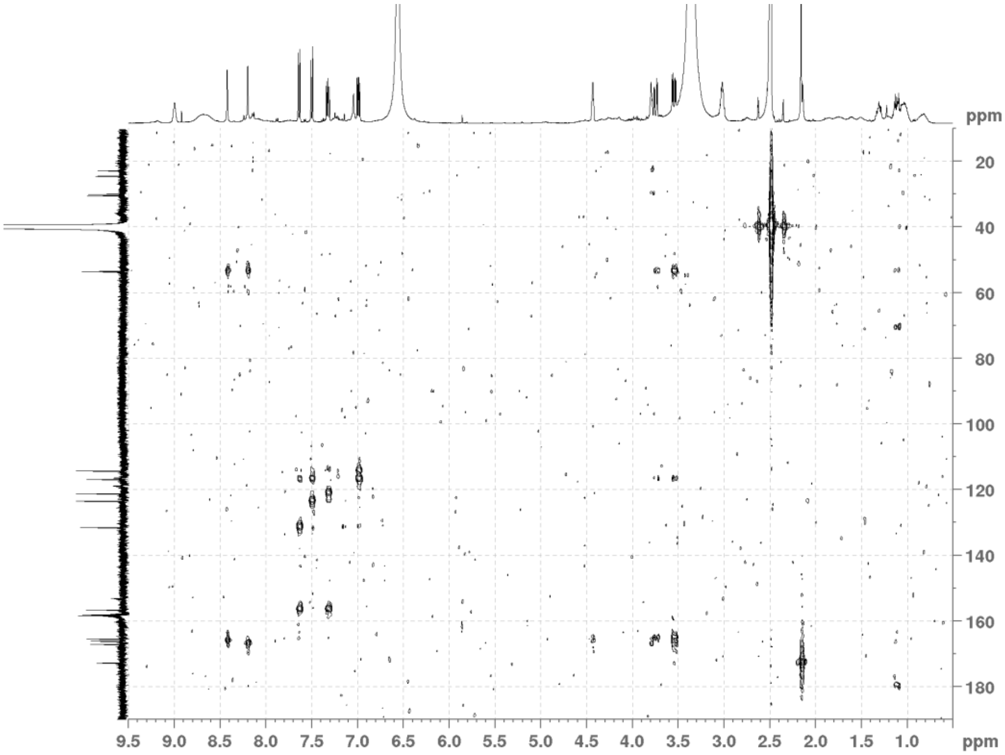
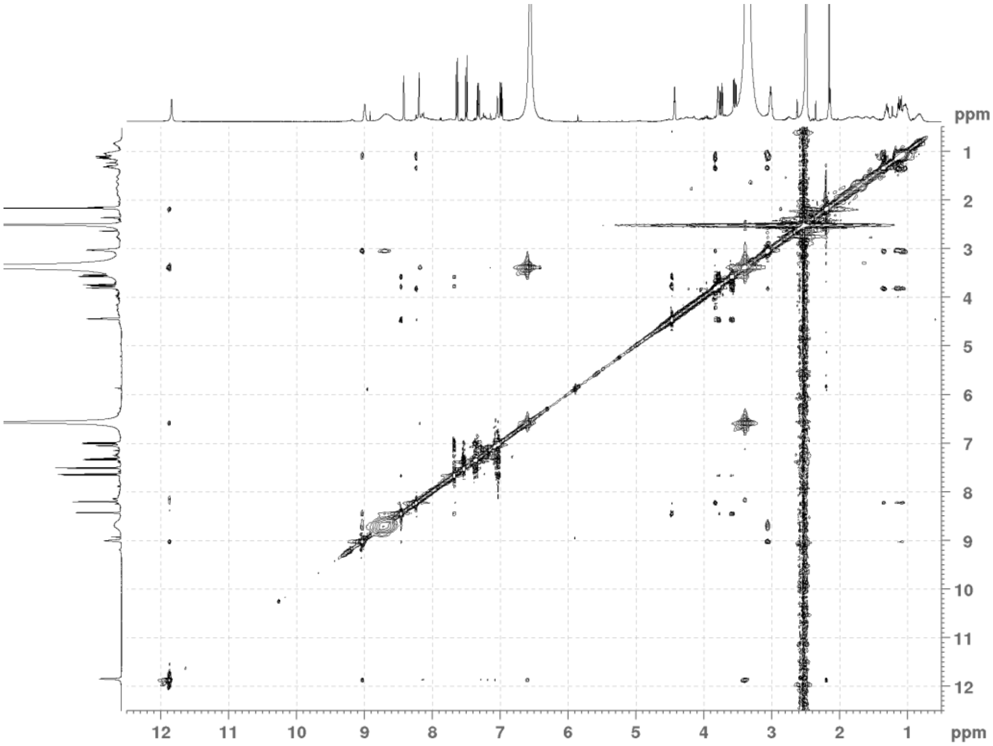
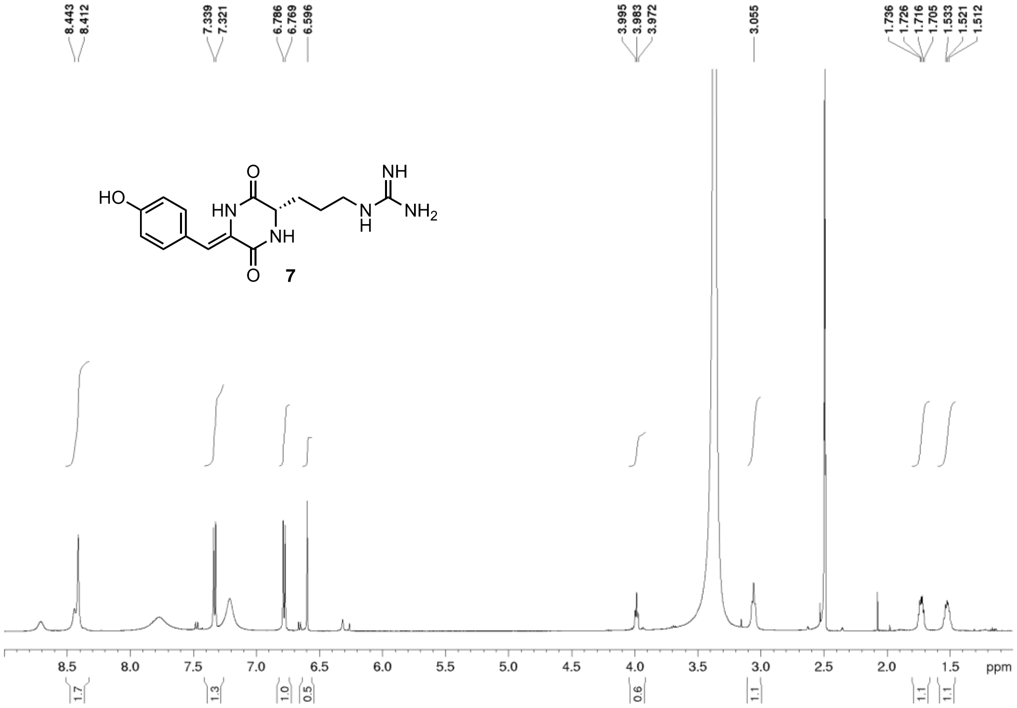


Figure S93. ROESY spectrum of compound **18** in DMSO-d_6 (500 MHz).



<u>^{1}H NMR spectra of previously isolated diketopiperazines</u>

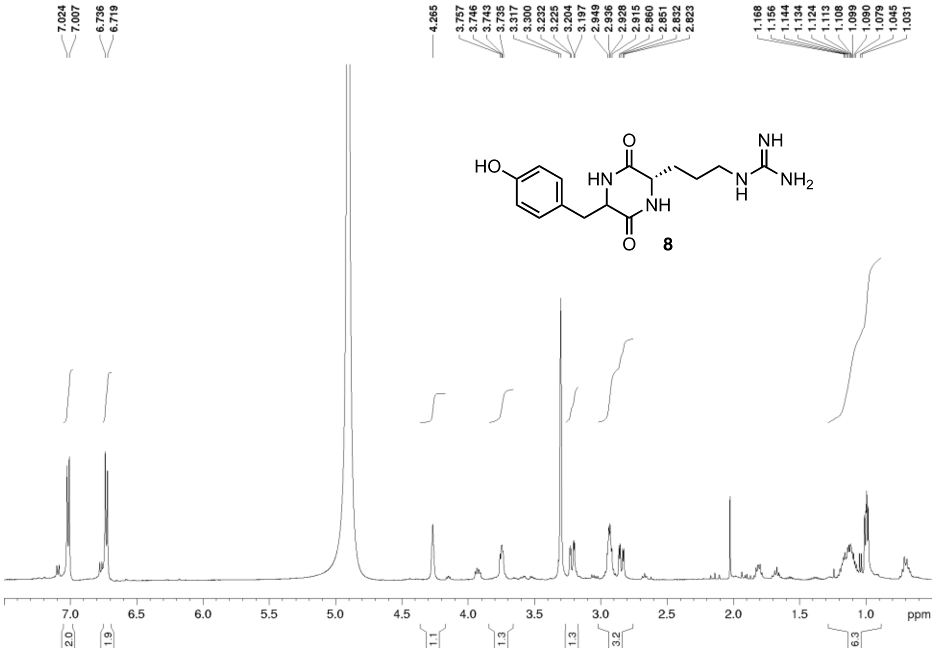
Figure S94. ^{1}H NMR spectrum of compound **7** in DMSO-d_6 (500 MHz).



Compound **7**: HRMS (ESI, M+H$^+$) calculated for $C_{15}H_{20}N_5O_3$ 318.1566; found 318.1557.

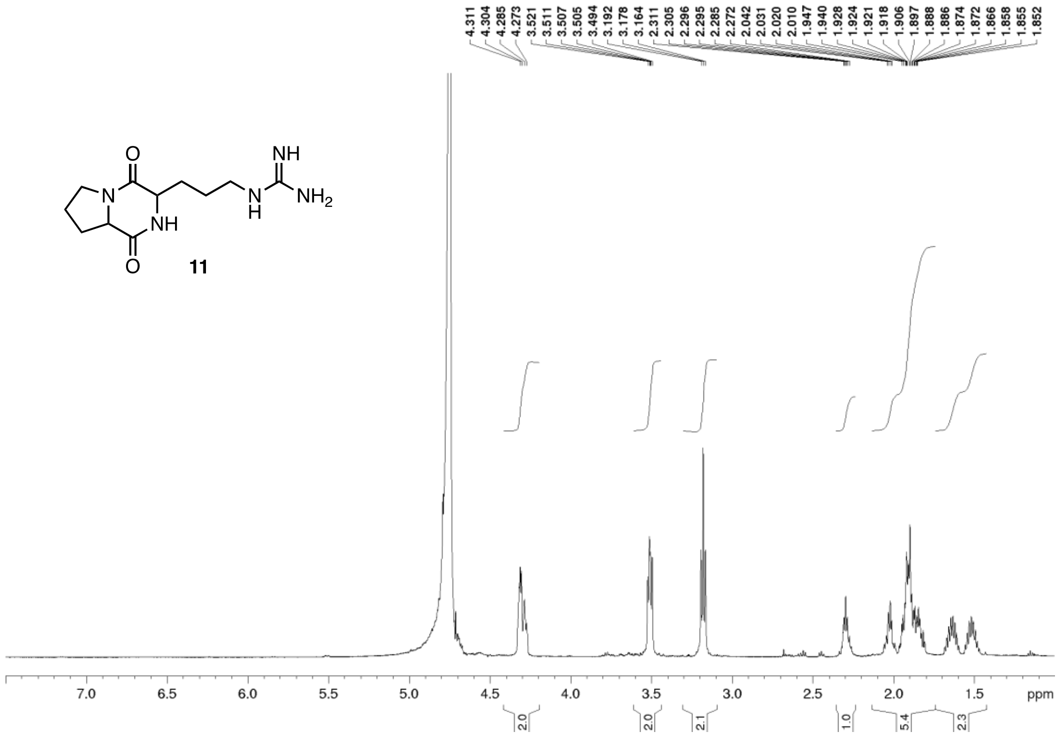
[α]$_D$: -48 (*c* 0.20, H_2O)

Figure S95. ^{1}H NMR spectrum of compound **8** in CD_3OD (500 MHz).



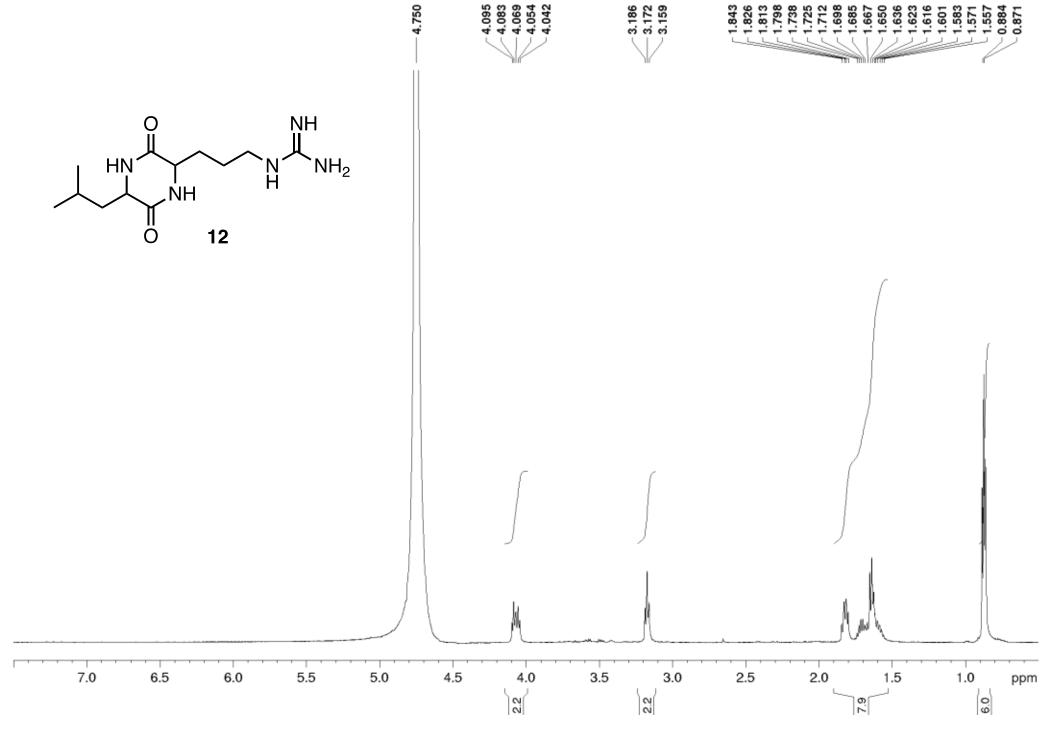
Compound **8**: HRMS (ESI, M+H$^+$) calculated for $C_{15}H_{22}N_5O_3$ 320.1723; found 320.1715.

Figure S96. ^{1}H NMR spectrum of compound **11** in D_2O (500 MHz).



Compound **11**: HRMS (ESI, M+H$^+$) calculated for $C_{11}H_{20}N_5O_2$ 254.1617; found 254.1595.

Figure S97. ^{1}H NMR spectrum of compound **12** in D_2O 0.5% TFA-*d* (500 MHz).



Compound **12**: HRMS (ESI, M+H$^+$) calculated for $C_{12}H_{24}N_5O_2$ 270.1930; found 270.1933.

REFERENCES

1. Tsukamoto, S., Kato, H., Hirota, H. & Fusetani, N. Pipecolate derivatives, anthosamines A and B, inducers of larval metamorphosis in ascidians, from a marine sponge Anthosigmella aff. raromicrosclera. *Tetrahedron* **51**, 6687–6694 (1995).
2. Tohyama, S. *et al.* Discovery and Characterization of NK13650s, Naturally Occurring p300-Selective Histone Acetyltransferase Inhibitors. *J. Org. Chem.* **77**, 9044–9052 (2012).
3. Li, X. *et al.* Determination of Absolute Configuration and Conformation of a Cyclic Dipeptide by NMR and Chiral Spectroscopic Methods. *J. Phys. Chem. A* **117**, 1721–1736 (2013).
4. Zhang, M., Meng, Y., Wu, Y. & Song, C. TfOH-Promoted Decyanative Cyclization toward the Synthesis of 2,1-Benzisoxazoles. *J. Org. Chem.* **86**, 7326–7332 (2021).
5. Mannam, S. & Sekar, G. An enantiopure galactose oxidase model: synthesis of chiral amino alcohols through oxidative kinetic resolution catalyzed by a chiral copper complex. *Tetrahedron Asymmetry* **20**, 497–502 (2009).
6. Bat-Erdene, U. *et al.* Iterative Catalysis in the Biosynthesis of Mitochondrial Complex II Inhibitors Harzianopyridone and Atpenin B. *J. Am. Chem. Soc.* **142**, 8550–8554 (2020).
7. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
8. Moutiez, M. *et al.* Unravelling the mechanism of non-ribosomal peptide synthesis by cyclodipeptide synthases. *Nat. Commun.* **5**, 5141 (2014).
9. Kreitler, D. F., Gemmell, E. M., Schaffer, J. E., Wencewicz, T. A. & Gulick, A. M. The structural basis of N-acyl-α-amino-β-lactone formation catalyzed by a nonribosomal peptide synthetase. *Nat. Commun.* **10**, 3432 (2019).

10. Koch, A. A. *et al.* A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macrocyclization Catalyst. *J. Am. Chem. Soc.* **139**, 13456–13465 (2017).
11. Vijayasarathy, S. *et al.* C3 and 2D C3 Marfey's Methods for Amino Acid Analysis in Natural Products. *J. Nat. Prod.* **79**, 421–427 (2016).