Electronic Supplementary Material (ESI) for Organic Chemistry Frontiers.

Discovery and Heterologous Production of Sarubicins and Quinazolinone C-Glycosides with Protecting Activity for Cardiomyocytes

Mingming Yu,‡[a,b] Jianying Luo,‡[a,b] Dan Luo,[c] Qiang He,[a] Yijun Yan,[a] Xu Ji,[d] and Sheng-Xiong Huang*[a]

[a] State Key Laboratory of Phytochemistry and Plant Resources in West China, and Yunnan Key Laboratory of Natural Medicinal Chemistry, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China.

[b] University of Chinese Academy of Sciences, Beijing 100049, China.

[c] College of Traditional Chinese Medicine, Yunnan University of Traditional Chinese Medicine, Kunming 650500, China.

[d] Key Laboratory of Medicinal Chemistry for Natural Resource of Ministry of Education, Yunnan University, Kunming 650091, China.

‡ M.Y. and J.L. contributed equally to this work.

*Corresponding Author. S.-X. Huang, Tel: +86-871-65215112. E-mail: sxhuang@mail.kib.ac.cn

Contents

Experiment Section

Tab. S1. Bacterial strains and plasmids used in this study

Strain and plasmid	Description	References or sources
E. coli		
DH5α	Host strain for cloning	Invitrogen
BW25113/pIJ790	Host strain for PCR-targeting	Ref. 1
ET12567/pUZ8002	Host strain for conjugation between *E. coli* and *Streptomyces*	Ref. 2
XL1-blueMR	Host strain for genomic library	Agilent Technologies
DH5α/BT340	Host strain for in-frame deletion	Ref. 3
Streptomyces		
S. sp. KIB-H91	Wild type sarubicins producing strain	This study
S. coelicolor M1154	Host strain for heterologous expression	Ref. 4
S. coelicolor M1154/p18F3	*Streptomyces coelicolor* M1154 integrated with plasmid 18F3 which contains *sar* biosynthetic gene cluster	This study
S. coelicolor 18F3 Δ*sarS3*	*sarS3* inactivation mutant of *S. coelicolor* 18F3	This study
S. coelicolor 18F3 Δ*sarS5*	*sarS5* inactivation mutant of *S. coelicolor* 18F3	This study
S. coelicolor 18F3 Δ*sarO*	*sarO* inactivation mutant of *S. coelicolor* 18F3	This study
Cosmid		
pJTU2554	Apr^r, Cosmid vector for genomic library construction	Ref. 1
pJTU6722	Ery^r, Vector for PCR targeting	Constructed by Prof. Meifeng Tao
p18F3	Apr^r, Cosmid which contains sar biosynthetic gene cluster	This study
p18F3_Δ*sarS3*	Apr^r, gene inactivation clone used for sarS3 mutant	This study
p18F3_Δ*sarS5*	Apr^r, gene inactivation clone used for sarS5 mutant	This study
p18F3_Δ*sarO*	Apr^r, gene inactivation clone used for sarO mutant	This study

Tab. S2. The primers used in this study

gene	Sequence (5′ to 3′)
For gene Screening	
gra-Orf14	14F: GCTGGAGCACATCGCGTTCA
	14R: CACGAGCTTGGGCATGTTGG
gra-Orf17	17F: TCGTGGTGAAGAGCGGGATC
	17R: GCGGCTGACCTTGGTGAAA
For gene disruption	
sarS3	5152F：GCCACGGCCGCACGGAACGCGGGGCACCAGGTCTTCATGattccggggatcc gtcgac
	5152R：TCTCGCAGGCGGCCACCACGGCGTCCTGCGAGTCCTCCCtgtaggctggag ctgcttc
sarS5	5158F：CGCCACCGTGGAACTGGACGCCGTGTACGTCCCGCTGCCattccggggatcc gtcgac
	5158R：CGGGTGAACGCCCGGTCCACCCGGACCGTGCCGTGTGCCCtgtaggctggag ctgcttc
sarO	5154F：TTGGAGACCCAGAACCTCATGGTCAAGACAGGCATCTCGattccggggatcc gtcgac
	5154R：CTCAGGTGCCCGGCCGGGCTCCGCCCGTTTGATCCCGCGCtgtaggctggag ctgcttc
For gene confirmation	
sarS3	5152YF：GAAGTTCCCTGGGTATTCC
	5152YR：CGCCCCGGCCAGCGCTTCG
sarS5	5158YF：CGCCGGATGCTGCCGGCC
	5158YR：GGTGCGCACGGCCGTCAC
sarO	5154YF：CCGTCGGAAGGTATTCGAC
	5154YR：CATGTACGTGATAGCCAAC

Tab. S3. Deduced functions of ORFs in *sar* BGC

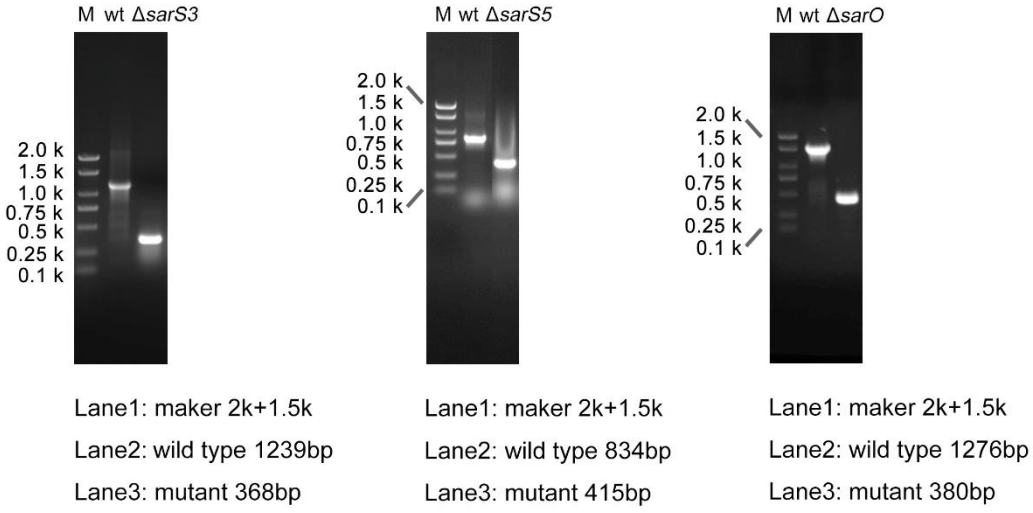
Gene NO.	Size aa	Proposed function by a BLAST search (blastp) Protein, Origin	Positives/ Identities	Accession No.
sarN	610	ATP-dependent amidotransferase, TtmN [*Streptomyces afghaniensis*]	66/55	ALJ49932.1
sarR1	391	MFS transporter, partial, *Streptomyces acidiscabies*	74/62	WP_040837276.1
sarA1	390	DHAP synthase, PhzC *[Streptomyces tendae]*	72/62	WP_028424979.1
sarA2	269	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase DhbX [*Klebsiella oxytoca*]	61/46	AFS18549.1
sarA3	228	Isochorismatase, PhzD [*Streptomyces tendae*]	71/64	AFS18573.1
sarA4	623	2-Amino-2-desoxyisochorismate Synthase PhzE [*Streptomyces tendae*]	73/62	AFS18572.1
sarU1	116	hypothetical protein, *Streptomyces gancidicus* BKS 13-15	74/63	WP_006135773.1
sarH	396	3-hydroxybenzoate-6-hydroxylase,3HB6H [*Rhodococcus jostii*]	46/33	WP_011594786.1
sarS1	323	dTDP-glucose 4,6 dehydratase, StrE *[Streptomyces griseus]*	72/66	WP_030951601.1
sarS2	329	dTDP-1-glucose thymidylyltransferase, StrD *[Streptomyces griseus]*	69/61	CAH94331.1
sarS3	344	*C*-glycosyltransferase,UrdGT2 [*Streptomyces fradiae*]	61/52	AAF00209.1
sarR2	200	SARP family transcriptional regulator, SrrZ [*Streptomyces sp*. F-3]	61/48	WP_003951745.1
sarO	311	Dehydrogenase, *Nocardia* sp. BMG51109	71/57	WP_036571061.1
sarU2	101	Hypothetical protein, *Amycolatopsis* sp. MJM2582	61/43	WP_037347854.1
sarR3	492	Arabinose ABC transporter permease, *Bacillus megaterium*	78/59	WP_034679354.1
sarS4	434	Putative NDP-deoxyglucose-2,3-dehydratase, SpnO [*Saccharopolyspora spinosa*]	60/47	AAG23276.1
sarS5	326	dTDP-3,4-diketo-2,6-dideoxy-D-glucose 3-ketoreductase, SpnN [*Saccharopolyspora spinosa*]	61/49	AAG23275.1
sarR4	151	MerR family transcriptional regulator, *Microbispora* sp. ATCC PTA-5024	80/67	WP_036330534.1
sarR5	384	Transposase, *Streptomyces himastatinicus*	61/51	WP_039941984.1
sarR6	652	Transcriptional regulator, *Streptomyces* sp. NRRL F-5065	62/49	WP_037858633.1

Tab. S4. Protective effects of compounds **1** and **2** against H_2O_2-induced cardiomyocytes injury

Compound	Concentration[a]	
	0 μM	10 μM
1	49.03±3.89	81.36±4.84**
2	49.03±3.89	89.43±2.58**
Carvedilol[b]	49.03±3.89	85.98±6.18**

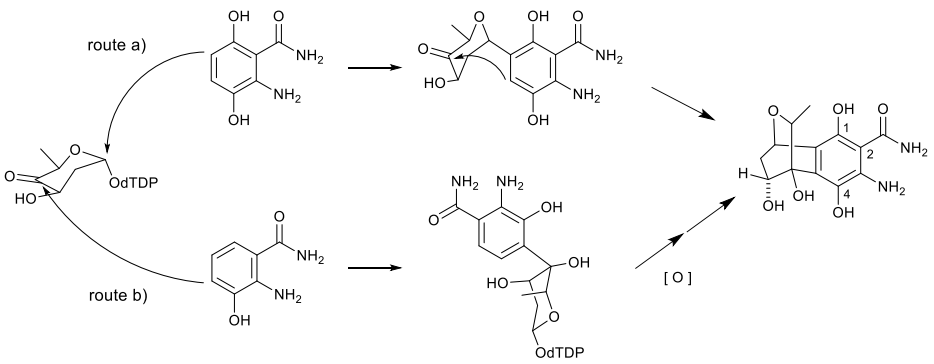
Date expressed as means ±SD (n = 3). [a] ** p<0.01 vs 0 μM group. [b] Positive control.

Fig. S1. Confirmation of mutants by PCR using the primers.



Confirmation of mutants *S. coelicolor* p18F3 Δ*sarS3*, p18F3 Δ*sarS5*, and p18F3 Δ*sarO* by PCR using the primers listed in supplementary Tab.S2.

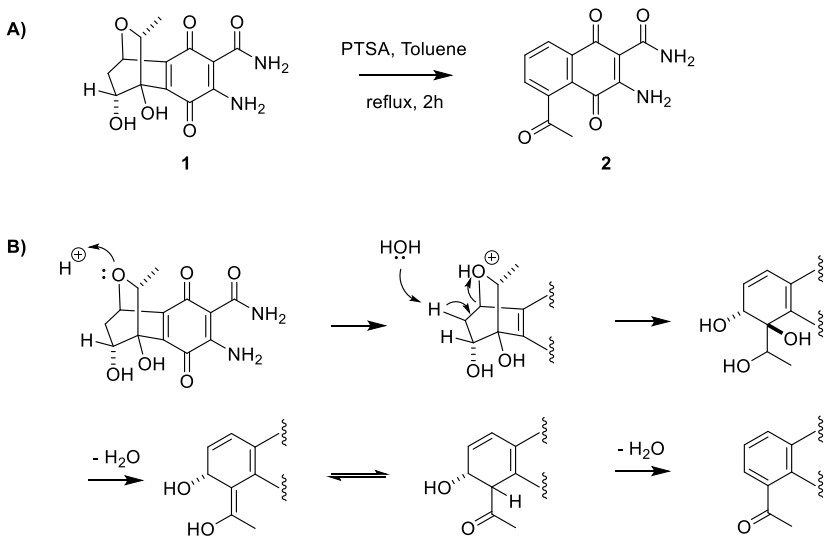
Fig. S2. Two possible routes of *C*-glycosylation in the biosynthesis of sarubicin A



Route a): glycosyl transfer at C-6 of hydroquinone; then intramolecular aldol condensation between C-5 and C-4'.

Route b): intermolecular aldol condensation between C-5 and C-4'; followed by hydroxylation at C-1 of 3HAA; then cyclization between C-6 and C-1'.

Scheme S1. A) Transformation of **1** to **2**; B) Proposed mechanism of transformation from **1** to **2**.



Tab. S5. 1D and 2D NMR data of compound **3** (H9101)

Position	H9101			
	C Signals	**H Signals**	**HMBC Correlations**	**ROESY Correlations**
	δ_C, type	δ_H, (int. mult. *J* in Hz)		
2	152.9 (s)			
2-Me	21.5 (q)	2.36 (3H, s)	2	
4	167.3 (s)			
4a	102.6 (s)			
5	147.1 (s)			
6	124.3 (s)			
7	117.9 (d)	7.18 (1H, s)	1', 5, 6, 8, 8a	2'-H_{aix}
8	143.4 (s)			
8a	135.5 (s)			
1'	70.2 (d)	4.68 (1H, d, 11.4)	5', 5, 6, 7	5', 3'
2'-aix	40.5 (t)	1.31 (1H, pseudo-q, 12.6)	1', 3', 7'	4', 7
2'-eq		2.09 (1H, dd, 12.6, 3.0)	3', 4'	
3'	71.8 (d)	3.48 (1H, overlapped)		1'
4'	77.2 (d)	2.85 (1H, t, 9.0)	3', 5', 6'	2'-H_{aix}, 6'
5'	76.1 (d)	3.32 (1H, dq, 9.0, 6.0)		1'
6'	18.5 (q)	1.22 (3H, d, 6.0)	5'	4'
Exchangeable		4.72-5.30 (2H)		
Protons		11.58 (5-OH, s)		

a) Data were recorded at 600 MHz for ^{1}H and 150 MHz for ^{13}C in DMSO-d_6 using the corresponding solvent residual signal as internal standard.

Tab. S6. 1D and 2D NMR data of compound **4** (H9102)

Position	H9102 [a]			
	C Signals	**H Signals**	**HMBC Correlations**	**ROESY Correlations**
	δ_C, type	δ_H, (int. mult. *J* in Hz)		
2	152.1 (s)			
2-CH_3	21.4 (q)	2.37 (3H, s)	2	
4	167.3 (s) [b]			
4a	106.1 (s)			
5	147.0 (s)			
6	125.0 (s)			
7	118.5 (d)	7.28 (1H, s)	1', 5, 6, 8, 8a	2'-H_{aix}
8	144.1 (s)			
8a	135.5 (s) [b]			
1'	70.4 (d)	4.63 (1H, d, 11.4)	2', 3', 5', 5, 6, 7	5', 3'
2'-aix	34.7 (t)	1.71 (1H, dd, 12.0, 2.4)	3'	7
2'-eq		1.56 (1H, brd, 12.0)	1', 6	
3'	69.3 (d)	3.69 (1H, brd, 11.4)		1'
4'	69.9 (d)	3.41 (1H, overlapped)	2', 3'	6'
5'	74.1 (d)	3.58 (1H, q, 6.0)		1'
6'	17.6 (q)	1.22 (3H, d, 6.0)	4', 5'	4'
Exchangeable		4.12 (1H, brs)		
Protons		4.30 (1H, brs)		
3-NH_b		8.93 (1H, brs)		
5-OH		11.50 (1H, s)		
3-NH_a		12.48 (1H, brs)		

a) Data were recorded at 600 MHz for ^1H and 150 MHz for ^{13}C in DMSO-d_6 using the corresponding solvent residual signal as internal standard; b) The signals were detected by the HMBC spectrum.

Tab. S7. 1D and 2D NMR data of compound **5** (H9103)

Position	H9103 [a]			
	C Signals	**H Signals**	**HMBC Correlations**	**ROESY Correlations**
	δ_C, type	δ_H, (int. mult. *J* in Hz)		
2	152.5 (s) [b]			
2-Me	21.4 (q)	2.37 (3H, s)	2	
4	167.1 (s)			
4a	106.4 (s)			
5	147.2 (s)			
6	122.7 (s)			
7	117.8 (d)	7.32 (1H, s)	1', 5, 8a	2'-H_{aix}
8	144.0 (s)			
8a	136.3 (s) [b]			
1'	70.2 (d)	4.92 (1H, dd, 12.0, 1.2)	5', 5, 6, 7	5'
2'-aix	47.0 (t)	2.73 (1H, pseudo-t, 12.6)	1', 3', 7'	4', 7
2'-eq		2.56 (1H, dd, 12.6, 1.2)	3'	
3'	206.4 (s)			
4'	78.4 (d)	3.88 (1H, d, 9.6)	3', 5', 6'	2'-H_{aix}, 6' (w)
5'	78.3 (d)	3.58 (1H, dq, 9.6, 6.0)		1'
6'	19.3 (q)	1.37 (3H, d, 6.0)	5'	4' (w)
Exchangeable		4.72-5.30 (2H)		
Protons		11.56 (5-OH, s)		

a) Data were recorded at 600 MHz for ^{1}H and 150 MHz for ^{13}C in DMSO-d_6 using the corresponding solvent residual signal as internal standard; b) The signals were detected by the HMBC spectrum.

NMR and HRESIMS spectra of compounds 3-5 and 8-10

Fig. S3. HRESIMS spectrum of **3**

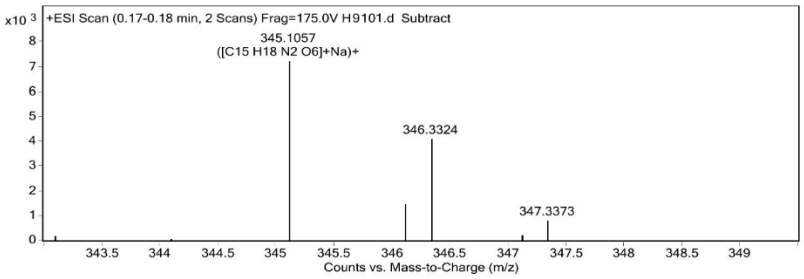
Qualitative Analysis Report

Data Filename	H9101.d	**Sample Name**	H9101
Sample Type	Sample	**Position**	P1-A3
Instrument Name	Instrument 1	**User Name**	
Acq Method	s.m	**Acquired Time**	3/11/2019 11:45:01 AM
IRM Calibration Status	Success	**DA Method**	Default.m
Comment			

Sample Group	**Info.**
Acquisition SW Version	6200 series TOF/6500 series Q-TOF B.05.01 (B5125.2)

User Spectra

Fragmentor Voltage	Collision Energy	Ionization Mode
175	0	ESI



Peak List

m/z	z	Abund	Formula	Ion
169.0111	1	8736.38		
274.2746	1	26216.83		
302.3058	1	6251.86		
318.301	1	19078.77		
345.1057	1	7243.87	C15 H18 N2 O6	(M+Na)+
437.1942	1	23759.44		
438.1971	1	5972.08		
453.1677	1	23368.36		
454.1712	1	4665.22		
659.289	1	4408.28		

Formula Calculator Element Limits

Element	Min	Max
C	3	60
H	0	120
O	0	30
N	0	5

Formula Calculator Results

Formula	CalculatedMass	CalculatedMz	Mz	Diff. (mDa)	Diff. (ppm)	DBE
C15 H18 N2 O6	322.1165	345.1057	345.1057	0.00	0.00	8.0000

--- End Of Report ---

Fig. S4. ^{1}H NMR (600 MHz) spectrum of **3** in DMSO-d_6

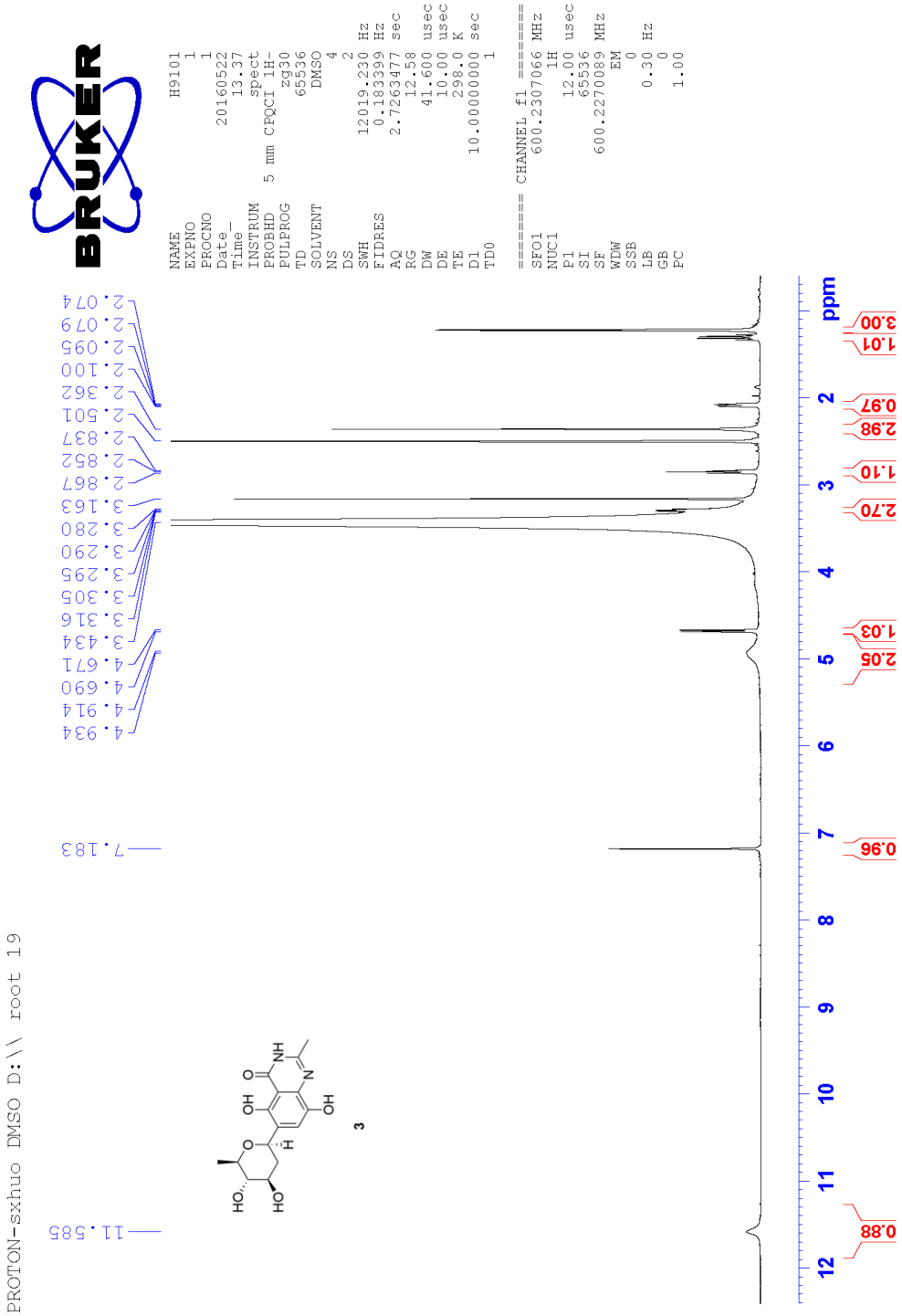


Fig. S5. ^{13}C NMR (150 MHz) spectrum of **3** in DMSO-d_6

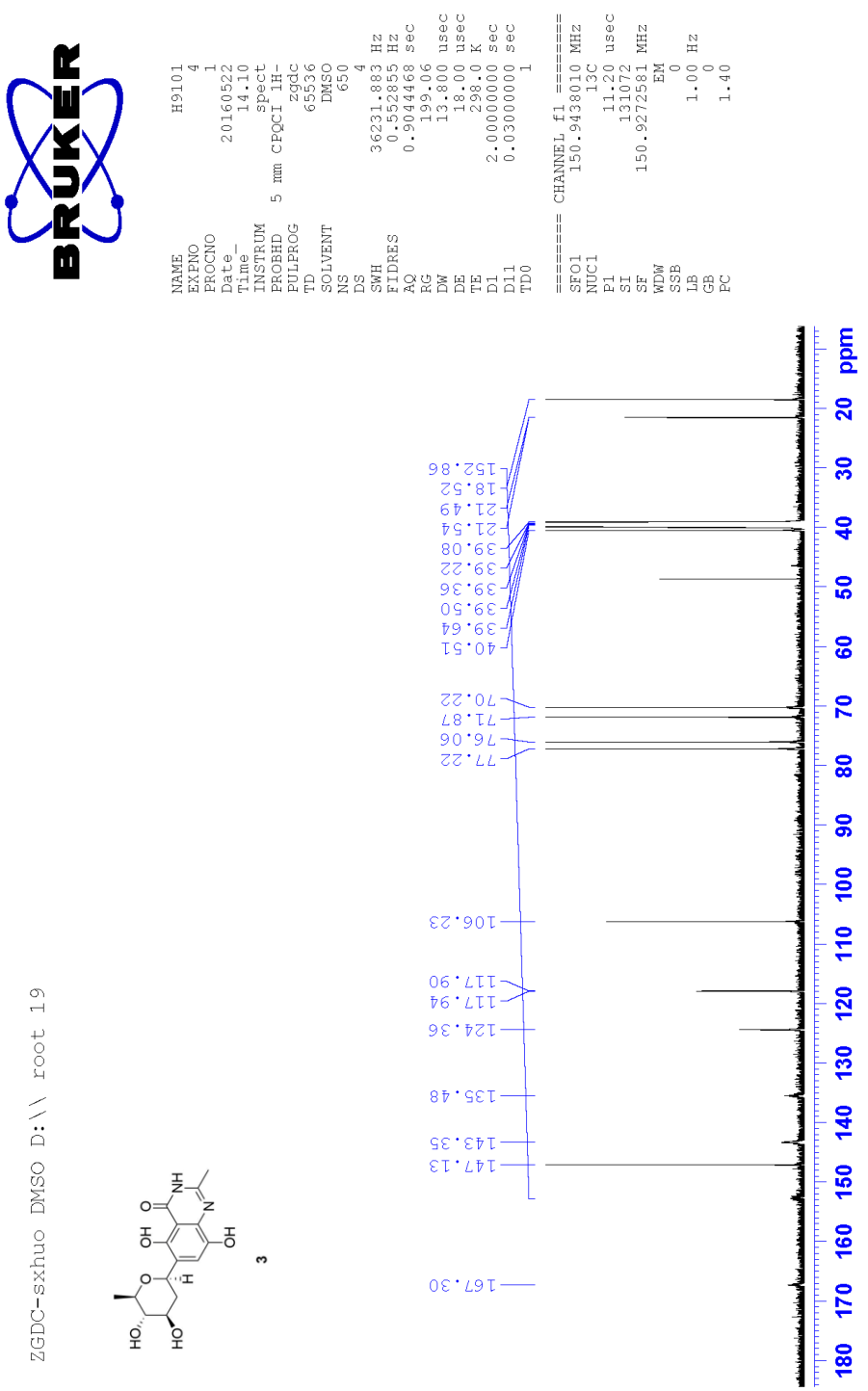


Fig. S6. HSQC (600 MHz) spectrum of **3** in DMSO-d_6

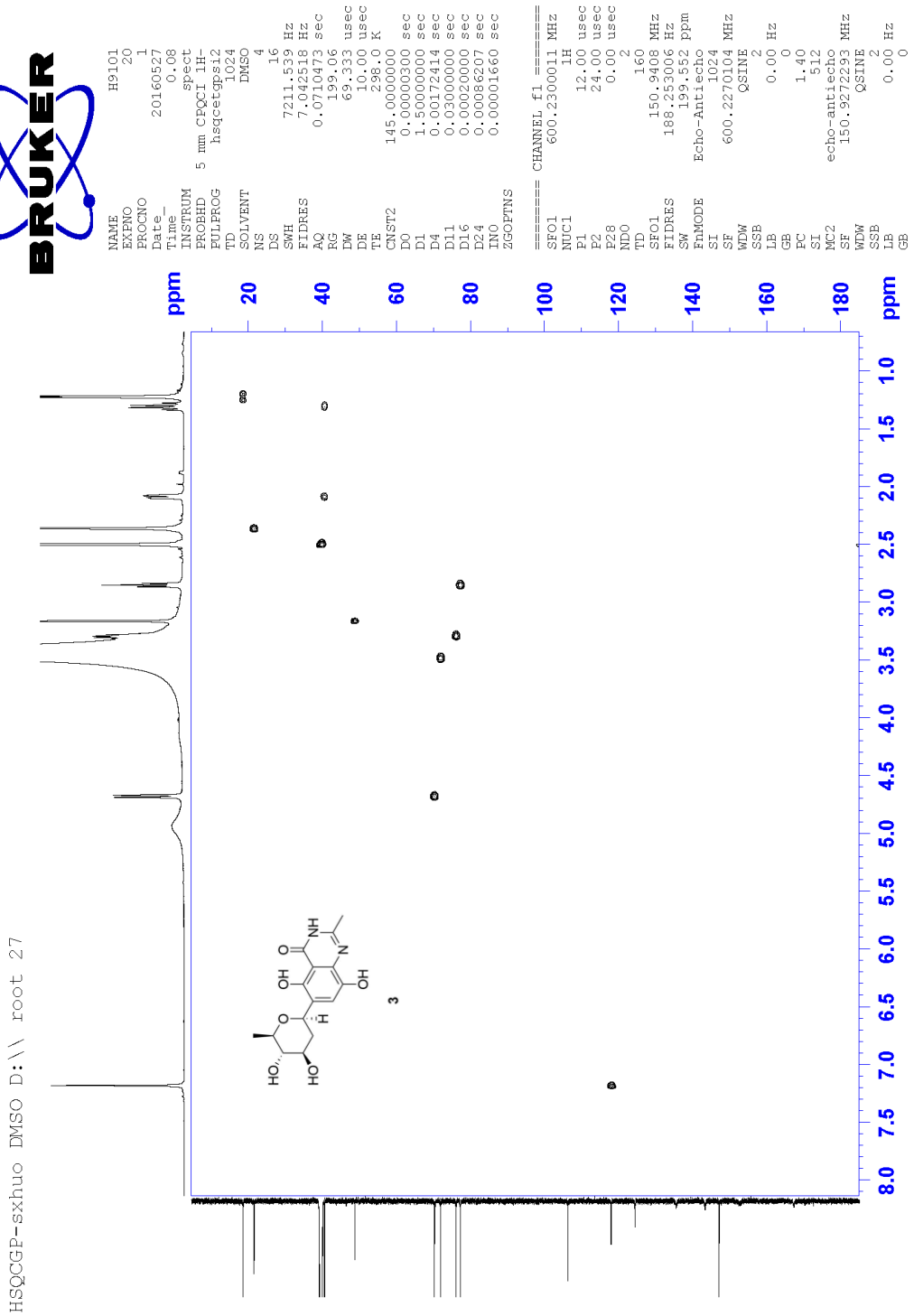


Fig. S7. HMBC (600 MHz) spectrum of **3** in DMSO-d_6

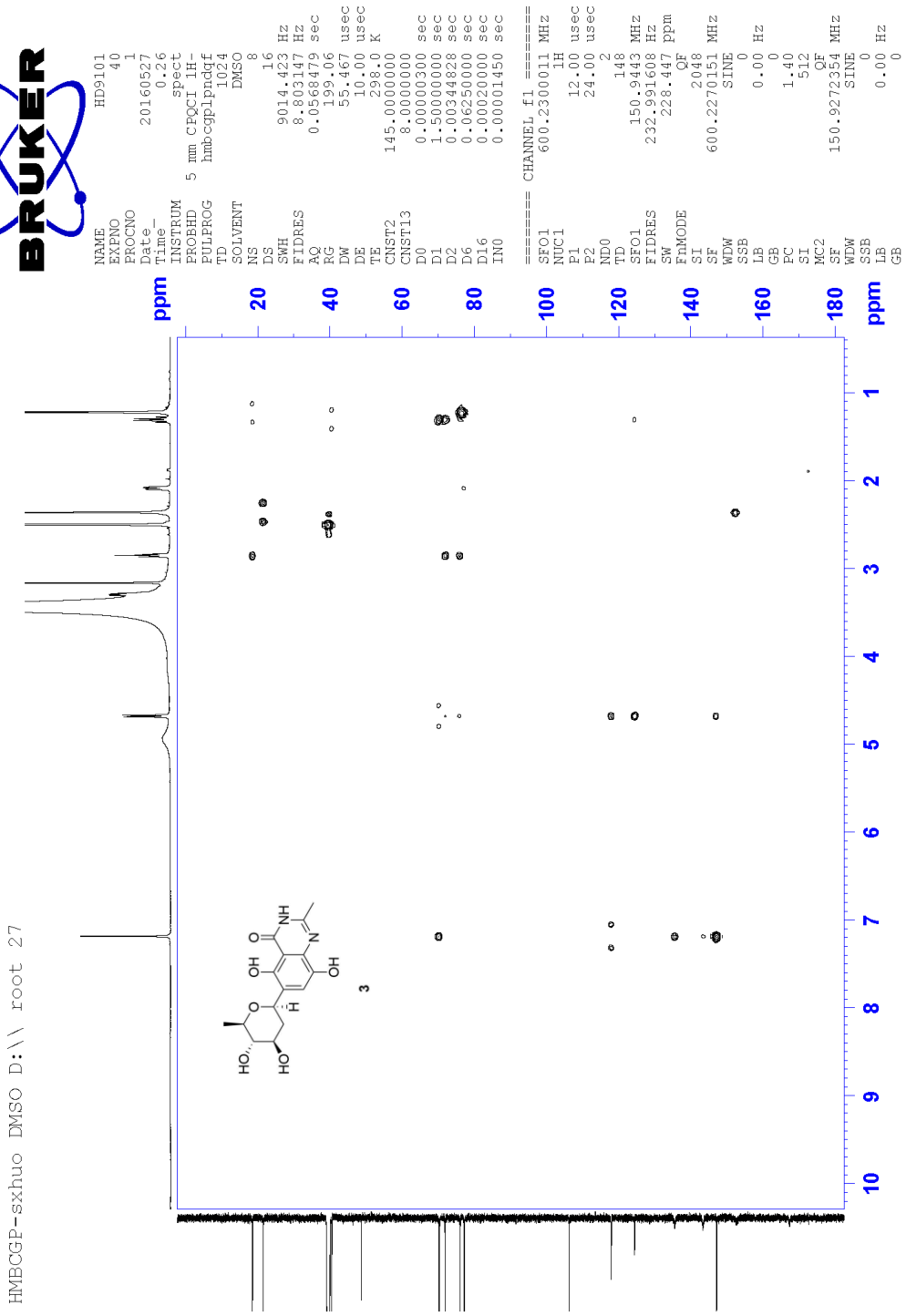


Fig. S8. ROESY (600 MHz) spectrum of **3** in DMSO-d_6

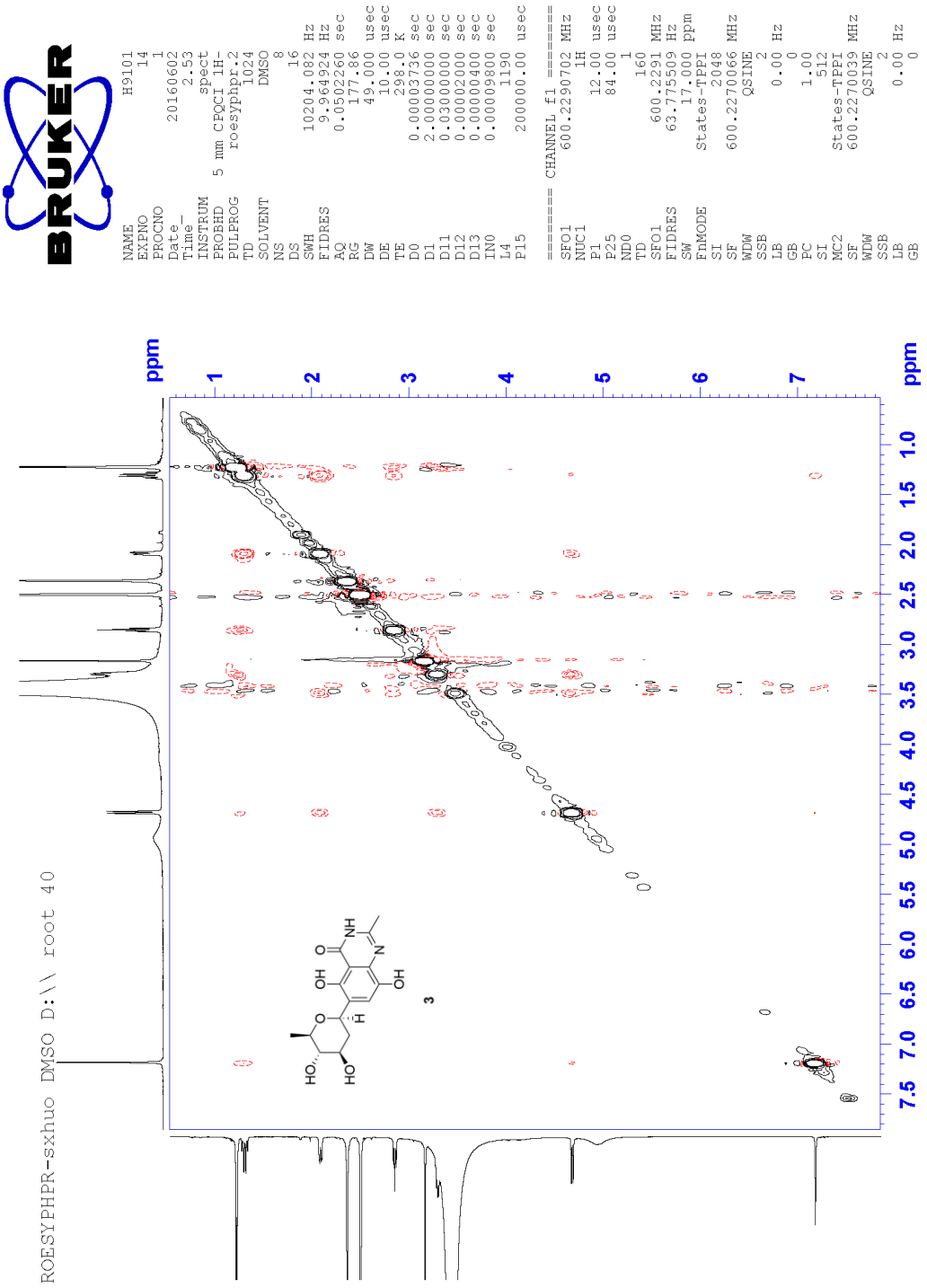


Fig. S9. HRESIMS spectrum of **4**

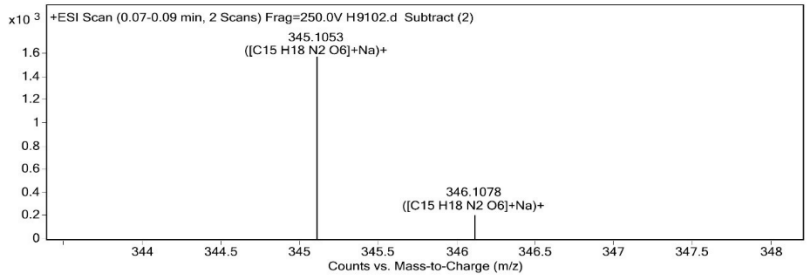
Qualitative Analysis Report

Data Filename	H9102.d	**Sample Name**	H9102
Sample Type	Sample	**Position**	P1-C2
Instrument Name	Instrument 1	**User Name**	
Acq Method	s.m	**Acquired Time**	3/11/2019 2:29:41 PM
IRM Calibration Status	Success	**DA Method**	Default.m
Comment			

Sample Group **Info.**
Acquisition SW Version 6200 series TOF/6500 series Q-TOF B.05.01 (B5125.2)

User Spectra

Fragmentor Voltage 250 **Collision Energy** 0 **Ionization Mode** ESI



Peak List

m/z	z	Abund	Formula	Ion
285.0241		218.58		
330.8971		287.57		
337.0899		270.67		
345.1053	1	1573.57	C15 H18 N2 O6	(M+Na)+
346.1078	1	215.08	C15 H18 N2 O6	(M+Na)+
367.0887		294.34		
399.1062		238.65		
437.1944	1	2455.96		
438.1994	1	618.82		
659.2844	1	302.24		

Formula Calculator Element Limits

Element	Min	Max
C	3	60
H	0	60
O	0	10
N	0	3

Formula Calculator Results

Formula	CalculatedMass	CalculatedMz	Mz	Diff. (mDa)	Diff. (ppm)	DBE
C15 H18 N2 O6	322.1165	345.1057	345.1053	0.40	1.16	8.0000

--- End Of Report ---

Fig. S10. ^{1}H NMR (600 MHz) spectrum of **4** in DMSO-d_6

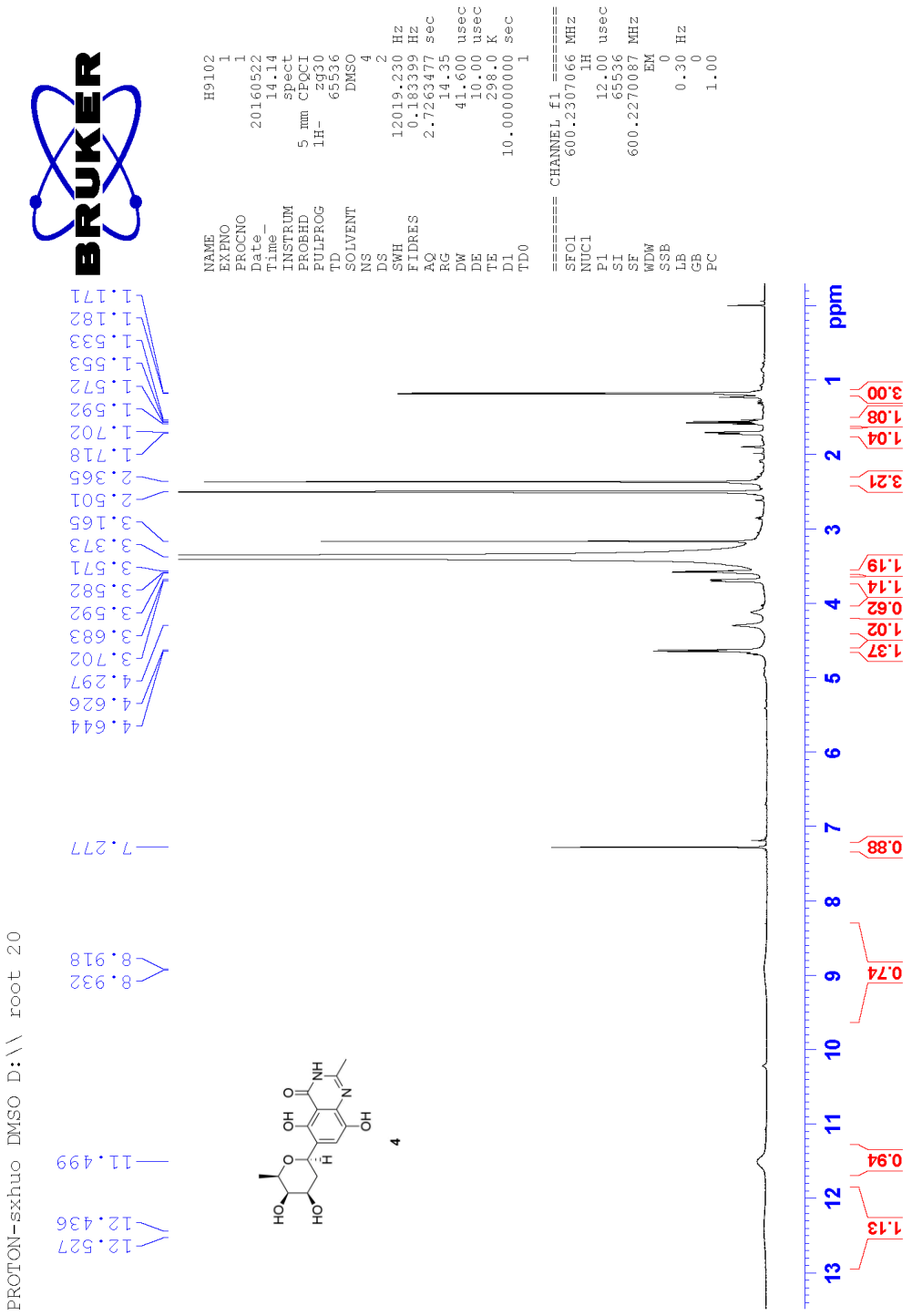


Fig. S11. ^{13}C NMR (150 MHz) spectrum of **4** in DMSO-d_6

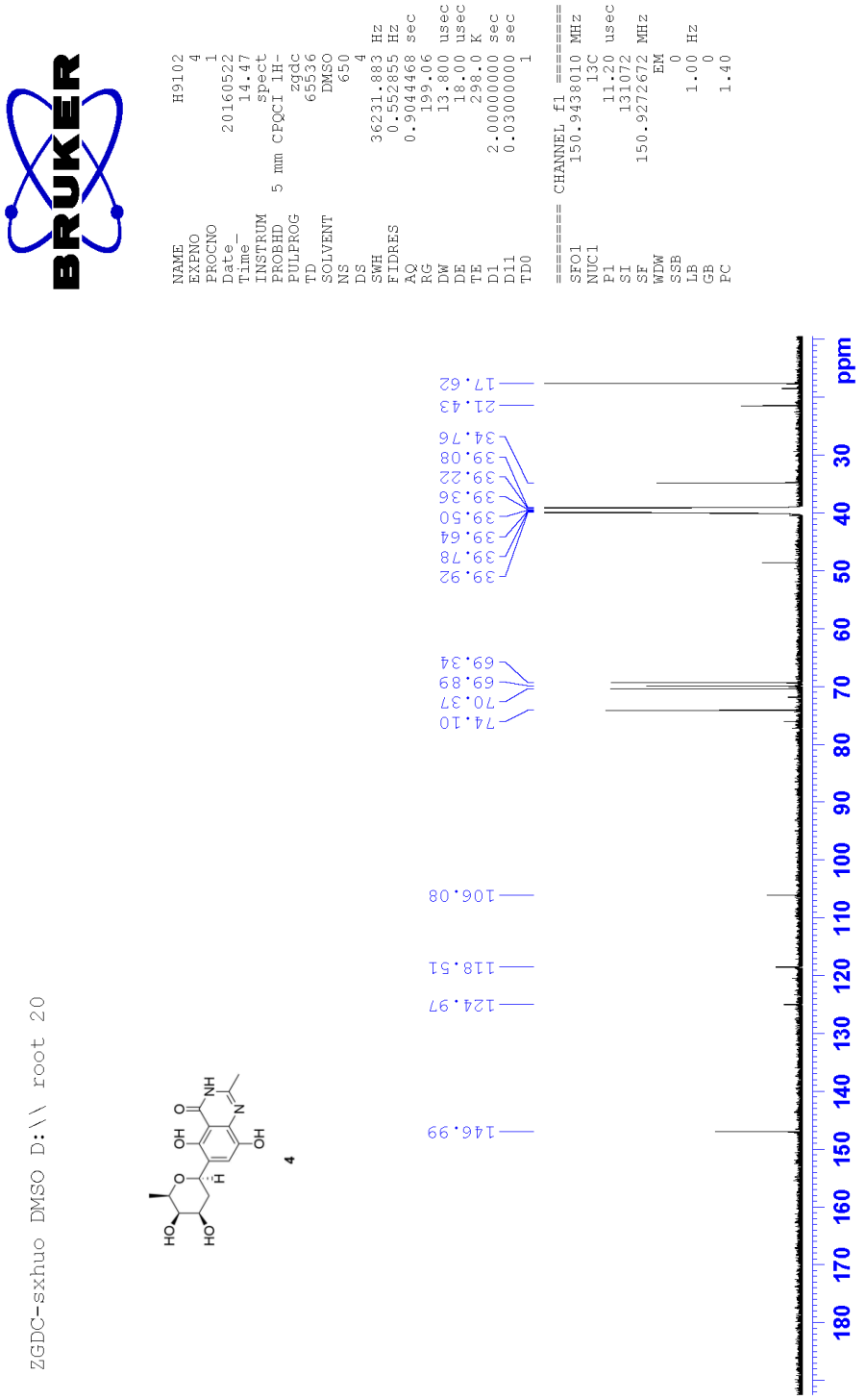


Fig. S12. HSQC (600 MHz) spectrum of **4** in DMSO-d_6

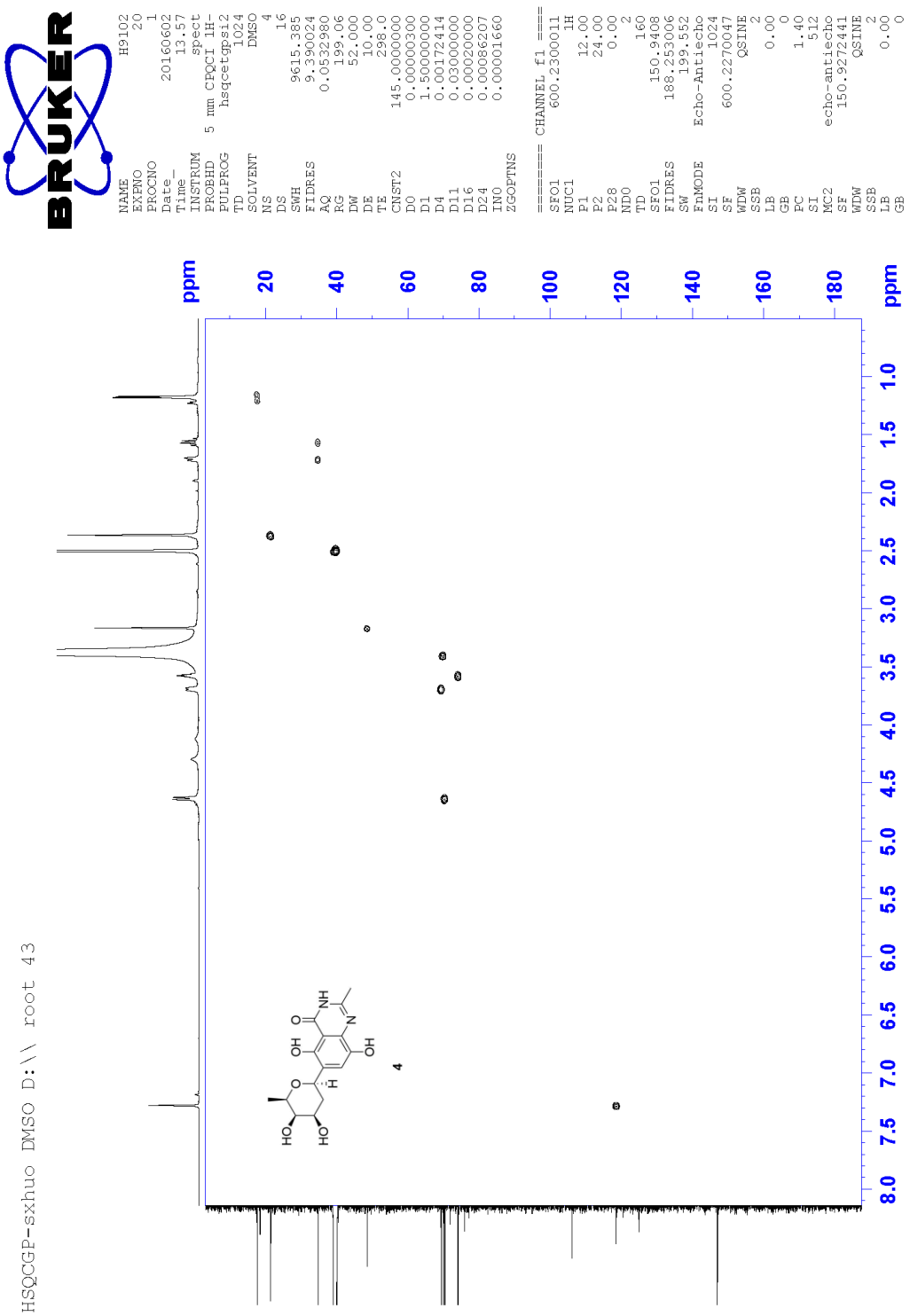


Fig. S13. HMBC (600 MHz) spectrum of **4** in DMSO-d_6

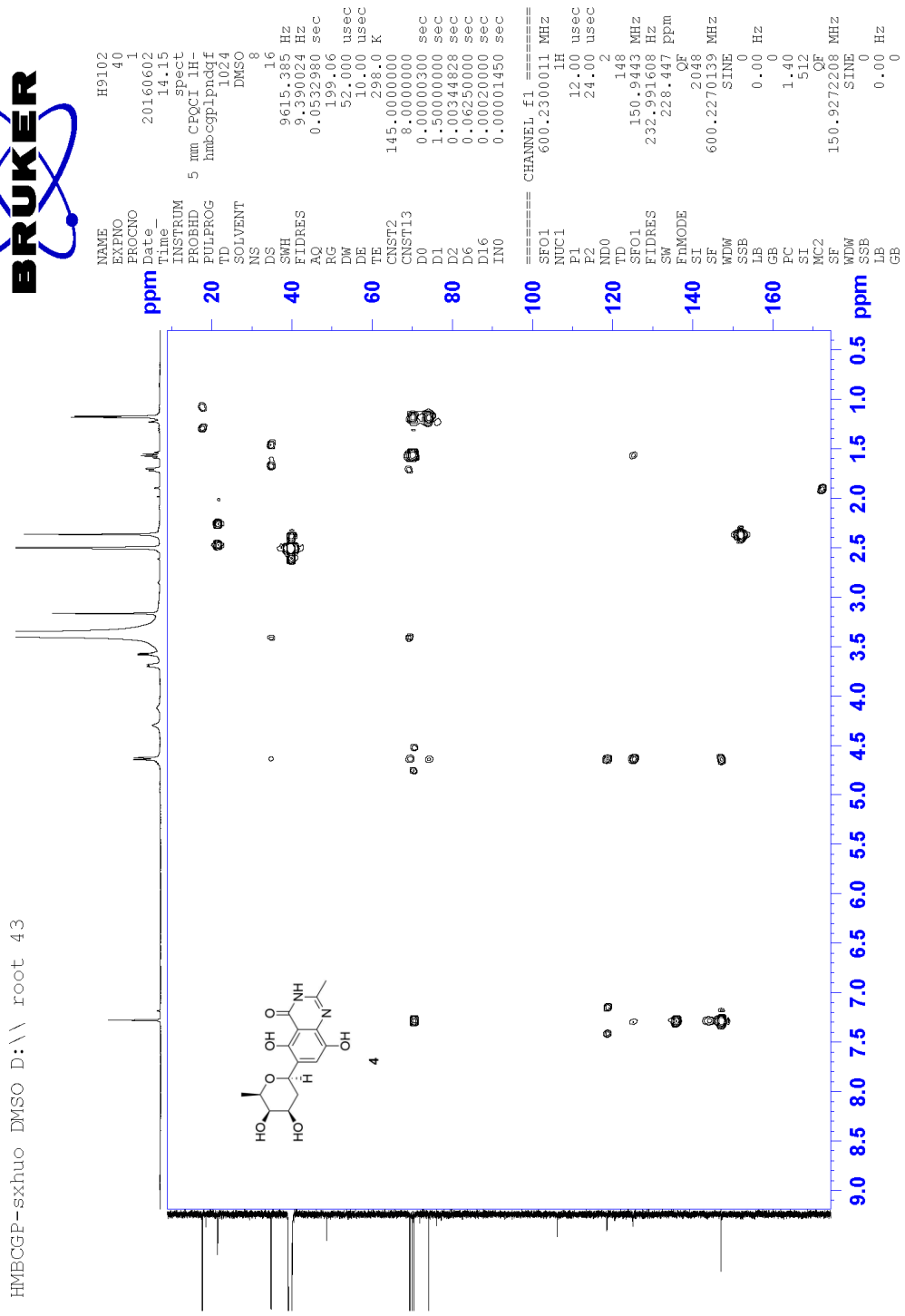


Fig. S14. ROESY (600 MHz) spectrum of **4** in DMSO-d_6

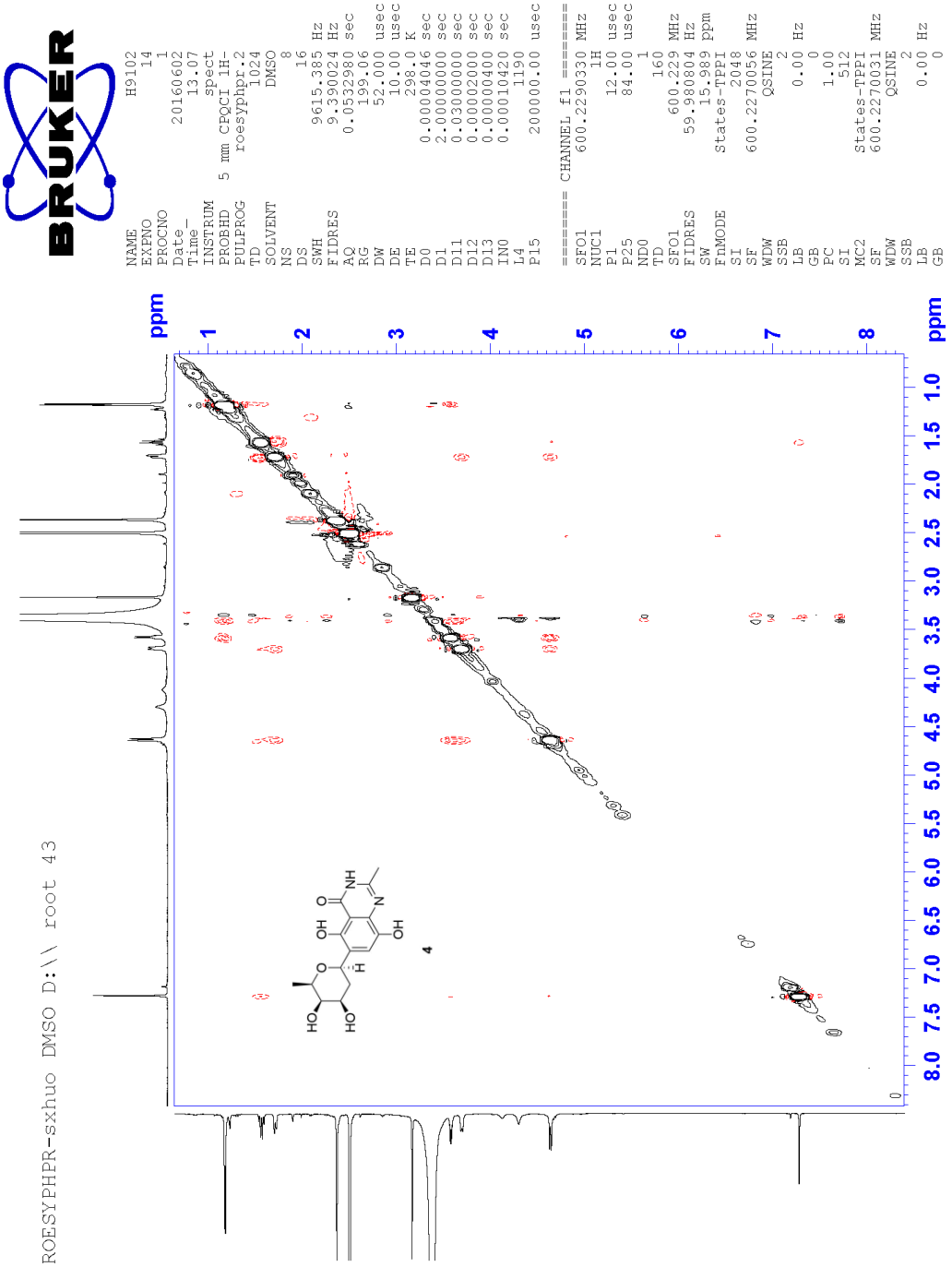


Fig. S15. HRESIMS spectrum of **5**

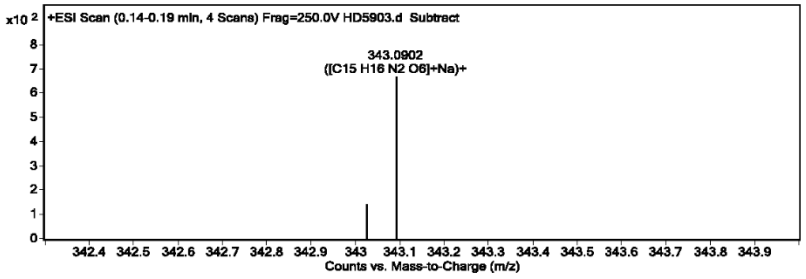
Qualitative Analysis Report

Data Filename	H9103.d	**Sample Name**	H9103
Sample Type	Sample	**Position**	P1-A1
Instrument Name	Instrument 1	**User Name**	
Acq Method	s.m	**Acquired Time**	3/20/2019 4:08:33
IRM Calibration Status	Success	**DA Method**	PM Default.m
Comment			

Sample Group	**Info.**
Acquisition SW Version	6200 series TOF/6500 series Q-TOF B.05.01 (B5125.2)

User Spectra

Fragmentor Voltage	Collision Energy	Ionization Mode
250	0	ESI



Peak List

m/z	z	Abund
285.0217	1	2646.84
299.1254	1	2326.36
345.1056	1	19489.03
346.1087	1	2973.38
371.1583	1	2529.09
437.1937	1	54439.48
438.1968	1	14210.92
453.1679	1	2502.27
659.2874	1	2891.54
675.2608	1	2747.1

Formula Calculator Element Limits

Element	Min	Max
C	3	60
H	0	60
O	0	10
N	0	10

Formula Calculator Results

Formula	CalculatedMass	CalculatedMz	Mz	Diff. (mDa)	Diff. (ppm)	DBE
C15 H16 N2 O6	320.1008	343.0901	343.0902	-0.10	-0.29	9.0000

--- End Of Report ---

Fig. S16. ^{1}H NMR (600 MHz) spectrum of **5** in DMSO-d_6

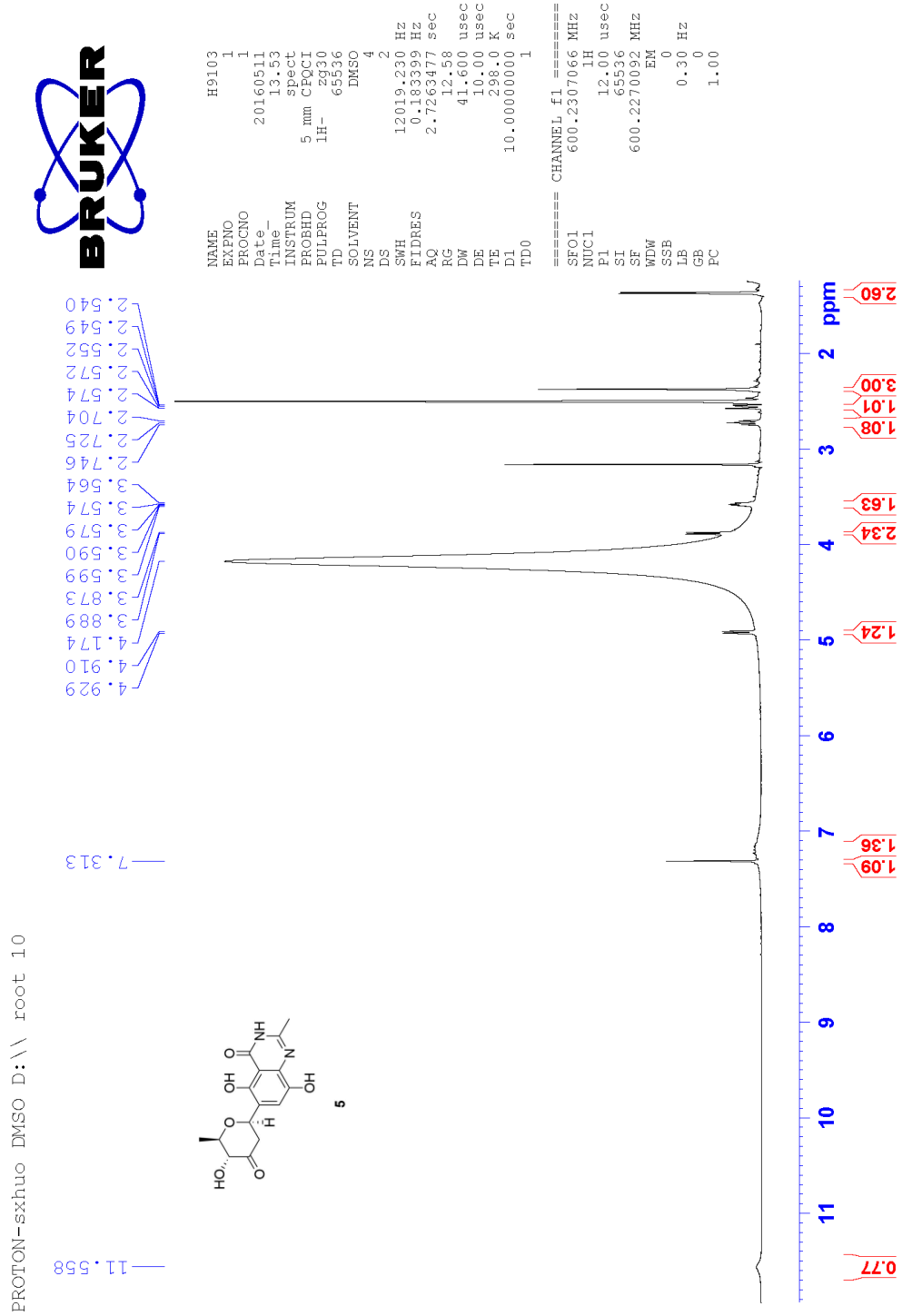


Fig. S17. ^{13}C NMR (150 MHz) spectrum of **5** in DMSO-d_6

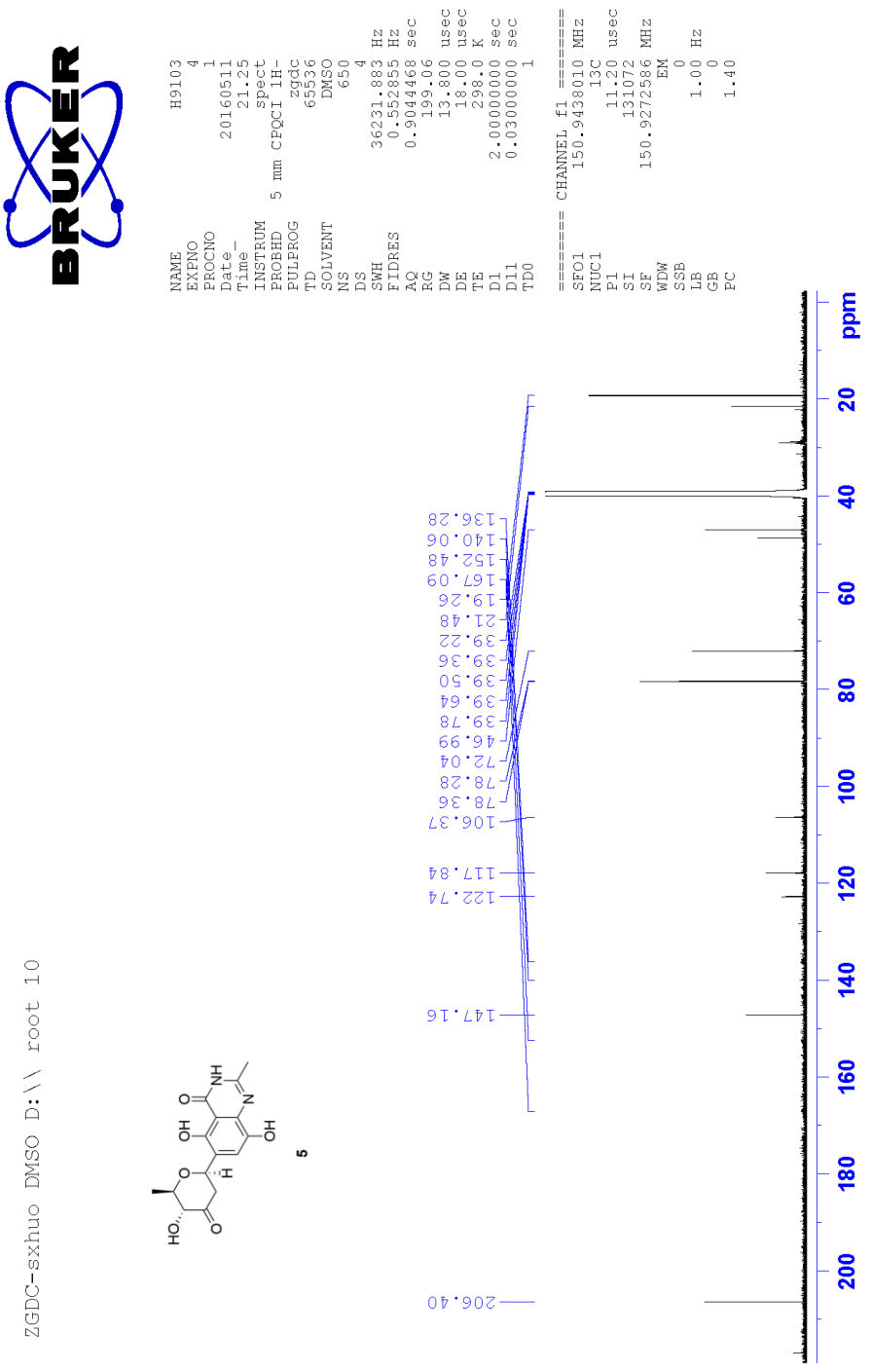


Fig. S18. HSQC (600 MHz) spectrum of **5** in DMSO-d_6

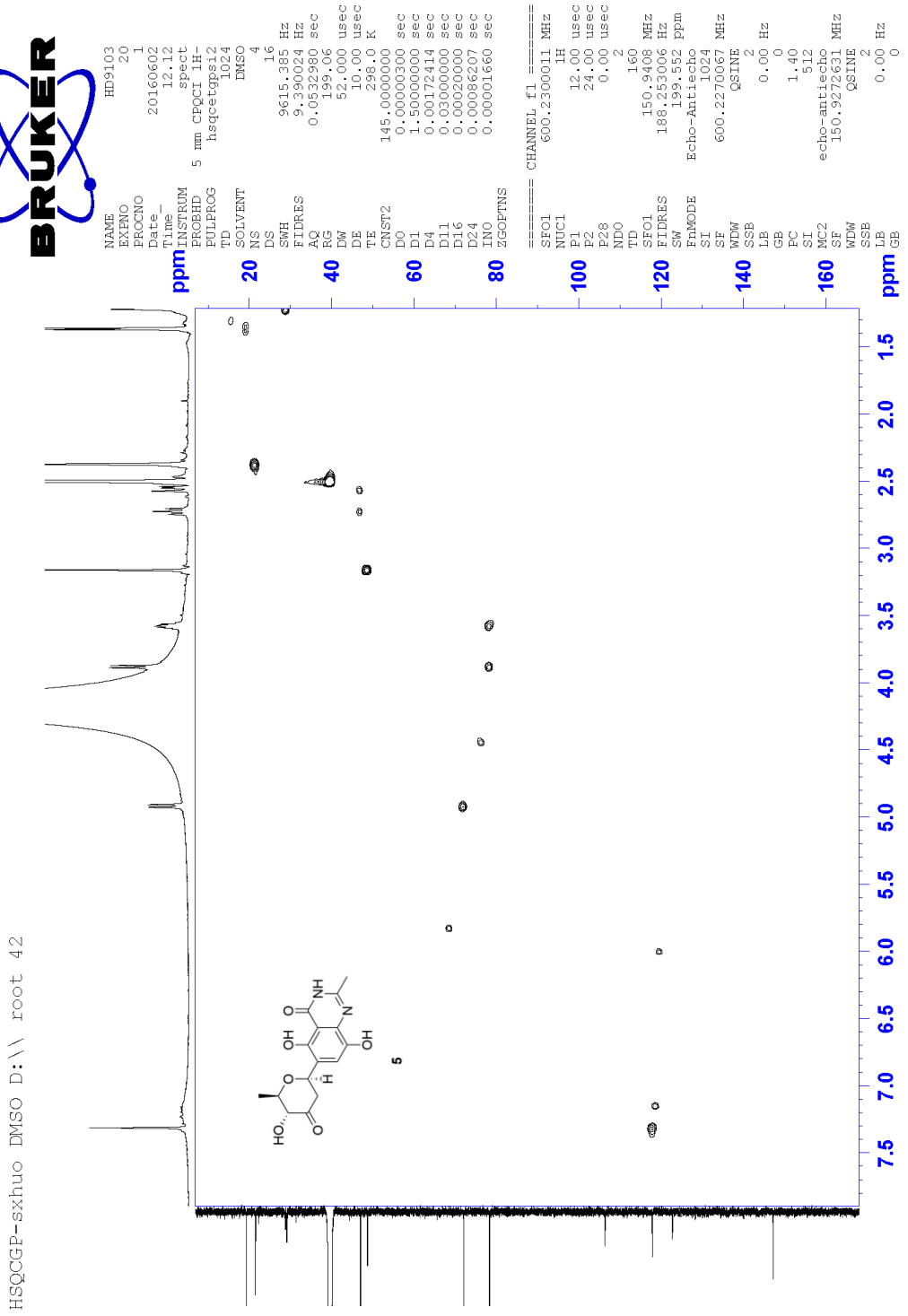


Fig. S19. HMBC (600 MHz) spectrum of **5** in DMSO-d_6

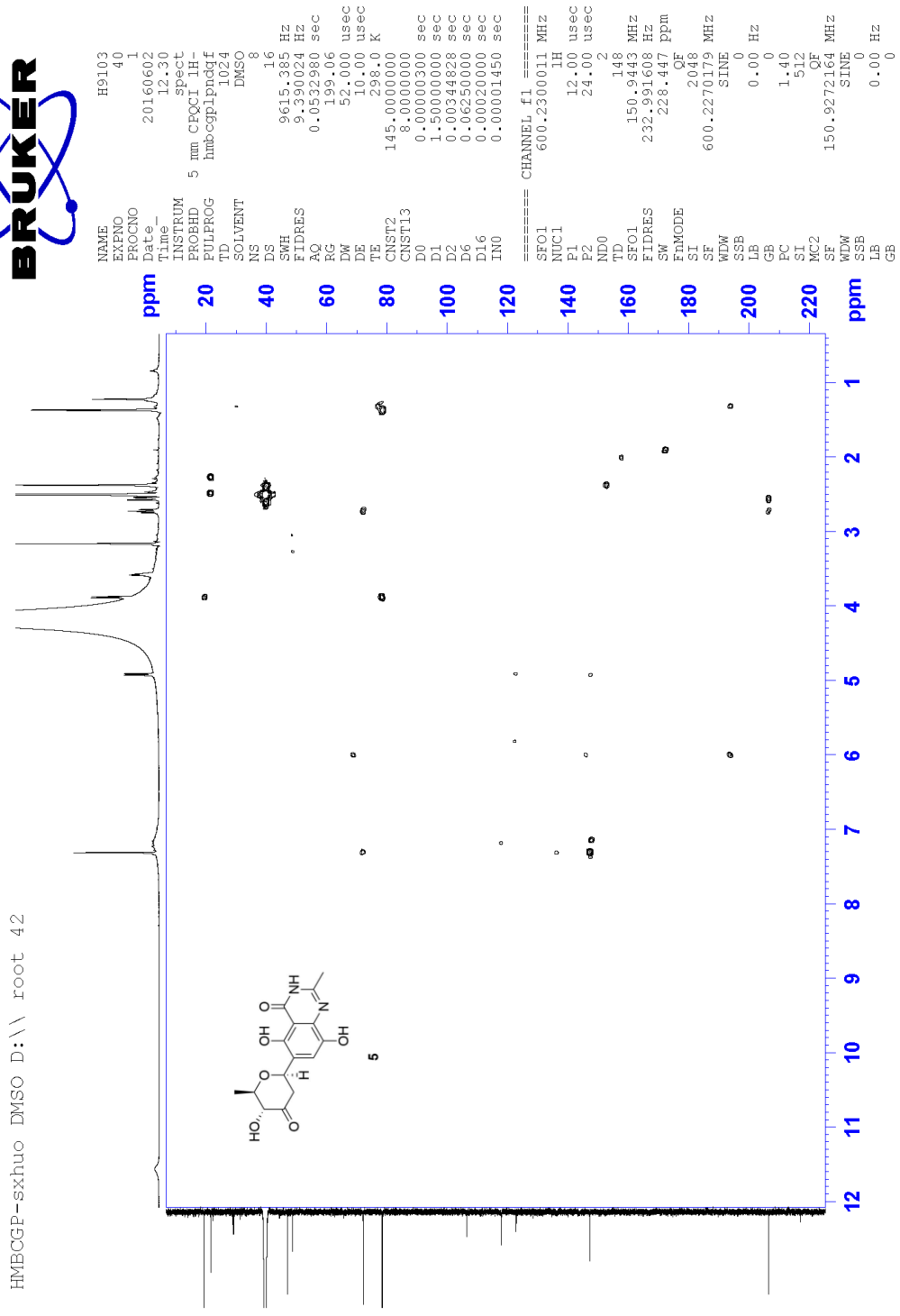


Fig. S20. ROESY (600 MHz) spectrum of **5** in DMSO-d_6

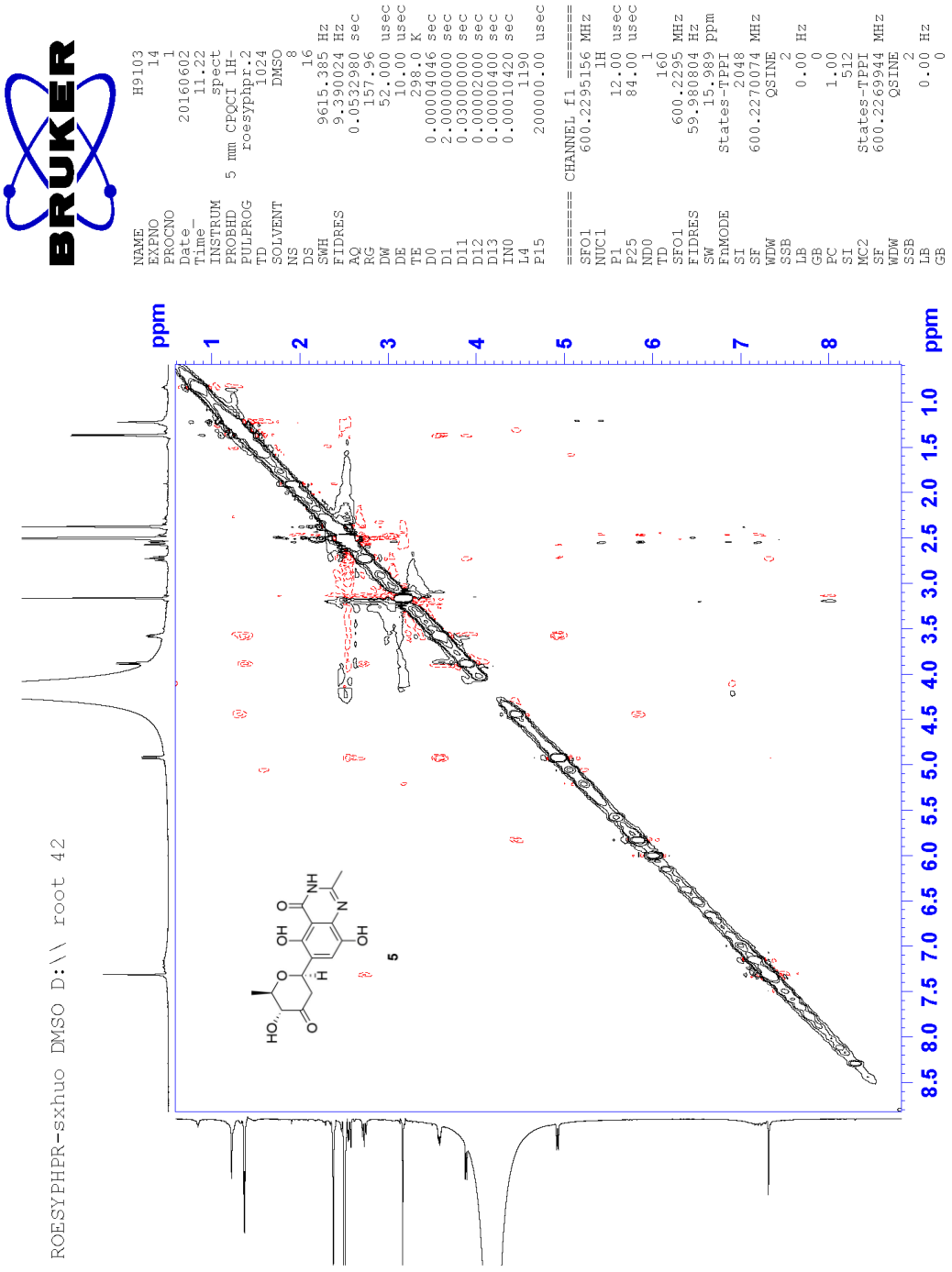


Fig. S21. ^{1}H NMR (600 MHz) spectrum of **8** in DMSO-d_6

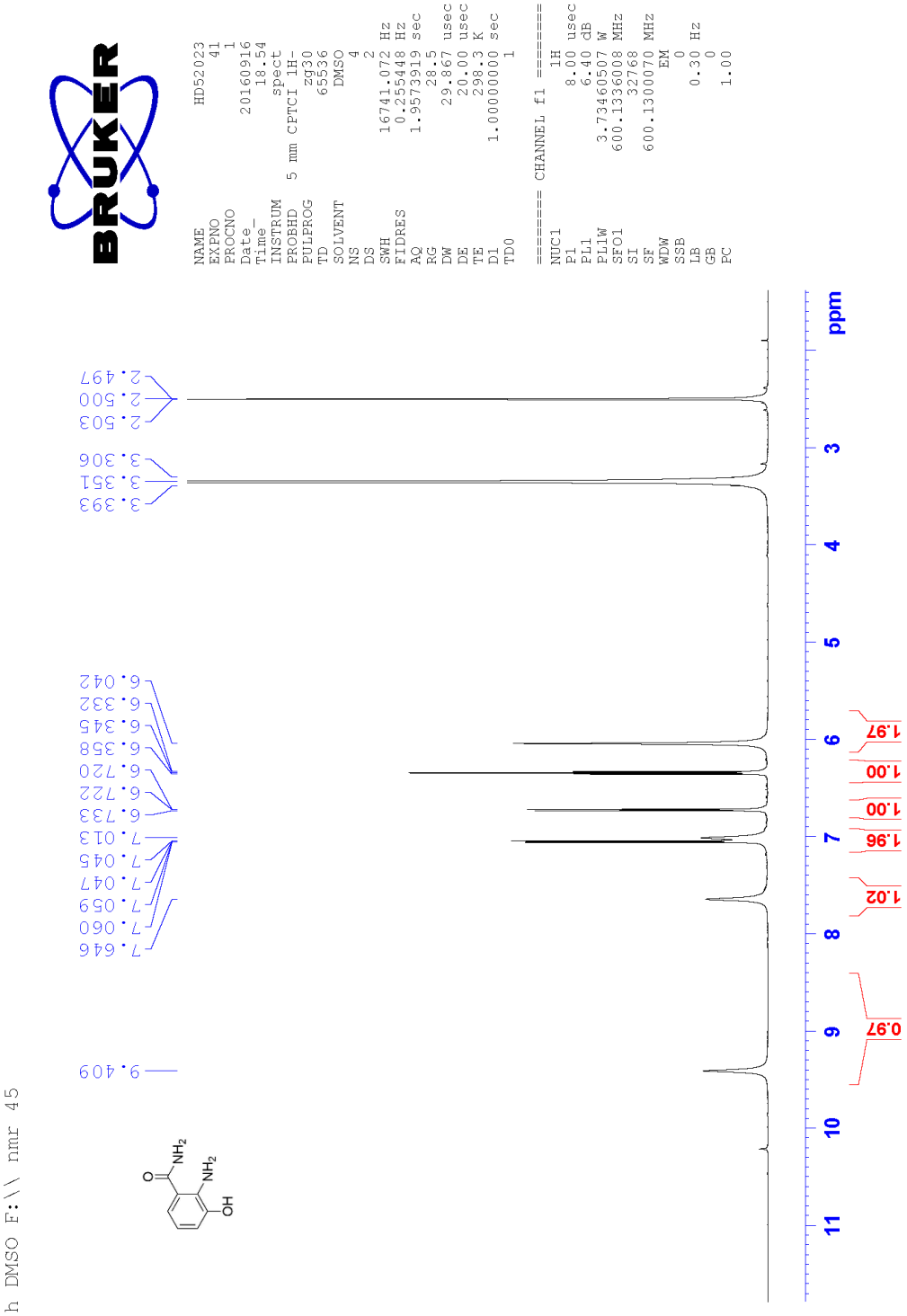


Fig. S22. ^{13}C NMR (150 MHz) spectrum of **8** in DMSO-d_6

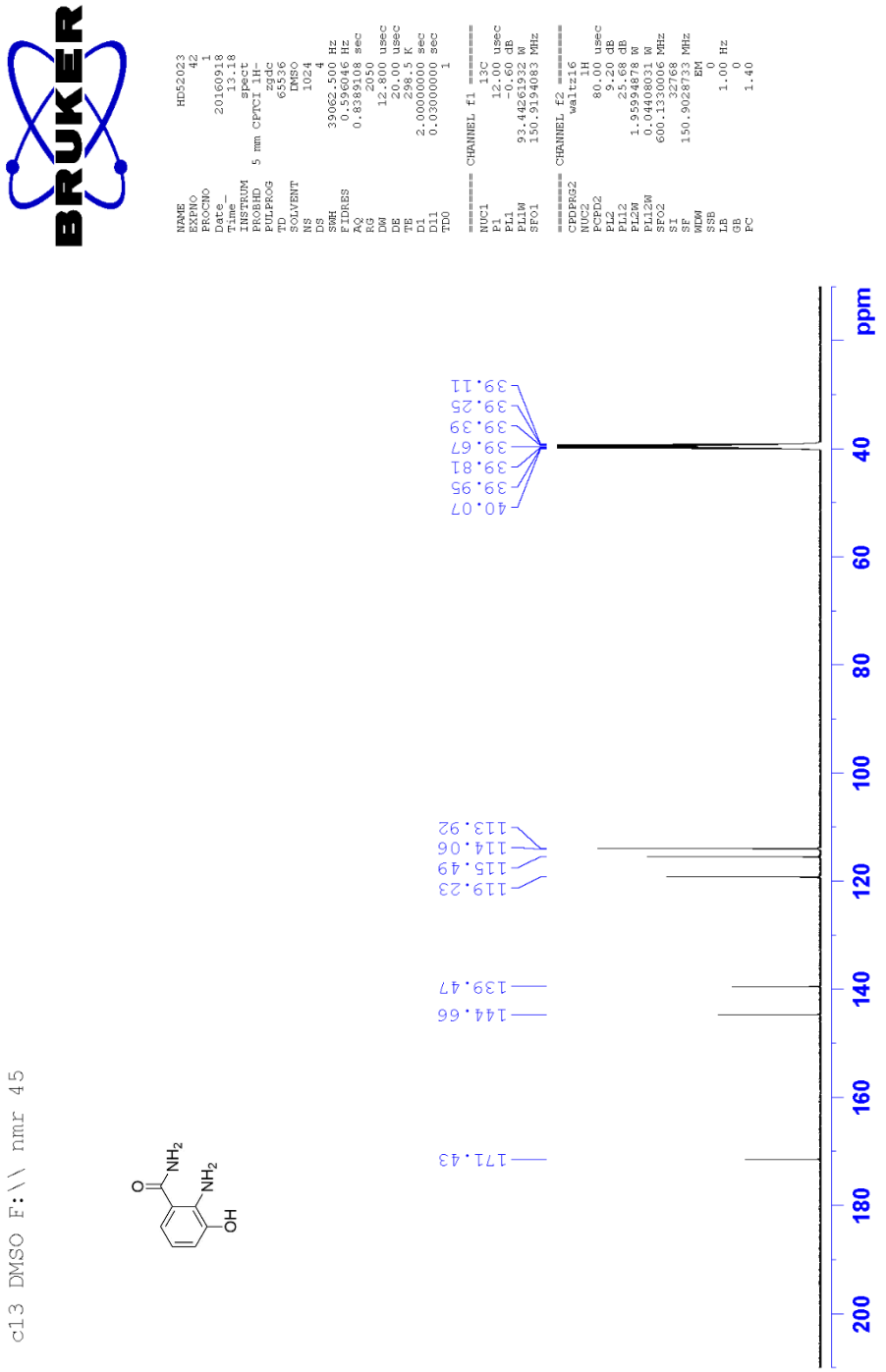


Fig. S23. ^{1}H NMR (400 MHz) spectrum of **9** in DMSO-d_6

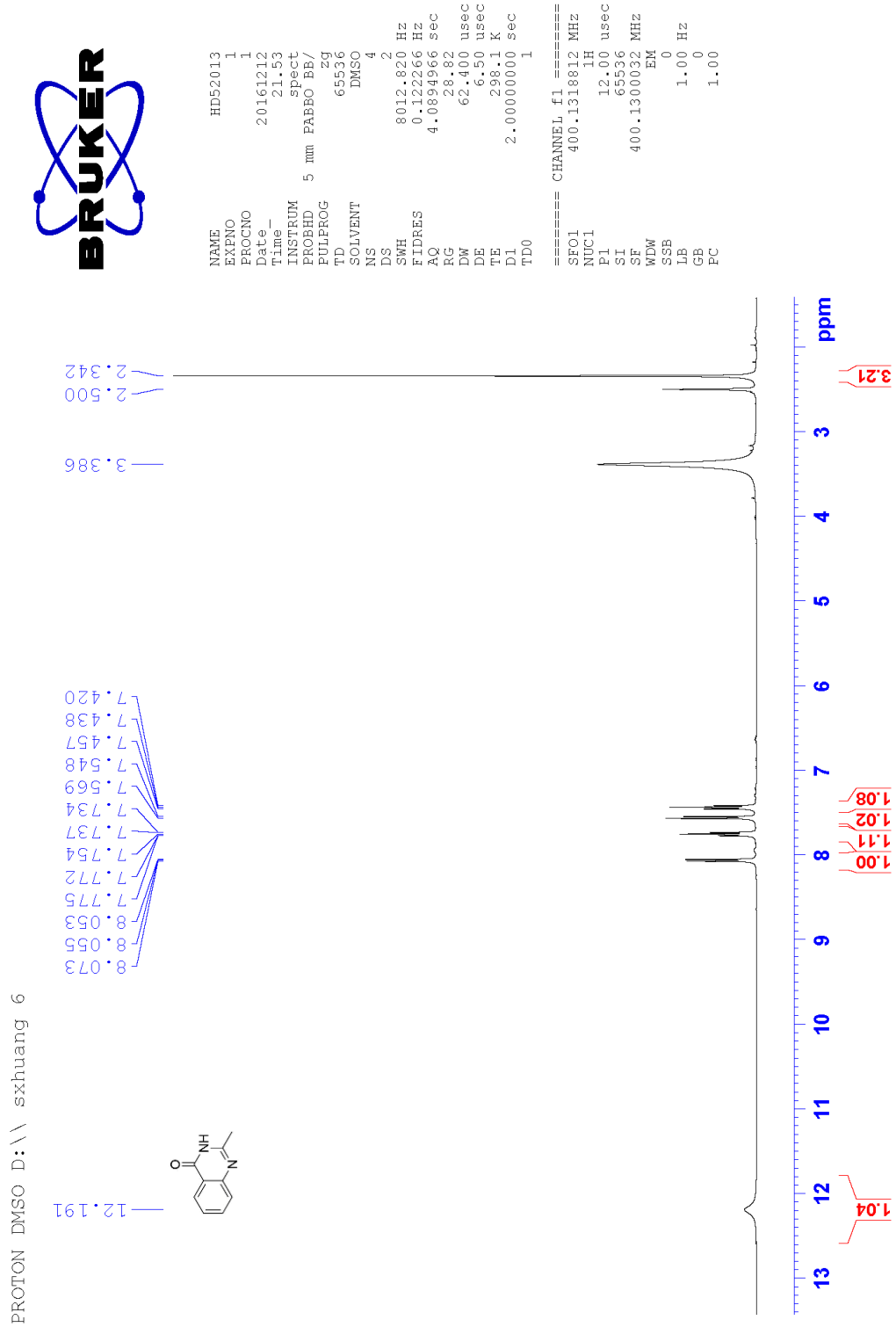


Fig. S24. ^{1}H NMR (100 MHz) spectrum of **9** in DMSO-d_6

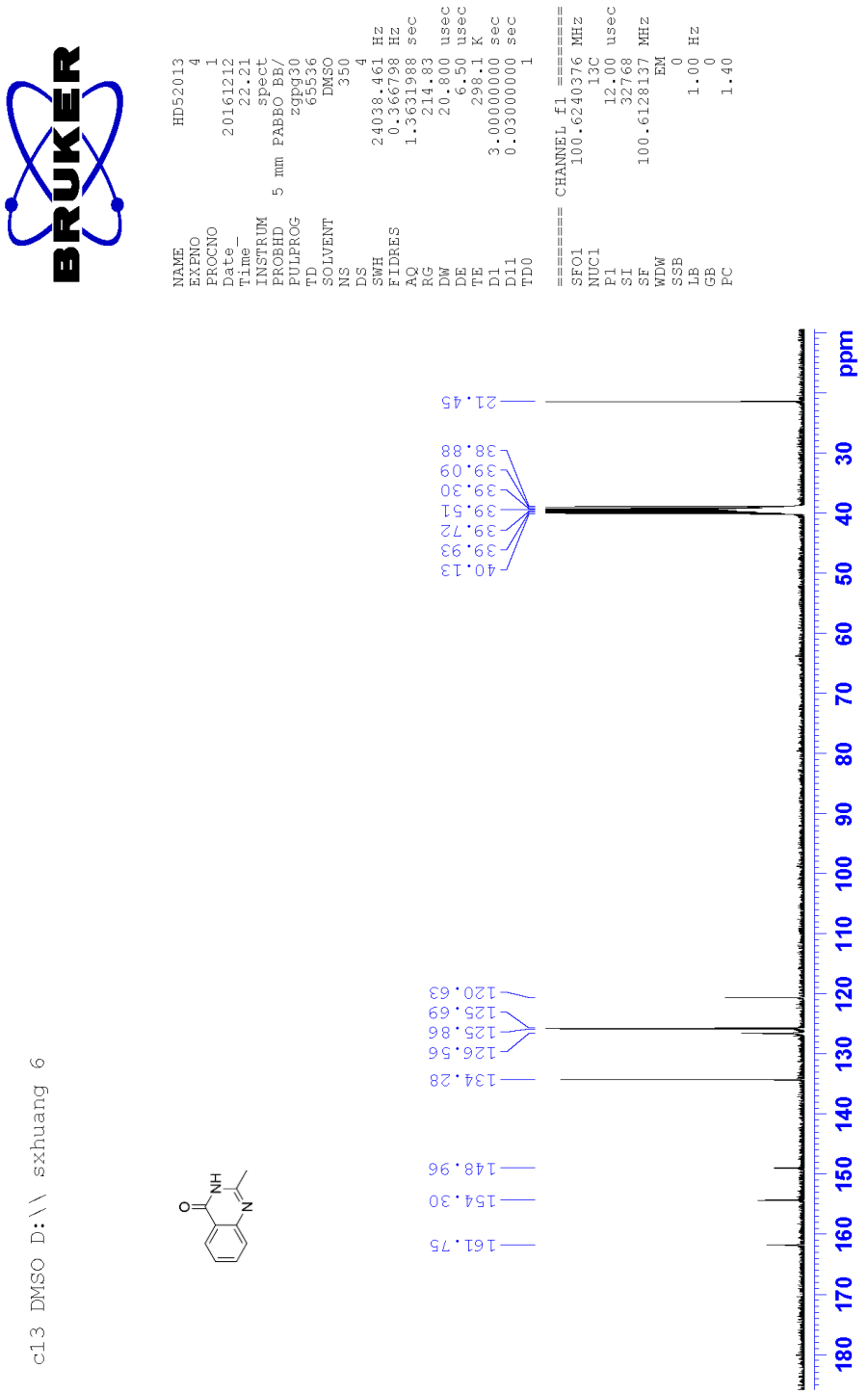


Fig. S25. ^{1}H NMR (400 MHz) spectrum of **10** in DMSO-d_6

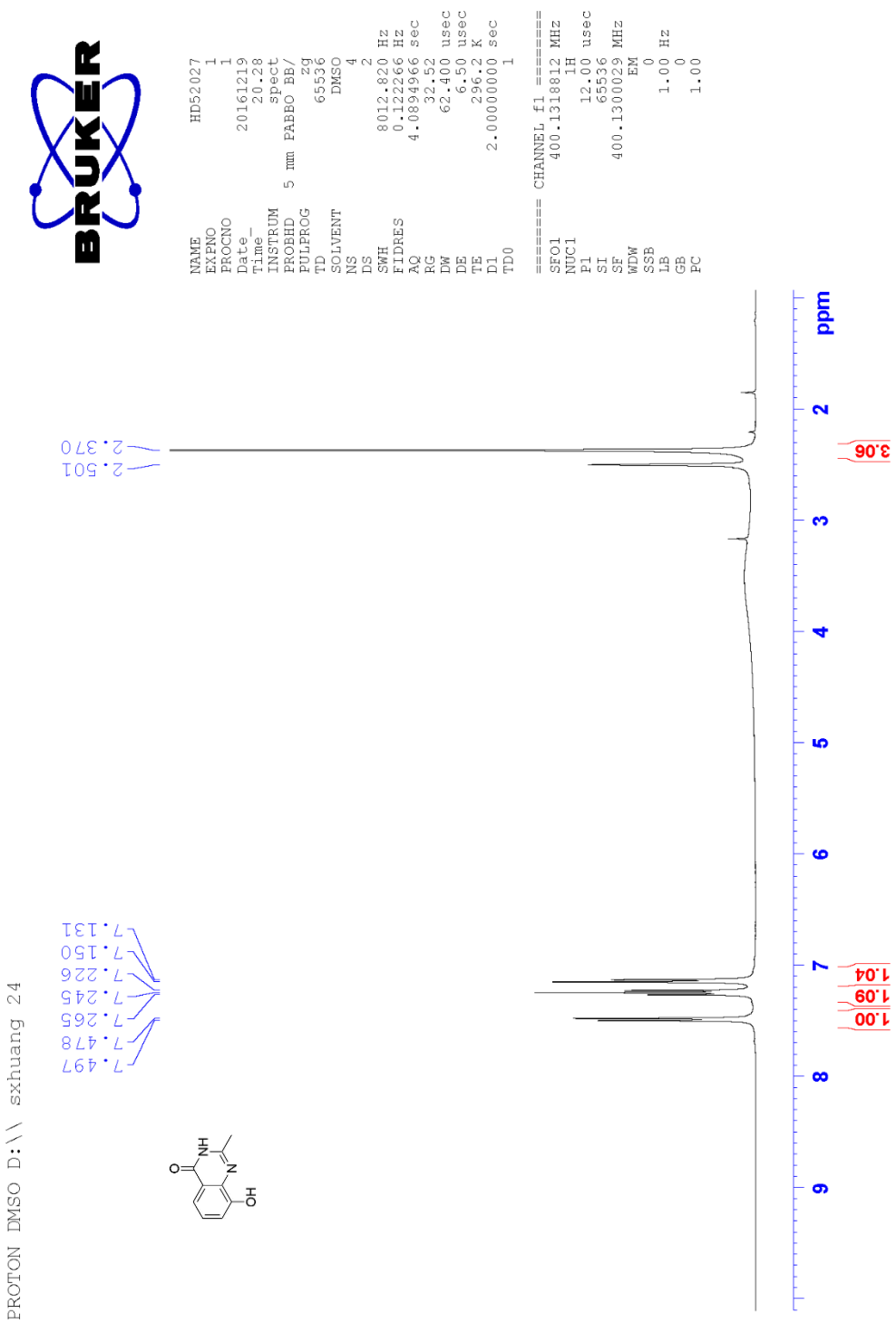
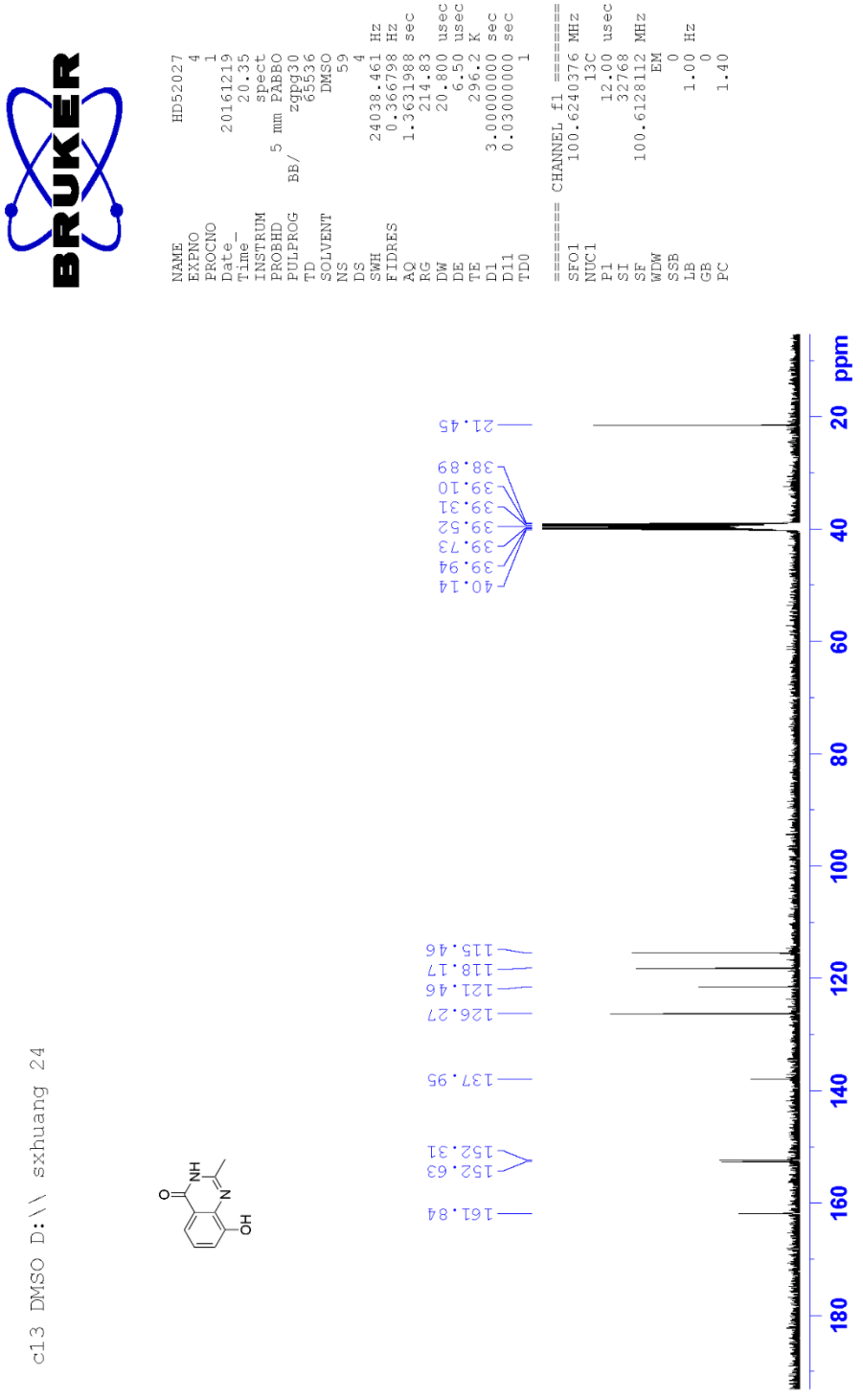


Fig. S26. ^{1}H NMR (100 MHz) spectrum of **10** in DMSO-d_6



Supplementary references

1. L. Li, Z. Xu, X. Xu, J. Wu, Y. Zhang, T. Zhang, X. He, T. Zabriskie and Z. Deng, The mildiomycin biosynthesis: initial steps for sequential generation of 5-hydroxymethylcytidine 5'-monophosphate and 5-hydroxymethylcytosine in *Streptoverticillium rimofaciens* ZJU5119. *ChemBioChem.,* 2008, **9**, 1286-1294.
2. M. S. B. Paget, L. Chamberlin, A. Atrih, S. J. Foster and M. J. Buttner, Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2).1999, *J. Bacteriol.,* **181**, 204–211.
3. K. A. Datsenko and B. L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A*, 2000, **97**, 6640-6645.
4. T. Hosaka, M. Ohnishi-Kameyama, H. Muramatsu, K. Murakami, Y. Tsurumi, S. Kodani, M. Yoshida, A. Fujie and K. Ochi, Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nat. Biotechnol.,* 2009, **27**, 462-464.