Article

Clausenain B, a Phenylalanine-rich Cyclic Octapeptide from *Clausena anisum-olens*

Yun-Song Wang,[a,b] Hong-Pin He,[b] Jing-Hua Yang,[a] Ying-Tong Di,[b] Ning-Hua Tan [b] and Xiao-Jiang Hao[b,]*

[a]Key Laboratory of Medicinal Chemistry for Natural Resource, Ministry of Education, School of Chemical Science and Technology, Yunnan University, Kunming 650091, P. R. China

[b]State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, P. R. China

Um novo octapeptideo cíclico nomeado clausenaina B foi isolado por processo cromatográfico de múltiplos passos a partir da *Clausena anisum-olens*. Sua estrutura foi estabelecida como sendo ciclo(-Phe[1]-Ser-Leu[1]-Phe[2]-Phe[4]-Gly-Leu[2]-Phe[3]-) (**1**) levando em conta extensivo estudo espectroscópico e evidências químicas. Clausenaina B (**1**) é um octapeptideo cíclico rico em fenilalanina.

A new cyclic octapeptide, named clausenain B, was isolated by a multi-step chromatography procedure from *Clausena anisum-olens*. Its structure was established as cyclo(-Phe[1]-Ser-Leu[1]-Phe[2]-Phe[4]-Gly-Leu[2]-Phe[3]-) (**1**) based on extensive spectroscopic studies and chemical evidence. Clausenain B (**1**) is a phenylalanine-rich cyclic octapeptide.

Keywords: Rutaceae, *Clausena anisum-olens*, cyclic octapeptide, clausenain B

Introduction

Plants of the Rutaceae family are widely distributed in the South of China.[1] Within this family, *Clausena* are shrubs widely distributed in South and Southeast Asia and many species are used in Chinese folk medicine for various indications.[1] Previous phytochemical studies have indicated that *Clausena* species contain structurally diverse and biologically active carbazole alkaloids and coumarins.[2-7]

Only a small number of cyclopeptides have been isolated from the plants of Rutaceae.[8] Cyclic peptides, which are widely produced in nature, are of considerable interest because these compounds exhibit various biological activities such as antibiotic, anti-inflammatory, and cytotoxic activities, and have often been used as models for studies of structural features of proteins.[9-12] Moreover, non-ribosomal cyclic peptides seems to be resistant to *in vivo* enzymic degradation, to present greater bioavailability than non-cyclic analogues and to show reduction of conformational flexibility.[13]

Clausena anisum-olens is a shrub growing in Hekou County of the Yunnan Province. The aerial parts of this plant have been used for the treatment of dysentery and arthritis.[1] As part of our continuing studies on the Chinese medicinal herbs as potential sources of new bioactive secondary metabolites, our previous studies on *Clausena anisum-olens* resulted in the isolation of a new cyclopeptide for the first time from the genus *Clausena*.[14] In order to search for other natural cyclic peptides from the genus *Clausena*, we further investigated the EtOH extract of *Clausena anisum-olens*. Chemical studies of this extract led to the isolation of a new cyclic peptide named clausenain B (**1**) from the EtOAc portion of whole plants. The new cyclic peptide shows highly repeated residue units with an unusually high content of phenylalanine ones. The ^{1}H NMR signals of amino acid residues, especially the signals of NH and α-H, heavily overlapped at room temperature. Using NMR with variable-temperature experiments, we obtained a well-resolved ^{1}H NMR spectrum, which displayed sharp proton signals. The structure of the new cyclic peptide was characterized as cyclo(-Phe[1]-Ser-Leu[1]-Phe[2]-Phe[4]-Gly-Leu[2]-Phe[3]-) (**1**) by analysis of extensive 2D-NMR spectroscopic methods and chemical evidences. Clausenain B (**1**) is a phenylalanine-rich cyclic octapeptide. This report describes the isolation and structure elucidation of compound **1**.

*e-mail: haoxj@mail.kib.ac.cn

Results and Discussion

The aerial parts of *Clausena anisum-olens*, collected from Hekou County, were extracted with 90% EtOH. The concentrated extract was in turn partitioned with petroleum ether, EtOAc and *n*-BuOH successively. The EtOAc extract was subjected to chromatography on silica gel, Sephadex LH-20 and RP C-18 to yield compound **1.**
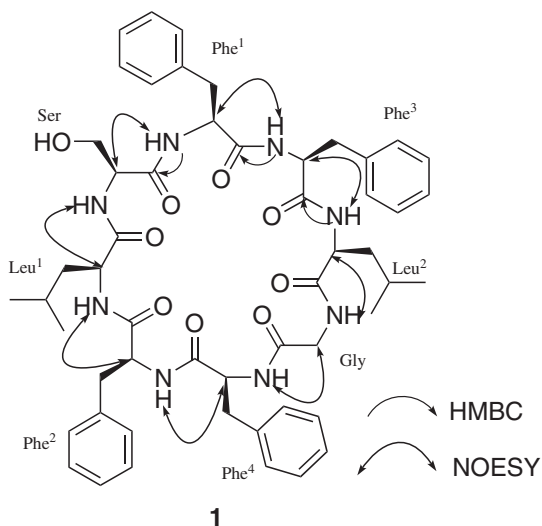

Figure 1. The structure, key HMBC (H→C) and NOESY (H→H) of **1**

Compound **1** was isolated as white amorphous powder. The compound did not react with ninhydrin, but hydrolysis with concentrated HCl, suggesting that **1** was a cyclic peptide. The molecular formula $C_{53}H_{66}N_8O_9$ was deduced from the HR-ESI$^+$-MS ([M+Na]$^+$ at *m/z* 981.4856, calc. 981.4850), corresponding to 25 degrees of unsaturation. IR bands between 1600-1730 and 3100-3500 cm^{-1} were characteristic of amide CO and NH groups, while UV absorption at 241, 252 and 260 indicated the presence of aromatic residues.

The ^{1}H NMR signals of the amino acid residues of **1**, especially the signals of NH and α-H, severely overlapped. Temperature change in the ^{1}H NMR experiments showed a pronounced effect on the ^{1}H NMR signals. A well-resolved ^{1}H NMR spectrum with sharp proton signals (see Table 1) was obtained at 315 K in DMSO-d_6. Assignment of the ^{1}H NMR signals to specific protons in individual residues was obtained by analysis of COSY and HSQC-TOCSY experiments, in order to assign the complete spin systems of the amino acid residues. The corresponding ^{13}C NMR assignments were determined on the basis of HSQC and HMBC experiments.

From 7.48 to 8.40 and 3.40 to 4.44, the ^{1}H NMR of **1** displayed the presence of 8 amide NH protons signals and

9 H-C(α), respectively, in agreement with an octapeptide structure including one glycine unit. At higher field, four methyl signals were observed. Analysis of the aromatic region of the ^{1}H and ^{13}C NMR spectra provided evidence for the presence of several aromatic ring systems, which were attributed to phenylalanine units. The ^{13}C NMR and DEPT spectra of **1** indicated the presence of eight amide carbonyl groups, four aromatic ring systems, eight CH(α) or CH$_2$(α) groups at δ_C 40 to 70 and four Me groups at δ_C 20-30. The δ_H and δ_C of the amino acid residues (except for the quaternary C-atoms) have been assigned by analysis of the HSQC-TOCSY spectrum. This technique provides not only total H-correlations in the F$_2$ dimension but also total C-correlations (except for the quaternary C-atoms) in the F$_1$ dimension.[15] Detailed analysis of the ^{1}H, ^{1}H COSY, HSQC-TOCSY, HMBC and NOESY data of **1** led to the complete structure assignments (Table 1). The sequence of the eight amino acids was elucidated mainly by detailed analysis of NOESY correlations between each amino acid residue α-H or β-H, and the next amino acid residue amide proton NH (as shown in the Figure 1) because long range HMBC correlations were not very informative in terms of sequencing due to the overlap of NMR signals in the region of the α-H protons in the spectrum but allowed the assignment of the amide carbonyl groups.

The nOe correlations Phe^1-NH/Ser-Hα, Ser-NH/Leu^1-Hα and Leu^1-NH/Phe^2-Hα established a segment Phe^1-Ser-Leu^1-Phe^2. The second segment Phe^4-Gly-Leu^2-Phe^3 was given by the nOe correlations Phe^4-NH/Gly-Hα, Gly-NH/ Leu^2-Hα, Leu^2-NH/ Phe^3-Hα. Besides, the two clear nOe correlations Phe^2-NH/ Phe^4-Hα and Phe^3-NH/ Phe^1-Hα gave the final sequence Phe^1-Ser-Leu^1-Phe^2-Phe^4-Gly-Leu^2-Phe^3. Taking into account the HMBC correlations between Phe^1-NH/Ser-CO, Leu^2-NH/Phe^3-CO and Phe^3-NH/Phe^1-CO, a segment Leu^2-Phe^3-Phe^1-Ser was ruled out, therefore confirming the amino acid sequence of **1**.

Hydrolysis of **1** at 120° with 6mol L^{-1} HCl, followed by standard amino acid analysis, confirmed the presence of Gly (1 equiv.), Ser (1 equiv.), Leu (2 equiv.) and Phe (4 equiv.). Considering that the eight identified amino acid residues accounted for 24 degrees of unsaturation, the extra unsaturation degree indicated a cyclic structure for **1**.

Therefore, the structure of **1** was assigned to be cyclo(-Phe^1-Ser-Leu^1-Phe^2-Phe^4-Gly-Leu^2-Phe^3-). The absolute configuration of the amino acid residues of **1** was determined by Marfey's method after hydrolysis of **1**.[16] All the amino acids were found to possess L-configurations. The new cyclic octapeptide **1** showed unusual high content of phenylalanine residues. The result of the present study suggested that cyclopeptides were characteristic constituents for the species *Clausena anisum-olens*.

Table 1. The ^1H and ^{13}C NMR spectra data for compound **1** (δ in ppm, in DMSO-d_6, T=315K)

Amino acid	residue		δ_H (*J* in Hz)	δ_C
Gly	CH$_2$(α)		3.83 dd (7.5, 17.0)	42.4 t
			3.40 dd (7.5, 17.5)	
	NH		8.17 t (5.6)	
	CO			169.1 s
Leu^1	H-C(α)		3.95 m	52.1 d
	CH$_2$ (β)		1.68 m	40.0 t
	H-C (γ)		1.58 m	24.0 d
	Me(δ)		0.80 d (6.3)	21.2 q
	Me(δ')		0.84 d (6.8)	23.0 q
	NH		8.04 d (7.2)	
	CO			172.5 s
Leu^2	H-C(α)		4.12 m	52.4 d
	CH$_2$ (β)		1.58 m	39.9 t
	H-C (γ)		1.52 m	24.2 d
	Me(δ)		0.92 d (6.2)	21.6 q
	Me(δ')		0.86 t (7.9)	23.4 q
	NH		7.83 t (6.2)	
	CO			171.0 s
Ser	H-C(α)		4.37 m	54.5 d
	CH$_2$ (β)		3.73 dt (5.5, 6.0, 11)	62.1 t
			3.62 dt (5.5, 6.0, 11)	
	-OH		5.43 br.s	
	NH		7.48 d (7.1)	
	CO			171.6 s
Phe^1	H-C(α)		4.27 m	55.6 d
	CH$_2$ (β)		2.95 dd (5.0, 16.0)	35.9 t
			2.78 dd (10.5, 14.5)	
	arom. H (δ)	1'	7.11-7.31m	137.9 s
		2', 6'		129.4 d
		3', 5'		128.3 d
		4'		126.4 d
	NH		8.40 d (9.2)	
	CO			171.4 s
Phe^2	H-C(α)		4.22 m	54.8 d
	CH$_2$ (β)		3.14 dd (5.0, 14.0)	36.8 t
			3.04 dd (4.5, 14.5)	
	arom. H (δ)	1'	7.11-7.31 m	137.9 s
		2',6'		128.8 d
		3', 5'		128.3 d
		4'		126.3 d
	NH		8.37 d (6.6)	
	CO			170.9 s
Phe^3	H-C(α)		4.40 m	54.0 d
	CH$_2$ (β)		3.07 dd (5.0, 14.0)	37.1 t
			2.87 m	
	arom. H (δ)	1'	7.11-7.31 m	137.9 s
		2',6'		129.0 d
		3', 5'		128.0 d
		4'		126.4 d
	NH		7.91 d (7.8)	
	CO			170.6 s
Phe^4	H-C(α)		4.44 m	55.6 d
	CH$_2$ (β)		3.00 m	35.9 t
			2.84 m	
	arom. H (δ)	1'	7.11-7.31 m	137.9 s
		2', 6'		129.1 d
		3', 5',		128.3 d
		4',		126.2 d
	NH		7.79 d (8.2)	
	CO			171.6 s

Experimental

General

Commercial Silica-gel plates (Qing Dao Marine Chemical Group Co.) were used for TLC analyses. UV Spectra was measured on Shimadzu UV-2401PC spectrophotometer; λ_{max} in nm. Optical rotation was obtained on Horiba SEAP-300 spectropolarimeter. IR spectra were measured on Bio-Rad FTS-135 infrared spectrophotometer, ν_{max} in cm^{-1}. ^1H and ^{13}C NMR as well as 2D NMR spectra were recorded on Brucker DRX-400 and DRX-500 spectrometer; chemical shifts δ in ppm related to TMS, coupling constant *J* in Hz. ESI-MS spectra were acquired on VG-Autospec 3000 mass spectrometers; in *m/z*. HPLC analyse was performed with an Agilent 1100 HPLC system.

Plant material

The leaves and twigs of *Clausena anisum-olens* were collected in Hekou County of Yunnan province, P. R. China, in May 2003 and identified by Prof. De-Ding Tao from Kunming Institute of Botany. A voucher specimen (No. 02041705) is deposited in State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences.

Extraction and isolation

The powdered plant material of *Clausena anisum-olens* (22.5 kg) was repeatedly extracted with EtOH at room temperature. The extract was then concentrated under reduced pressure to give brown syrup, which was sequentially partitioned between H$_2$O and petroleum ether, EtOAc and n-BuOH. The EtOAc extract (110.5 g) were subjected to a silica gel column chromatography eluting with PE-EtOAc (4:1-2:3), EtOAc, EtOAc-MeOH (8:2-1:1), MeOH, by which nineteen fractions (I-XIX) were obtained. Fraction XIV was subjected to additional separation steps by silica gel column chromatography (CHCl$_3$-MeOH 9:1-5:5), Sephadex LH-20 (MeOH) and RP C-18 (MeOH-H$_2$O 1:1) to yield compound **1** (52 mg).

Clausenain B (cyclo(-Phe^1–Ser–Leu^1–Phe^2–Phe^4–Gly–Leu^2–Phe^3-) (1)

White amorphous powder; [α]$_D^{21.8}$ −91.08 (MeOH, *c* 0.71,). IR ν_{max}/cm^{-1} (KBr): 3433, 1640; UV λ_{max}/nm (MeOH): 241, 252, 260; ^1H NMR and ^{13}C NMR: see Table1; HRESIMS: *m/z* 981.4856 ([M+Na]$^+$) (calcd. for C$_{53}$H$_{66}$N$_8$O$_9$, 981.4850); FAB$^+$MS: *m/z* 959 ([M+H]$^+$).

Absolute configuration of amino acids

A solution of **1** (1 mg) in 6 mol L^{-1} HCl was heated at 110 °C for 24 h in a sealed tube. After removal of HCl by evaporation in *vacuo*, the hydrolyzate was dissolved in water and treated with 1-fluro-2,4-dinitrophenyl-5-L-alanine amide (Marfey's reagent) and 1 mol L^{-1} NaHCO$_3$ at 35 °C for 1 h. After cooling, it was treated with 2 mol L^{-1} HCl and then concentrated to dryness. This residue was subjected to reversed-phase HPLC, at a flow rate 1 mL min^{-1}, detection at 340 nm, solvent: 10-80% MeOH (80 min gradient added on 10 min to be eluted with 100% MeOH)/50 m mol L^{-1} triethylamine phosphate (TEAP) buffer (pH 3.2).

Acknowledgments

Supplementary Information

MS, ^{1}H NMR, ^{13}C NMR, ^{1}H-^{1}H COSY, HSQC, HMBC, HSQC-TOCSY, NOESY NMR spectra of compound **1** are available free of charge at http://jbcs.org.br, as PDF file.

References

1. Institute Botanicum Kunmingenge Academiae Sinicae; *Flora Yunnanica*, Tomus 6 (Spermatophyta); Science Press: Beijing, 1995, p.767 (in Chinese).

2. Chakraborty, A.; Chowdhury, B. K.; Bhattacharyya, P.; *Phytochemistry* **1995**, *40*, 295.

3. Wu, T. S.; Huang, S. C.; Wu, P. L.; *Tedrahedron Lett.* **1996**, *37*, 7819.

4. Wu, T. S.; Huang, S. C.; Wu, P. L.; *Heterocycles* **1997**, *45*, 969.

5. He, H. P.; Shen, Y. M.; He, Y. N.; Yang, X. S.; Zhu, W. M.; Hao, X. J.; *Heterocycles* **2000**, *53*, 2067.

6. Ito, C.; Itoigawa, M.; Katsuno, S.; Omura, M.; Tokuda, H.; Nishino, H.; Furukawa, H.; *J. Nat. Prod.* **2000**, *63*, 1218.

7. Nakamura, K.; Takemura, Y.; Ju-ichi, M.; Ito, C.; Furukawa, H.; *Heterocycles* **1998**, *48*, 549.

8. Mastumoto, T.; Nishimura, K.; Takeya, K.; *Chem. Pharm. Bull.* **2002**, *50*, 857.

9. Auvin-Guette, C.; Baraguey, C.; Blond, A.; Xavier, H. S.; Pousset, J. L.; Bodo, B.; *Tetrahedron* **1999**, *55*, 11495.

10. Laupacis, A.; Keown, P. A.; Ulan, R. A.; McKenzie, N.; Stiller, C. R.; *Can. Med. Assoc. J.* **1982**, *126*, 1041.

11. Kirkpatrick, P.; Raja, A.; LaBonte, J.; Lebbos, J.; *Nat. Rev. Drug Discov*ery **2003**, *2*, 943.

12. Sandhu, P.; Xu, X.; Bondiskey, P. J.; Balani, S. K.; Morris, M. L.; Tang, Y. S.; Miller, A. R.; Pearson, P. G.; *Antimicrob. Agents Chemother.* **2004**, *48*, 1272.

13. Hamada, Y.; Shioiri; T.; *Chem. Rev.* **2005**, *105*, 4441.

14. Wang, Y. S.; He, H. P.; Yang, J. H.; Shen, Y. M.; Zhou, J.; Hao, X. J.; *Helv. Chim. Acta* **2005**, *88*, 2345.

15. Teng, R.W.; Ding, Z.T.; He, Y. N.; Yang, C. R.; Wang, D. Z.; *Chin. J. Magn. Reson.* **2003**, *20*, 397.

16. Marfey, P.; *Carlsberg Res. Commun.***1984**, *49*, 591.

Supplementary Information

Clausenain B, a Phenylalanine-rich Cyclic Octapeptide from *Clausena anisum-olens*

Yun-Song Wang,[a,b] Hong-Pin He,[b] Jing-Hua Yang,[a] Ying-Tong Di,[b] Ning-Hua Tan [b] and Xiao-Jiang Hao[b,]*

[a]*Key Laboratory of Medicinal Chemistry for Natural Resource, Ministry of Education, School of Chemical Science and Technology, Yunnan University, Kunming 650091, P. R. China*

[b]*State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, P. R. China*

Figure S1. HRESIMS spectrum of compound **1**.

*e-mail: haoxj@mail.kib.ac.cn
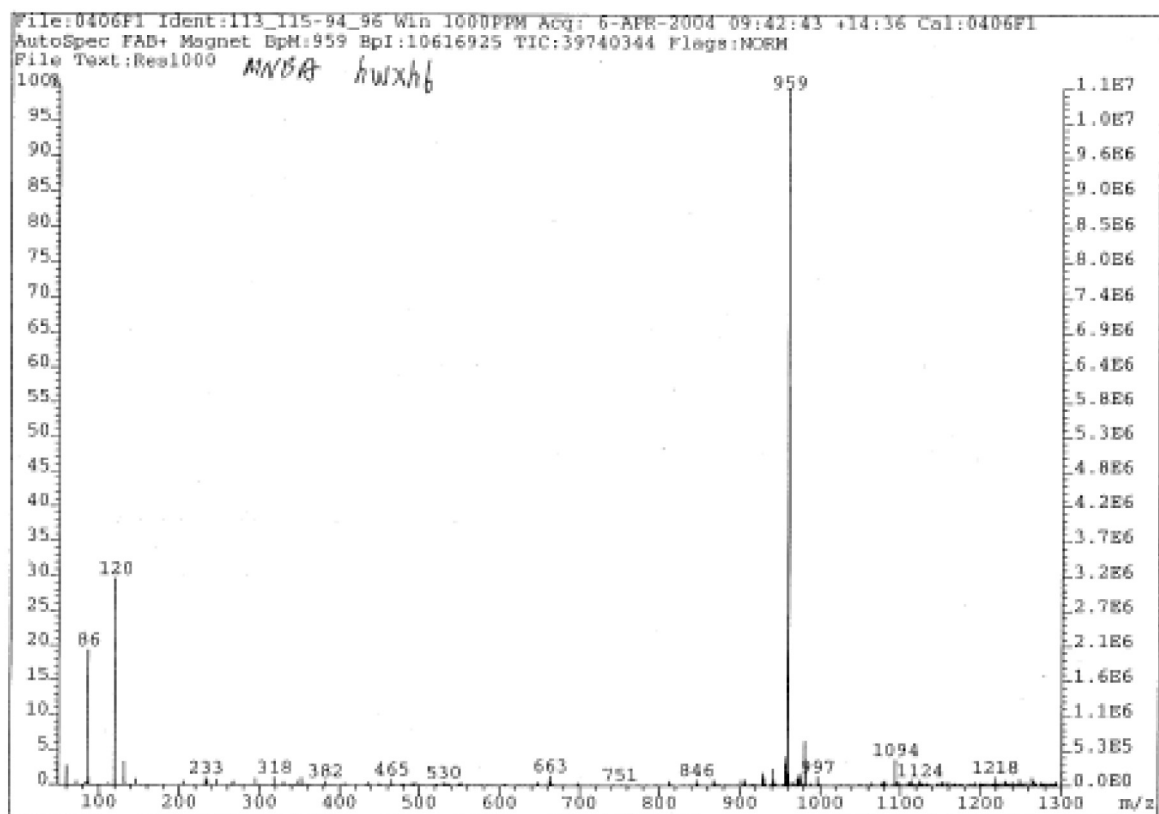
Figure S2. FAB[+]MS spectrum of compound **1**.
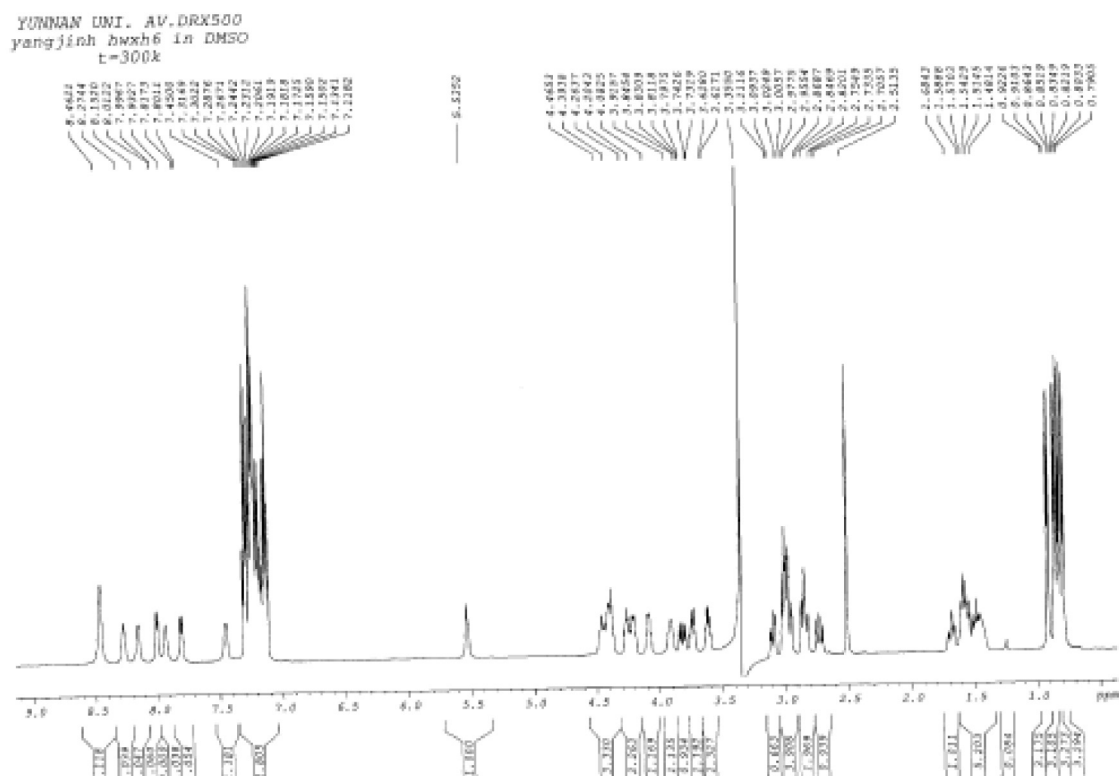
Figure S3. ^{1}H NMR spectrum (500 MHz, DMSO-d_6, T=300K) of compound **1**.
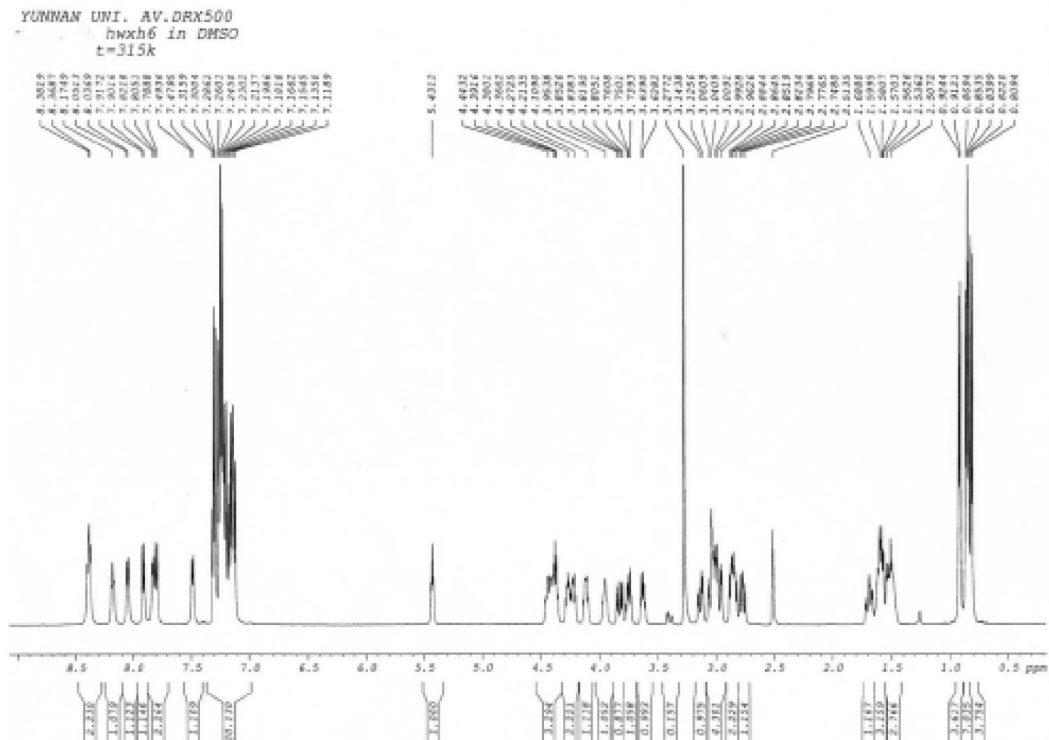
Figure S4. ¹H NMR spectrum (500 MHz, DMSO-d_6, T=315K) of compound **1**.
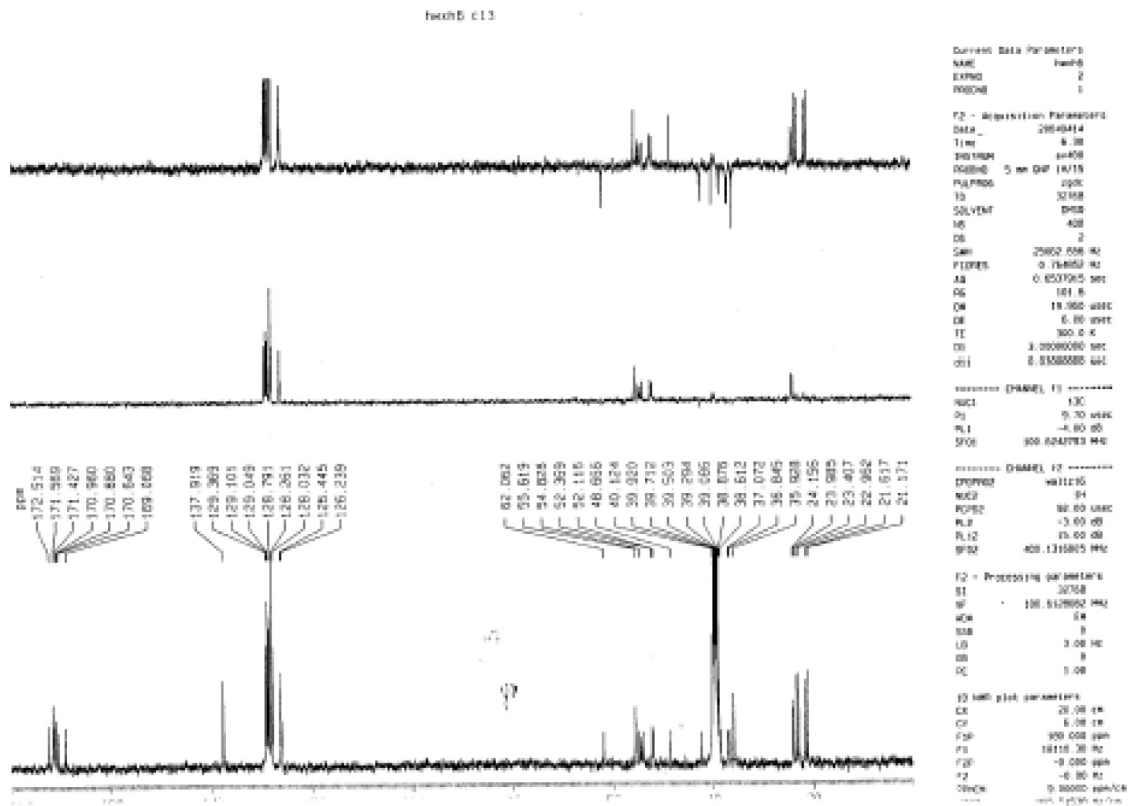
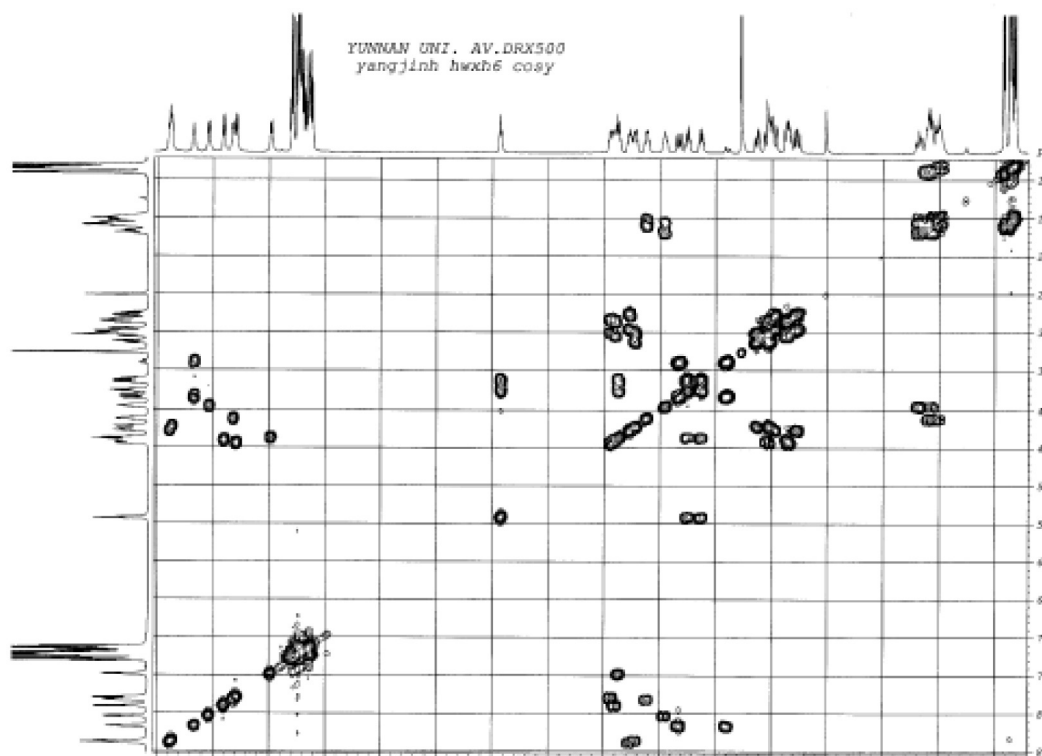**Figure S5.** ¹³C NMR spectrum (400 MHz, DMSO-d_6) of compound **1**.

Figure S6. ^{1}H-^{1}H COSY NMR spectrum (500 MHz, DMSO-d_6, T=315K) of compound **1**.
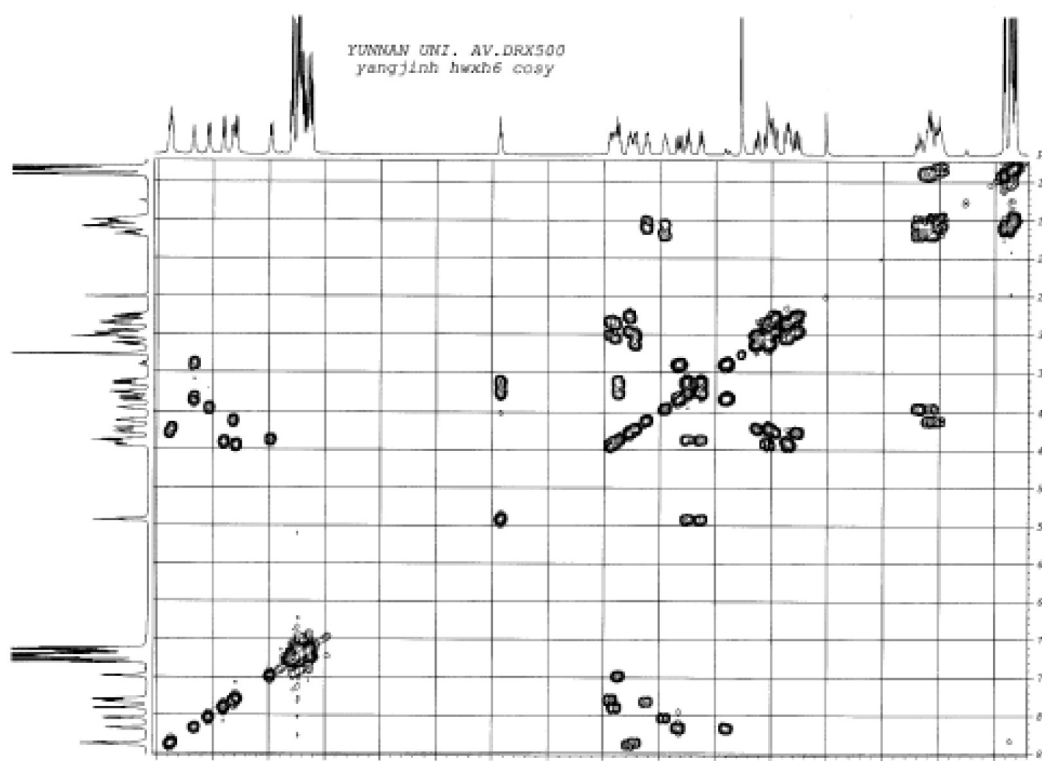
Figure S7. HSQC spectrum (DMSO-d_6) of compound **1**.
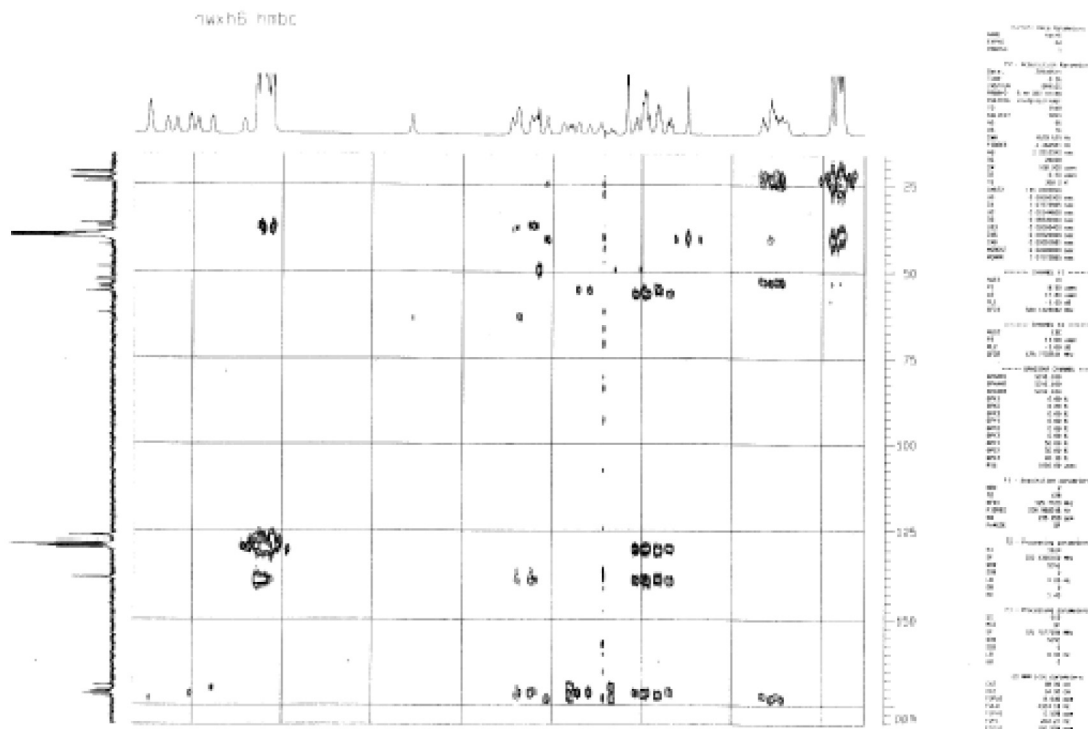
Figure S8. HMBC spectrum (DMSO-d_6) of compound **1**.
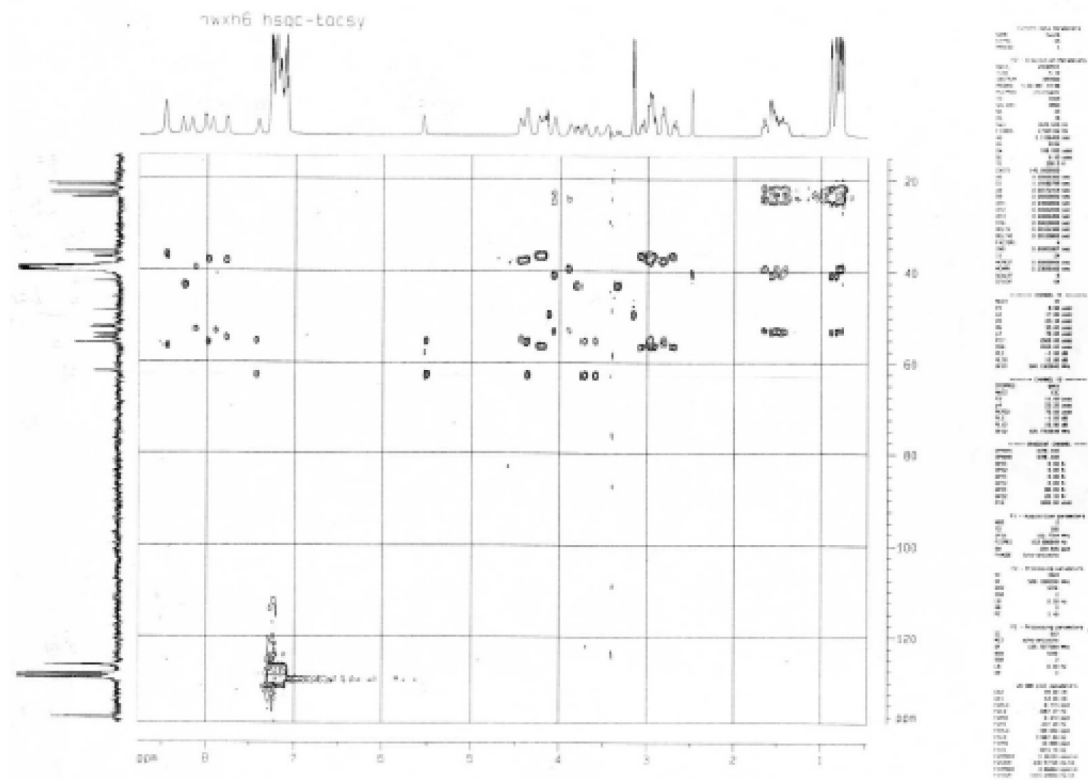
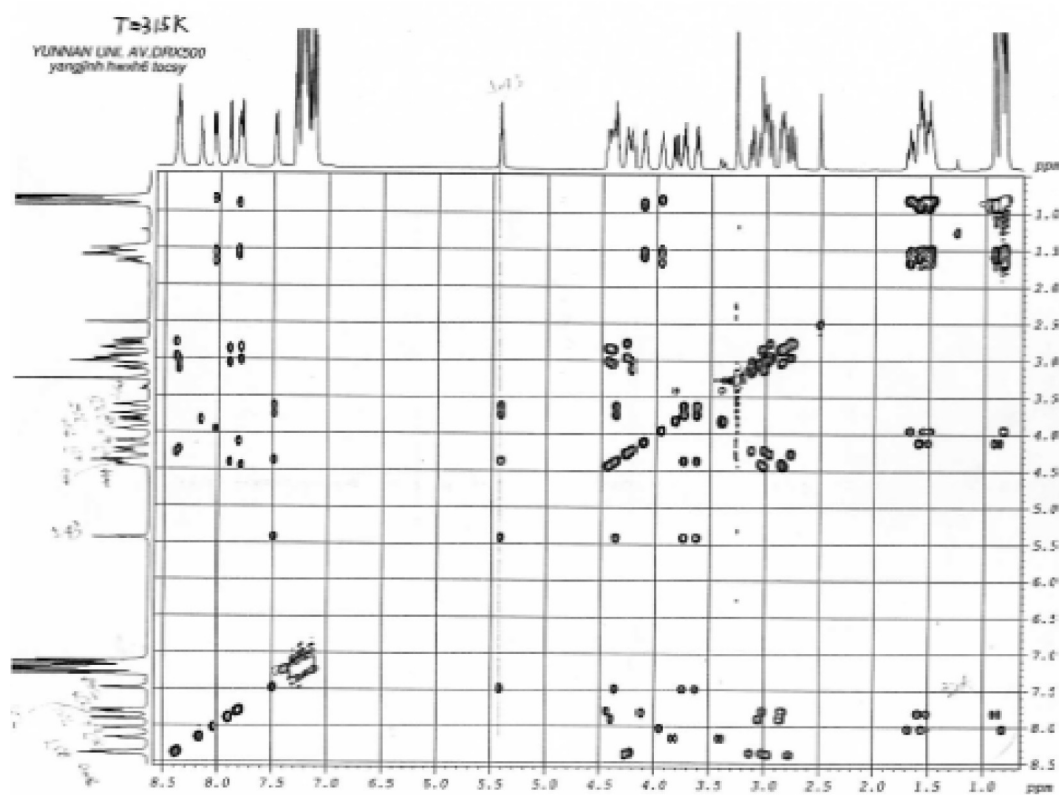**Figure S9.** HSQC-TOCSY spectrum (DMSO-d_6) of compound **1**.

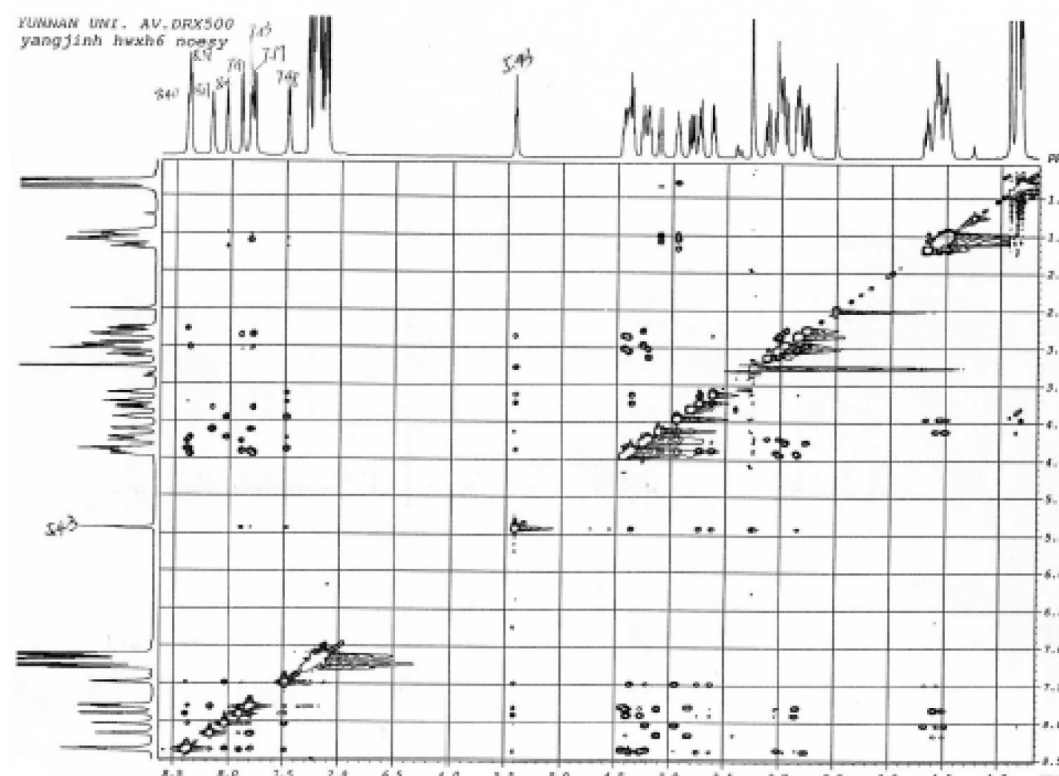**Figure S10.** TOCSY spectrum (DMSO-d_6) of compound **1**.

Figure S11. NOESY spectrum (DMSO-d_6) of compound **1**.