

Supporting Information

Heterologous expression of a single HR-PKS leads to the formation of diverse 2-alkenyl-tetrahydropyrans in model fungi

Hai-Ning Lyu ^{a, b‡}, Jinyu Zhang ^{a, c‡}, Shuang Zhou ^a, Hong-Wei Liu ^a, Wen-Ying Zhuang ^a, Shu-Ming Li ^d, Wen-Bing Yin ^{a,c,*}

^a *State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, People's Republic of China*

^b *Artemisinin Research Center, Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, People's Republic of China*

^c *Savaid Medical School, University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China*

^d *Institut für Pharmazeutische Biologie und Biotechnologie, Philipps-Universität Marburg, Marburg 35037, Germany*

*Corresponding authors:

Wen-Bing Yin,
E-mail: yinwb@im.ac.cn.
Tel: 86-10-64806170.

‡ These authors contributed equally to this work.

Content

1. Supplementary tables

Table S1. Fungal strains and plasmids used in this study

Table S2. PCR primers used in this study

Table S3. HR-ESI-MS data for isolated compounds

2. Supplementary figures

Figure S1. The *app* biosynthetic gene cluster and its homologues in *Trichoderma* (A), and gene function predictions (B)

Figure S2. Transformant verification of App1 in *A. nidulans* by diagnostic PCR

Figure S3. ¹H NMR spectrum of **1** in CDCl₃ (500MHz)

Figure S4. ¹³C NMR spectrum of **1** in CDCl₃ (125MHz)

Figure S5. ¹H-¹H COSY spectrum of **1** in CDCl₃

Figure S6. HSQC spectrum of **1** in CDCl₃

Figure S7. HMBC spectrum of **1** in CDCl₃

Figure S8. ¹H NMR spectrum of **2** in CDCl₃ (500MHz)

Figure S9. ¹³C NMR spectrum of **2** in CDCl₃ (125MHz)

Figure S10. ¹H-¹H COSY spectrum of **2** in CDCl₃

Figure S11. HSQC spectrum of **2** in CDCl₃

Figure S12. HMBC spectrum of **2** in CDCl₃

Figure S13. ¹H NMR spectrum of **3** in CDCl₃ (500MHz)

Figure S14. ¹³C NMR spectrum of **3** in CDCl₃ (125MHz)

Figure S15. ¹H-¹H COSY spectrum of **3** in CDCl₃

Figure S16. HSQC spectrum of **3** in CDCl₃

Figure S17. HMBC spectrum of **3** in CDCl₃

Figure S18. ¹H NMR spectrum of **4** in CDCl₃ (500MHz)

Figure S19. ¹³C NMR spectrum of **4** in CDCl₃ (125MHz)

Figure S20. ¹H-¹H COSY spectrum of **4** in CDCl₃

Figure S21. HSQC spectrum of **4** in CDCl₃

Figure S22. HMBC spectrum of **4** in CDCl₃

Figure S23. ROESY of **4** in CDCl₃ (500MHz)

Figure S24. ¹H NMR spectrum of **5** in CDCl₃ (500MHz)

Figure S25. ¹³C NMR spectrum of **5** in CDCl₃ (125MHz)

Figure S26. ¹H-¹H COSY spectrum of **5** in CDCl₃

Figure S27. HSQC spectrum of **5** in CDCl₃

Figure S28. HMBC spectrum of **5** in CDCl₃

Figure S29. ¹H NMR spectrum of **6** in CDCl₃ (500MHz)

Figure S30. ¹³C NMR spectrum of **6** in CDCl₃ (125MHz)

Figure S31. ¹H-¹H COSY spectrum of **6** in CDCl₃

Figure S32. HSQC spectrum of **6** in CDCl₃

Figure S33. HMBC spectrum of **6** in CDCl₃

Figure S34. ¹H NMR spectrum of **7** in CDCl₃ (500MHz)

Figure S35. ¹³C NMR spectrum of **7** in CDCl₃ (125MHz)

Figure S36. ¹H-¹H COSY spectrum of **7** in CDCl₃

Figure S37. HSQC spectrum of **7** in CDCl₃

Figure S38. HMBC spectrum of **7** in CDCl₃

Figure S39. ¹H NMR spectrum of **8** in CDCl₃ (500MHz)

Figure S40. ¹³C NMR spectrum of **8** in CDCl₃ (125MHz)

Figure S41. HSQC spectrum of **8** in CDCl₃

Figure S42. HMBC spectrum of **8** in CDCl₃

Figure S43. ¹H NMR spectrum of **9** in CDCl₃ (500MHz)

Figure S44. ¹³C NMR spectrum of **9** in CDCl₃ (125MHz)

Figure S45. ¹H-¹H COSY spectrum of **9** in CDCl₃

Figure S46. HSQC spectrum of **9** in CDCl₃

Figure S47. HMBC spectrum of **9** in CDCl₃

Figure S48. ROESY of **9** in CDCl₃ (500MHz)

3. Supplementary nucleotide sequence data

4. References

1. Supplementary tables

Table S1. Fungal strains and plasmids used in this study

Strain/plasmid	Description	Reference
<i>Aspergillus nidulans</i> LO8030	<i>pyroA4, riboB2, pyrG89, nkuA::argB, sterigmatocystin cluster (AN7804-AN7825)Δ, emericellamide cluster (AN2545-AN2549)Δ, asperfuranone cluster (AN1039-AN1029)Δ, monodictyphenone cluster (AN10023-AN10021) Δ, terrequinone cluster (AN8512-8520)Δ, austinol cluster part 1 (AN8379-AN8384)Δ, austinol cluster part 2 (AN9246-9259)Δ, F9775 cluster (AN7906-7915)Δ, asperthecin cluster (AN6000-AN6002)Δ</i>	¹
TYZS7	pYZS4 in <i>A. nidulans</i> LO8030	This study
<i>Saccharomyces cerevisiae</i> BJ5464-NpgA	<i>MATα ura3-52 his3-Δ200 leu2-Δ1 trp1 pep4::HIS3 prb1 Δ1.6R can1 GAL</i>	²
pXW55	<i>2μ, URA3, ADH2p::ACPC, Amp</i>	³
pRGAMA1	<i>pyrG, AMA1</i>	⁴
pYZS4	<i>AMA1::T7792_04478::pyr4</i>	This study
pYLV2	<i>URA3::ADH2p::T7792_04478 ORF</i>	This study

TXX = original transformant

pXX = plasmid

Table S2. PCR primers used in this study

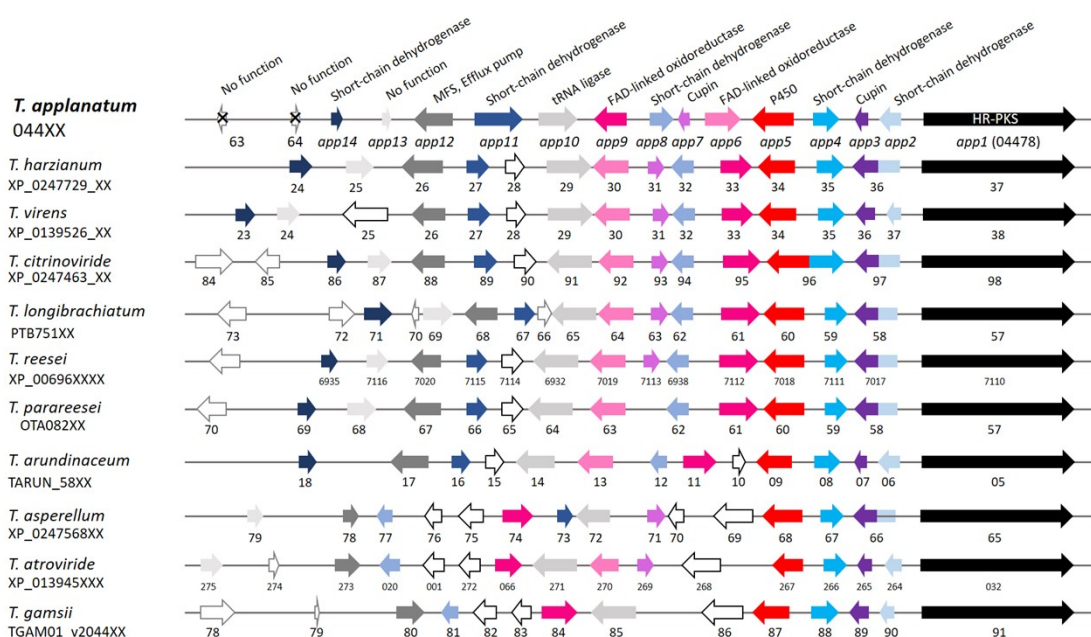
Name	Oligonucleotide sequence (5'-3')	Use
XmaI-7792-C40 for pRG-AMAI-F	ccccccgggcaagccgtgtgtaagaagtgg	Construction of pYZS4
NotI-7792-C40 for pRG-AMAI-R	gaatgcggccgcctcctgacattatgcattcggc	
detect F	ctgcatgaagagtgaccgc	Confirmation of TYZS7
detect R	gactcgaggataacgtgtccg	
4478_1 F	ggctagcggattataaggatgatgatgataagactagtagtgcactcttgagcctgc	Construction of pYLV2
4478_1 R	ggtatcttgctcgtgatgagctcc	
4478_2 F	gtcagagaatgctgactctcaacc	
4478_2 R	gtgagtgaattcggacctcctc	
4478_3 F	cgatgcctgctccaagca	
4478_3 R	ccggtagcagccatgatatcc	
4478_4 F	ccatcctgatccactctggc	
4478_4 R	gtcatttaaattagtgatggatggatggatgcacgtggctagcccagctggcaacaag	

Table S3. HR-ESI-MS data for isolated compounds

Compound	Formula	Calculated	Ion	Observed
1	C ₁₈ H ₃₀ O ₄	311.2222	[M+H] ⁺	311.2225
2	C ₁₈ H ₃₀ O ₅	327.2171	[M+H] ⁺	327.2170
3	C ₁₈ H ₃₀ O ₅	327.2171	[M+H] ⁺	327.2169
4	C ₁₆ H ₂₆ O ₅	299.1858	[M+H] ⁺	299.1856
5	C ₂₀ H ₃₄ O ₄	339.2535	[M+H] ⁺	339.2533
6	C ₂₀ H ₃₄ O ₄	339.2535	[M+H] ⁺	339.2532
7	C ₁₈ H ₂₈ O ₅	325.2015	[M+H] ⁺	325.2017
8	C ₁₈ H ₃₀ O ₃	295.2273	[M+H] ⁺	295.2269
9	C ₁₈ H ₃₀ O ₃	295.2273	[M+H] ⁺	295.2276

2. Supplementary figures

A



B

<i>T. applanatum</i>	<i>T. harzianum</i>	<i>T. virens</i>	<i>T. citrinoviride</i>	<i>T. longibrachiatum</i>	<i>T. reesei</i>	<i>T. parareesei</i>	<i>T. arundinaceum</i>	<i>T. asperellum</i>	<i>T. atroviride</i>	<i>T. gamsii</i>	Probable function
App14 (04465)	XP_024772924 98/89	XP_013952623 98/90	XP_024746386 99/84	PTB75171 98/84	XP_006966935 86/82	OTA08269 99/82	TARUN_5818 98/92	-	-	-	Short-chain dehydrogenase
App13 (04466)	XP_024772925 99/89	XP_013952624 99/86	XP_024746387 99/86	PTB75169 99/76	XP_006967116 99/76	OTA08268 99/76	-	XP_024756879 99/86	XP_013945275 99/82	-	No function
App12 (04467)	XP_024772926 80/83	XP_013952626 80/90	XP_024746388 84/88	PTB75168 80/88	XP_006967020 84/88	OTA08267 99/81	TARUN_5817 99/83	XP_024756878 79/57	XP_013945273 79/61	TGAM01_v2044 80 83/59	MFS, Efflux pump
App11 (04468)	XP_024772927 55/81	XP_013952627 55/84	XP_024746389 55/75	PTB75167 55/72	XP_006967115 55/76	OTA08266 49/77	TARUN_5816 49/80	XP_024756873 42/75	-	-	Short-chain dehydrogenase
App10 (04469)	XP_024772929 99/92	XP_013952629 99/89	XP_024746391 99/90	PTB75165 99/90	XP_006966932 99/90	OTA08264 99/90	TARUN_5814 99/93	XP_024756872 99/91	XP_013945271 99/91	TGAM01_v2044 85 99/91	tRNA ligase
App9 (04470)	XP_024772933 99/77	XP_013952633 90/83	XP_024746395 99/80	PTB75161 99/80	XP_006967112 99/80	OTA08261 99/79	TARUN_5811 99/82	XP_024756874 99/59	XP_013945066 99/59	TGAM01_v2044 84 99/60	FAD-linked oxidoreductase
App8 (04471)	XP_024772932 99/74	XP_013952632 99/75	XP_024746394 95/76	PTB75162 99/76	XP_006966938 94/76	OTA08262 99/76	TARUN_5812 99/82	XP_024756877 98/58	XP_013945020 99/57	TGAM01_v2044 81 99/56	Short-chain dehydrogenase
App7 (04472)	XP_024772931 98/75	XP_013952631 98/76	XP_024746393 98/74	PTB75163 98/73	XP_006967113 98/77	-	-	XP_024756871 98/75	XP_013945269 98/75	-	Cupin
App6 (04473)	XP_024772930 96/80	XP_013952630 99/83	XP_024746392 99/79	PTB75164 96/84	XP_006967019 99/83	OTA08263 95/83	TARUN_5813 99/83	-	XP_013945270 98/72	-	FAD-linked oxidoreductase
App5 (04474)	XP_024772934 98/79	XP_013952634 98/78	XP_024746396 97/77	PTB75160 97/77	XP_006967018 97/79	OTA08260 89/87	TARUN_5809 89/83	XP_024756868 88/76	XP_013945267 88/74	TGAM01_v2044 87 99/72	P450
App4 (04475)	XP_024772935 99/79	XP_013952635 99/79	-	PTB75159 99/77	XP_006967111 99/77	OTA08259 99/77	TARUN_5808 80/78	XP_024756867 98/80	XP_013945266 99/80	TGAM01_v2044 88 99/80	Short-chain dehydrogenase
App3 (04476)	XP_024772936 91/94	XP_013952636 93/94	XP_024746397 85/96	PTB75158 85/97	XP_006967017 85/98	OTA08258 85/98	TARUN_5807 99/95	XP_024756866 94/89	XP_013945265 93/88	TGAM01_v2044 89 99/82	Cupin
App2 (04477)	XP_024772937 99/74	XP_013952637 99/74	XP_024746398 99/83	PTB75157 92/97	XP_006967016 92/76	OTA08257 92/76	TARUN_5806 99/81	55/77	XP_013945264 99/71	TGAM01_v2044 90 99/77	Short-chain dehydrogenase
App1 (04478)	XP_024772937 97/89	XP_013952638 97/90	XP_024746398 99/87	PTB75157 99/86	XP_006967110 99/86	OTA08257 99/86	TARUN_5805 99/90	XP_024756865 98/86	XP_013945032 98/86	TGAM01_v2044 91 99/86	HRPKS

Figure S1. The *app* biosynthetic gene cluster and its homologues in *Trichoderma* (A), and gene function predictions (B)

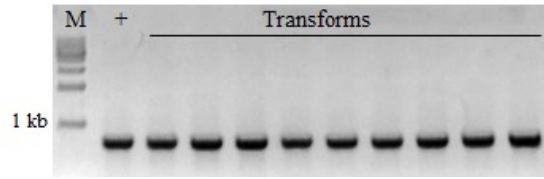


Figure S2. Transformant verification of App1 in *A. nidulans* by diagnostic PCR

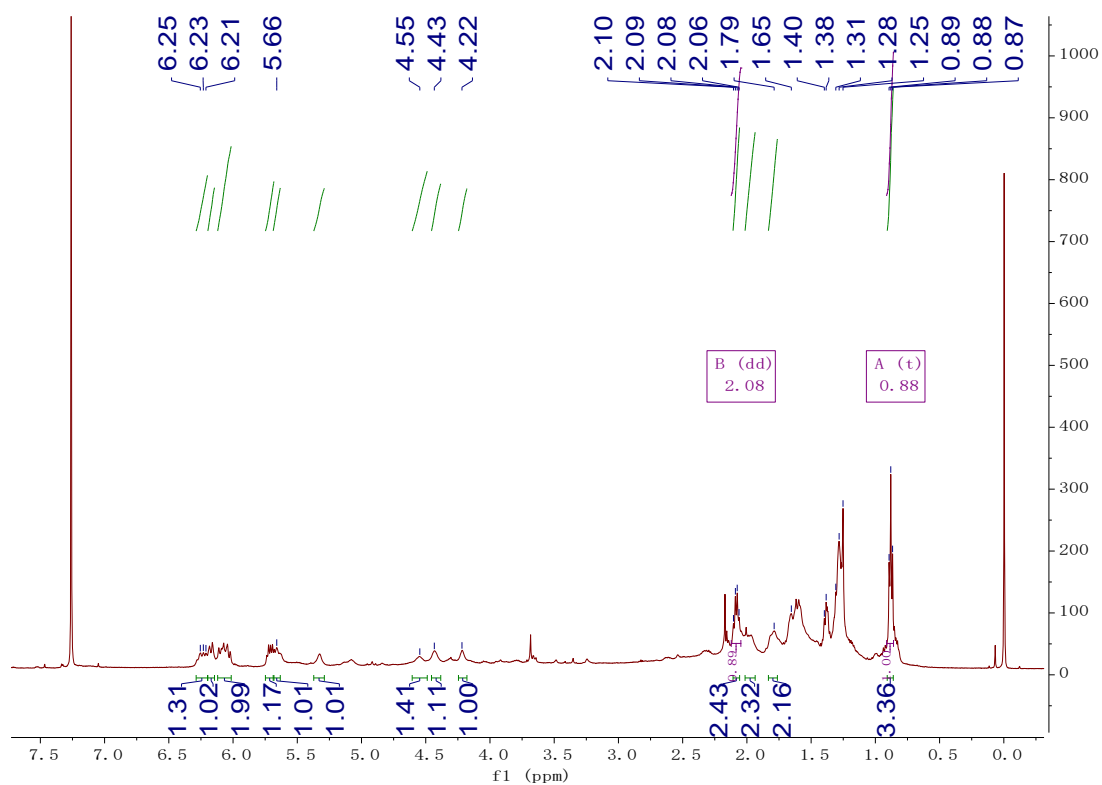


Figure S3. ¹H NMR spectrum of **1** in CDCl₃ (500MHz)

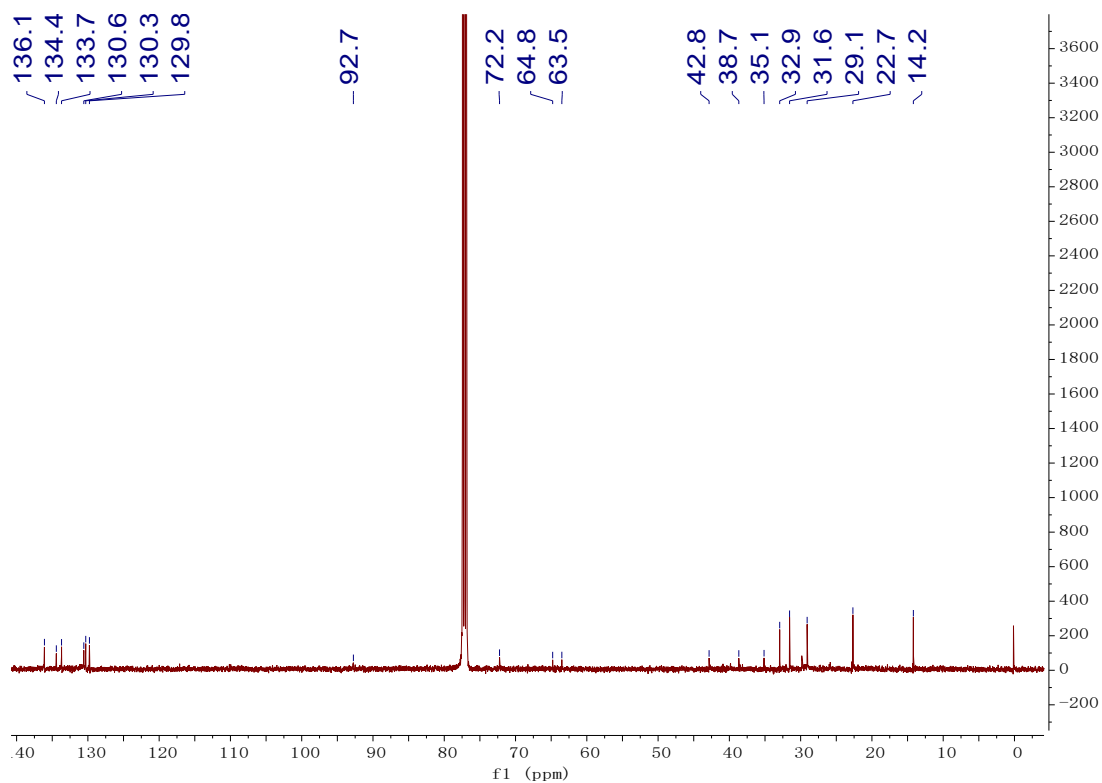


Figure S4. ¹³C NMR spectrum of **1** in CDCl₃ (125MHz)

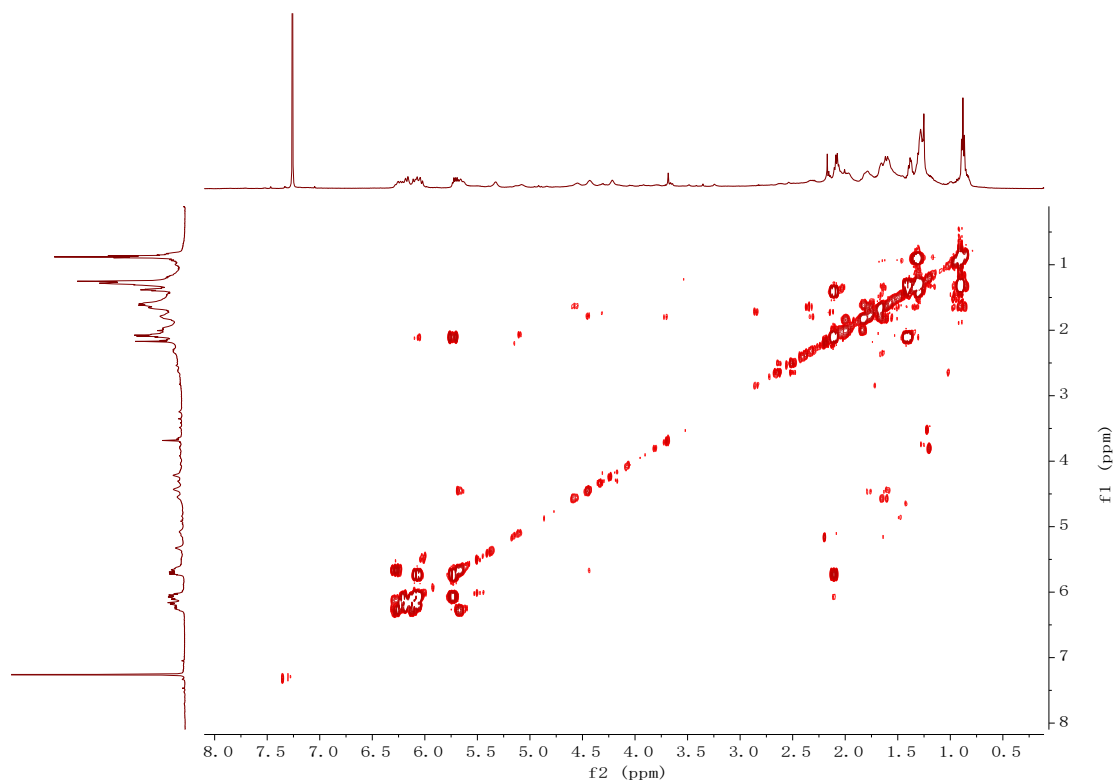


Figure S5. ^1H - ^1H COSY spectrum of **1** in CDCl_3

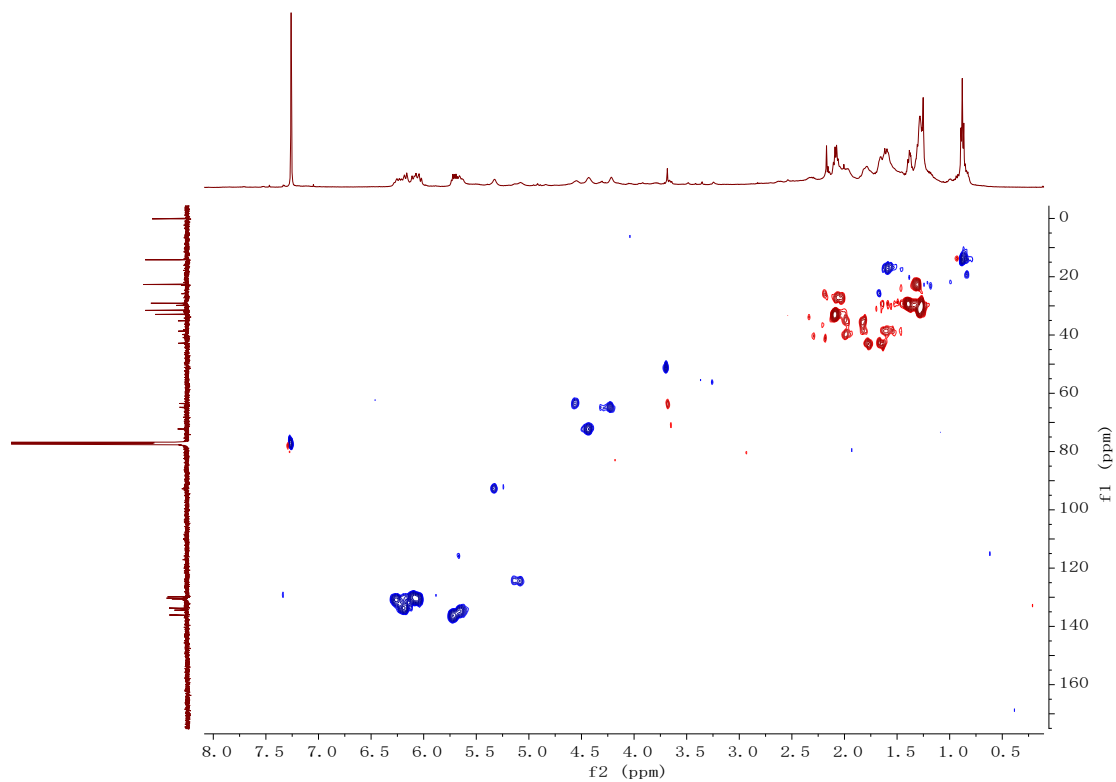


Figure S6. HSQC spectrum of **1** in CDCl_3

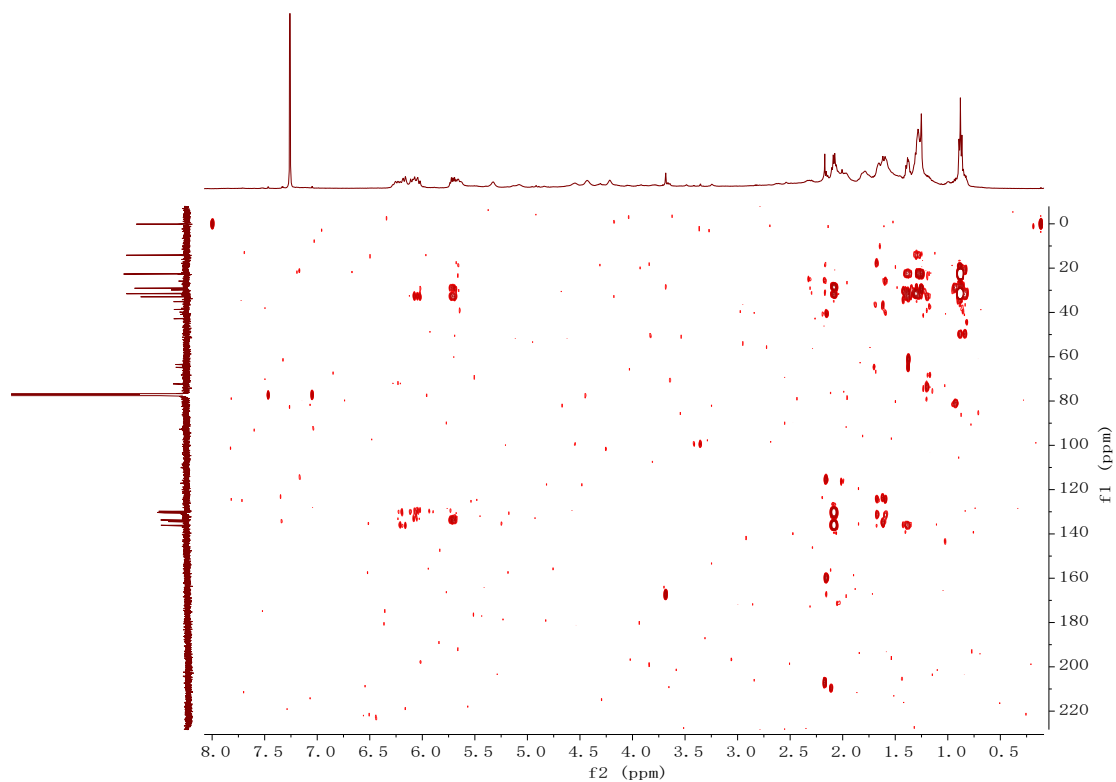


Figure S7. HMBC spectrum of **1** in CDCl_3

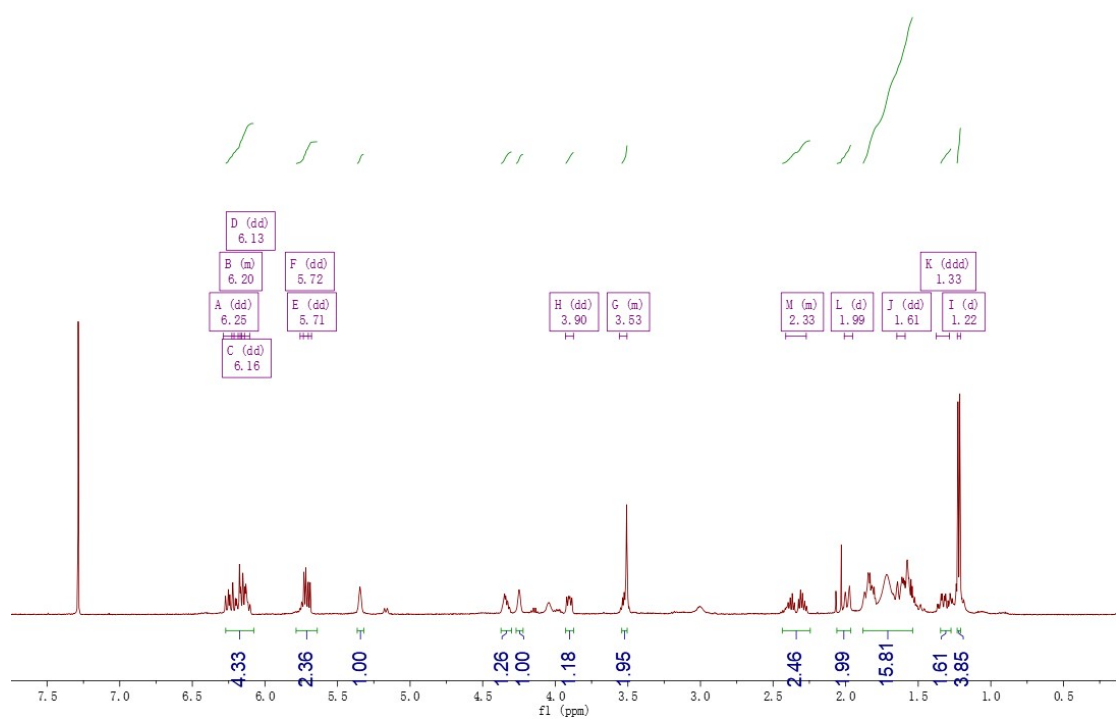


Figure S8. ^1H NMR spectrum of **2** in CDCl_3 (500MHz)

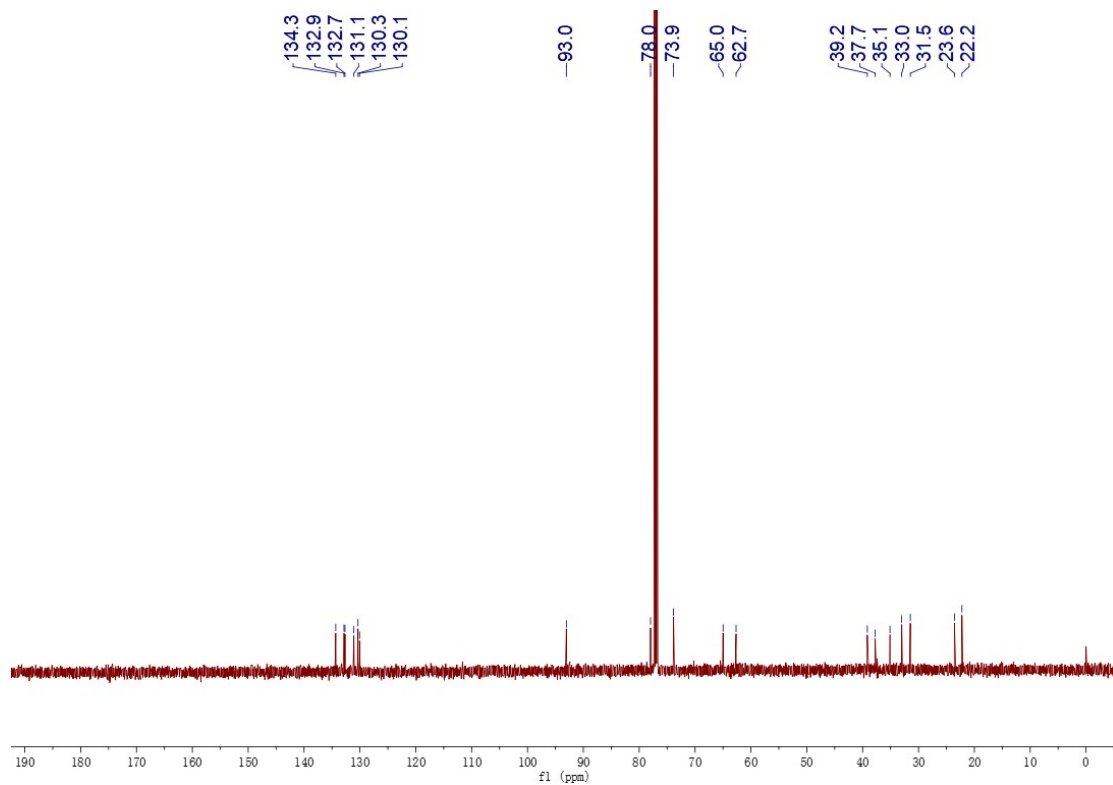


Figure S9. ^{13}C NMR spectrum of **2** in CDCl_3 (125MHz)

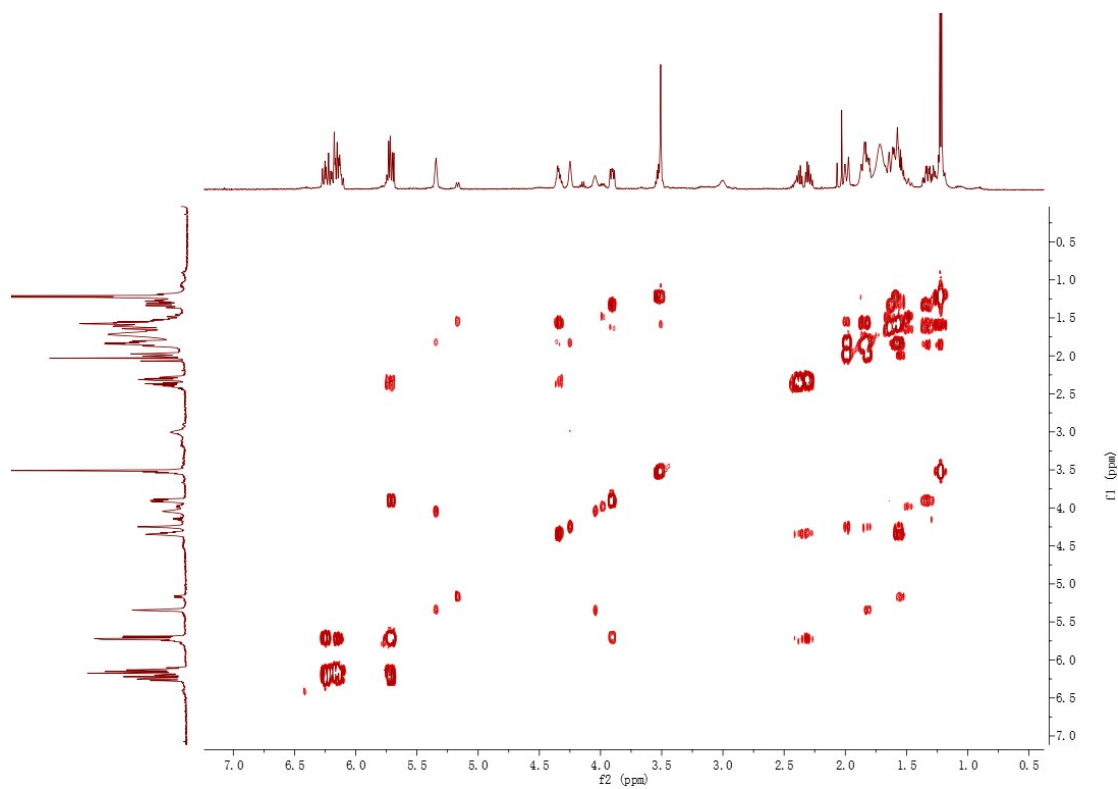


Figure S10. ^1H - ^1H COSY spectrum of **2** in CDCl_3

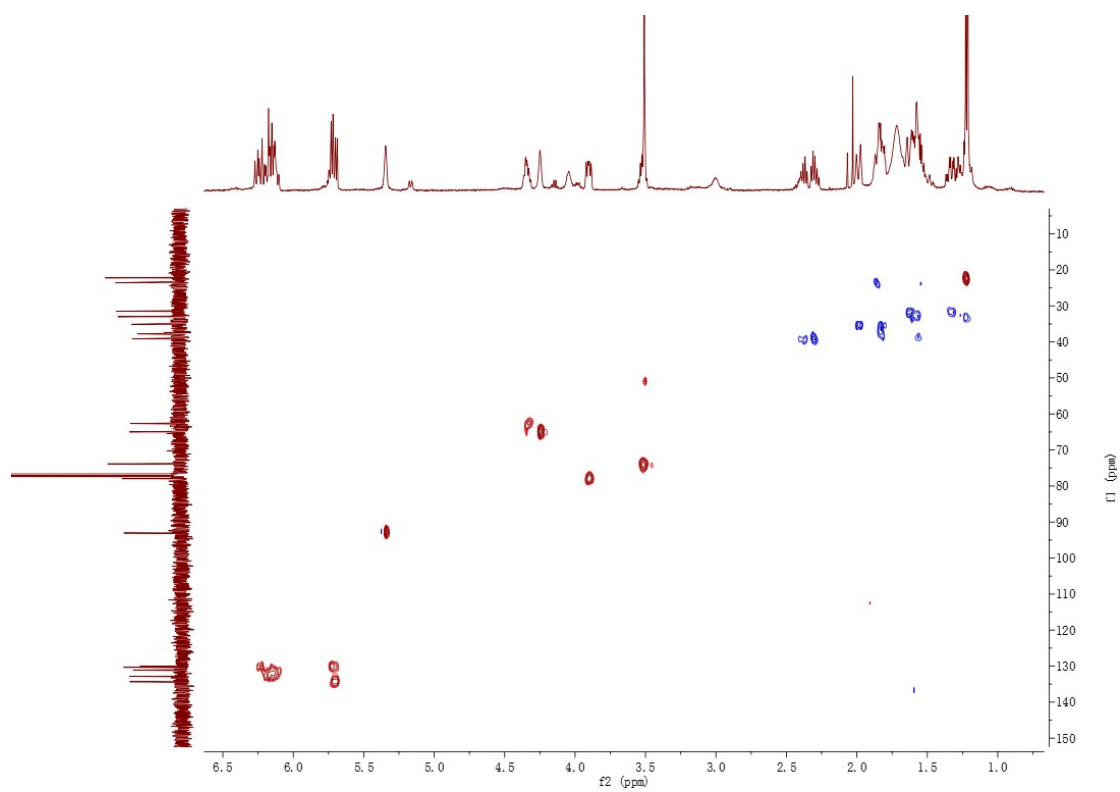


Figure S11. HSQC spectrum of **2** in CDCl_3

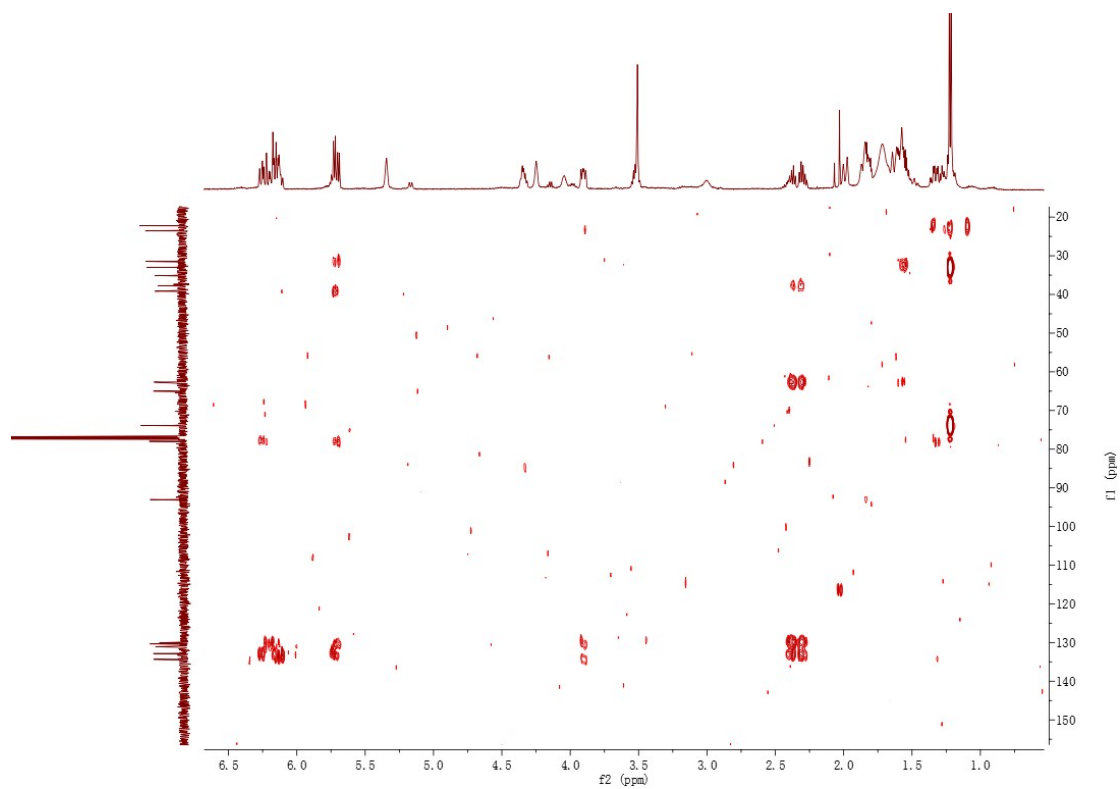


Figure S12. HMBC spectrum of **2** in CDCl_3

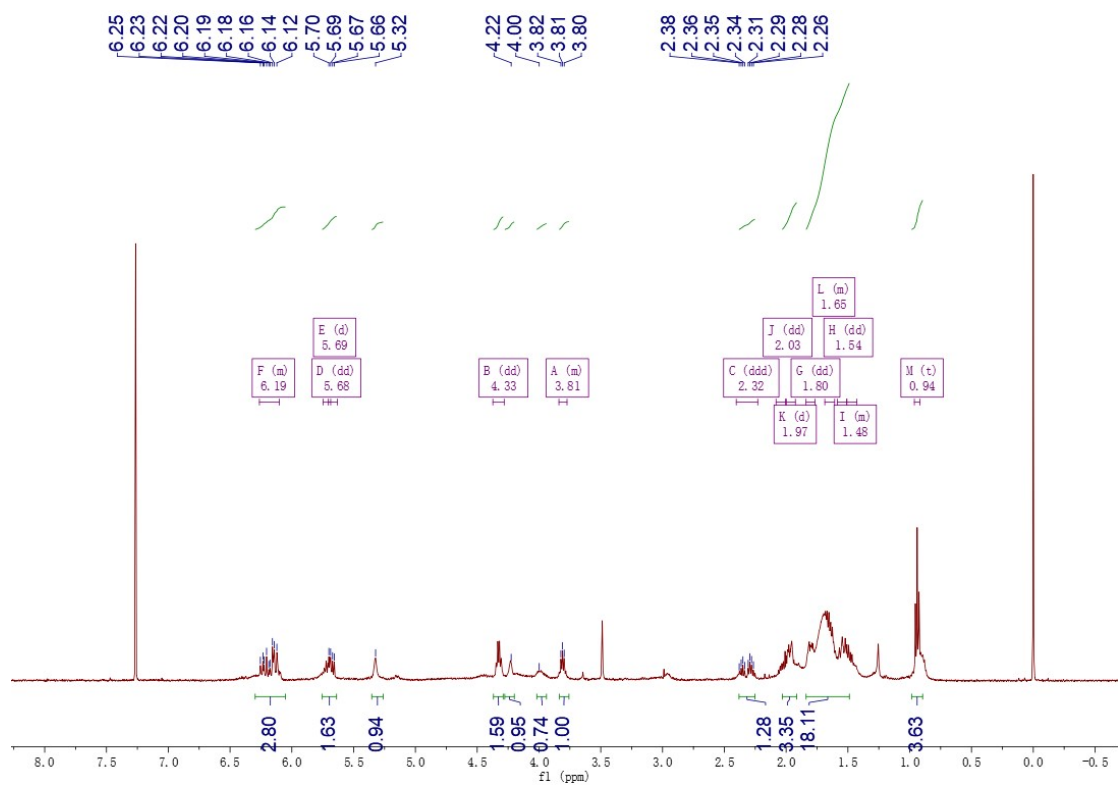


Figure S13. ^1H NMR spectrum of **3** in CDCl_3 (500MHz)

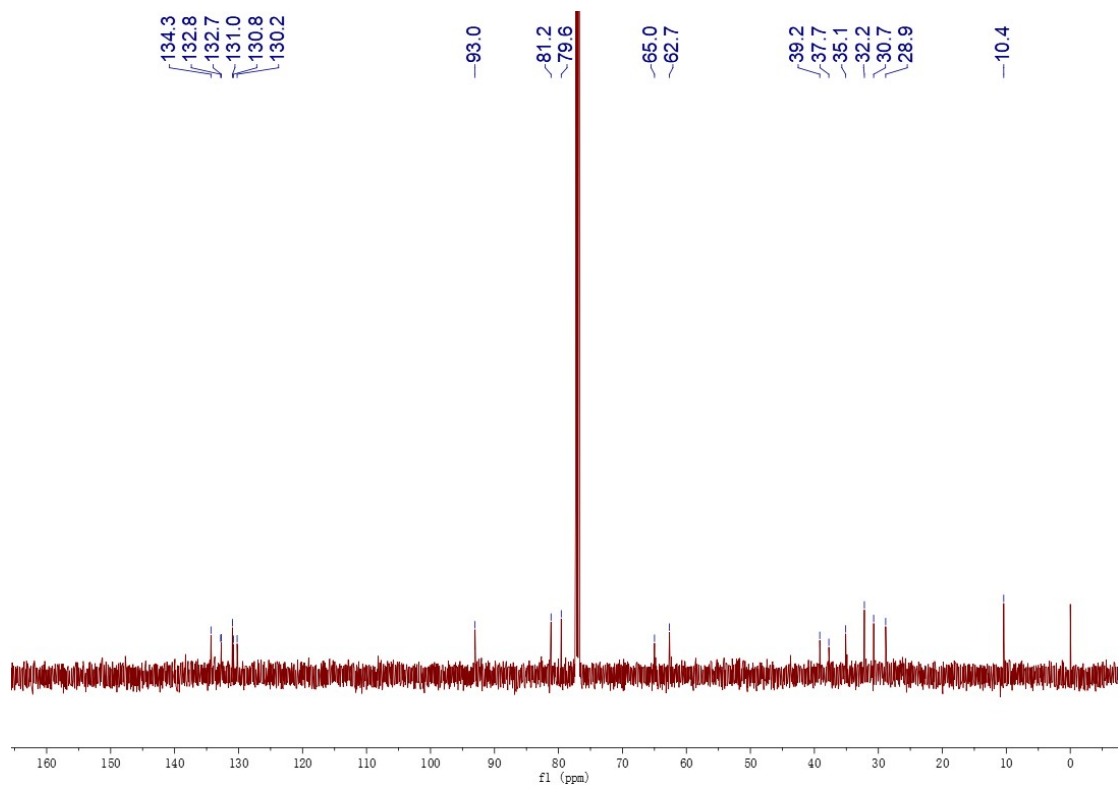


Figure S14. ^{13}C NMR spectrum of **3** in CDCl_3 (125MHz)

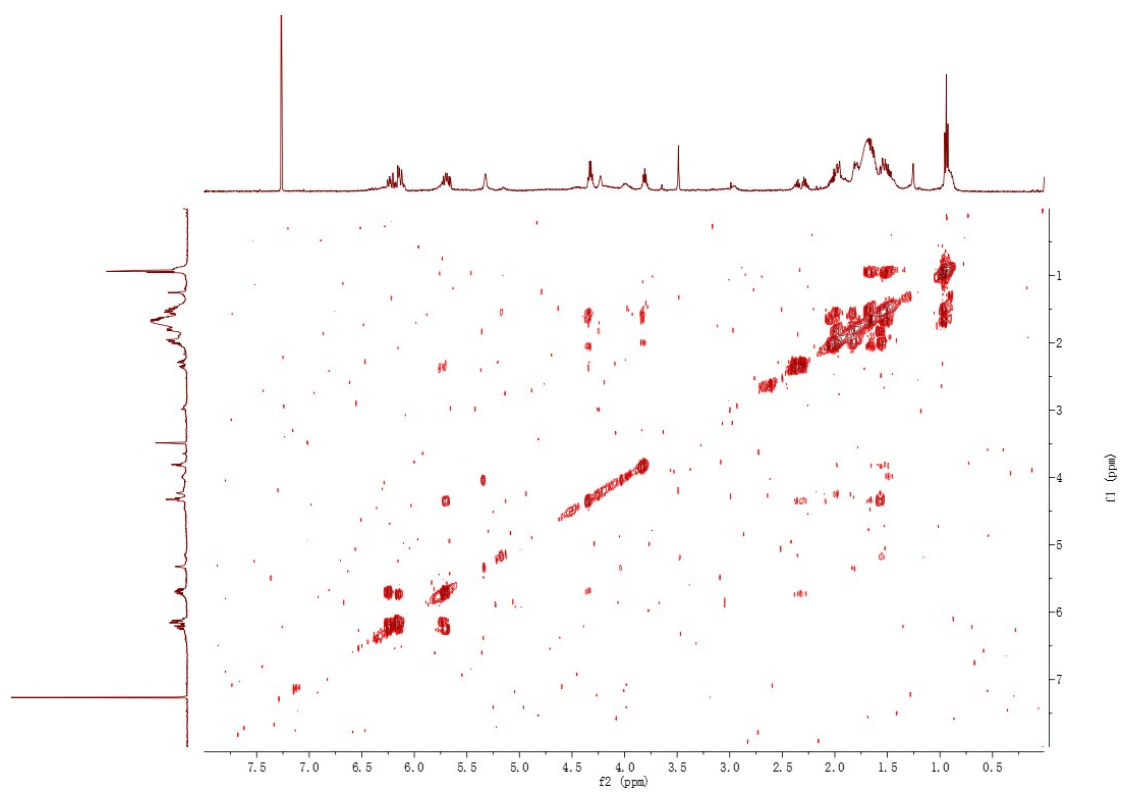


Figure S15. ^1H - ^1H COSY spectrum of **3** in CDCl_3

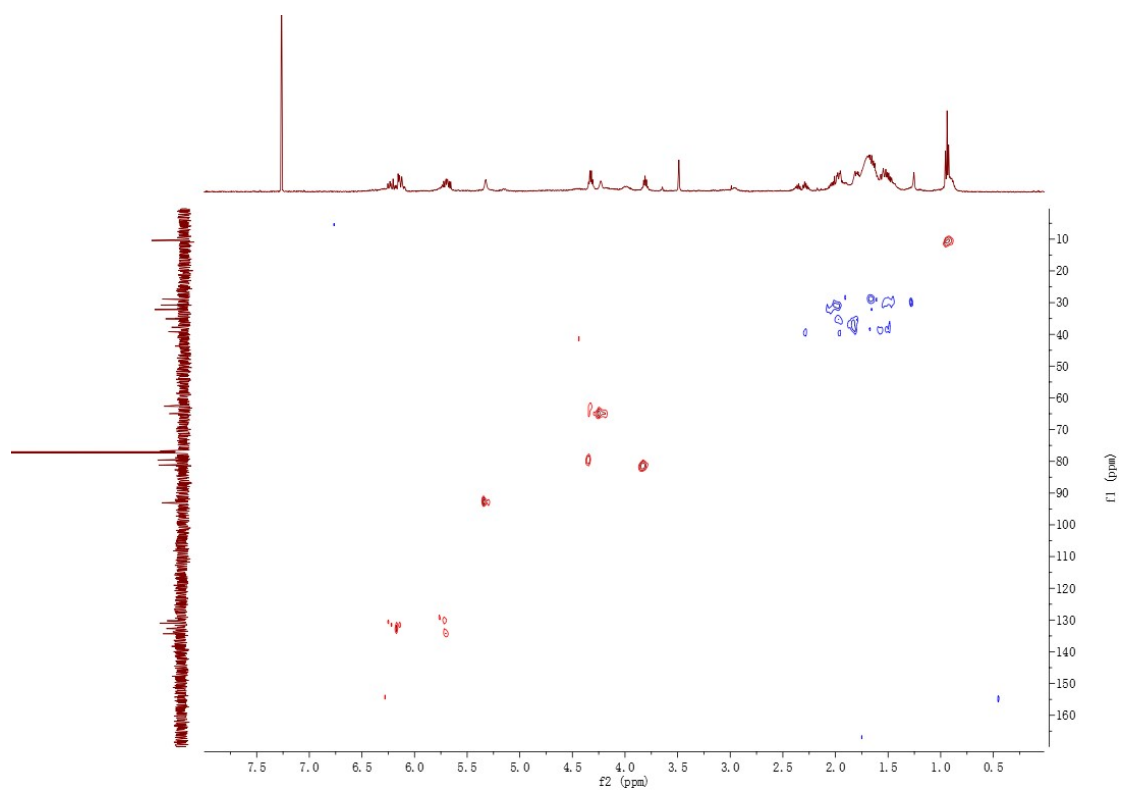


Figure S16. HSQC spectrum of **3** in CDCl_3

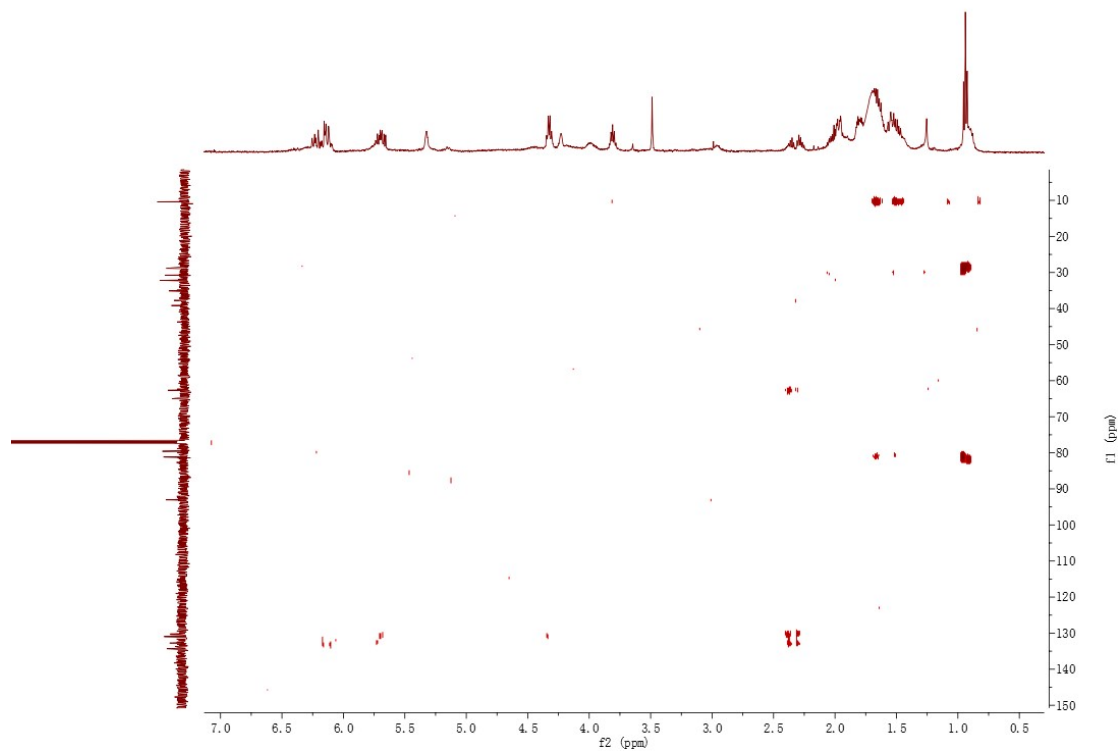


Figure S17. HMBC spectrum of **3** in CDCl_3

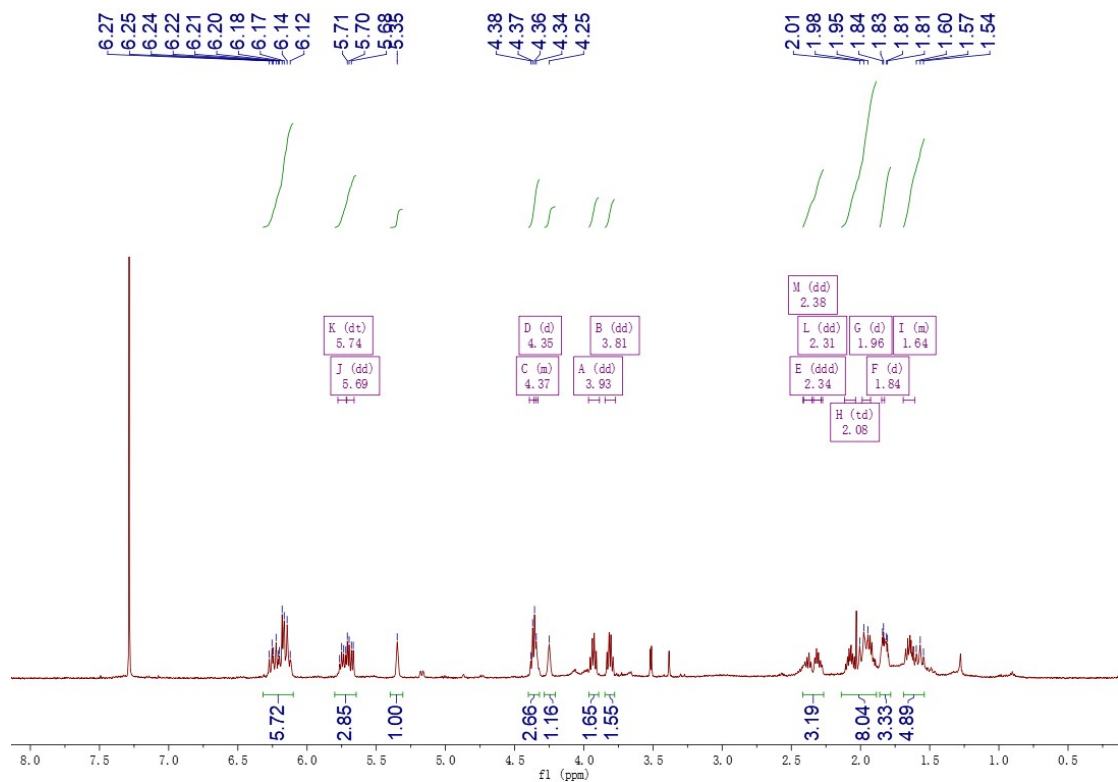


Figure S18. ^1H NMR spectrum of **4** in CDCl_3 (500MHz)

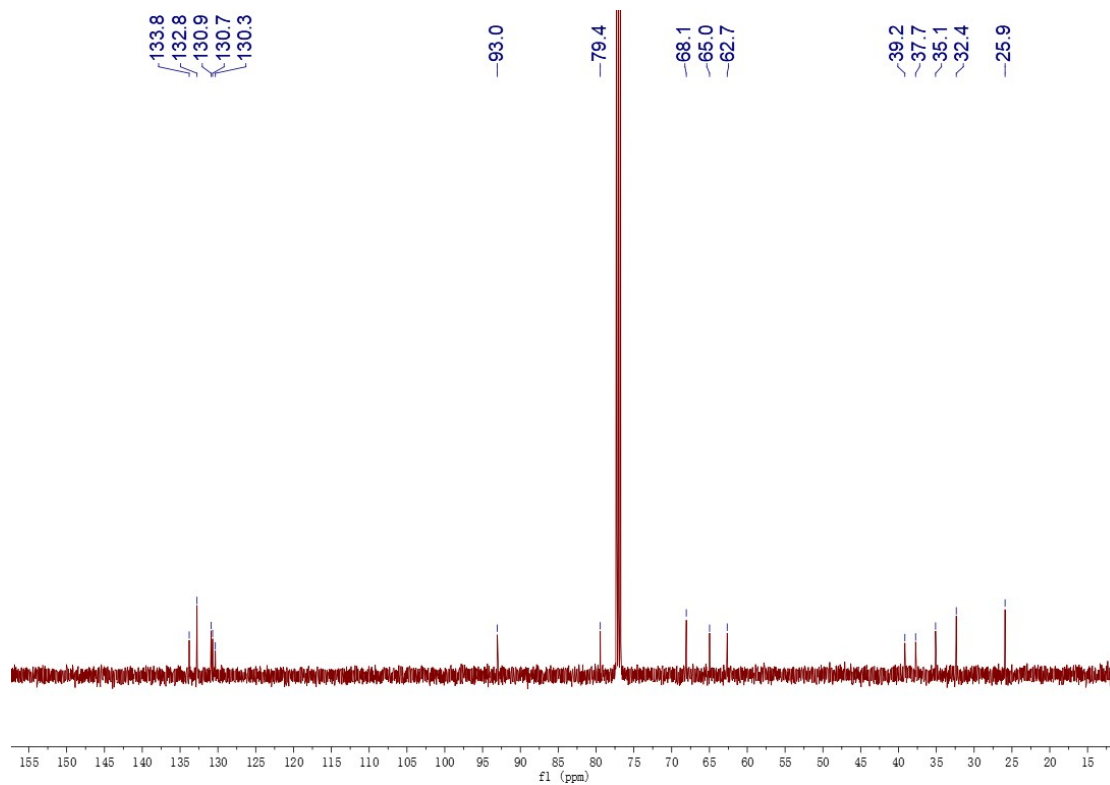


Figure S19. ^{13}C NMR spectrum of **4** in CDCl_3 (125MHz)

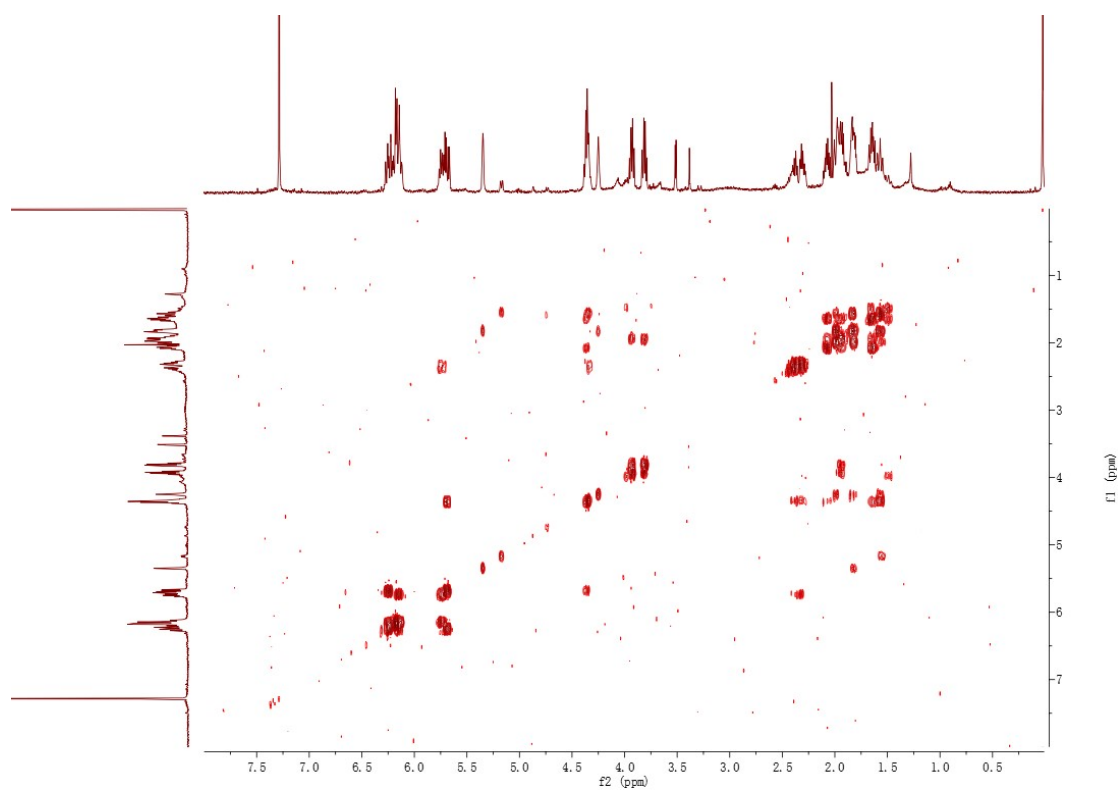


Figure S20. ^1H - ^1H COSY spectrum of **4** in CDCl_3

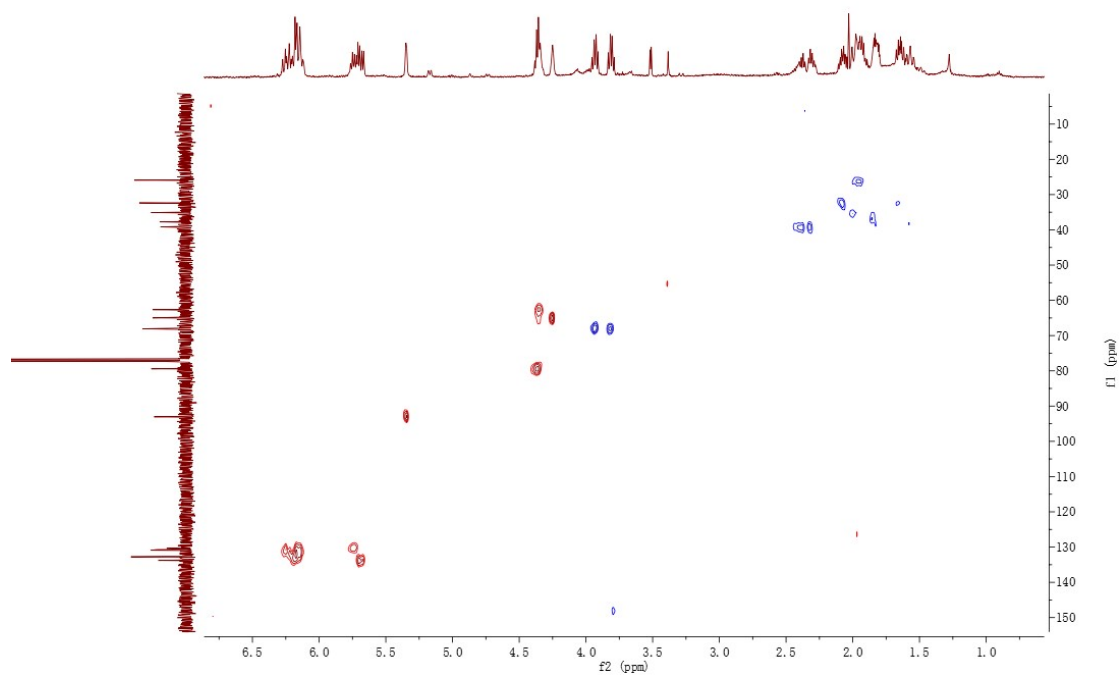


Figure S21. HSQC spectrum of **4** in CDCl_3

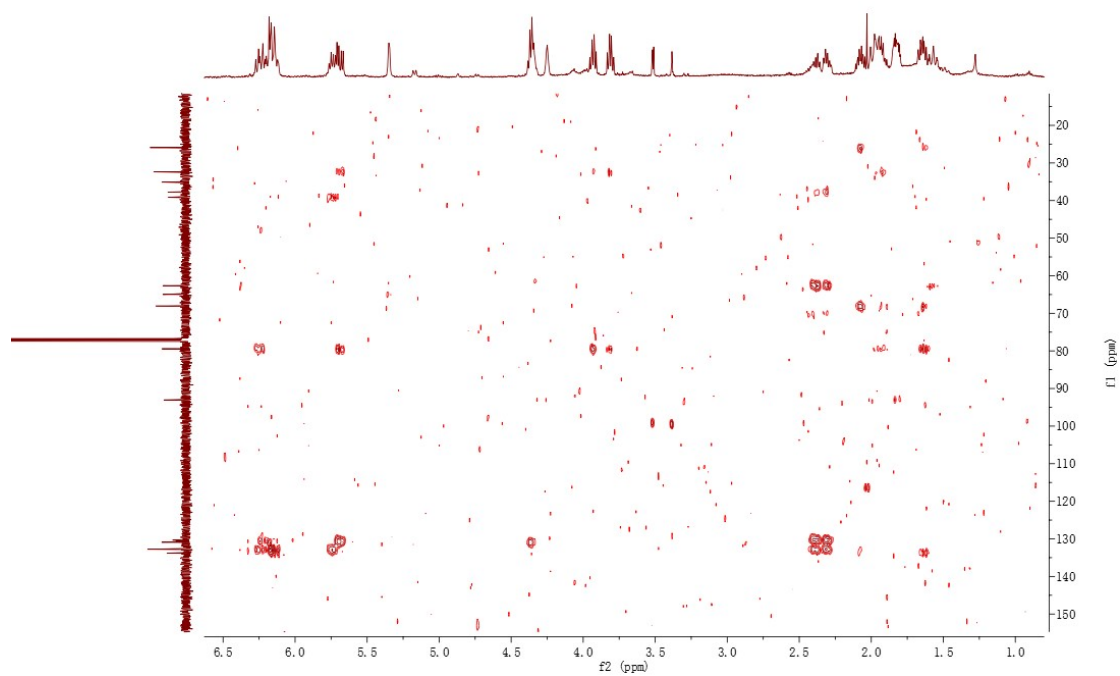


Figure S22. HMBC spectrum of **4** in CDCl_3

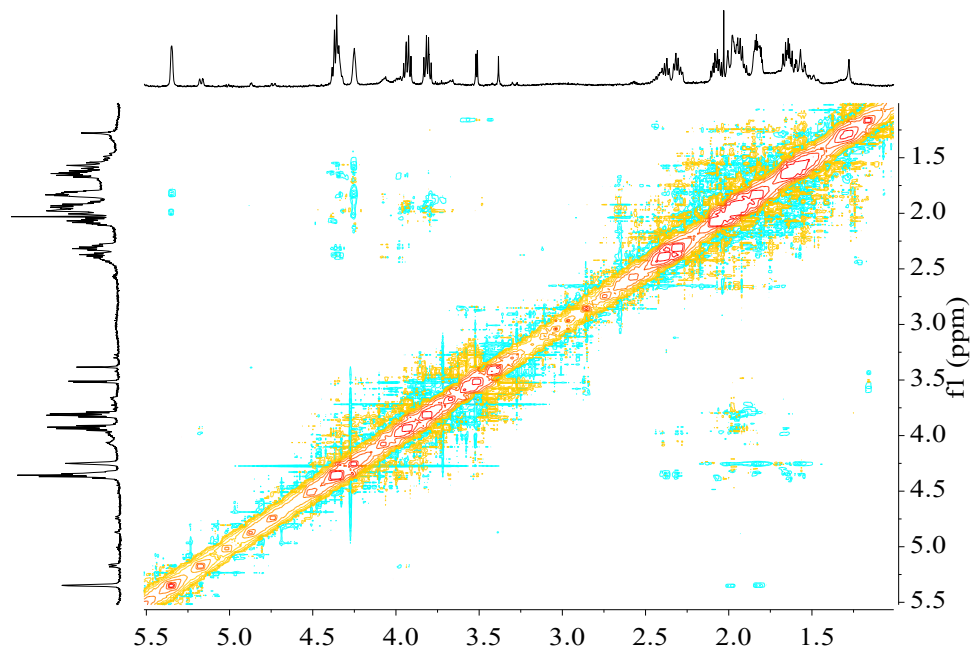


Figure S23. ROESY of **4** in CDCl_3 (500MHz)

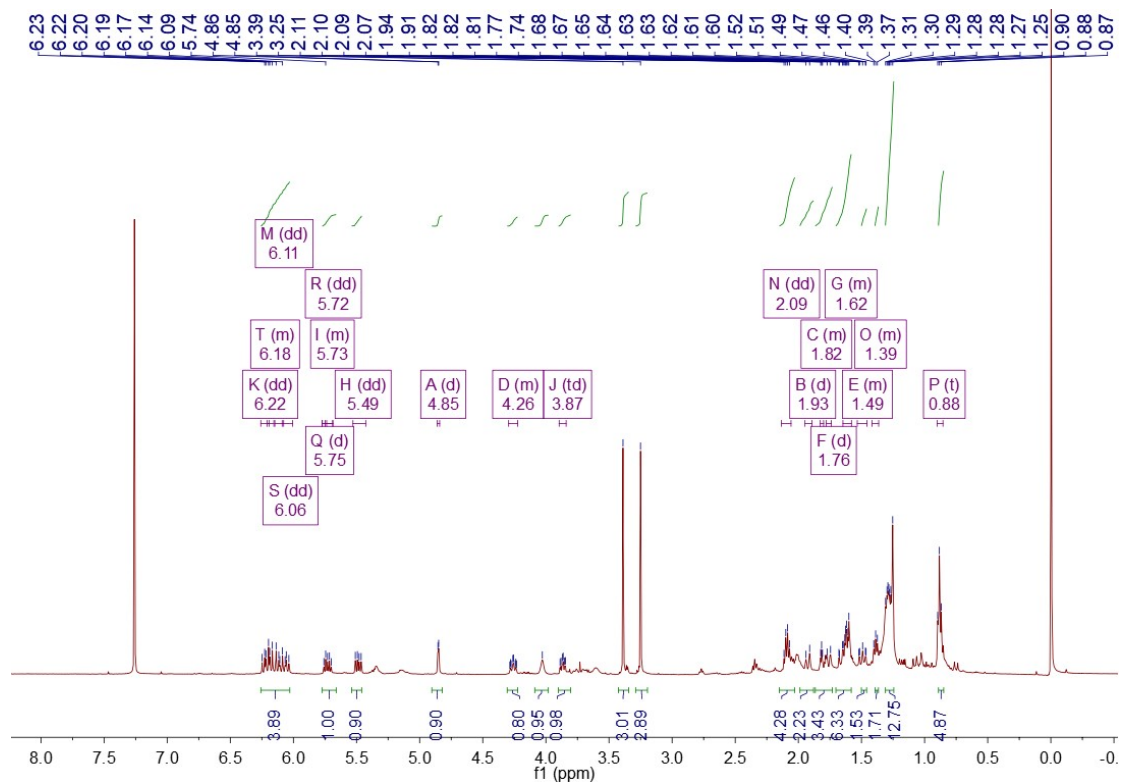


Figure S24. ^1H NMR spectrum of **5** in CDCl_3 (500MHz)

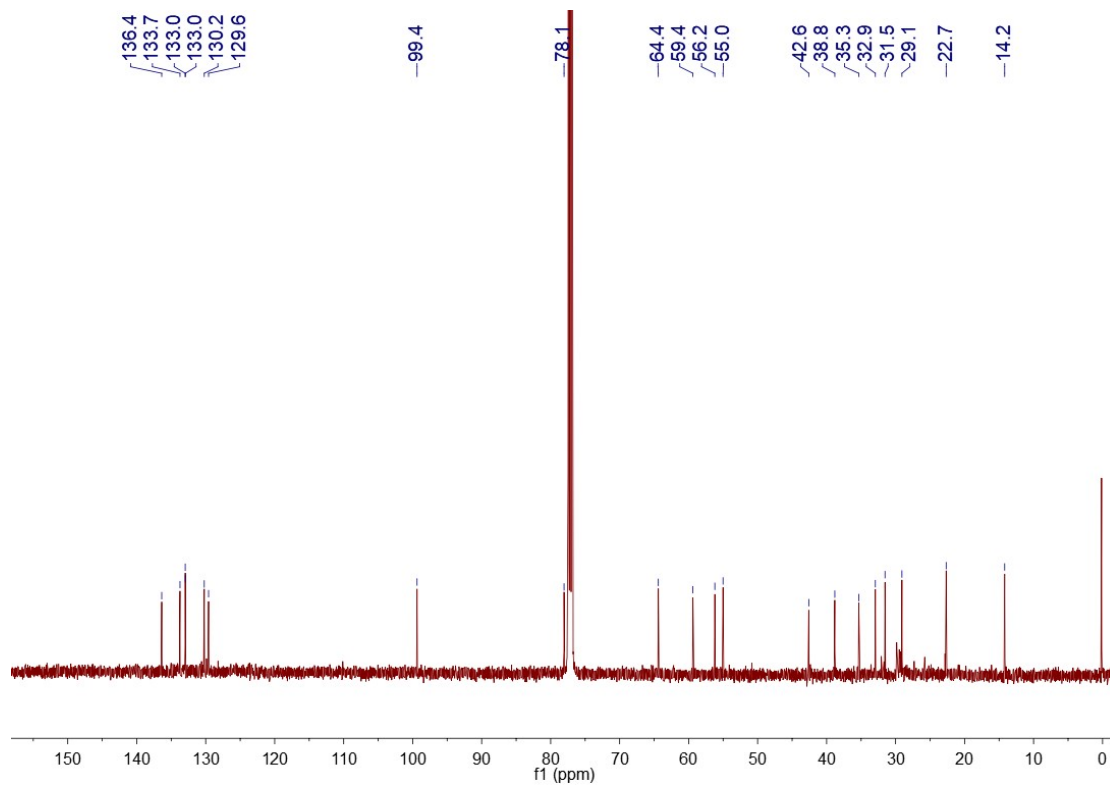


Figure S25. ^{13}C NMR spectrum of **5** in CDCl_3 (125MHz)

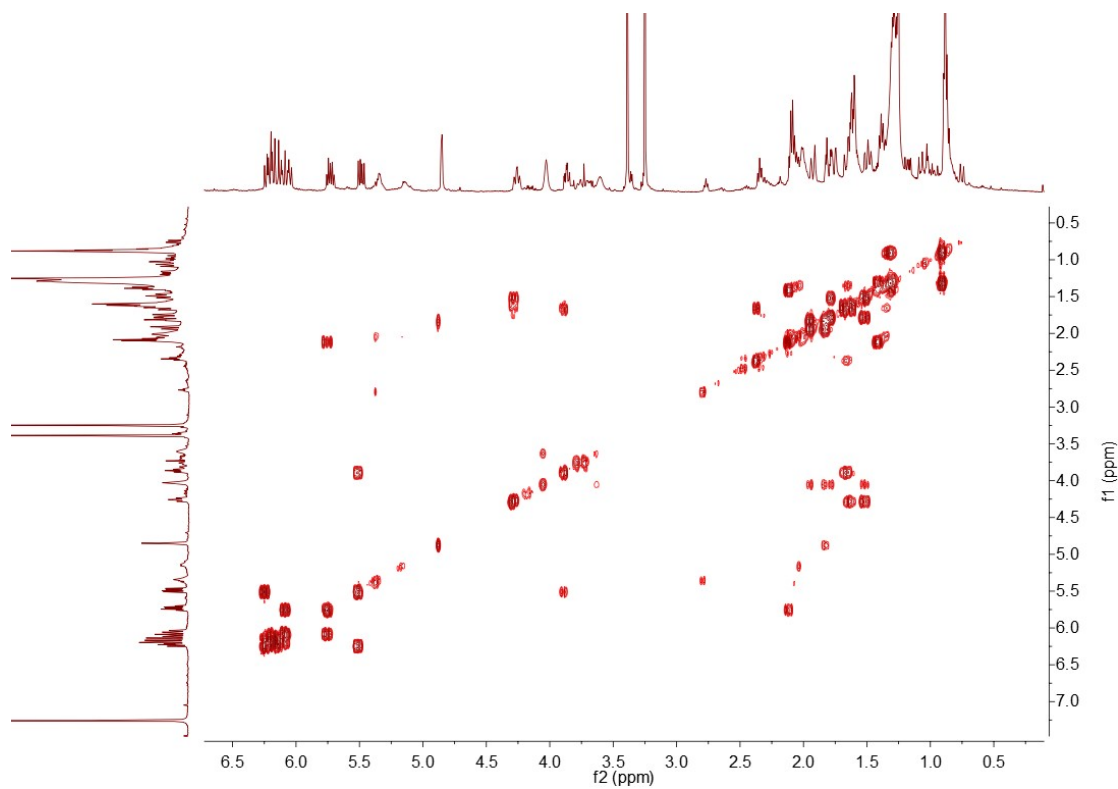


Figure S26. ^1H - ^1H COSY spectrum of **5** in CDCl_3

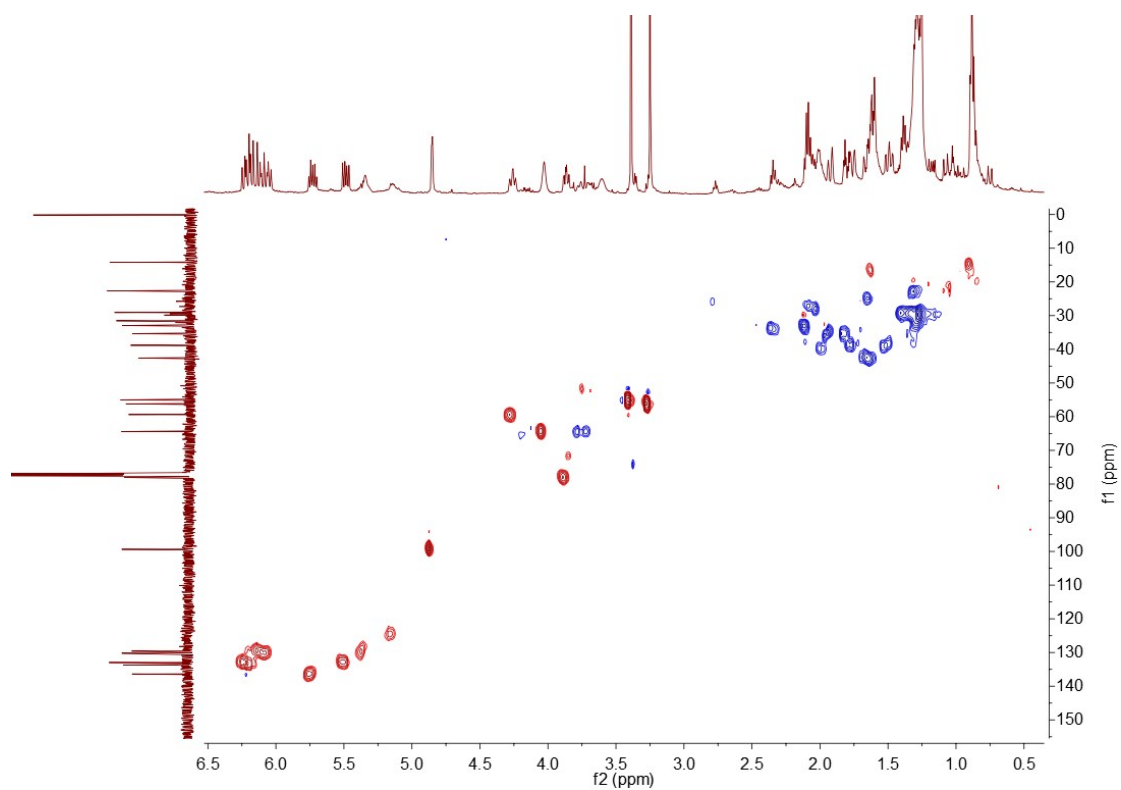


Figure S27. HSQC spectrum of **5** in CDCl_3

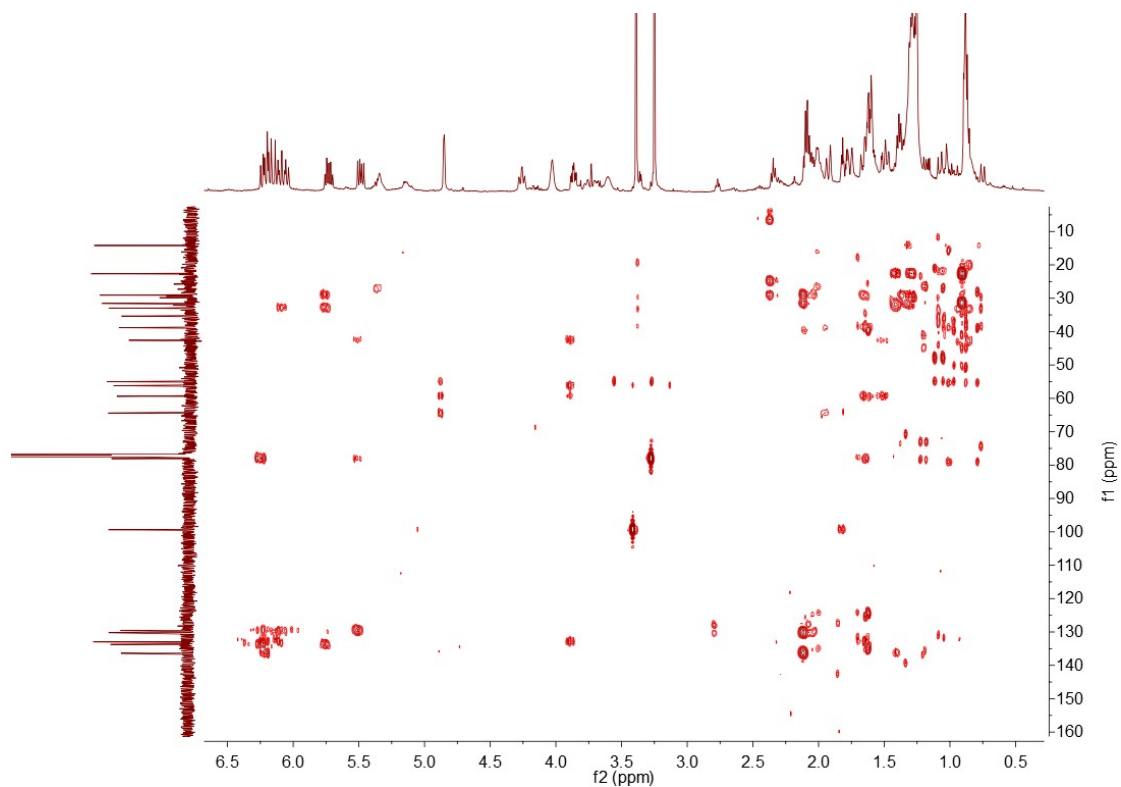


Figure S28. HMBC spectrum of **5** in CDCl_3

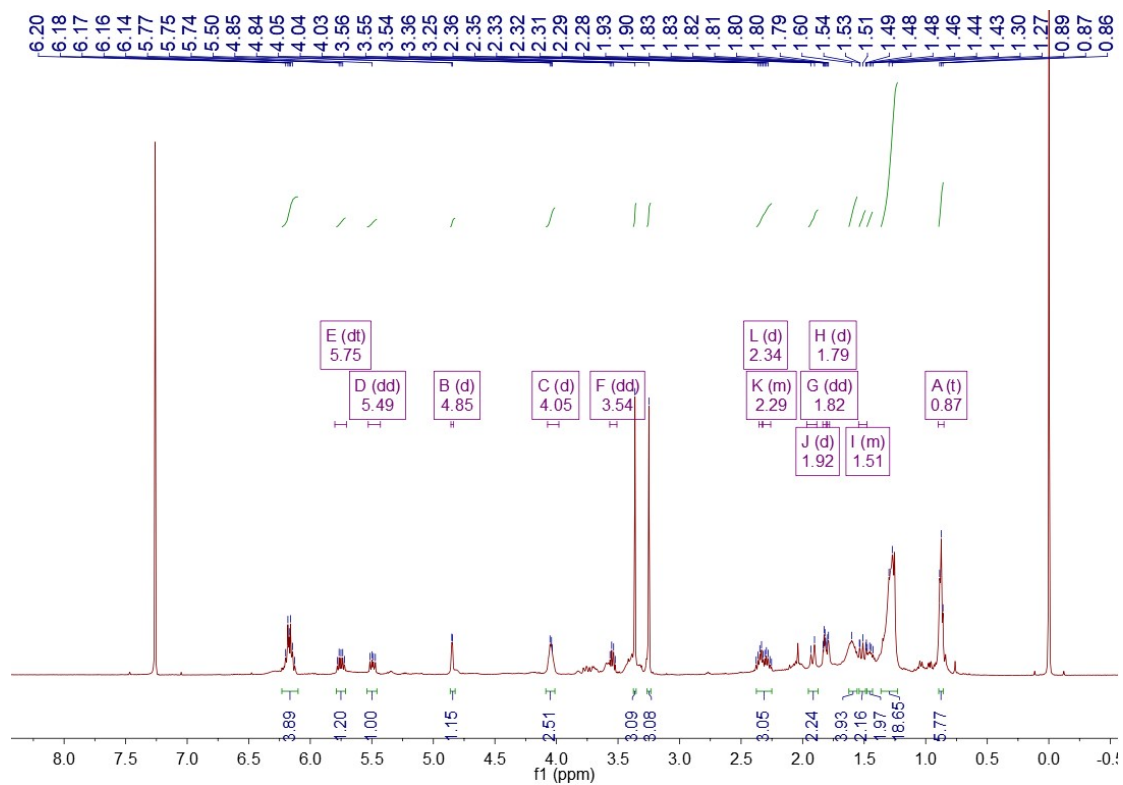


Figure S29. ^1H NMR spectrum of **6** in CDCl_3 (500MHz)

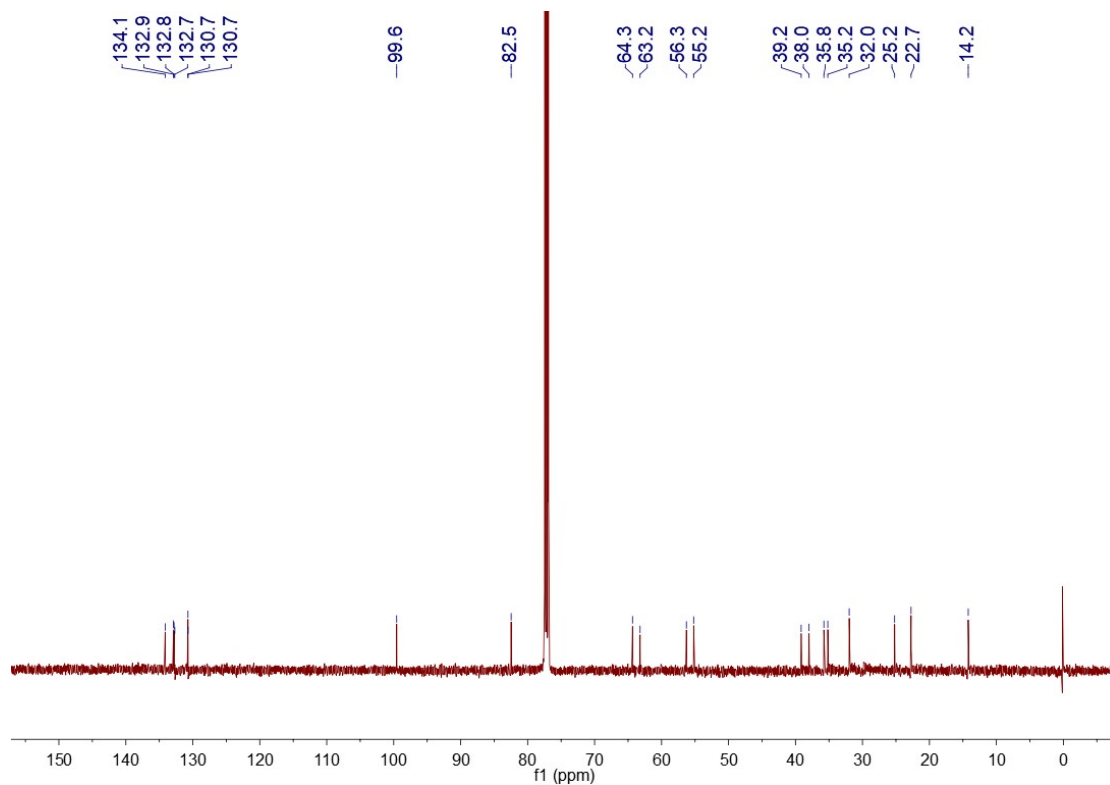


Figure S30. ^{13}C NMR spectrum of **6** in CDCl_3 (125MHz)

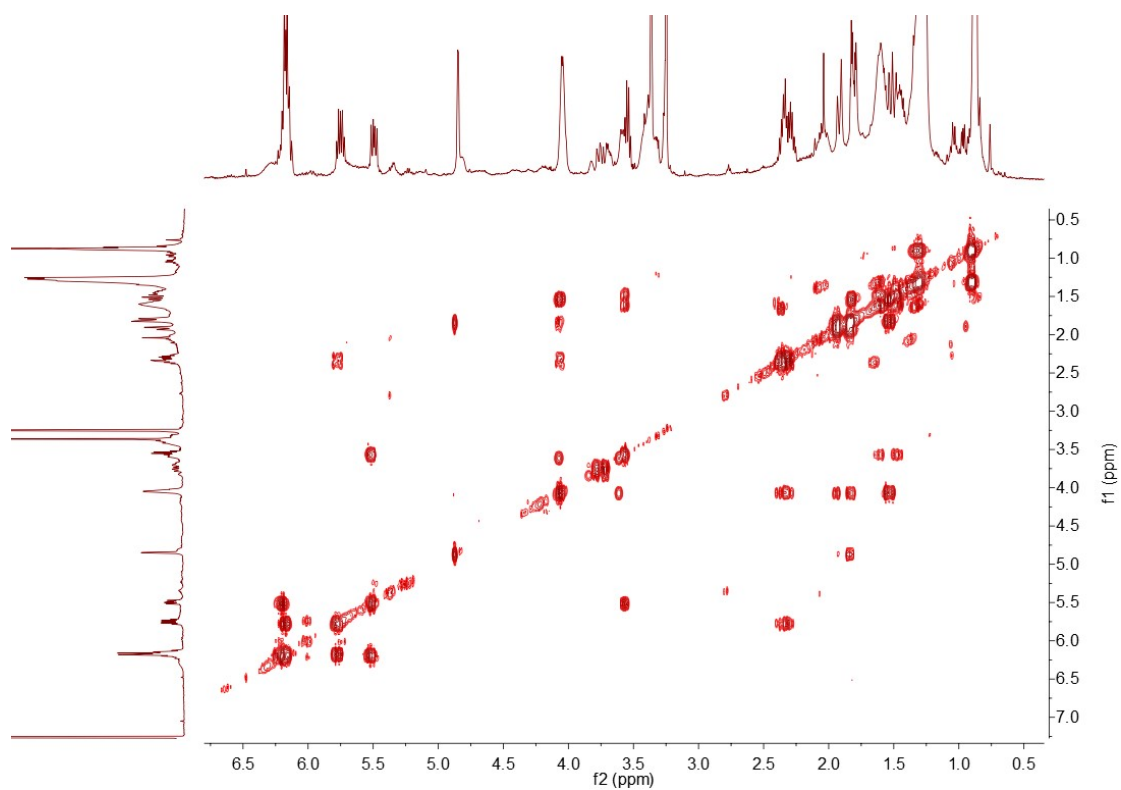


Figure S31. ^1H - ^1H COSY spectrum of **6** in CDCl_3

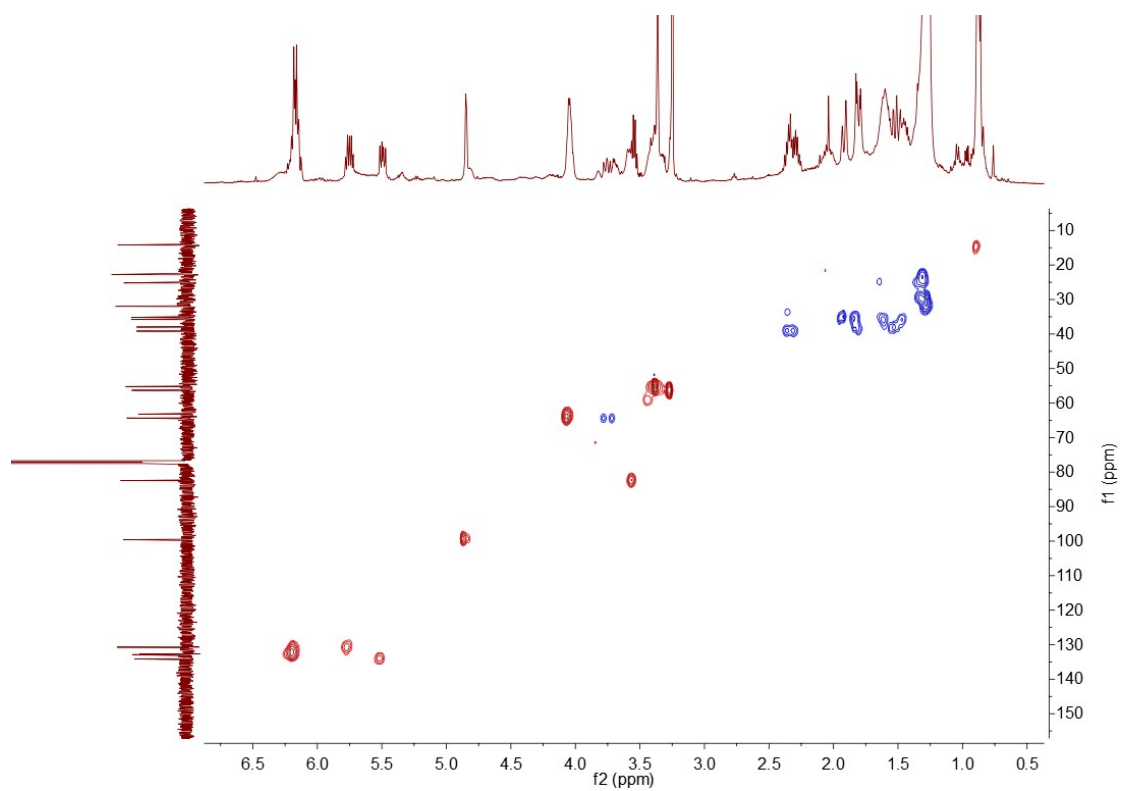


Figure S32. HSQC spectrum of **6** in CDCl_3

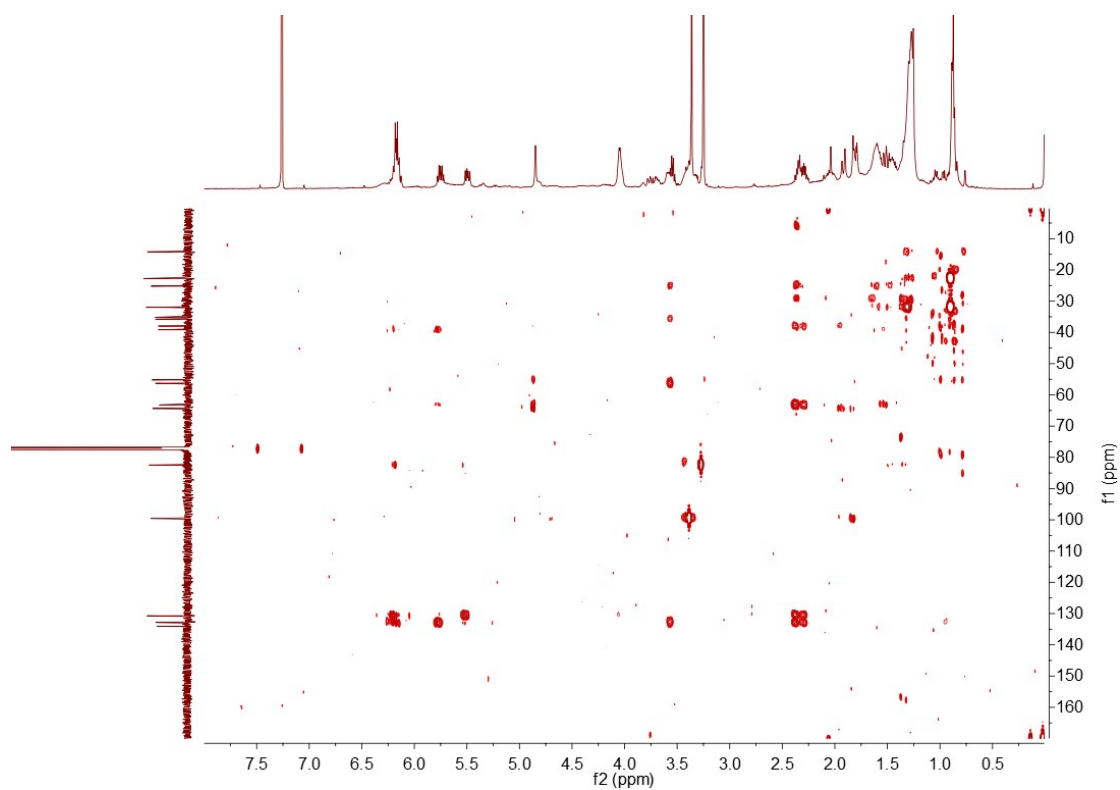


Figure S33. HMBC spectrum of **6** in CDCl_3

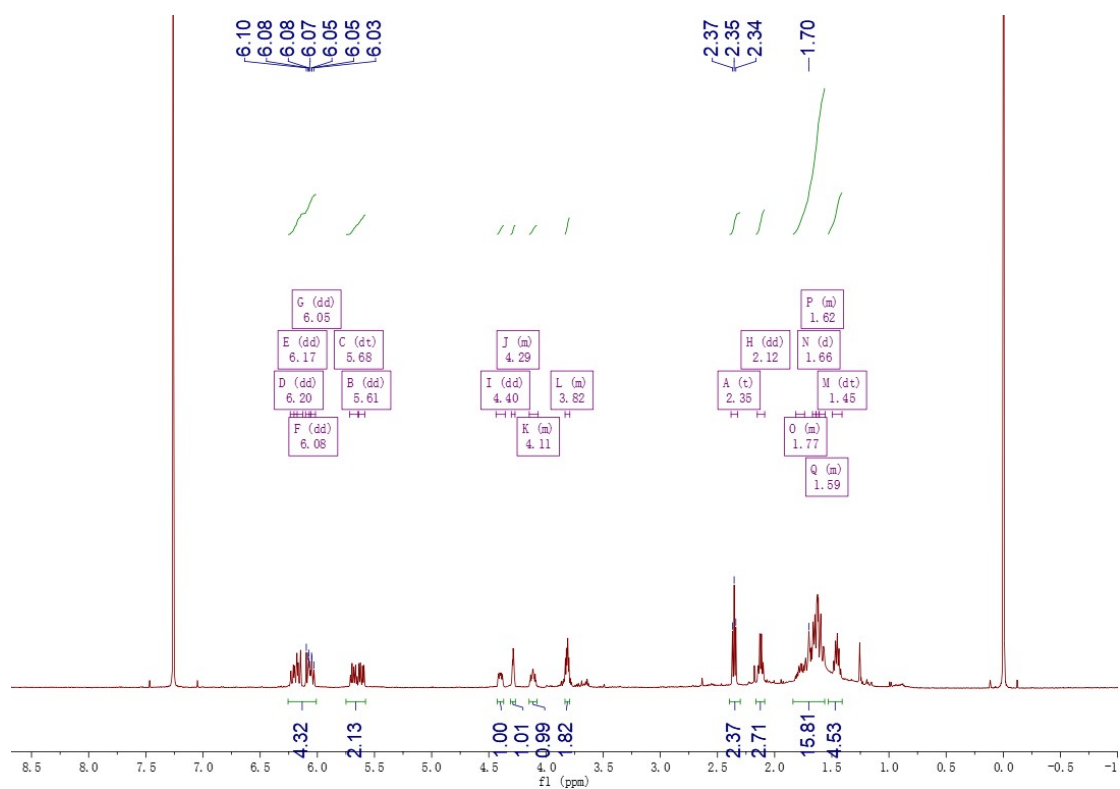


Figure S34. ^1H NMR spectrum of **7** in CDCl_3 (500MHz)

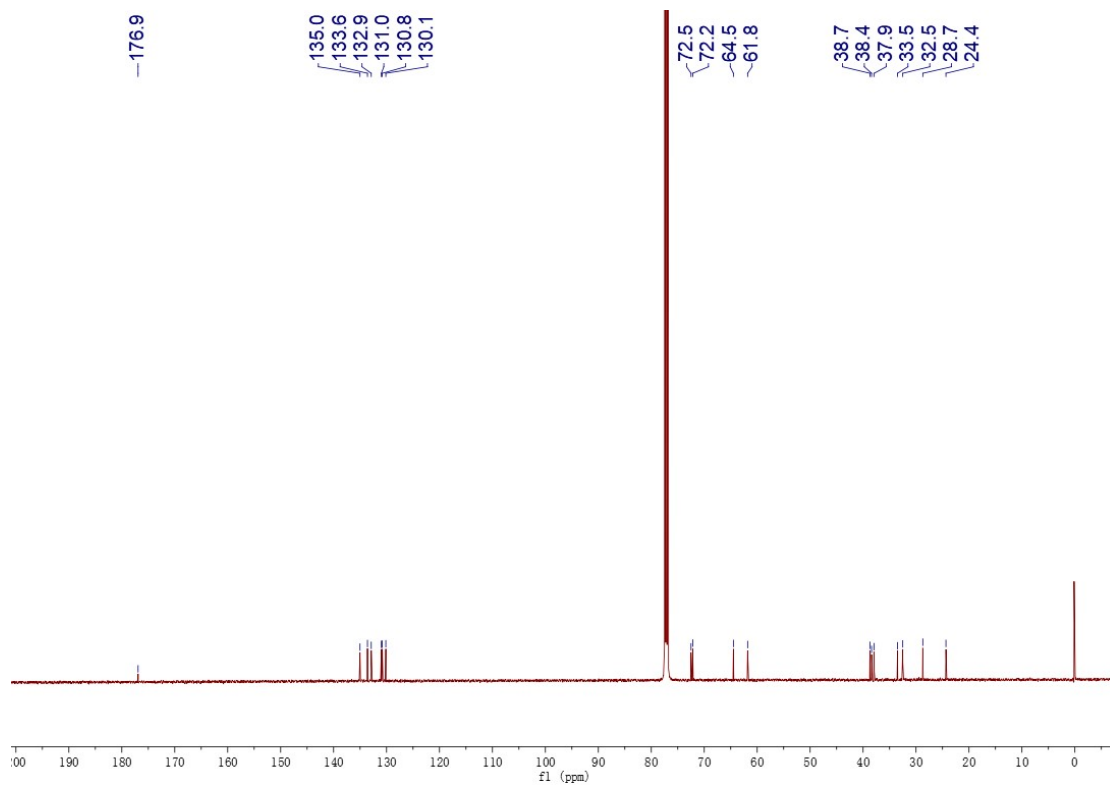


Figure S35. ^{13}C NMR spectrum of **7** in CDCl_3 (125MHz)

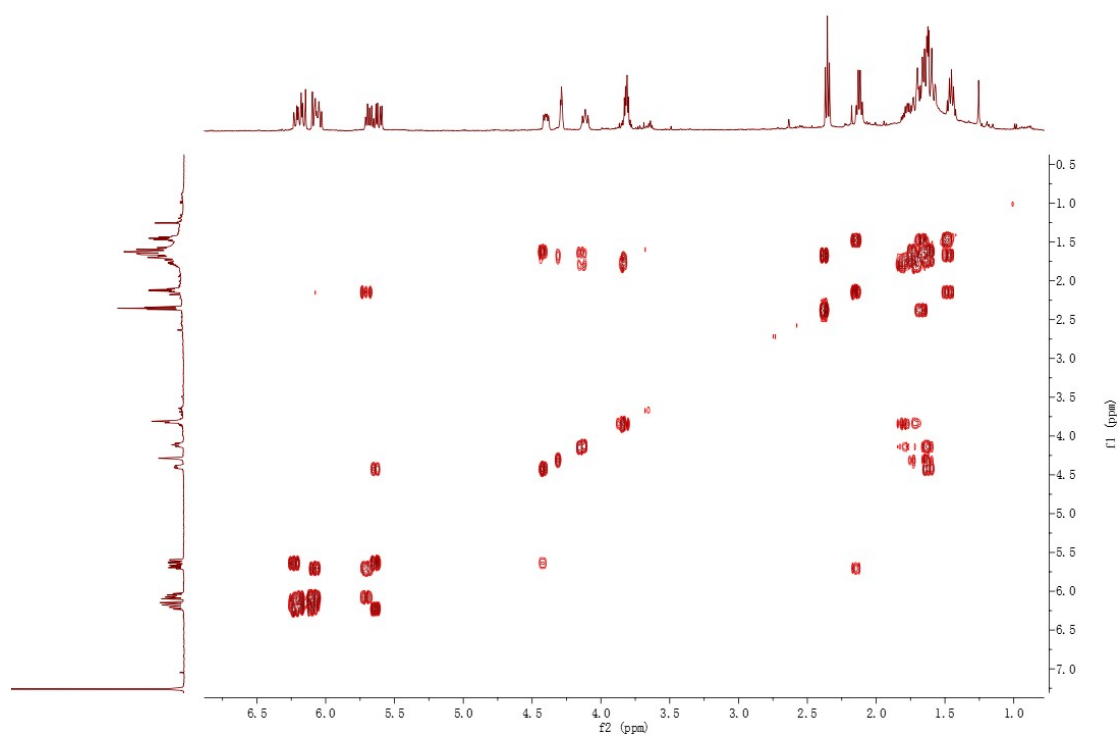


Figure S36. ^1H - ^1H COSY spectrum of **7** in CDCl_3

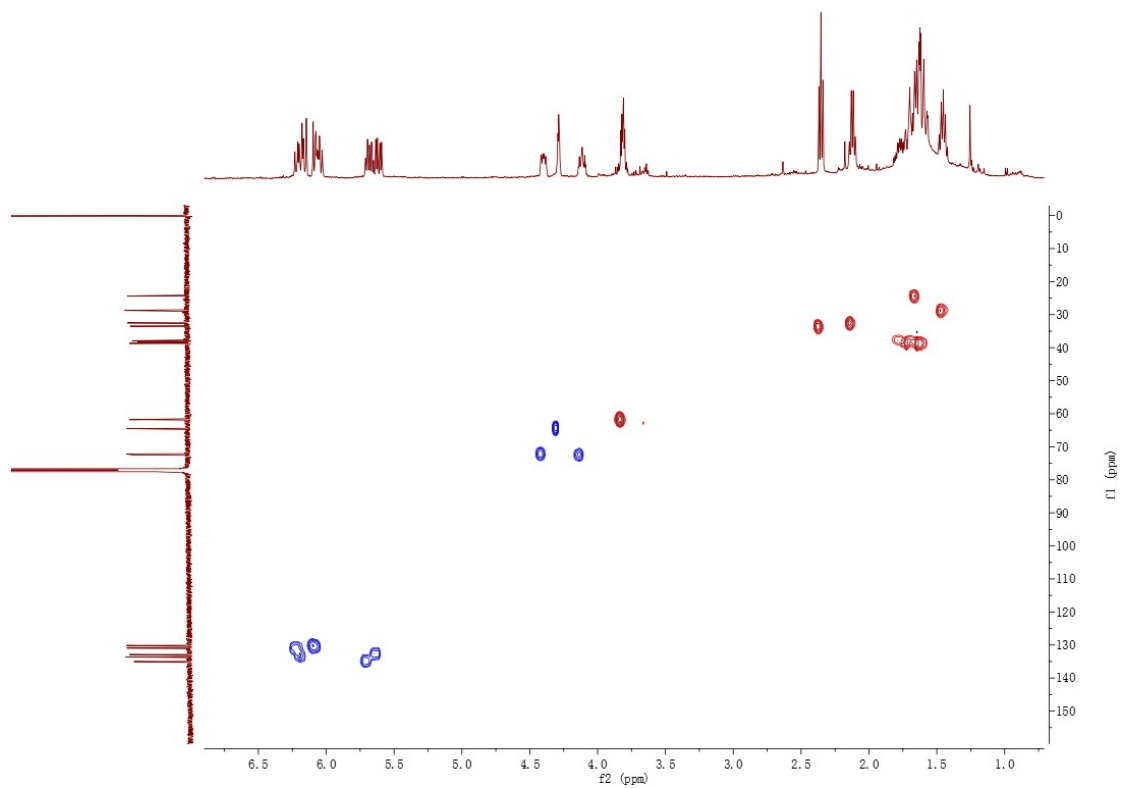


Figure S37. HSQC spectrum of 7 in CDCl_3

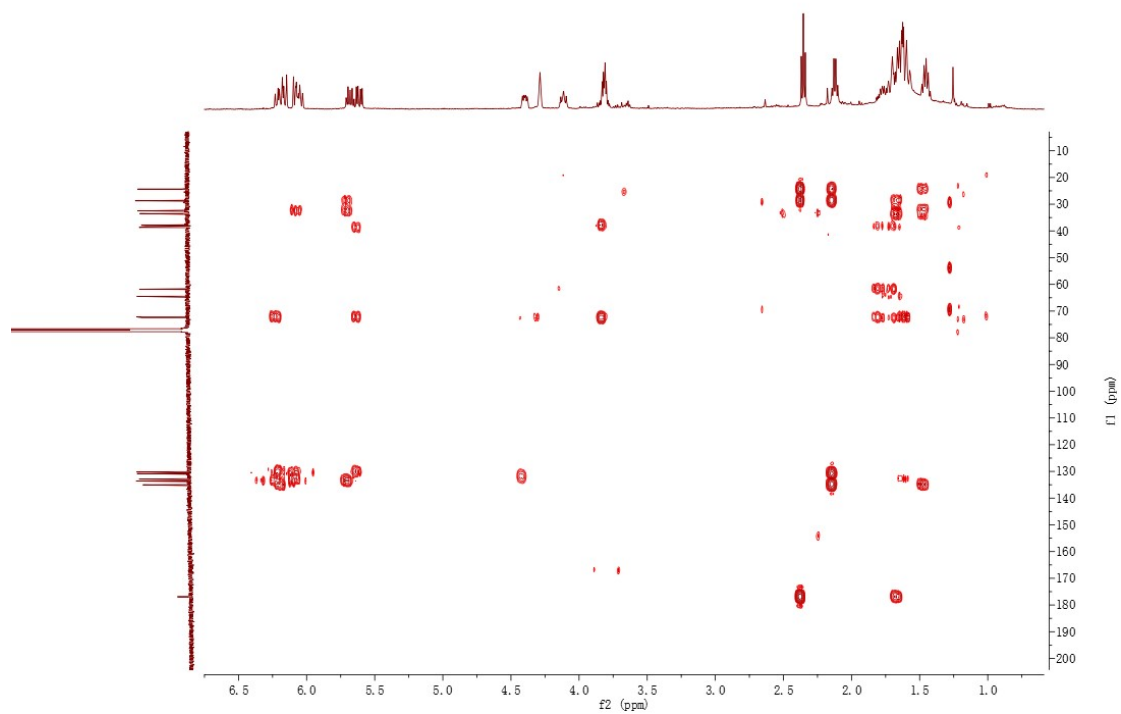


Figure S38. HMBC spectrum of 7 in CDCl_3

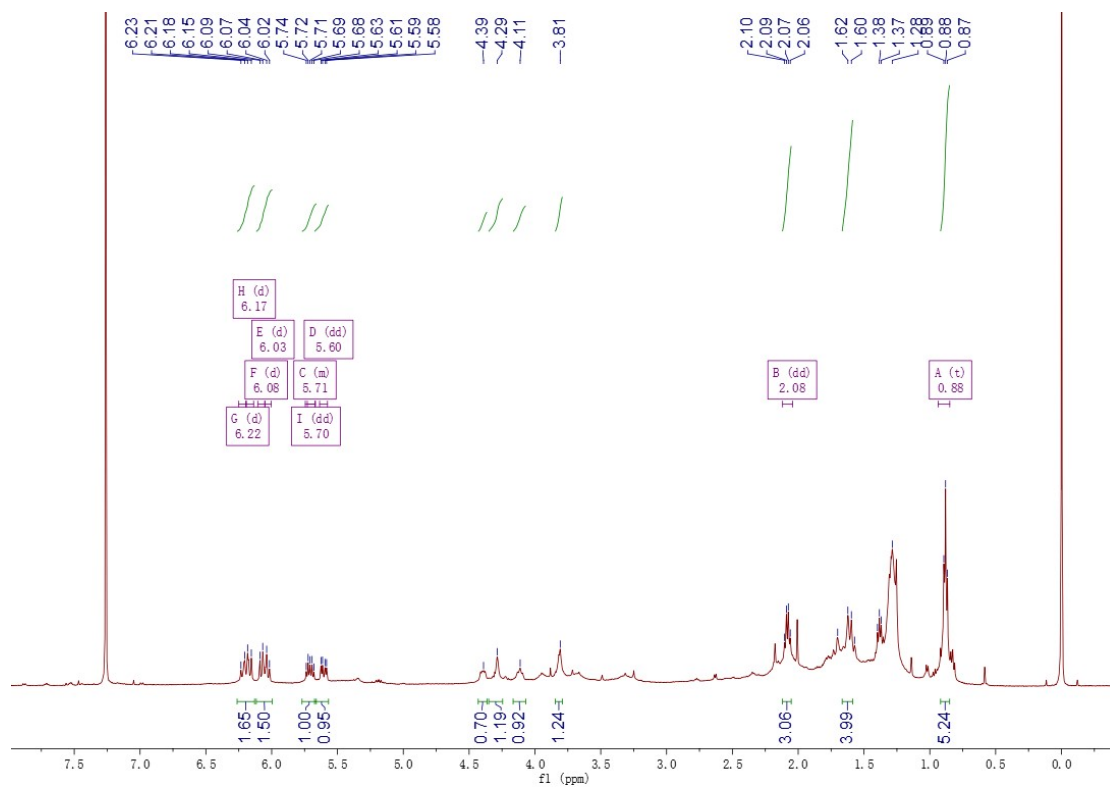


Figure S39. ^1H NMR spectrum of **8** in CDCl_3 (500MHz)

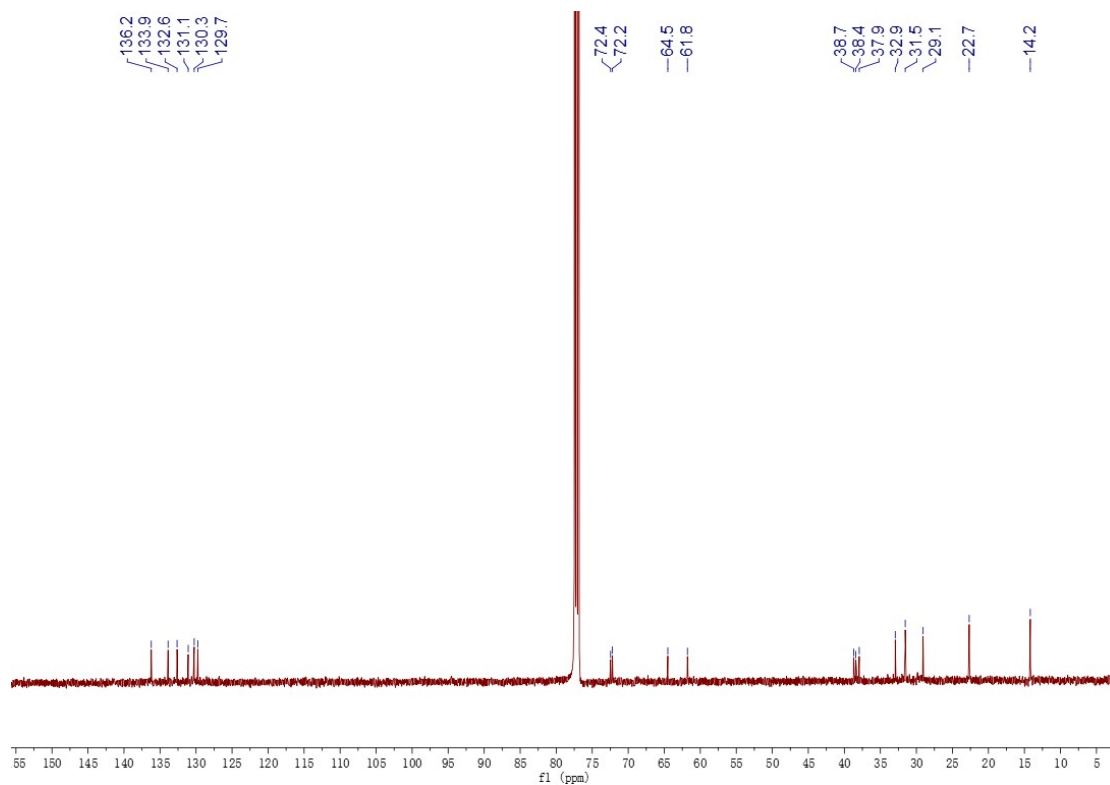


Figure S40. ^{13}C NMR spectrum of **8** in CDCl_3 (125MHz)

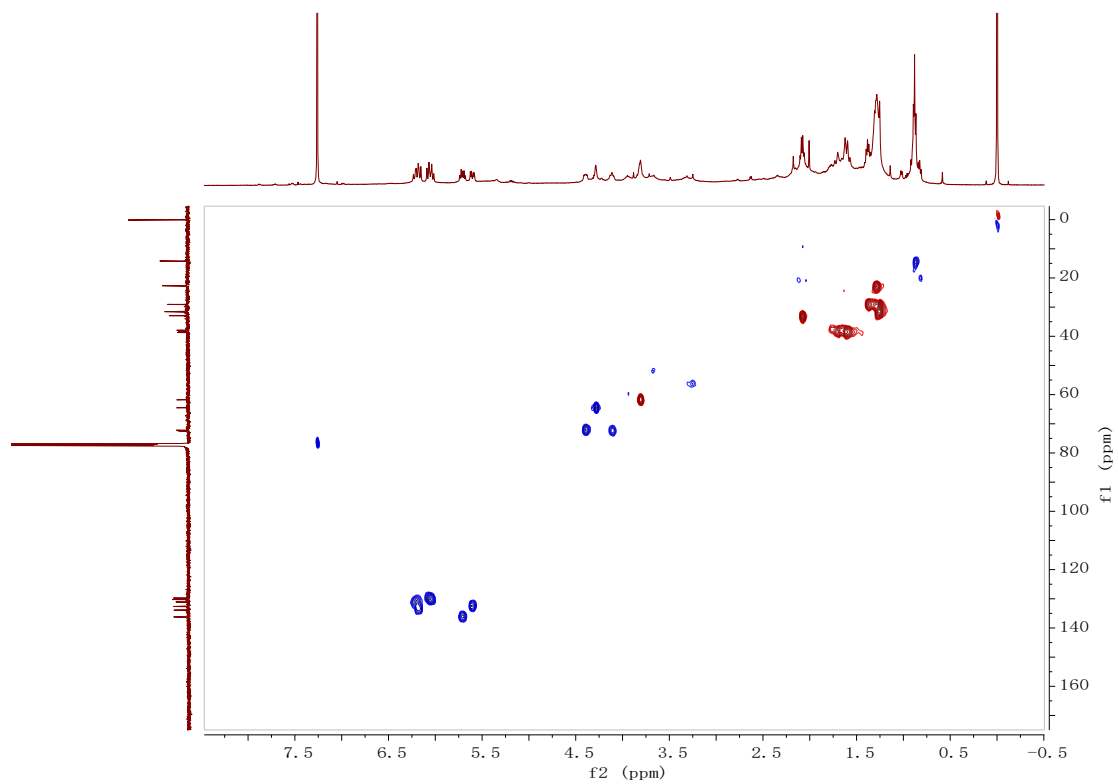


Figure S41. HSQC spectrum of **8** in CDCl_3

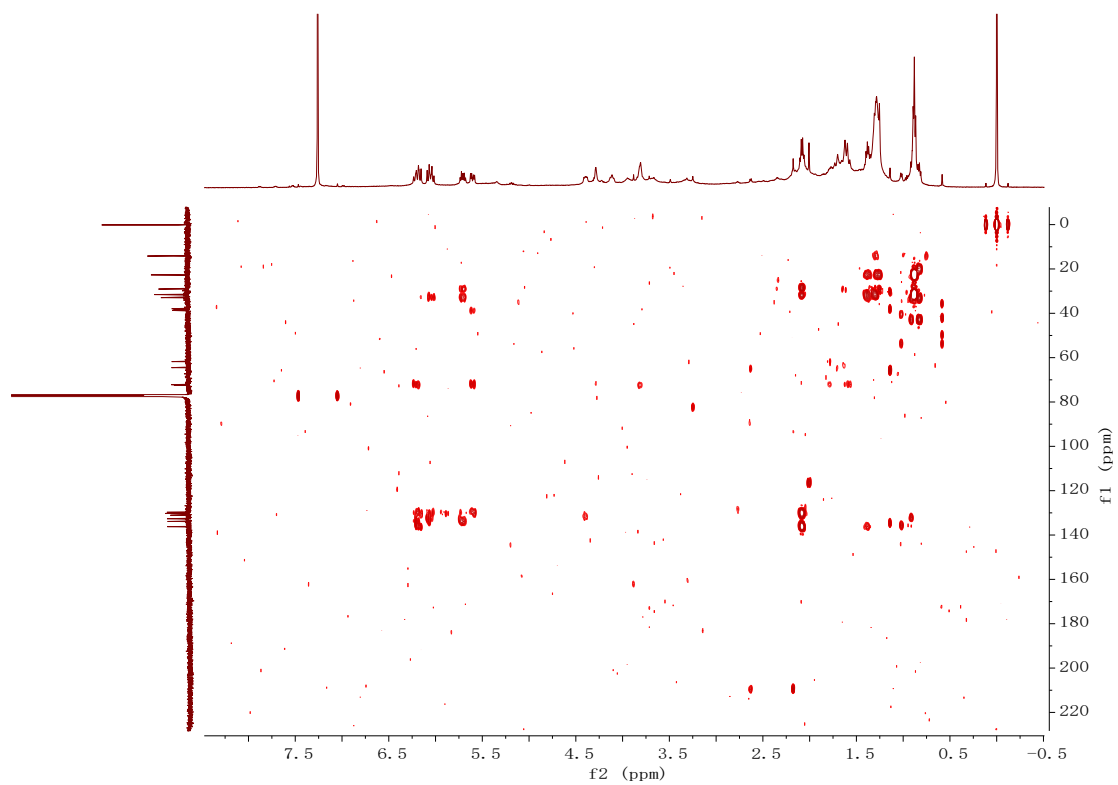


Figure S42. HMBC spectrum of **8** in CDCl_3

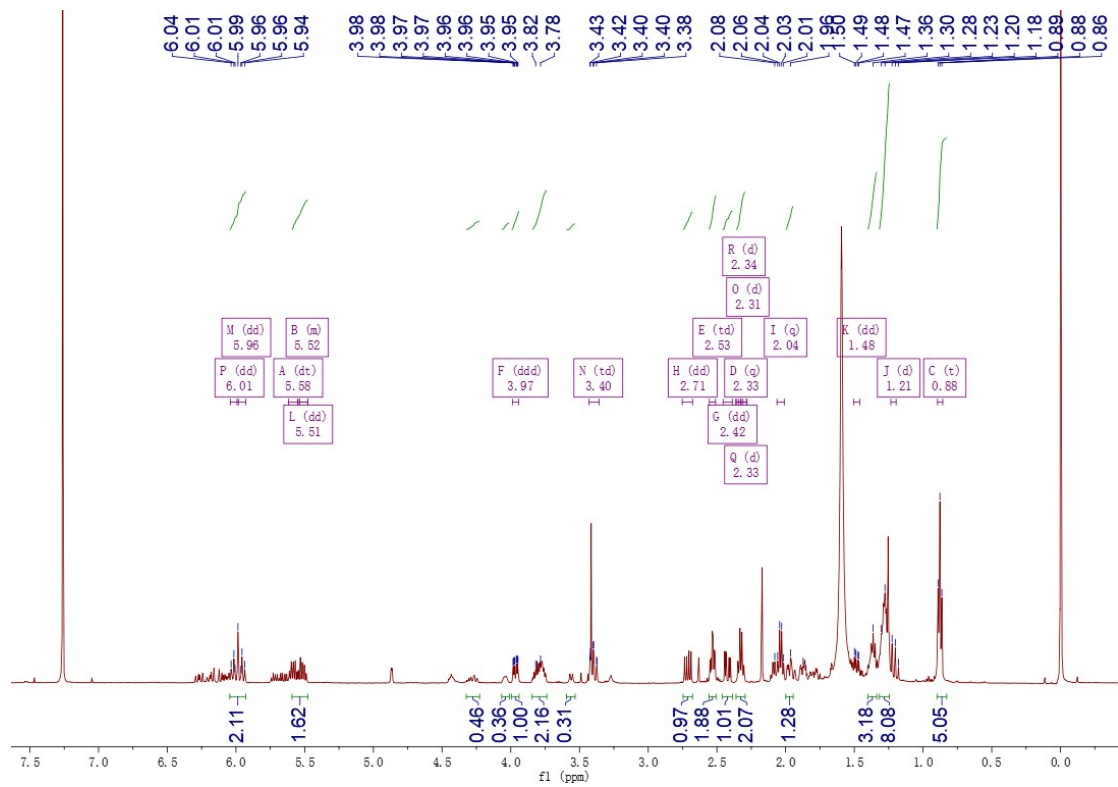


Figure S43. ^1H NMR spectrum of **9** in CDCl_3 (500MHz)

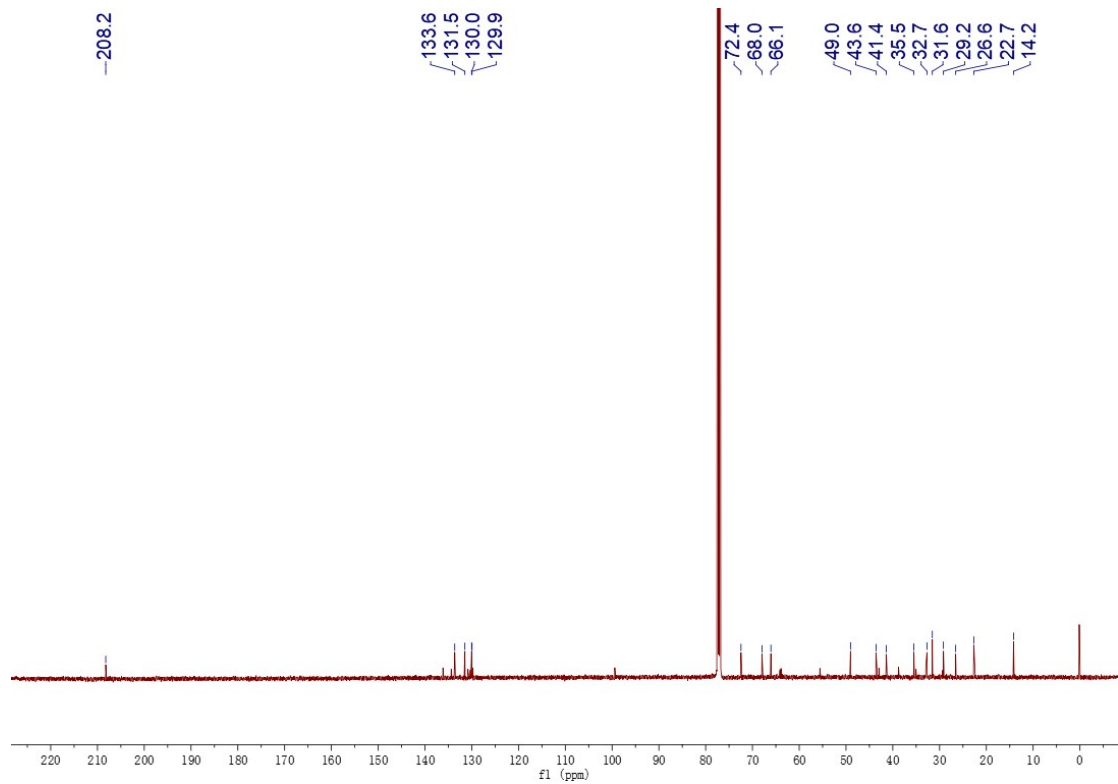


Figure S44. ^{13}C NMR spectrum of **9** in CDCl_3 (125MHz)

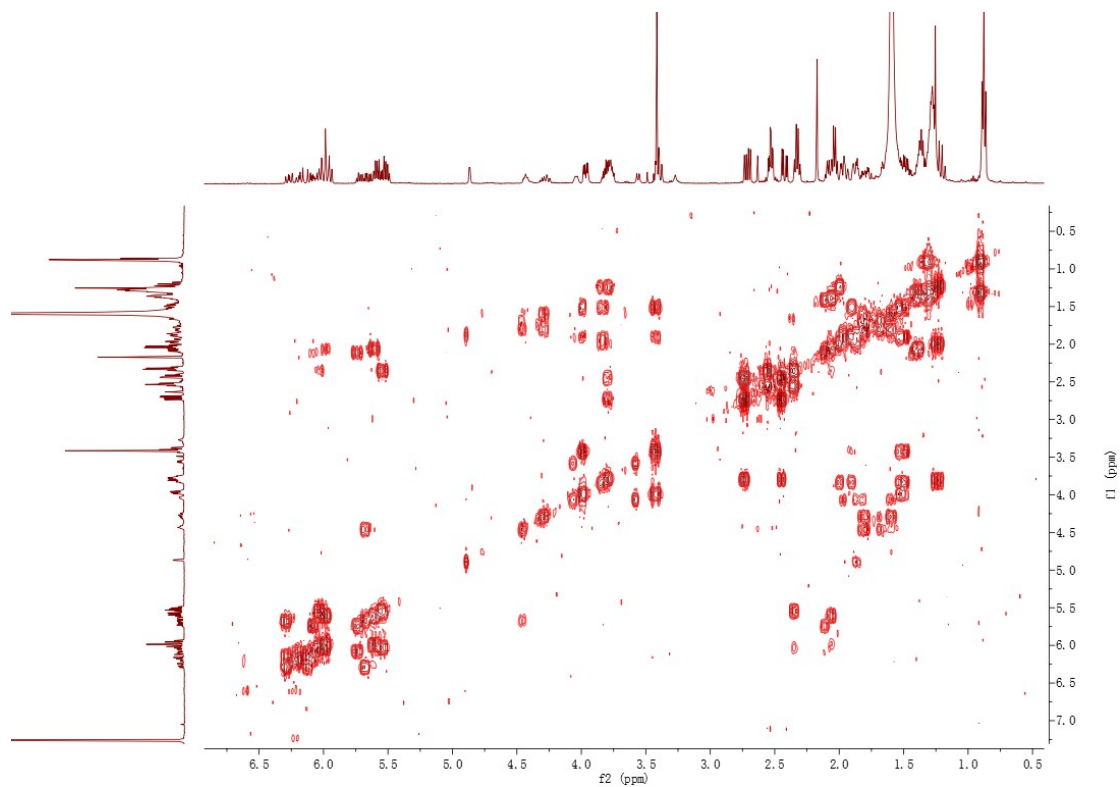


Figure S45. ^1H - ^1H COSY spectrum of **9** in CDCl_3

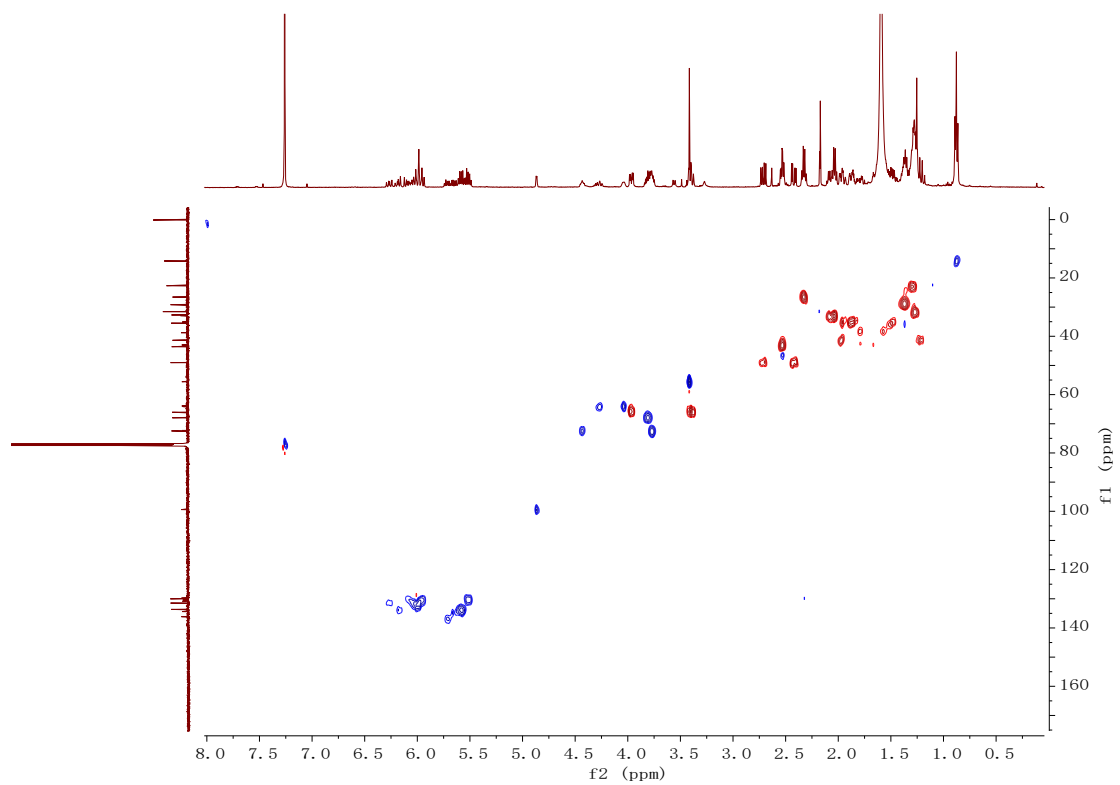


Figure S46. HSQC spectrum of **9** in CDCl_3

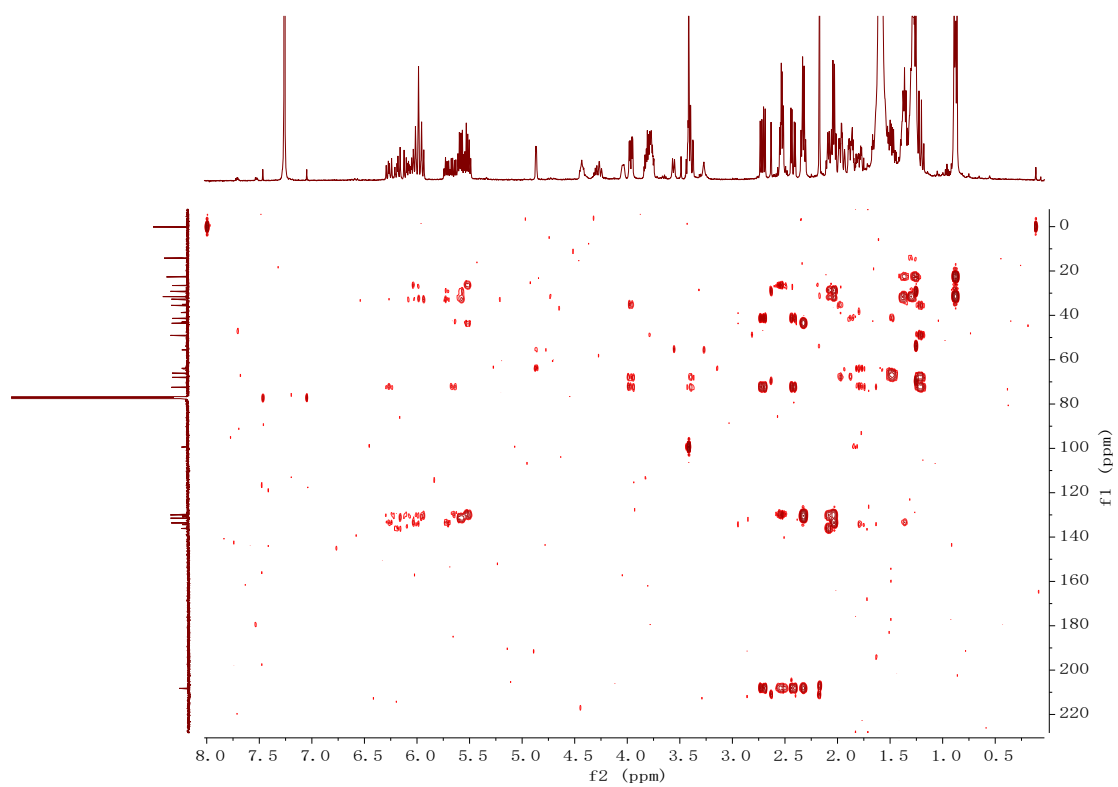


Figure S47. HMBC spectrum of **9** in CDCl_3

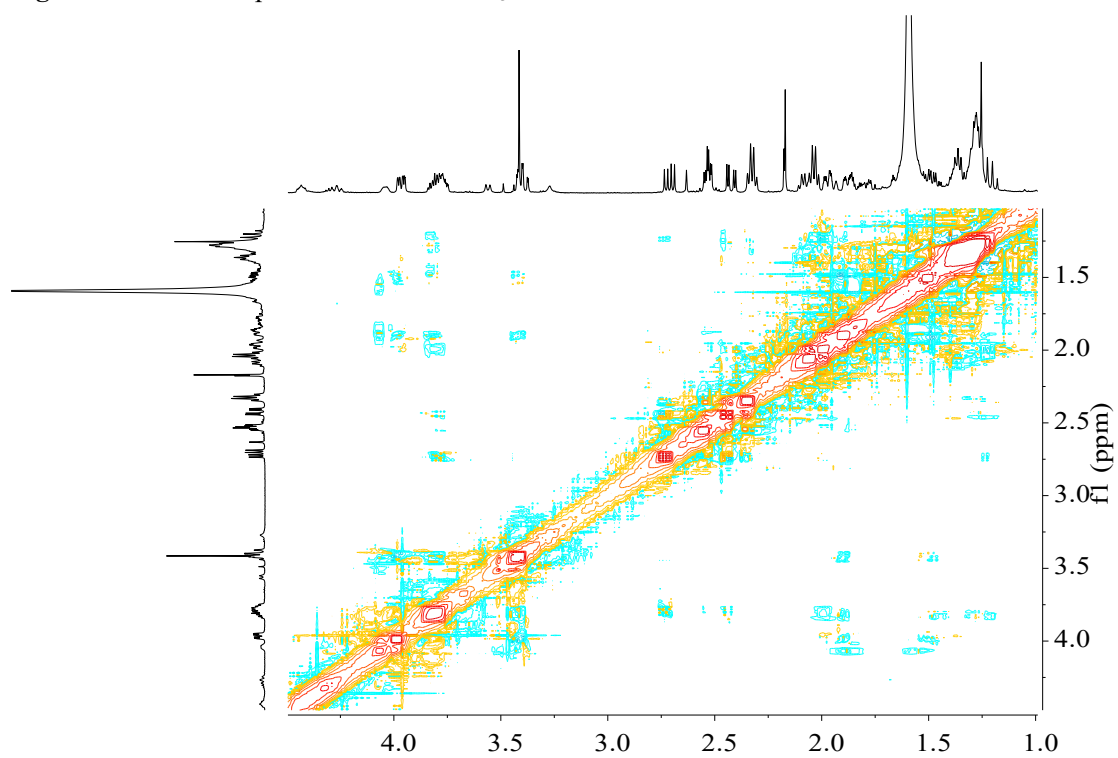


Figure S48. ROESY of **9** in CDCl_3 (500MHz)

3. Supplementary nucleotide sequence data

Genomic sequence: *app1*

Size: 7634 bp

ATGTTCACTCTTGAGCCTGCTTTGGACATCCCTCCGACTCAGAGTCGTTGCTTCGACGCCATGCGAAGGACGCCATCGTTGTCACTGAAGCATCTATTTGAATGATGATT
TGCAGGAAACTTCGGGCCCTTAGTGCCGTGTAACAATTGAAGACAATACCGTGTGCAITGTTGGCATGGGTCACTCTCAAAAAACCCCTTTAACACGACGCTGTTTTC
TAATACATTCGGTGATTTACACAGCTTGCATCTCCCCGGCGGGAATCACTCAACATCGGGGCTTTGGGACTATGTTTACAACAAAAAACCCGCGCAAGGCACTGTGCCG
TGGATCGGTATAACATCGAGGGCTTACCCACAAGGATGGTAGCAGGGGCTGGCGTCAAGCGTCAACGGCGGATACTTCTTCAAGAAGATGTCGAAAAGTTGATCCC
TCGTTCTTTGGTATCAATAACCTTGAAGCGTCTTACATGGATCTCAGCAGCGAAAGCTGCTTGAAGTTGTCACGAATGCTTTGAAAATGCTGGTCTATCTATGGAAGAT
GTGTCGGCACTAACACGGCTGTGATGTCGGCAACTCACTGTCGACTCTGTCTATGCAGTCTCTGTGATACAGACTACGTTACAGAGACTGCTGCTACTGGTGGTGGC
ACTTCTATCATGTCCAACCGTATACGCCATGCTTCAATTTGATGGCCCTAGGTTAGTTACTTCTGCTCTCGGCTTCTCCGGAAGAGGGGCTATTGCGGAACTTCTGAA
CGCTCGTGCAGAGGCTTTAACCGTATCTACATCAGCGATTGTGCTGCTGCAATGAAGAGTGACCCGAAGATCCATGCAGTCACTCGTGGAAAGGGCCATCAACCGGTAAG
TCGAGTATCTCAGCTCAACGGAAATAAATGGCTAATAAGACAAAGCAATGGCAAACTCTGGCATTACCTTCCCGATGCAAGATGCAAGGCTGCTGTAATCAGAAA
AGCATACCAGAATGGCCGCTTAACTTTGGCGATACCAGTATGTTGAGTGGCAGGCACTGTTACACCAAGTGGTGATCCGATAGAGGGTGGATGCTATGCGGGCTGTT
TGACGGCTGAGAGGAGCCCTTGAGAAATGGATCAGTATGTGTCATCCATGATGTTGCTCCATGTCGCATATACGTTATGATGCAAGCCATCAACGGTGAAGAACGCTG
AACATGGGTCAAGTGAAGCCGCAAGCGGTTTGGATCCATCAAAAGTTGGCTGCTTCTTTGAGCAGGGATAATCCCAACCCACTTATGGCTGGAAGAACGCTGAACCC
CAAGCGTAAGACCCATCTTCAATCACCACAATGACATCGAGAGGCAATTCGCTGACCAATTTATAATTATAGTGAACCTCAAGGAGCGAAATGAAAGTATTGACCGA
CGTCGAGTCTGGCCAGAGCCCTCGCCGCGCTGGTGTAAATTCCTTTGGTTACGGCGGTGCCAACGGACAGCTTACTCTCGAGTCTATCGAATAGCTACTCTGTTGGCT
TGGCTTACTCTCCCACTCAGAGGACTTGTCAAGCCTATGATTTTCGAAAGGATGATTTGAGTTCCTTTGCTTCTTCTGCTCTTCAAGCAAGCTTCAAGGTTGCTG
GTAAGCAAGCTATTTCCAGCATGGAAGAAGTGAAGGCAAGTCTCTCAAGGCCCTTGTGCTGCTATGAGCAAGCGACAGGCAAGATGCGCTCGCCGACTACGTACTC
CGTGCAGAGAATGGCCGCTTAACTTGGATACTGGATGTCGATGTCGATGTCGGCGCAACTGATGGCTGCTGGGCTCACTGATGCTGCAATGCTTGAAGCAAGCTTGGC
GGGCGACAATACGCAAAATGGAAGAGGCTATCGAGCAAGATACCGGCTTTTGGCCAGCACTTGTGATCTCGATGAGGTTTGTACTCTTACCCGCGAAATGCA
AGCCCTCTGGGCTTTGGAGCAACTTCTGACAAAGCTGCCACGAGTGAAGTAAATACGAGCTCACTCGTAGCCAGCCGATCTGACCCGAGTTCAAGTGGCTTCCGTA
ACATGCTGGCAAGCTGGGTGTGACGCCGAAGGCTGTTATCGGTCATTTCTTGGAGAGATTGCTGACGATACGCCCGCCGCTCTTACTGCTCTGAAAGTATTCTGG
CAGCTATTTCGCGGGCTTTCAGTTGGTCAACTTCGGCTTCGCGGCAACTGATGGCTGCTGGGCTCACTGATGCTGCAATGCTTGAAGCAAGCTTGGCCTGG
TGAAGTGGGAGTCCGATGTGTCACCGCCCTGAGAGGCTTACTATTAGTGGAGTGTCAAAGATATTGATGCTTCTCAGTCTGAACTCCAGAAAGAGGCAAAATTTGCT
CGCAAGCTTGAAGTGGTGGCCGTGCCCTACCCTCCTACATGATGGTTCGAAATCGGCGACCTGTACGAGTCACTTGTGACGCTTATTTCGCGCAAAAGTTGCTGATGA
ATTGAAGCAAGATGTTTCCAGGTCGGTCACTGCTAATGCCATTTGGCGTGTGATCTCCACCAATATGGCATCTTACTTCCGGAAGAACTTGAAGCAACTGTT
CAATTCAGTGGCGCTTGAACAGCACTTCTCAGGCGACAAGTATCAATTTGATGAAATTTGGTCTTACTTGTCTCAAAGGCTTCAAGGCTTCCGCTTGAAGTGA
AAGCGGCAAGGACCGCTCCCTACTCTCCACTTGTGCGAAAGGAGAACCGGATATCTGCTTGAAGAAGCTTGGCGGCTCTTCCCTATGAGCCACGACCTC
GACTGCTGAGGCTTGAACAGCTCCGCGATACCGTGGCTCCCTAATCTCTCTTCTTATCTTGGGACTACTCAAGCTCTGCGGCTTCCGCTGCTGCTGAG
GTGTTGAGCAGCTTCCGAAAGCTCCCGCTCAGAGCTCTTGAACACAAGTAAACAGCAGGATTAACGGTATCGAGTGGTGTGCGGCAAACTCCCCCATGCTGTGAG
ATGCTTGGCTGCGTGGCAACAGCTTGGCAGTCAACAAGTGGTTTGGCGCACTGATGGCAGTCAATGGCAGTGGTATCGAAGCTCTCAAGCTCTCAAGCTTCAAGC
TAAGCTGGTGGCGGAGCGCTTCTTTCGATTTGAGAATATCAATAAGCGGAGCAATTTGCTGCTGACGAGAGCGGCTGAGGCTGACAGCCTGAGCTT
ACAGCTCATGCTCTCCAGGATCTTCCGCAAACTCAGCTGCGCAATTTGATGAAATTTCCATCTCTCTTGGGTGCTGGTGGCTTGTGACTGCTGAGTGGT
CATCCGAGTCAAGGATGAAGCCAGTACGGATCTGCCCTTATTCAGGACAAGGCAAGGCTGGGATGAGCGCTGGTATACCAAGGCAAGGAG
GAAGGCTCAACTTTGGCCCACTTCCAGTCTGACAAAGTGTGACACAGGATGGCAACAGAACTTCAACAGATTCTATCGCGCACTCTTCTGGAACCTCCGCTCCGA
CTGGCCACCGCATGTTCTACGGCTGCACTATCACTATCGATGCTGCTTCAAGCAGCATATGGCGGTACTGCTGTAATATCAATACGCTTCGAGCTTTTGA
CCAGTCTTCGTTCCAGCTGCTTATTCAGATTCTCGGGGAGGAGTCCAGTCTGGGCAAGAGGAGGTCGGAATCTACTACCGGATGGAGAAAGCGGCTTTTCTAC
CCGTGCTGTGCTTTACTCTCGAATTTGCCGATGGCACACTGCTGCCATGATGATGCTGCTGTTTCAAGAAAGCTCTGGAGCAAGGCAAGGAGATGGCTCTTGAATTTCCGA
CATCTATCTCAACGCCAGCTTGTCTTGTGTCATGGAAGCCAGATGTTCTGCGCTCCAGCCAGGCTGAAAGGCTATTCAGGAGTATATTTCTCTTTGCTGCT
GAACAGTCTGACGATTTGAATGACAACCGGGCCCTTATTGATTTGCAAGTCTTCTGCTTCTCCGCAAAAGGTTCCCGCTATGAGCGTTTGGAACCTCGGGCAAGAG
CGTCAATGGACCCGAAAGGATGGCAATCCATTTGCGGCAAGGACATGCAATCCCTCGTTCCGAGTTTGAACGATGGCAAGCTTGGTGGAGCAATAAGATGGCAT
TGAGAATGCCAAGGAATCCAATCCCTACGATGCTGTTGATTTCTCATCTGCTGTTTCAAGAAAGCTCTGGAGCAAGGCAAGGCAAGGAGATGGCTCTTGAATTTCCGA
CGACGGATCATTTACTCGGAAAGCAAGGATGCCATCTCCGCTTGAAGGCTACTGGCTTGTGACTTATCTCTCCGCAAGGACTCTATCGAGTGAAGAGCC
CCAAGAGACTGTTTGAAGGCAAGGAAGTGAATTTGTCAAAGCAAACTTCTTCTCTCCATCGATTCTTGGCTGCTGCTGACTGCTCACTGAAGAGCGCTG
GTGTTGCTCAGCTCAGCATCGCTCTATCGCCAAATTTGAGTCTGTCAATTCACAGCAAGTGTGTTGATATCTTCTAGAGATGGAACCGCAATTTTGGCCACAT
CAGCGGGAAGACATGGATCGTCCGTAAGATCACAGACAATGTCGGTATCTTTTGGTCAAGGAGCAACATGCTTTTGCACCCCAACCCTGATCTCACTTGTG
TAGCGGCTCTCTCAGGCTCATGCTAGAGCAGCGGCAATTGAGGATGCTGTTTGGACATTTGGAGTGAATACCAAGCCCACTCACTGCAACTATGCAACAA
CGTCTCGGAGCCTTGAACATTCGGTACGAAACAGACGATAACGAAATTTATCCAGAAAGATGGAATCCTTACGTCAGCCGTTTCGCTGATGTTGAGTTGAACTCCCT
GTCCGACAGCGTATGACTCCGATCACTAAGAGTGTGCTTGAAGAGGTTGGCCAGCCAGCTGCTATTTGGTCAAGTGGAAAGGACGACTATTCACCTTCA
GCAAGTCTGAGCGCAAGGACCGCTCCAGCTGGTTTTGGGATGTTGATCTCCGAGCTGTTGGGCTCAAATGCCAAAGACATCTACGCCCTCAATGGTCTGCGCGAGA
CTGATCGGCTACTACTGCTTGACTTTGGTGGTGTATCTCGGCCGTTGGACTGATGTCAGCAGCTTAAAGTGGAGACCGGCTGGCTGGTTCGTTCTCAACCACT
TGGCACTAGGAAACCGCTACTGTCAGCTGTCCATAAGATGCTGCCAGAGGAGGATTTCACTGCTGCTGCTACGCTTCTGACAGTCTACTGTACTGCTGGTGGCT
TCGTGACCGGCTCATCTTGTGCGCGGCAATCCATCTGATCCACTTGGCGCTGGTGGCTTTGGTCTGGCCCATCACTATGGCCAAGCACATGGGCGCTCAGTATT
GCAACTGTTGGTCTCAATCAAGCGCAATATCTTGTCAAGGAGATGGCGTCCCAACGGAAACATCTTCAACTCTCGAATGCTTTCATGGAGGATATCATGGCT
GCTACCGCGCGGAGTGTCAATGTCAATGTCAACTCTTGTGGAGACTGTGATGCTGCTTGGCTTGCATTGCTCTTTGGAGCATTCGTGGAGTTGGCAAGC
GTGAGTGTGATGACGCTGGCAAGCTTGATATGCGCATCTTCTGAAGAACGCTACATCTACTGCTTTGACTTGTGCGGAGTTTTCACGCGCAAGATAACTACTACCAAG
GTGTTGCTCAAACTACACGGCTGAAGTGAATGAGATGACCGCGCGGATCAACAGGCTCCGCTATTGCAACCTTTGATGTTGGCGGAGATGGCCAAAGCTTACCGAT
ACTTTGGCAACAAGATCGGCTCGGCAAGTGTGATCTCAATGGAGAATCCCAAGAGTCTTGGCAGTCAAGGCAAGTCAAGGCAAGTCAAGGCAAGTCAAGGCAAGT
CTTATTGCTGTTGGTGTCTCGGTTGCTCGGCCGAGTCTCAGTCTGGATGATGTTCTCGCGGTGCCCGCAAAATCTGCTTCTTGGAGCTTGGATGCGATAAGCC
CAGGCGAGCTGAGCTTGTCAACCGTCTCCGAGATGCGCGTCCAGCGTCACTGTTGCTGCGGATGTTGTCACAGGAGTCAAGGTTCTGTAAGCTGTCGCGACTTCTGC
CAAGGATGGCCCAATGGCGGTGTGCTCCAGCGGCTATGGGCTCAGTGAAGCTTTGTTCTCTGTCATGACGAACAAGGCTGGCACACTGGCATAAGGCAAGTGA
AGGGCTCTTGAACCTTCAAAATGCGCTAGAGGGCCATGATGCGGCACTGATTTCTTCTCTGACTTCACTGATTTCTGGCAGCTGTTGGCAGCTACCGAGATAACT
ACTGCTGCTGAACCGCTTCCGATTCGTTCTGCTAGATGCGCGGCTCTCAAGGCAAGCTGCGGCTTCTGTTGGCTTGGGCAATGATTTCGGAAGTGGTATCTCATGA
AAACCTGAGATCGAGGCTATGCTTCCGCAAGGATTCAGCGCTCAACAGGATGAGTTCCTCAAGTCTTAGACTACGGTATTGGCGCCAGGCAAGTGTGCGG
AATTTGCTCGCGGTGCTCCATGAACAGAGAAGCCGCTCATATCTTGAACGGTCTGAGGATCTTACGGTGTCCGCAAGCTGATGGCTCAGGCTTCAAGGTAACAACCG
GTCATGATGAGTCCCGACGTCATTTCTGCTGCTCTGCTCTGAGAAAGGATGCAAGGAGGAAAGAAAGGTTGGCAGTGTGGACAGCTCTCGCCGCTGCGCA
ATGGGTCAAGGAGCTTCCCAAAAGTGTCTCCCTATGTTGATGCCCCAGGCGACGCCACCCAGTATGTTGATGCTTACGCTTACGCTTACGAAAGGAGGCTTCAAGCAAC
TGATCTGATGACAGTATGACAGTGTGACGAGCTGCCCCCTCCCATCTTGGTGTGACAGCATGCTTGTGCTGAGAGTTCGCAACTTGGTCTTCAACAGCTCAAGAT
TGATGTACCTTCTCGATATTGTGACCCCAAAAAGTCTGACACAGCTGGCAGAATTTGTTGAGGAGAAGCTTGTGGCAGCTGGGCTAGCTAG

mRNA sequence: *app1*

Size: 7305 bp

ATGTTCACTCTTGAGCCCTGCTTTGGACATCCCTCCGACTCAGAGTCGTTGCTTCGACGACATGCGAAGGACGCCATCGTTGCTACTGAAGCATCTATTTGAATGATGATT
TGCAGGAAACTTCTGGGGCCCTTATGTCGCTGTACAATTTGAAGACAATACCGTGTGCAITTTGGCATGGCTTGCATCTCCCGGGGGAATCACCTCACCACCGGGGCTTT
GGGACTATGTGTACAAACAAAACCGCCGCAAGGCACCTGTGCCCTGGATCAACATCGAGGGCTTCTACCAAAAGGATGGTAGCAGGGCTGGCTGATAGGCGT
CAACGGCGGATCTCTTCAAGAAAGTTCGCAAAAGTTGATCCCTCGTTCTTTGGTATCAATAACCTTGAAGCGTCTACATGGATCTCAGCAGGAAAGCTGCTTGA
AGTTGTCTCGAATGCTTTGAAAATGCTGGTCTATCTATGGAAGATGTGTGGCACTAACACGGCTGTGATGTCGGCAACTTCACTGTGCACTCTGTCAATGCGACT
TCGGTATACAGACTACGTTACAGACTAGTGTCTACTGGTGGTGGCACTTCTATCATGTCCAACCGTATCAGCCATGTCTCAATTTGACGGCCFAGCTTTACCTTTGGAT
ACTGCTGTCTTCCACCATTATGCACTTCCAGGCGTGAACGCCATCAAGAATGGTGTGACTGGCACTTGCATTTGGCGGAGCCAACTTGTACAGCTCCCGGAG
CAGCACTTTGGAACAGCAAAAGGGTGGCTTCTCTTCCAAACCTCGGCTTGCATACATTCGACACGTCAGCAGATGGTTACGCTGTGTCAGAGGCTTAAACGCTATCTAC
ATCAGCGGATTTGCTGTCCATGAAGAGTGAACCGCAAGATCCATGCACTTCTGGAACGCCATCAACGCCAATGGCAAACCTCTGGCATTACCTTGGCCGATGC
CAAGATGCAGGCTGTGTAATCAGAAAAGCATACCAGAATGCCGGCTTCACTTGGCCGATACCGACTATGTTGAGTGGCCACGGCCTGGTACACCAAGTTGGTATCCGA
TAGAGGTGGATGGTATCGCGGCTGTTTTCAGGTCGTGAGGGGAGAGCTCTGAGAATTTGATCAGTGAACCAAGCAATGGGTACAGTGAAGCCGCCAGCGGTTTTCAG
ATCCATCATCAAAGTTGGCTCCCTTTTGGACACGGAGTAATCCCAACCTTATGGCGTGAAGAACGTAACCCCAAGCTGAAACTCAAGGACGAAATATGAAGGTAT
TGCCAGCTCGAGCTGTGGCCAGAGCCCTTCCGCGGCTGGTGTAAATTCCTTGGTACGGCGGTGCCAAGCCAGCTTATCTCGAGTCTATCGATAGCTACTTCCG
TTGGCTCTCTGCTTCACTCCCATCACGAGGACTTTGTCAAGCCCTTATGATTTGAGAAAGGATCAAGTCTTTGTGCTTCTTCTCAAGCAAGTCTGGTGA
GGTTCGCGGTAAGCAAGCTATTTCCGACATGGAGAAGACTGAGGCCAAGTCTCTCAAGGCCCTTGTGCTGTATGAGCAAGCGACAGCCAAAGTACGCTCCGCGACT
ACGTACTCGGCTCAGAGAATGCTGCTCAACAAAATGATTGAGATGCTGTGAACACAAGTAAACAGCAAGTAAACAGCAAGTAAACCGTATCGAGTGGTTCGCGAAACATCCCCGCA
TGCTAGATGGCTTGGCTGGTGACCAAGCTTGGCGAGTCAAGTGGTGTGGCCAGGAGCGGCATACATGGCGATTGCTATCGAAGCTCTCTCAGAGTCTCAGAAC
ATAAAGGCTAAGCTCGTTCGGGAGAGCGCTTCTCTTTCGAGTTTGAAGATATCAATATAAGCGCAGCATTTGTCGTCCCTGACGAGAGCGGAGCTGACGAGCAC
TGAGTTCACACGGTTCATGCTCTCGCAAGATCTACCGCAATACATTTGGCAATTTGGCATGAATTTCCATCTCTTGGGTGTGGTGTGCTACTTTGCACTG
TTTCTACCCGTGCTGTGCTTTTACTTTCGATTGCCGATGGCACACCTGCCATTGACATGCCGACCTACCGCATGAAACCGGTATACCTGGTAAAGCTCCCGTGCAGCCTG
AGACAAGCATCTATCTCAACGCCAGCCATGCTTCTGTTCCAATGGAAGCCAGATGTTTGGCGCTCCAGCCAGGCGTGAAGCGCTATTCAGGAGTATATTTCTCTCT
TTGCTGTGAACAGTCTGACGATTGGAATGACAACGGGGCCCTTATTGATTTTGGCAGCTTCTCGATCTCTCGGCCACAAGGTTCCGCTATGAGCGTTTGGAACTCGG
GCAAGAGCGTCAATGGACCCGGAAGGATGCAATCCATCTTGGGCAAGGACACTGCTATCCCTCGTTCCGAGTTTGAACAGTGGCAAGCTTGGTGGAGACAATAAGA
TTGCCATTGAGAAATGCCAAGGAATCAATCCCTACGATGCTGTTCTGATTCTCATCTGTTGTTTGAAGAACTTGGAGCAAGGCAGCAGGAGATTGCCCTTAG
TTCCGACGACGGGATCAATATCACTCGGAAAGCAAGGATGCCATCTCCGCTGTAAGGCTACTGGCTTTGACTCTATCTCTCCCAAGCAGACTCTCATCGCAGTGA
AGAGCCCAAGAAAGACTGGTTTGAAGGCAAGGAAGTAGTATTGTAAGCCAACTCTTCTCTCTCCATCGATTCTTGGCTGCTGCTGTGACTGCTCACTGAAAGA
GGCTGTGTTGCTCAGTCAAGTCCGCTATCGCCAAATATTGAGTCTGTCAATCTCAACGACCAAGTGTGTCATCTCTTCAAGAGTGAACCGCAATTTCTGTC
CACCATAGCGGCGAAGCATGGATCCGCTTCCGTAAGATCACAGCAATGTCGGTGTCTTTTGGCTCAGCGGAGCAAACTGCTTTTGCACCAACCTGATCTCA
CTTGTCTAGCGGCTCTCTCGAGCTCTCATGCTAGAGCAGCCGGCATGGAGTATGCTGTTTGGACATTTGAAGTGTATTAACCAAGCCCACTACACTGCAACTATCT
GCAACAACGCTCGGAGCTTTGACTTCCGTTACGAAACAGACGATAACGAATTTACCAAGAAAGATGGAATCTTTACGTCAGCGCTTTCCGCTCGATGTTGAGTGA
ACTCTTGTCCGACAGCGTATGACTCCGGATACTATGAAGTGTGCTCTTGAAGAGGTTGGCCAGCCAAAGCTGTCTATTGTCAGGTTGGAATGACCGACTATTC
ACTTTCAGCAAGTCTCGAGCGCAAGACACCGCTCCAGCTGGTTTGTGGATGTTGATCTCCGAGCTGTGGGCTCAATGCCAAAGACATCTACCGCCTCAATGGTCTGTG
CCGAGACTGATCGGCTACTACTGCTTGTACTTGGTGGTGTATCTCGCGGCTTGGACCTGATGTCACGACCTTAAGGTAGGAGACCGCGGGCTGGTGTTCGTTCTTA
ACCAGTCTGGCACTACGGAACCGGCTACTGTGCAAGCTGCCAATGAATGCTGCCAGAGGAGGATTTCACTGTCTGCCTACGCTTCTGACAGTCTACTGTACTGTCTGG
TTGCCCTTGTGACCGGCTCATCTTCTGTCGGGCAATCCATCTGATCACTTGGCGTGGTCCCTTGGTCTTGGCCCATCACTATGGCCAAAGACATGGGCGCTAC
AGTATTGCAACTGTTGGTTCTCAATCTAAGCGCGATTATCTGTCAAGGAGATGGGCGTCCCAACGGAAAACATCTTCAAACCTCGCAATGCTCTTTTCAAGGAGATAT
CATGGCTGTACCAGCGCGCGAGGTTGCAATGTCATTGCAACTCTCTTGTGGAGACTGATGTCATGCCTTGGTCTTGCATTTGTCCTTTTGGACGATTCGTGAGAT
GGCAAGCGTGAAGTGTGATGACGCTGGCAAGCTTGAATGCGCATCTCTGAAAGAACGCTACATCTCACTGCCTTGTACTGTGCGGAGTTTTTCTACCGCAAGATAACTAC
TACCAGGTTGGTGTACAACACTACCGGCTGAAGTATTGAGATGACCGCGCGGATATCAAGGCTTCGCTATTGCAACCTTTGATGTTGGGAGATTGGCCAAAGCT
TACCGATACTTTGGCAACAAAGATCGCGTCCGCAAAAGTTGTCATCTCAATGGAAGATCCCAAGAGTCTTGTCCAGGCTATGCCAGCATCTTACCAATCCGTTTCCATCT
GAGAAGACTATTTCGCTGTTGTTGCTCGGTTGTTCTCGGTTGTTCTCGGCGCAGTCTCAGTGTGATGATGTCGCGGTGCCGCAAAATTCGTTCTTGGAGCTTCTGGATGCG
ATAAGCCAGCGCAGTGTGAGTGTCAACCGTCTCCGAGATGCCGTTGCCAGGCTCACTGTTGTTGTTGCGGATGTCCTCAACGAGGATCAGGTTGTTGAAAGTGTGCGCA
GCTTCTGCAAGGATGGCCCAATGGCGGTGTCGTCAGGCGGCTATGGGCTCAGTGAAGCTTTGTTCTGTCTGATGACGAACAAGGCTGGCACACTGGCATAACAGCC
CAAGTGAAGGGCTCTTGAACCTTCAATGCCCTAGAGGGCCATGATGCCAGCTTGAATTTCTCTCTGACTTCACTGATTTTGGCAGCTGTGGCACCGCTACCGA
GAGTAACTACTGCTGCTCAACGGCTTCTCGATTGCTTCCGATGATGGCGCCCTCAAGGCAAGCTGCCGTTTCTGTTGGCTTGGCATGATTTGGAAGTTGGTAT
CTCATCAAAACCTGAGATCGAGGCCATGCTGCTCCGCAAGGATATTACGCGCTCAACAGGATGAGTCTTCAAGTTCTAGACTACGGTATGGCGGCCAGGCGAG
TGATGGCAATTTGCTCGGCTGTTCTCAATGAACAGAGAAGCGCTCATCTTTGACGGGCTGGAGTCTTACGGTGTCCGCAAGCTGATGGCTTCCGAGCTTCCAGGTTGA
ACAACGGTGTATGATGAGTCCGCAAGTCAATTTCTGCTGCTCTGCTCTGAGAAGGATGCCAAGGAGGAGAGAAGGTCGGATGTTGGAACAGCTCTCCGCC
GCTCGAATGGGTCAAGGACTTCCCAAGTGTCTCCCTATGTTGATGCCAGGCCAGCCACCAAGTCTGAGTGGATGCCATCTTACGCTTCAAGAAAGCGGTT
CAGCAACCTGATCTGATGACGCTAGATGACGTTGACGAGCGTCCCTCTCCATCTTTGGTGTGACAGCATGCTGTGTCAGAGTCCGAAGTGGTCTTCAACAC
GTTCAAGATTGATGATCTTCTCGATATTGTCAGCCCAAAAAGTCTCTGACACCGCTGGCAGAAITTTGTTGAGGAGAAGCTTGTTCGACGCTGGGCTAGC

4. References

- (1) Chiang, Y. M.; Ahuja, M.; Oakley, C. E.; Entwistle, R.; Asokan, A.; Zutz, C.; Wang, C. C.; Oakley, B. R. *Angew. Chem. Int. Edit.* **2016**, *55*, 1662-1665.
- (2) Yin, W. B.; Chooi, Y. H.; Smith, A. R.; Cacho, R. A.; Hu, Y.; White, T. C.; Tang, Y. *ACS Synth. Biol.* **2013**, *2*, 629-634.
- (3) Yu, D.; Xu, F.; Zi, J.; Wang, S.; Gage, D.; Zeng, J.; Zhan, J. *Metab. Eng.* **2013**, *18*, 60-68.
- (4) Aleksenko, A.; Clutterbuck, A. J. *Fungal Genet. Biol.* **1997**, *21*, 373-387.