Reconstruction of cladoniamide biosynthesis reveals non-enzymatic routes to bisindole diversity

Yi-Ling Du,[†] David E. Williams,[†] Brian O. Patrick,[†] Raymond J. Andersen,[†,‡] and Katherine S. Ryan[†,*]

[†]Department of Chemistry and [‡]Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, Canada.

SUPPLEMENTARY INFORMATION

CONTENTS:

Table S1. Bacterial strains and vectors used in this study.

Strains or vectors	**Description**	**Source**
Strains		
S. uncialis L72	Wild-type cladoniamides producer	[1]
S. uncialis YD8	*S. uncialis* L72 (*ΔclaY*::*aac(3)IV*)	This study
S. coelicolor M1146	Heterologous host for cladoniamides production	[2]
S. coelicolor YD52	*S. coelicolor* M1146 + *cla* cluster (*ΔclaM1*)	This study
E.coli DH5*α*	General cloning host	Laboratory stock
E.coli BW25113/pIJ790	Host used for PCR-targeting system	[3]
E.coli ET12567/ pUZ8002	Methylation-deficient strain	[4]
E.coli BL21 (DE3)	Host for protein expression and pathway reconstruction	Laboratory stock
E.coli cX1	*E.coli* BL21 (DE3) / pET28a-claX1	This study
E.coli cX2	*E.coli* BL21 (DE3) / pET28a-claX2	This study
E.coli aX1	*E.coli* BL21 (DE3) / pET28a-abeX1	This study
E.coli X1	*E.coli* BL21 (DE3) / pACYC-abeX1	This study
E.coli X1M1	*E.coli* BL21 (DE3) / pACYC-abeX1claM1	This study
E.coli X1M1cX2	*E.coli* BL21 (DE3) / pACYC-abeX1claM1+ pCOLA-claX2	This study
E.coli X1M1aX2	*E.coli* BL21 (DE3) / pACYC-abeX1claM1+ pET22b-abeX2	This study
E.coli X1M1X2M3	*E.coli* BL21 (DE3) / pACYC-abeX1claM1+ pCOLA-claX2 + pET22b-claM3	This study
E.coli Ctrl1	*E.coli* BL21 (DE3) /pET28a	This study
E.coli Ctrl2	*E.coli* BL21 (DE3) /pACYCDuet-1	This study
E.coli Ctrl3	*E.coli* BL21 (DE3) /pACYCDuet-1+pCOLADuet-1	This study
Vectors		
pYD1	Integrated vector harboring the *cla* biosynthetic gene cluster	[5]
pYD130	pYD1 (*ΔclaM1::scar*)	This study
pYD132	pYD1 (*ΔclaM1::aac(3)IV*)	This study
pYD134	pYD1 (*ΔclaM1::scar, Δneo::aac(3)IV*)	This study
pHY773	FRT-*aac(3)IV*-FRT	[6]
pET22b, pET24b, pET28a pACYCDuet-1 pCOLADuet-1	Vectors for protein expression and pathway reconstruction in *E. coli*	Novagen
pET24b-claM1	Vector for ClaM1 expression, cloning site NheI/XhoI	This study
pET28a-claX1	Vector for ClaX1 expression, cloning site NdeI/XhoI (*claX1* coding region contains a NdeI site, two DNA fragments were	This study

	thus cloned one after another)	
pET28a-claX2	Vector for ClaX2 expression, cloning site NdeI/XhoI	This study
pET22b-abeX1	Vector for AbeX1 expression, cloning site NdeI/XhoI	This study
pET22b-abeX2	Vector for AbeX2 expression, cloning site NdeI/XhoI	This study
pET28a-abeX1	Vector for AbeX1 expression, cloning site NdeI/XhoI	This study
pET28a-abeX2	Vector for AbeX2 expression, cloning site NdeI/XhoI	This study
pACYC-abeX1claM1	Vector for AbeX1 and ClaM1 coexpression, cloning site NdeI/XhoI for *abeX1* and BamHI/HindIII for *claM1*	This study
pCOLA-claX2	Vector for ClaX2 expression, cloning site BamHI/XhoI	This study
pET22b-claM3	Vector for ClaM3 expression, cloning site NdeI/XhoI	This study
pYD36	pYD1 (*ΔclaY*::*aac(3)IV*)	This study
pYD17	A 8kb fragment (KpnI-digested) containing the claY knock out cassette (from pYD36) was inserted into the KpnI site of pMRD400	This study

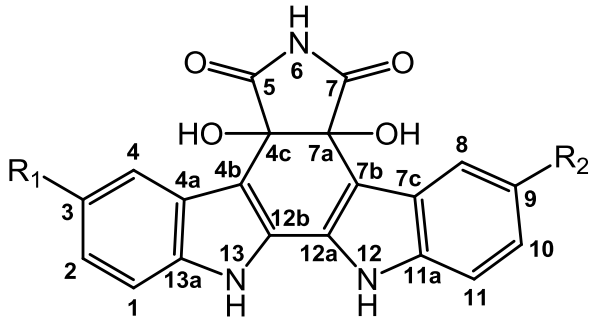
Table S2. Oligonucleotide primers used in this study.

Primers	Sequence (5'→ 3')	Description
claM1-KOF	caccacagccgtacttcgacgacctcgcggacctgtacgattccggggatccgtcgacc	Primers for *claM1* inactivation in pYD1
claM1-KOR	atggaagacggcgttgacggacatcaccaggtcgaaggttgtaggctggagctgcttc	
claX1-NdeI-F	agcagccatatgcgtgacgcacacgattgtg	Primers for cloning of *claX1*
claX1-XhoI-R	agcagcctcgagccctgcggctgcgctgacctc	
claX2-BamHI-F	agcagcggatccgatgagcggcccggctcccacc	Primers for cloning of *claX2*
claX2-NdeI-F	agcagccatatgagcggcccggctcccacc	
claX2-Xho-R	agcagcctcgagtgtcgggtccttctcgtggtc	
abeX2-NdeI-F	agcagccatatgatgtgtattgcaggtggtggtc	Primers for cloning of *abeX2*
abeX2-XhoI-R	agcagcctcgaggcgtacagtaccttccggcg	
abeX2-XhoI-R2	agcagcctcgagtcagcgtacagtaccttccggcg	
abeX1-NdeI-F	agcagccatatgatgcgcgatgaacgcgattgtg	Primers for cloning of *abeX1*
abeX1-XhoI-R	agcagcctcgagggcaggcgctttcacacgcgtc	
abeX1-XhoI-R2	agcagcctcgagtcaggcaggcgctttcacacgcgtc	
claM3-NdeI-F	agcagccatatgagcgcagccgaccccg	Primers for cloning of *claM3*
claM3-XhoI-R	agcagcctcgagtcagctcccggccgttacc	
claM1-BamHI-F	agcagcggatccgatgaccgaccaggcccccgtc	Primers for cloning of *claM1*
claM1-HindIII-R	agcagcaagcttgcacgtcgagcacgctcacggg	
claM1-NheI-F	gcatcattagctagcatgaccgaccaggcccc	
claM1-XhoI-R	aggacattcctcgagcggggcgtcccagac	
claY-KOF	cggagtgttgtgatgtccgatgcgaccctgccggcgaacattccggggatccgtcgacc	Primers for *claY* inactivation
claY-KOR	acggaaggtcgcccaggaacgtcacccaggggcgtcgtctgtaggctggagctgcttc	
claY-NdeI-F	agcagccatatgtccgatgcgaccctgcc	Primers for the comfirmatino of *claY* inactivation mutant
claY-XbaI-R	agcagctctagagaaggtcgcccaggaacgtcac	

Table S3. Mass spectrometry data for compounds generated in this study.

Compound(s)	**Molecular Weight (observed ions through LC-MS or HRESIMS)**
10	393 ($[M+H]^+$=394, $[M+TFA-H]^-$ = 506)
11	427 ($[M+H]^+$=428, $[M+TFA-H]^-$ = 540)
12	359 ($[M+H]^+$=360, $[M+TFA-H]^-$ = 472)
13/14	375 ($[M+H]^+$=376, $[M+TFA-H]^-$ = 488)
15	341 ($[M+H]^+$=342, $[M+TFA-H]^-$ = 454)
16	363 ($[M+H]^+$=364)
17	397 ($[M+H]^+$=398)
18	423 ($[M+Na]^+$=446, $[M-H]^-$ = 422)
19	395 ($[M+Na]^+$=418, $[M-H]^-$ = 394)
20	395 ($[M+Na]^+$=418, $[M-H]^-$ = 394)
22	373 ($[M+H]^+$=374, $[M+TFA-H]^-$ = 486)
23	355 ($[M+H]^+$=356, $[M+TFA-H]^-$ = 468)
3,9-dichloroarcyriaflavin A	393 (HRESIMS of $[M-H]^-$= 391.9980)

Table S4. ^{1}H NMR data (for compound **10**, **11**, **12**) in DMSO-d_6 at 600 MHz (δ is given in ppm, the DMSO-*d6* signals (2.50 ppm for ^{1}H and 39.5 ppm for ^{13}C) were used as references).



	^{1}H NMR of **12**	^{1}H NMR of **10**	^{1}H NMR of **11**
1	7.47 (d, *J*=7.8 Hz, 1H)	7.55 (d, *J*=8.7 Hz, 1H)	7.56 (d, *J*=8.7 Hz, 1H)
2	7.13 (t, *J*=7.8 Hz, 1H)	7.13 (dd, *J*=8.7, 2.1 Hz, 1H)	7.15 (dd, *J*=8.7, 2.1 Hz, 1H)
3	7.06 (t, *J*=7.8 Hz, 1H)		
4	7.94 (d, *J*=7.8 Hz, 1H)	7.95 (d, *J*=8.7 Hz, 1H)	7.96 (d, *J*=2.1 Hz, 1H)
4a			
4b			
4c			
5			
6	11.62 s	11.64 s	11.72 s
7			
7a			
7b			
7c			
8	7.94 (d, *J*=7.8 Hz, 1H)	8.00 (d, *J*=7.8 Hz, 1H)	7.96 (d, *J*=2.1 Hz, 1H)
9	7.06 (t, *J*=7.8 Hz, 1H)	7.08 (t, *J*=7.8 Hz, 1H)	
10	7.13 (t, *J*=7.8 Hz, 1H)	7.14 (t, *J*=7.8 Hz, 1H)	7.15 (dd, *J*=8.7, 2.1 Hz, 1H)
11	7.47 (d, *J*=7.8 Hz, 1H)	7.52 (d, *J*=7.8 Hz, 1H)	7.56 (d, *J*=8.7 Hz, 1H)
11a			
12	11.19 (s)	11.22 (s)	11.41 (s)
12a			
12b			
13	11.19 (s)	11.36 (s)	11.41 (s)
13a			
OH (C4c)	6.31 (s)	6.42 (s)	6.50 (s)
OH (C7a)	6.31 (s)	6.36 (s)	6.50 (s)

Table S5. ^{13}C NMR data of compound **10**, **11** and **12** in DMSO-d_6 at 150 MHz.

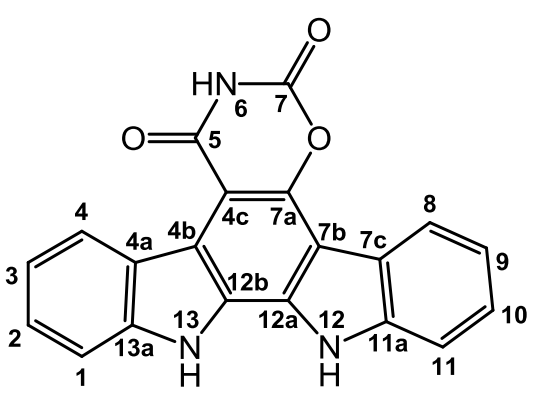
	^{13}C NMR of **12**	^{13}C NMR of **10**	^{13}C NMR of **11**
1	111.99	113.34	113.52
2	122.22	121.51	121.90
3	120.26	124.49	124.60
4	121.36	119.97	120.19
4a	127.15	128.04	127.87
4b	108.80	108.22	108.93
4c	77.95	77.30	77.10
5	177.23	176.84	176.80
7	177.23	176.81	176.80
7a	77.95	77.42	77.10
7b	108.80	109.36	108.93
7c	127.15	126.78	127.87
8	121.36	121.36	120.19
9	120.26	119.96	124.60
10	122.22	122.11	121.90
11	111.99	111.81	113.52
11a	137.08	136.89	135.41
12a	NO[b]	126.30[a]	NO[b]
12b	NO[b]	128.49[a]	NO[b]
13a	137.08	135.29	135.41

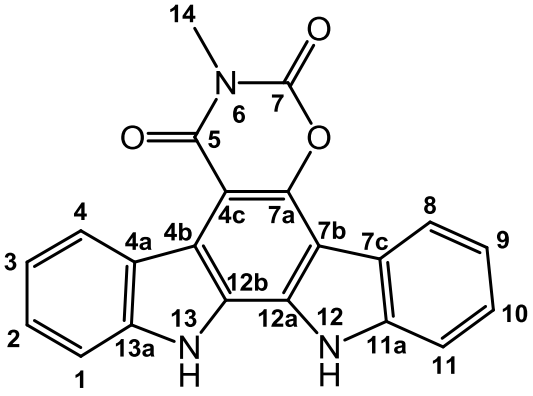
[a] assignments may be interchanged
[b] not observed

Table S6. X-ray crystallography data on compound **10**

I. *Crystal Data*	
Crystal Colour, Habit	Yellow, blade
Crystal Dimensions	0.02 x 0.10 x 0.33 mm
Space Group	$P\ 2_12_12$ (#18)
Lattice Parameters	a = 21.4118(13) Å b = 25.061(2) Å c = 7.4126(5) Å V = 3977.6(5) Å^3
Z value	8
D_{calc}	1.308 g/cm^3
F_{000}	1600.00
μ(Cu-Kα)	19.67 cm^{-1}
II. *Intensity Measurements*	
Diffractometer	Bruker APEX DUO
Radiation	Cu-Kα (λ = 1.54178 Å)
Data Images	2076 exposures @ 20.0, 60.0 s
Detector Position	49.85 mm
$2\theta_{max}$	112.3°
No. of Reflections Measured	Total: 45628 Unique: 11300 (R_{int} = 0.084)
Corrections	Absorption (T_{min} = 0.716, T_{max}= 0.967) Lorentz-polarization
III. *Structure Solution and Refinement*	
Structure Solution	Direct Methods (SIR97)
Refinement	Full-matrix least-squares on F^2
Function Minimized	$\Sigma w\ (Fo^2 - Fc^2)^2$
Least Squares Weights	$w=1/(\sigma^2(Fo^2)+(0.1623P)^2+ 12.3005P)$
Anomalous Dispersion	All non-hydrogen atoms
No. Observations (I>0.00 σ (I))	11300
No. Variables	510
Reflection/Parameter Ratio	22.16
Residuals (refined on F^2, all data): R1; wR2	0.134; 0.324
Goodness of Fit Indicator	1.05
No. Observations (I>2.00 σ (I))	8438
Residuals (refined on F^2): R1; wR2	0.109; 0.301
Max Shift/Error in Final Cycle	0.00
Maximum peak in Final Diff. Map	0.77 e$^-$/Å^3
Minimum peak in Final Diff. Map	-0.40 e$^-$/Å^3

Table S7. ^{1}H NMR data (for compound **15** and **23**) in DMSO-d_6 at 600 MHz



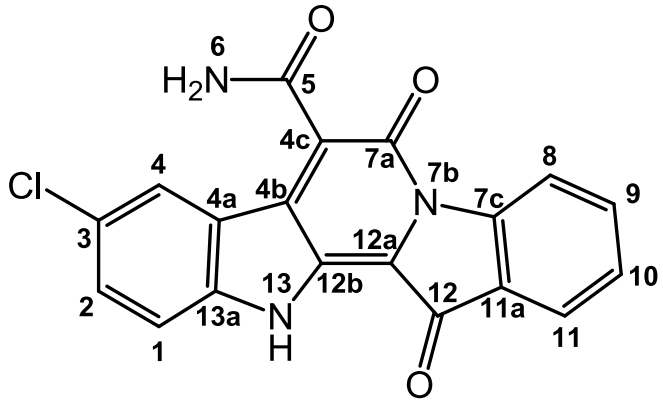


	^{1}H NMR of **15**	^{1}H NMR of **23**
1	7.76 (d, *J*=7.8 Hz, 1H)	7.77 (d, *J*=7.8 Hz, 1H)
2	7.47 (t, *J*=7.8 Hz, 1H)	7.48 (t, *J*=7.8 Hz, 1H)
3	7.24 (t, *J*=7.8 Hz, 1H)	7.25 (t, *J*=7.8 Hz, 1H)
4	9.39 (d, *J*=7.8 Hz, 1H)	9.43 (d, *J*=7.8 Hz, 1H)
4a		
4b		
4c		
5		
6	11.99 (s)	
7		
7a		
7b		
7c		
8	8.33 (d, *J*=7.8 Hz, 1H)	8.34 (d, *J*=7.8 Hz, 1H)
9	7.40 (t, *J*=7.8 Hz, 1H)	7.40 (t, *J*=7.8 Hz, 1H)
10	7.53 (t, *J*=7.8 Hz, 1H)	7.54 (t, *J*=7.8 Hz, 1H)
11	7.84 (d, *J*=7.8 Hz, 1H)	7.85 (d, *J*=7.8 Hz, 1H)
11a		
12	11.85 (s)	11.87 (s)
12a		
12b		
13	11.57 (s)	11.62 (s)
13a		
14		3.48 (s)

Table S8. ^{13}C NMR data of compound **15** and **23** in DMSO-d_6 at 150 MHz.

	^{13}C NMR of **15**	^{13}C NMR of **23**
1	111.65	111.32
2	125.56	125.41
3	119.06	118.60
4	126.14	125.82
4a	122.81	122.73
4b	115.41	115.24
4c	100.91	100.68
5	162.44	161.44
7	147.93	148.54
7a	147.51	145.95
7b	107.17	106.77
7c	121.53	121.38
8	121.66	121.51
9	120.95	120.95
10	125.74	125.56
11	112.36	112.08
11a	139.15	139.11
12a	130.58	130.48
12b	123.00	123.06
13a	139.95	139.87
14		28.57

Table S9. ^{1}H NMR data for **16** in DMSO-d_6 at 600 MHz and ^{13}C NMR data for **16** in DMSO-d_6 at 150 MHz.



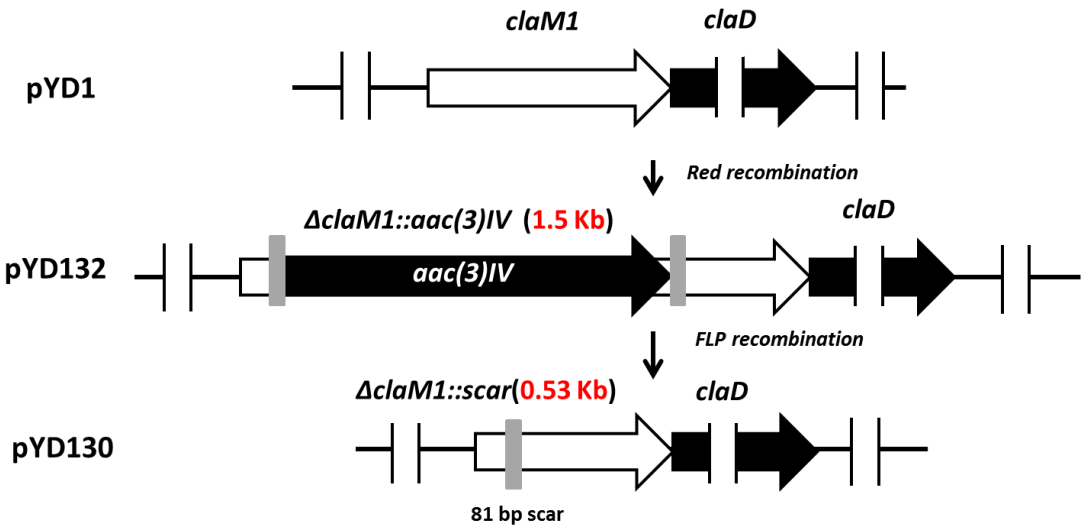
	^{1}H NMR of **16**	^{13}C NMR of **16**
1	7.41 (d, *J*=8.6 Hz, 1H)	114.01
2	7.60 (dd, *J*=8.6, 1.8 Hz, 1H)	132.21
3		124.95
4	7.94 (d, *J*=1.8 Hz, 1H)	124.45
4a		119.70
4b		136.95
4c		129.17
5		165.17
6	8.30 s, 8.08 s	
7a		154.81
7c		136.12
8	8.60 (d, *J*=7.9 Hz, 1H)	117.73
9	7.77 (t, *J*=7.9 Hz, 1H)	124.49
10	7.43 (t, *J*=7.9 Hz, 1H)	126.74
11	7.84 (d, *J*=7.9 Hz, 1H)	123.95
11a		146.08
12		180.37
12a		115.55
12b		128.50
13	11.94 s	
13a		146.40

Table S10. Summary of IC_{50} values against HCT116.

Compounds	IC50 (μM)
13/**14**[1]	2.97
15	5.65
16	>100
17	>100

[1]Mixture of **13**/**14**(1:1) tested

A



B

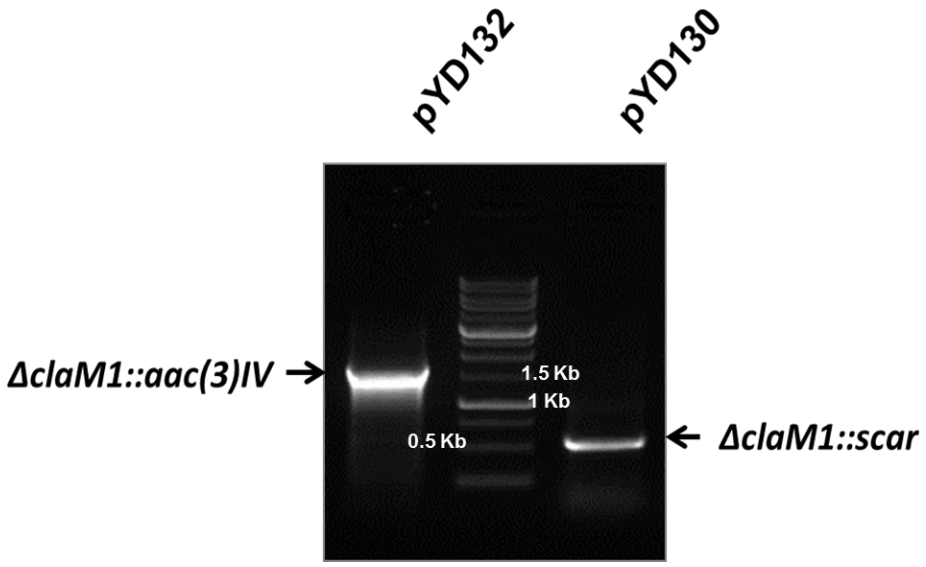


Figure S1. Generation of the *cla* (*ΔclaM1::scar*) gene cluster. (A) Design of vectors and (B) confirmation by PCR (primer pair: claM1-NheI-F/ claM1-XhoI-R, which is also used to amplify the *claM1* coding region for *claM1* expression in *E. coli*).

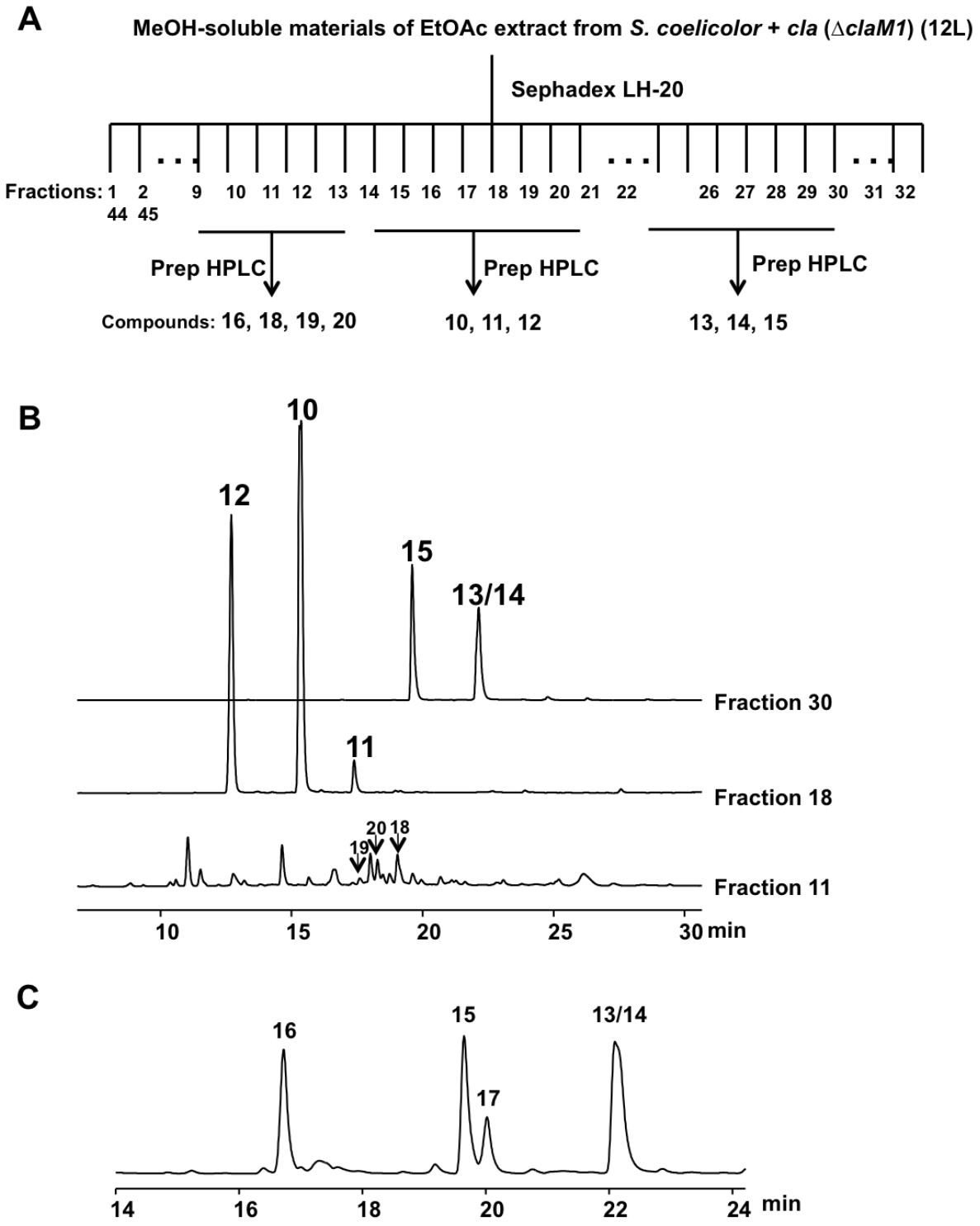


Figure S2. Clone-specific metabolites isolated from *S. coelicolor* + *cla* (Δ*claM1*). A. Purification of clone-specific metabolites by Sephadex LH-20 and semi-preparative HPLC from the methanol-soluble materials of the EtOAc extract. B. HPLC analysis of fractions 11, 18 and 30 after Sephadex LH-20 fractionation in (A). C. HPLC analysis of methanol-insoluble materials of the ethyl acetate extraction. Detection wavelength: 348 nm.

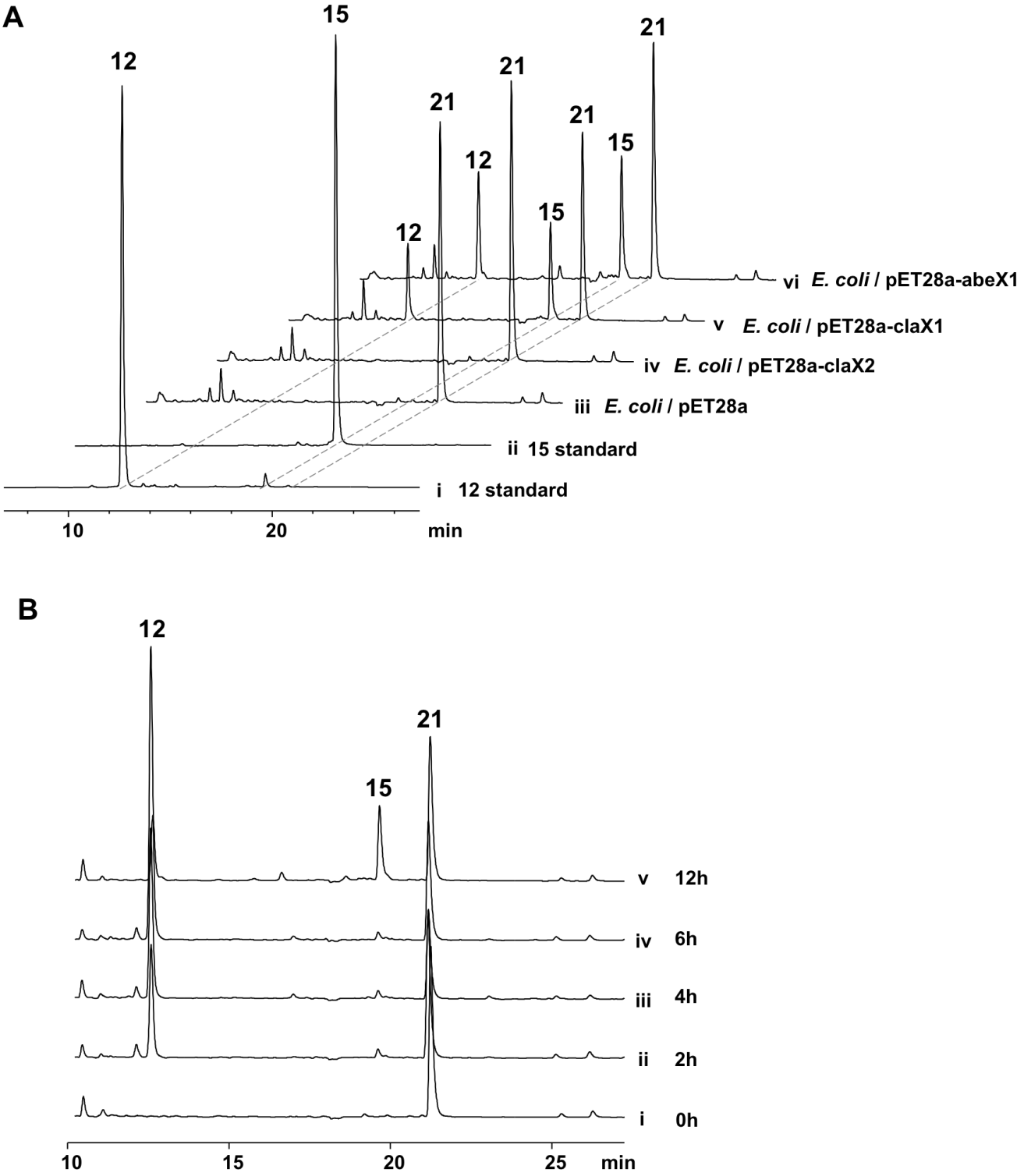


Figure S3. Biotransformation with *claX1* and *abeX1*. A. *In vivo* bioconversion of **21** to **12** and **15** by *E. coli* host carrying *claX1* or *abeX1*. B. Time course of the bioconversion assay of **21** to **12** and **15** by *E. coli* host carrying pET28a-*claX1*. Detection wavelength: 348 nm.

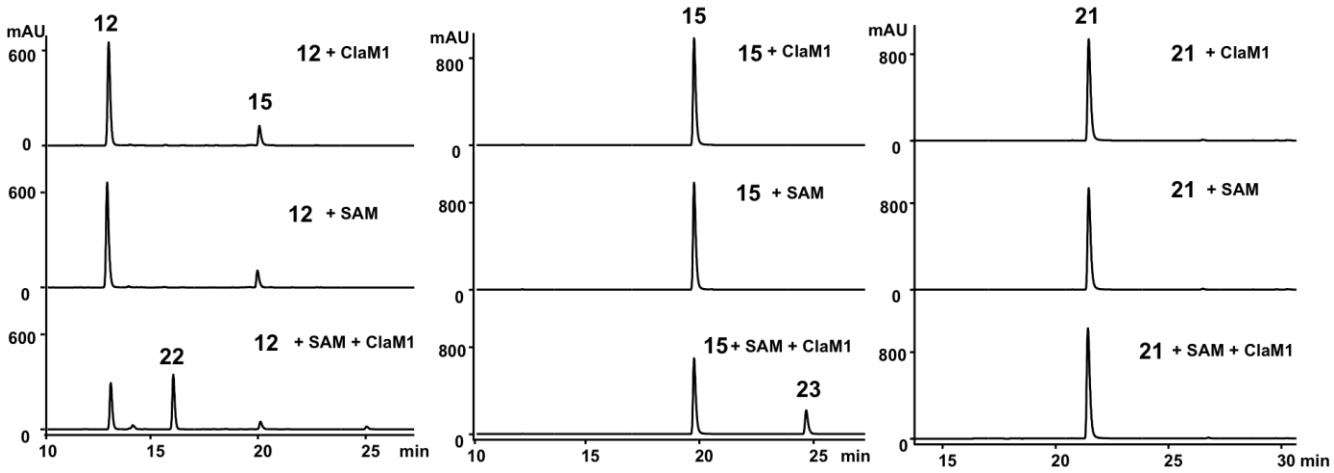


Figure S4. *In vitro* biochemical assay of ClaM1. A) Substrate is **12**, B) substrate is **15**, and C) substrate is **21**. Detection wavelength: 348 nm

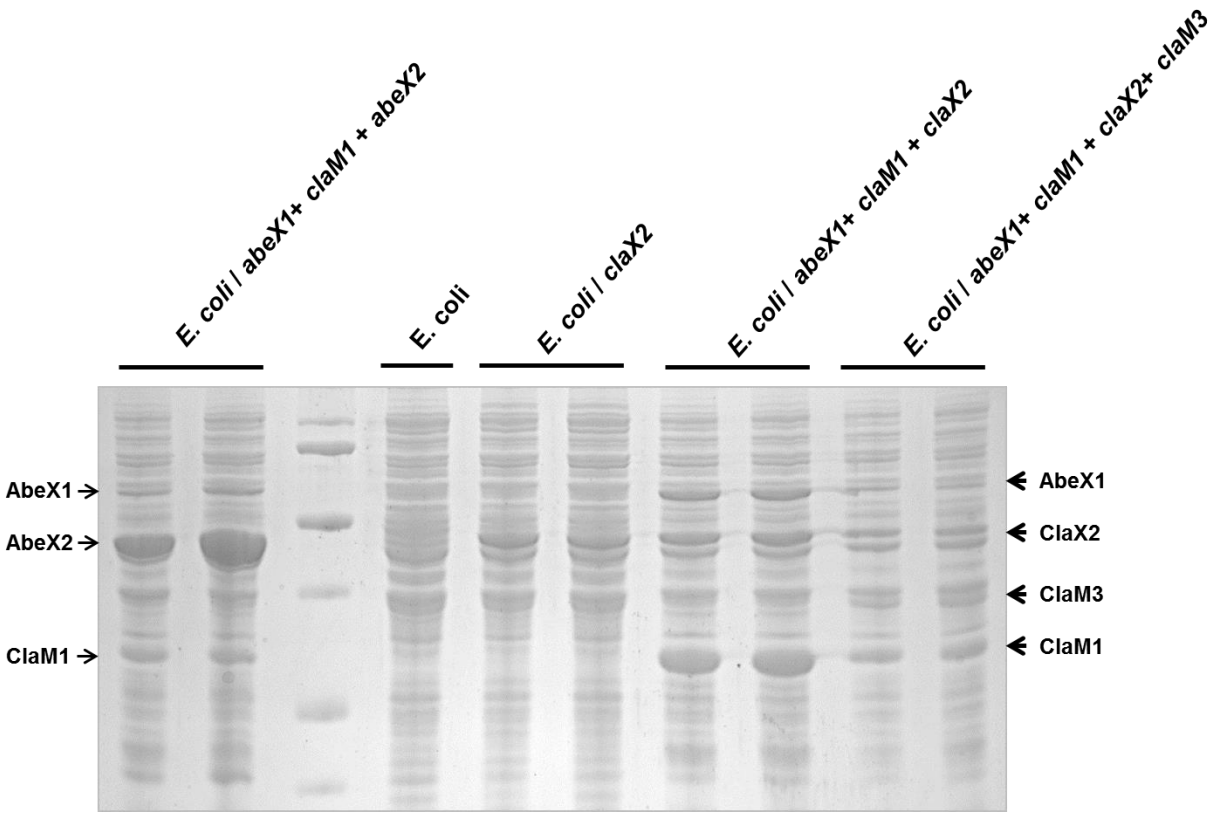


Figure S5. Co-expression of *cla*/*abe* genes in *E.coli* BL21(DE3).

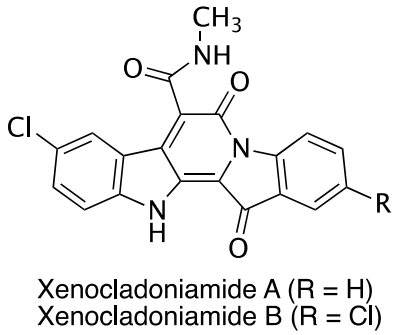


Figure S6. Structures of xenocladoniamides A and B.

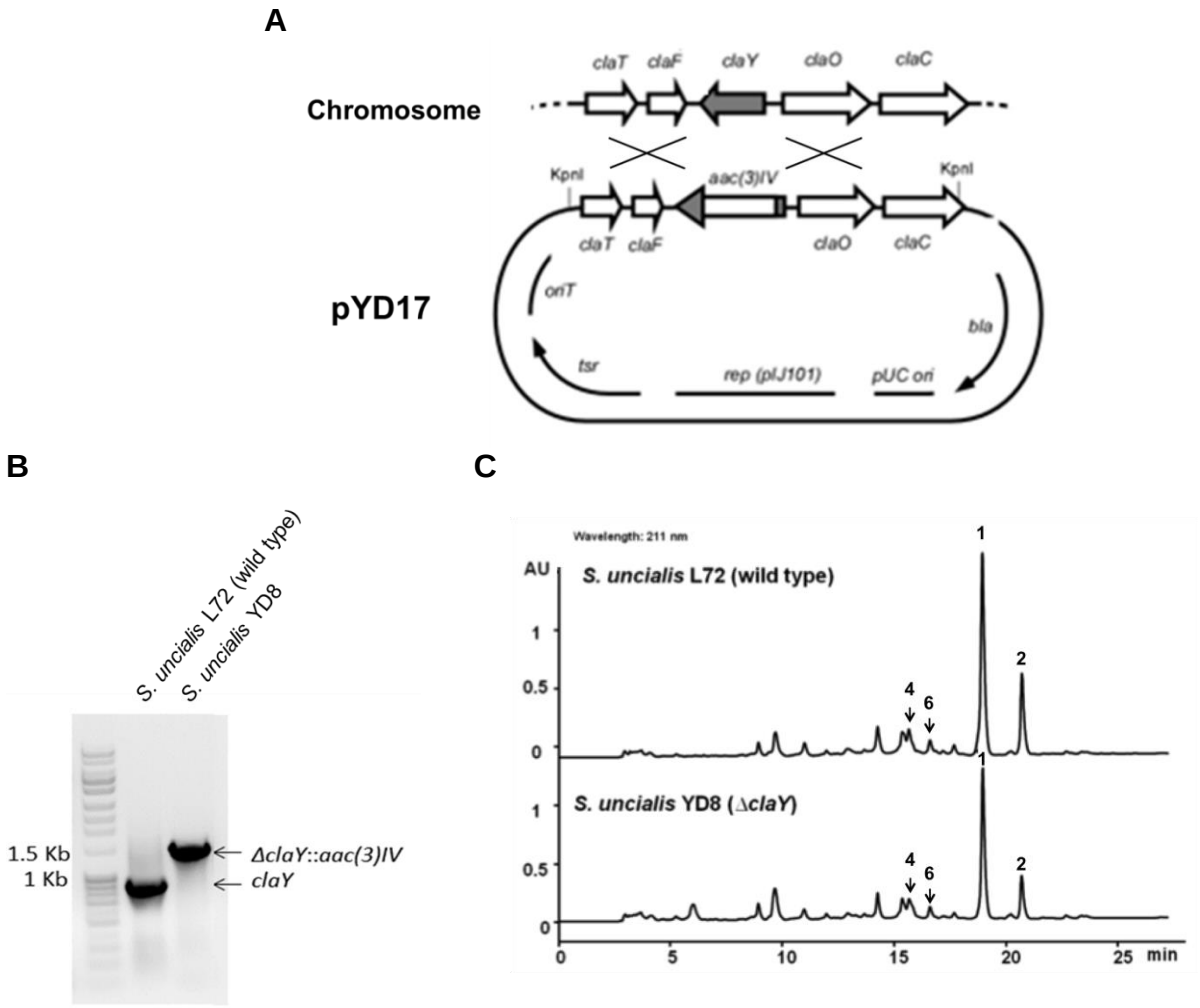


Figure S7. Gene inactivation of *claY* in *S. uncialis* has no effect on cladoniamide production. A. Construction of *claY*-disrupted mutant. B. Confirmation of of *claY*-disrupted mutant *S. uncialis* YD8 by PCR with the primers for *claY*. C. Metabolic analysis of wild type and *claY*-disrupted strain by HPLC. Detection wavelength: 348 nm.

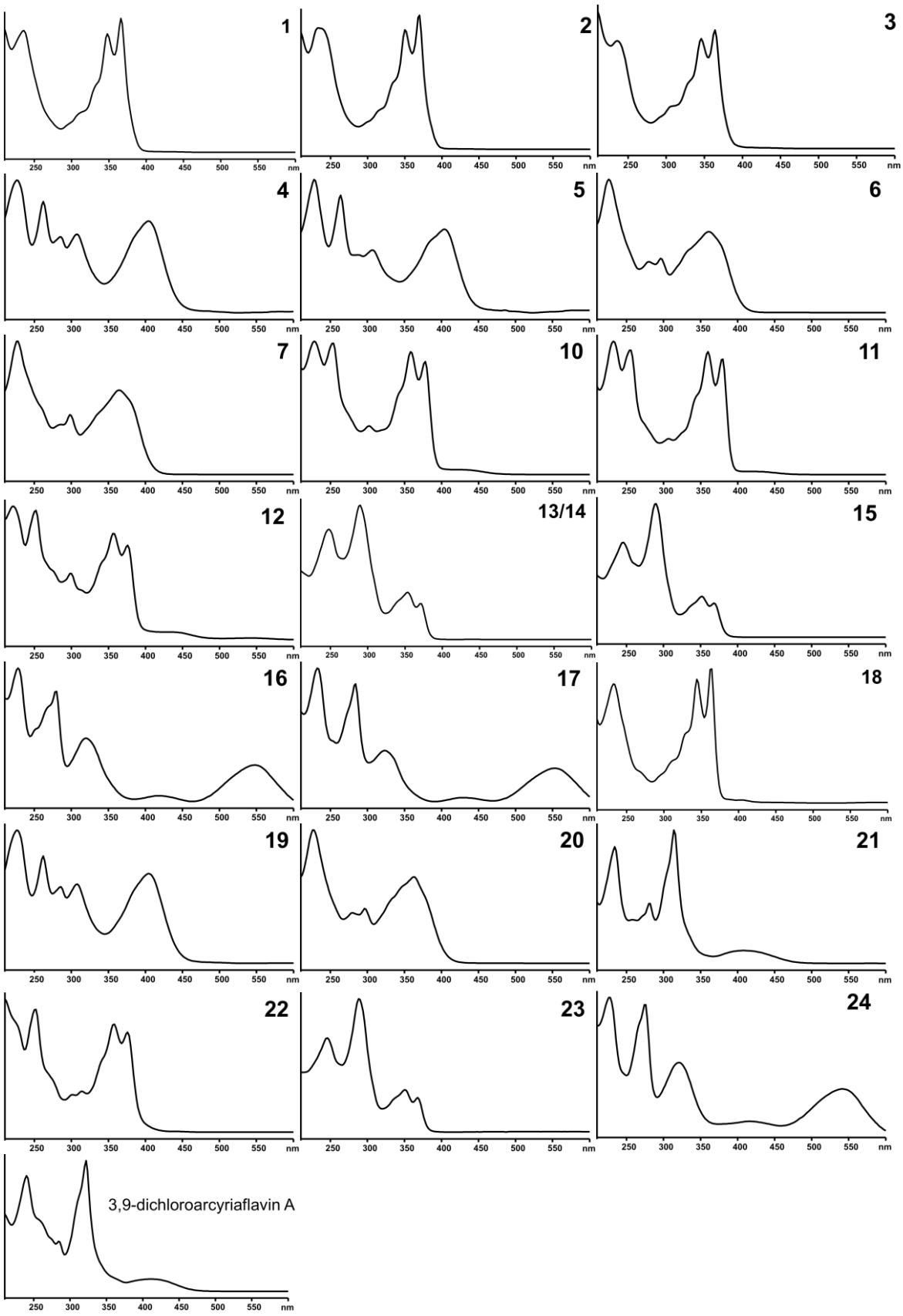


Figure S8. UV-Vis spectra of compounds **1-7, 10-24**, and 3,9-dichloroarcyriaflavin A.

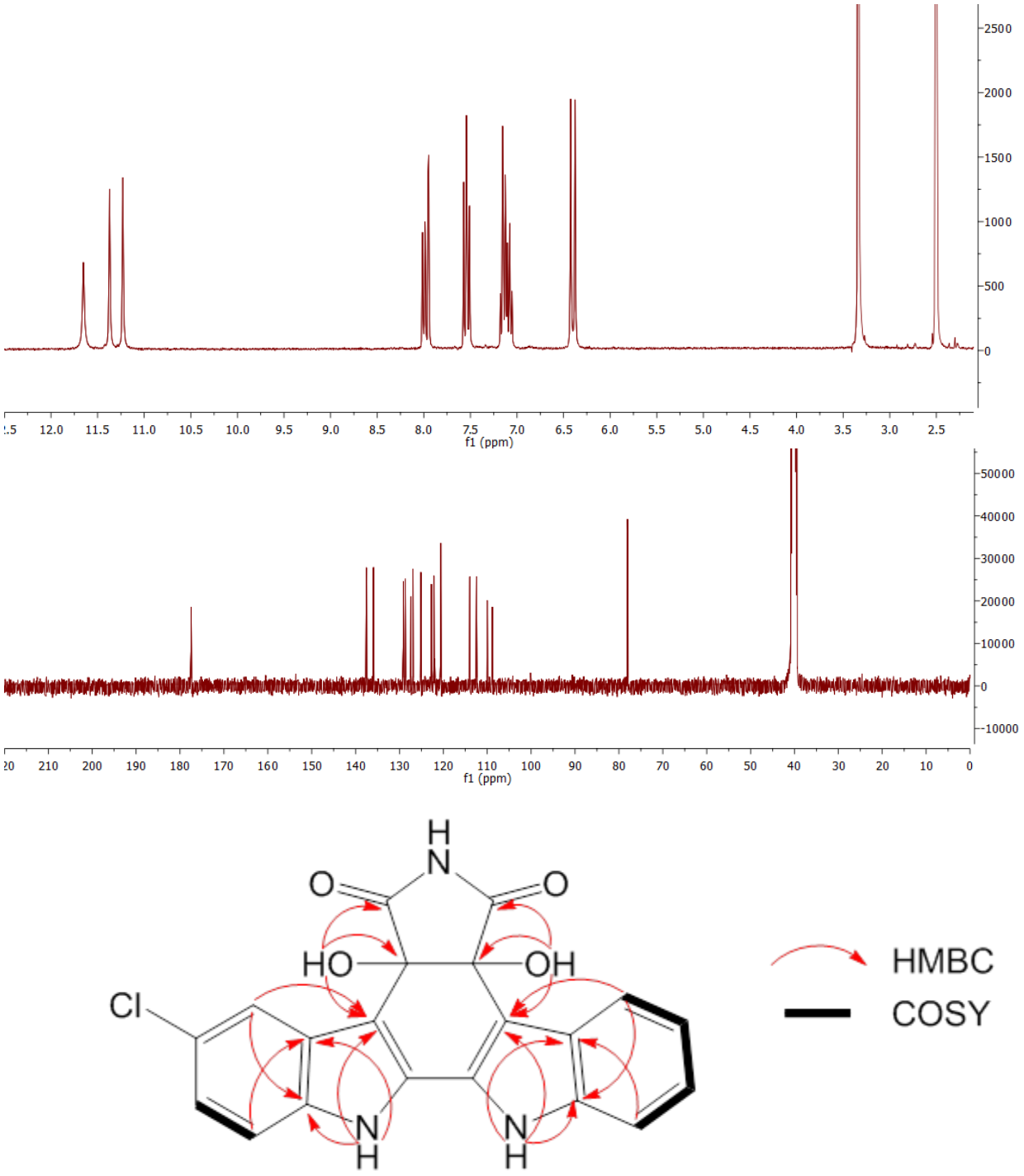


Figure S9. ^{1}H NMR spectrum of **10** recorded at 600 MHz in DMSO-d_6 (top), ^{13}C NMR spectrum of **10** recorded at 150 MHz in DMSO-d_6 (middle) and COSY and key HMBC correlations of **10** (bottom).

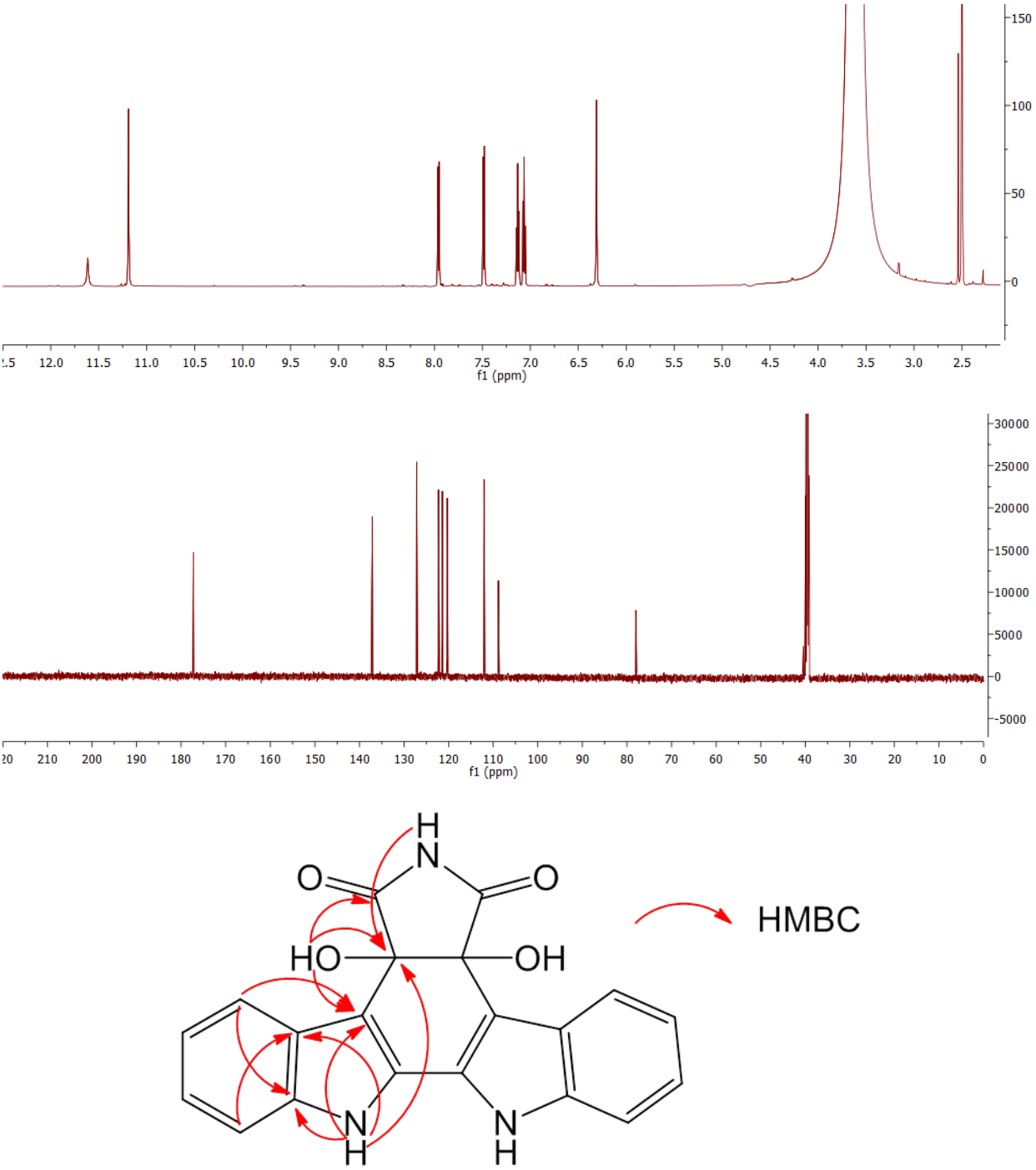


Figure S10. ^{1}H NMR spectrum of **12** recorded at 600 MHz in DMSO-d_6 (top), ^{13}C NMR spectrum of **12** recorded at 150 MHz in DMSO-d_6 (middle) and key HMBC correlations of **12** (bottom).

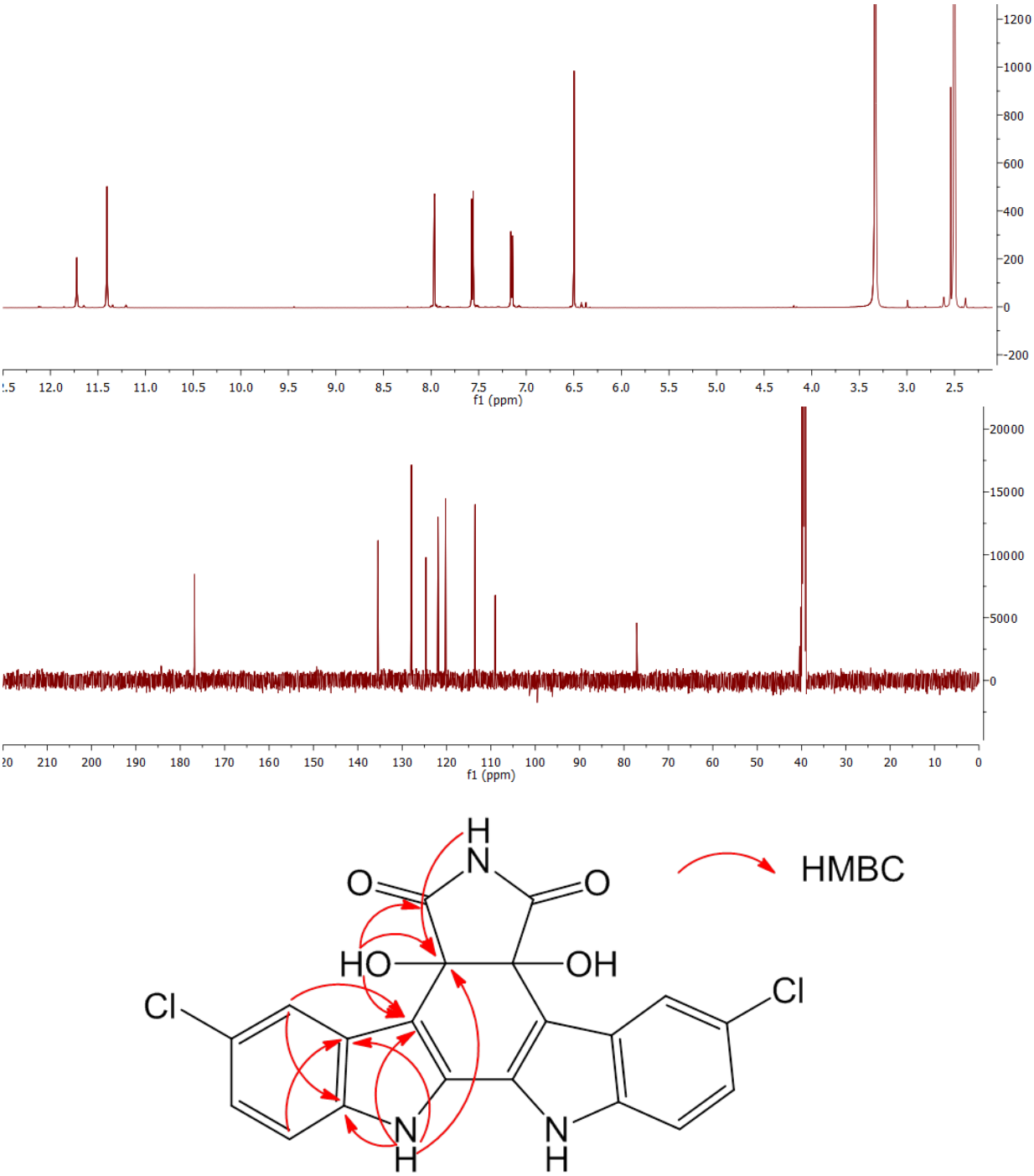


Figure S11. ^{1}H NMR spectrum of **11** recorded at 600 MHz in DMSO-d_6 (top), ^{13}C NMR spectrum of **11** recorded at 150 MHz in DMSO-d_6 (middle) and key HMBC correlations of **11** (bottom).

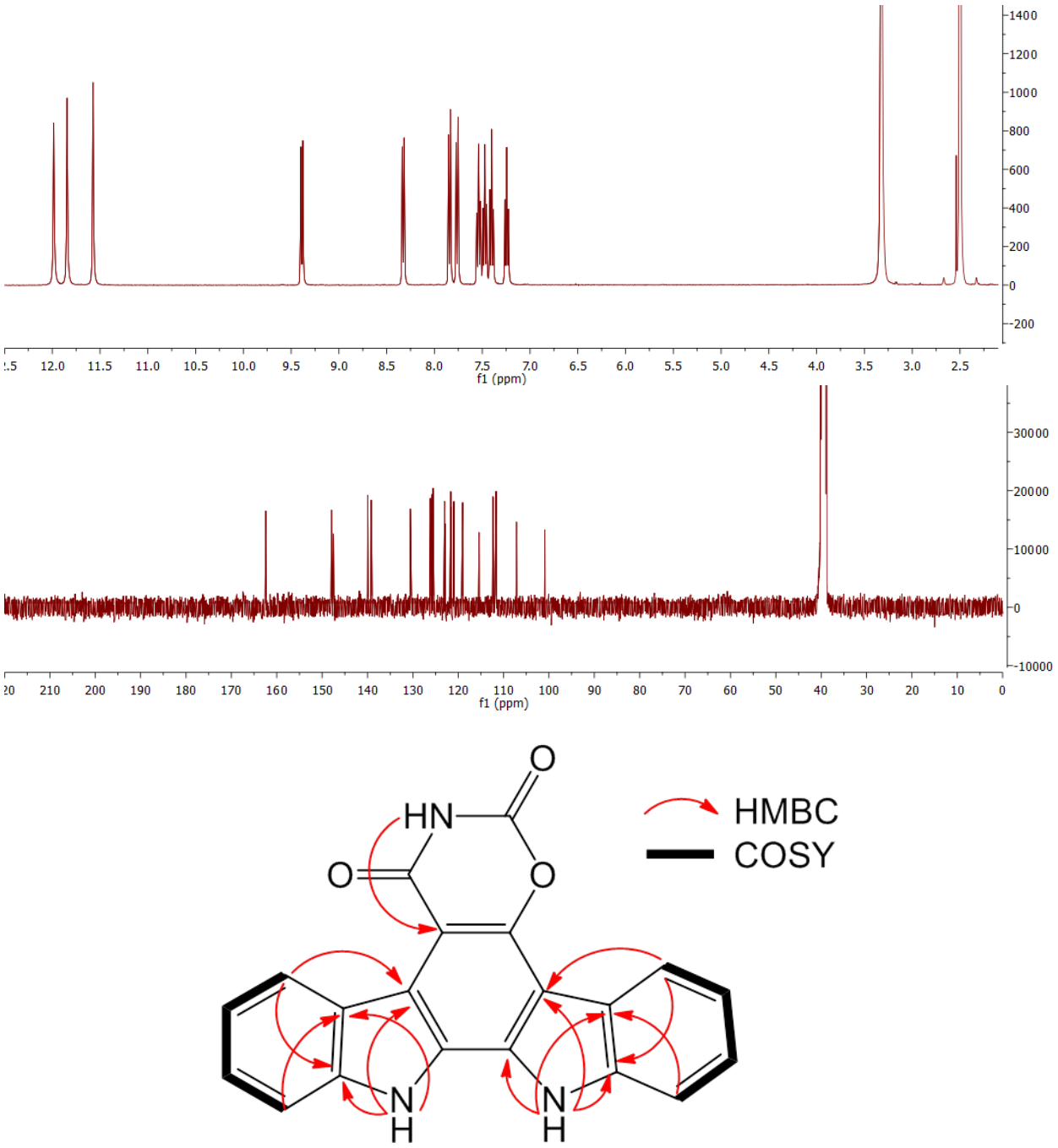


Figure S12. ^{1}H NMR spectrum of **15** recorded at 600 MHz in DMSO-d_6 (top), ^{13}C NMR spectrum of **15** recorded at 150 MHz in DMSO-d_6 (middle) and COSY and key HMBC correlations of **15** (bottom).

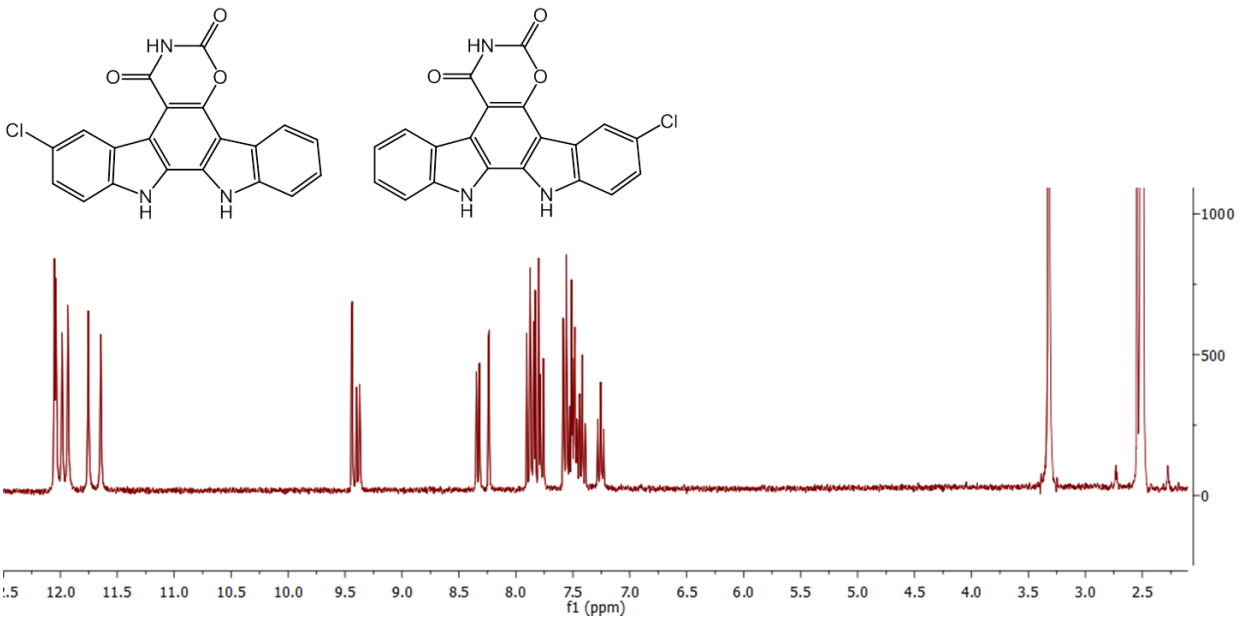


Figure S13. ^{1}H NMR spectrum of **13**/**14** recorded at 600 MHz in DMSO-d_6

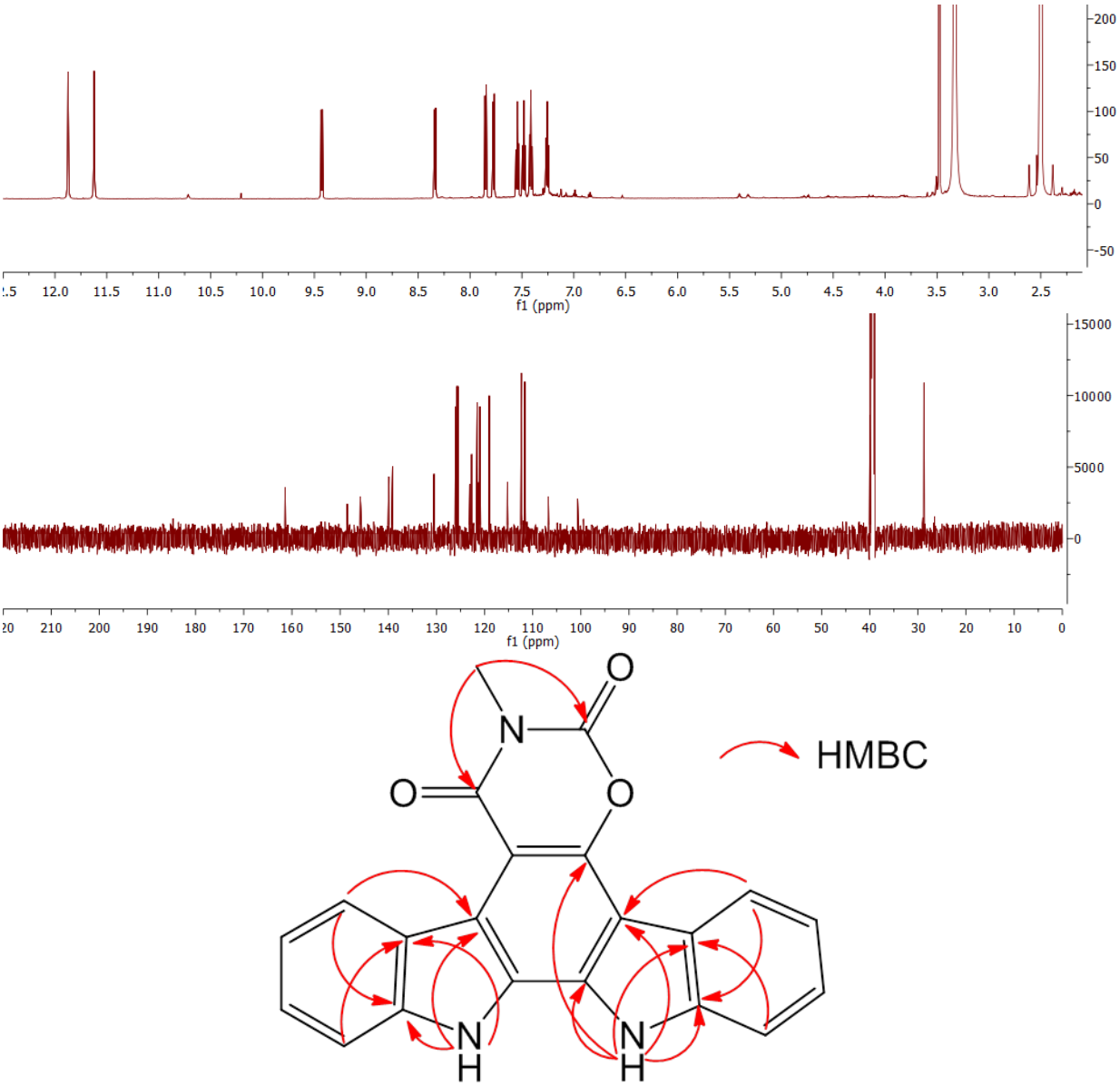


Figure S14. ^{1}H NMR spectrum of **23** recorded at 600 MHz in DMSO-d_6 (top), ^{13}C NMR spectrum of **23** recorded at 150 MHz in DMSO-d_6 (middle) and key HMBC correlations of **23** (bottom).

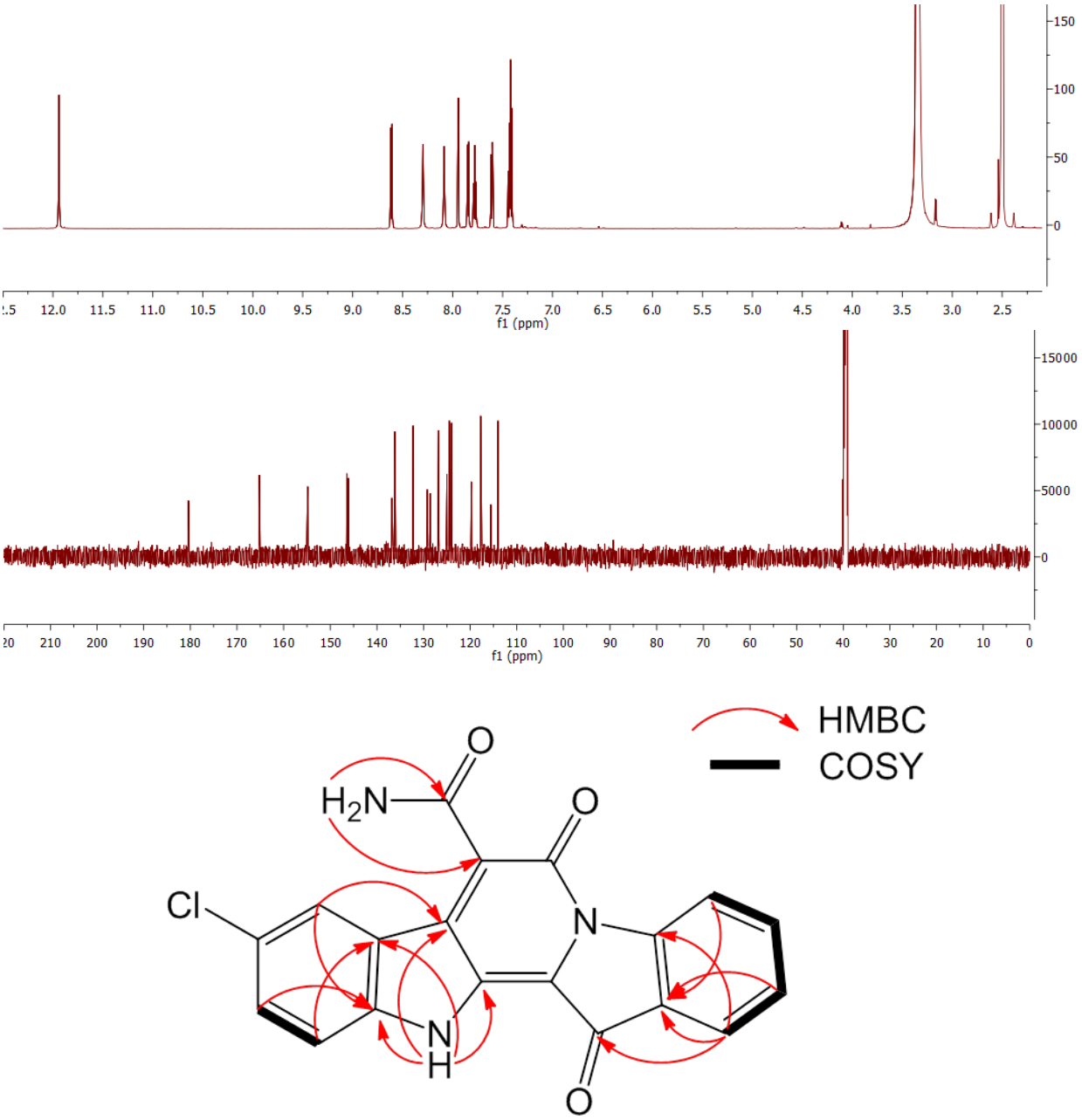


Figure S15. ^{1}H NMR spectrum of **16** recorded at 600 MHz in DMSO-d_6 (top), ^{13}C NMR spectrum of **16** recorded at 150 MHz in DMSO-d_6 (middle) and COSY and key HMBC correlations of **16** (bottom).

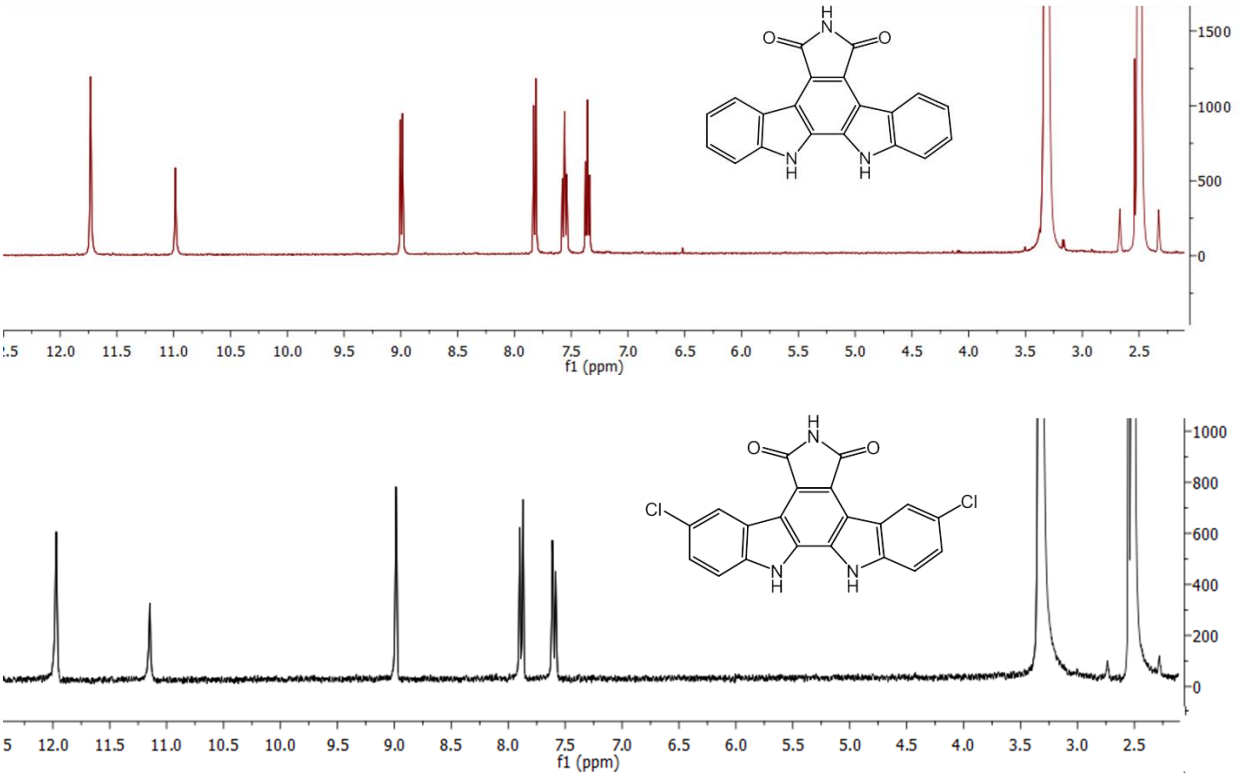


Figure S16. ^{1}H NMR spectrum of **21** (top) and 3,9-dichloroarcyriaflavin A (bottom) recorded at 600 MHz in DMSO-d_6.

MATERIALS AND METHODS

General Bacterial strains and vectors used or generated in this study are listed in Table S1. All oligonucleotide primers used are presented in Table S2. Primers were purchased from Integrated DNA Technologies, and all sequencing was performed at the DNA Sequencing Laboratory at the University of British Columbia. Reagents were purchased from Fisher Scientific Canada, Sigma-Aldrich, Invitrogen, New England Biolabs, P212121, and Santa Cruz Biotechnology. DNA manipulations in *Escherichia coli* and *Streptomyces* were carried out according to standard procedures.[74] *E. coli*-*Streptomyces* conjugation was performed as described previously.[5] Kanamycin (50 μg/mL), apramycin (50 μg/mL), ampicillin (100 μg/mL) and nalidixic acid (25 μg/mL) were used for selection of recombinant strains.

Gene deletion of *claM1* in heterologous host Integrative vector pYD1, carrying the whole *cla* gene cluster, was previously constructed for heterologous production of cladoniamides.[5] Putative *N*-methyltransferase gene *claM1* was first inactivated in pYD1 by insertion of an *aac(3)IV* gene amplified from pHY773 into its coding region, using the λ-RED-mediated PCR targeting method. The targeted plasmid was then transformed into *E. coli* DHα/BT340 in order to excise the *aac(3)IV* cassette as described previously.[3] The resulting plasmid was designated pYD130. The *neo* gene of the backbone of pYD130 was then replaced by an *aac(3)IV* gene, resulting in pYD134. After confirmation with restriction and PCR analysis, pYD134 was passaged through *E. coli* ET12567/pUZ8002 before being introduced via conjugation into *Streptomyces coelicolor* M1146 to afford *S. coelicolor* YD52. Apramycin resistant colonies were confirmed for integration by PCR with the primer pairs for *claM1* (Table S2).

Production and analysis of metabolites from engineered *Streptomyces* strain Cultivation of *Streptomyces* mutants was performed as described previously.[5] At the end of fermentation, the cultures were extracted twice with 1 volume of ethyl acetate, and combined extracts were dried under vacuum and then dissolved in methanol. Insoluble material was separated by centrifugation, washed three times with methanol, and then dissolved in DMSO for further analysis.

Analysis of bisindole compounds production was performed with an Agilent 1260 high-pressure liquid chromatography (HPLC) apparatus with a diode array UV detector, fitted with an Phenomenex Luna C18(2) column (5 μm, 4.6 mm ID x 250 mm). Elution was performed at 1 mL/min with a mobile-phase mixture consisting of a linear gradient of water and acetonitrile (85:15 to 20:80, 0 to 25 min; 0:100, 25 to 30 min), both of which contain 0.05% (v/v) trifluoroacetic acid. LC-MS was performed under the same conditions (except that the flow rate was set at 0.8 mL/min) on an Agilent 6120 Quadrupole LC/MS system operated in both positive and negative ion electrospray modes.

Isolation and purification of bisindoles from *Streptomyces* mutants Ethyl acetate extracts from *S. coelicolor* YD52 cultures (12 L total) were dried under vacuum and then dissolved in methanol. Insoluble material was separated by centrifugation, washed three times with methanol, and then dissolved in DMSO for further purification. The methanol-soluble fraction was fractionated on Sephadex LH-20 with MeOH elution.

Metabolites of interest, as tracked by analytical HPLC as described above, were purified from these fractions by reversed-phase semi-preparative HPLC (Phenomenex Luna C18(2), 5 µm, 10 mm ID x 250 mm). Compound 3,9-dichloroarcyriaflavin A was isolated as a minor product (< 0.1 mg/L) from the previously reported YD2, a heterologous cladoniamide producer.[5]

Structural elucidation The ^{1}H- and ^{13}C-NMR spectra were recorded on a Bruker AV-600 MHz spectrometer with a 5 mm CPTCI cryoprobe. The DMSO-d_6 signals (2.50 ppm for ^{1}H and 39.5 ppm for ^{13}C) were used as references. High-resolution mass spectrometry data were acquired using a Micromass LCT time-of-flight (TOF) mass spectrometer equipped with an electrospray ion source. X-ray diffraction data was collected on a Bruker APEX DUO diffractometer with cross-coupled multilayer optics using Cu-Kα radiation.

Cloning, expression, and purification of recombinant ClaM1. The *claM1* coding region was amplified from cosmid pYD1 by using Q5 DNA polymerase and primer pair claM1-NheI-F/claM1-XhoI-R. After digestion with NheI and XhoI, the PCR products were ligated into the corresponding sites of pET24b. After confirmation by DNA sequencing, pET24b-claM1 was introduced into *E. coli* BL21 (DE3) for recombinant protein expression. The culture was grown at 37°C to an optical density at 600 nm of 0.4 to 0.6. Isopropyl-β-D-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.1 mM. After 12 h of further incubation at 16°C, the cells were harvested by centrifugation and disrupted by sonication on ice, and the supernatant was recovered by centrifugation (13,000 × g for 20 min). His-tagged ClaM1 was separated using HisTrap HP column (GE Healthcare). After being washed with washing buffer (300 mM NaCl, 50 mM Tris-HCl, 50 mM imidazole, pH 8.0), protein was eluted with elution buffer (300 mM NaCl, 50 mM Tris-HCl, 300 mM imidazole, pH 8.0). Protein-containing fractions were confirmed by SDS-PAGE, and then dialyzed overnight into a buffer containing 20 mM Tris-HCl, 50 mM NaCl, 10% glycerol (pH 8.0). The dialyzed protein was concentrated by using a Millipore Amicon Ultra Centrifugal filter (10,000 MWCO) and stored at -80°C for further use. Protein concentration was determined by using the Bradford Protein assay (BioRad).

Biochemical assays of ClaM1 The assay for the ClaM1 activity was performed at 100 µL scale, 30°C for 4 hours. The assay mixture contained 50 mM Tris-HCl (pH 8.0), 1 mM *S*-adenosyl-L-methionine (SAM) and 100 µM substrate and 4 µM ClaM1. All the reactions were stopped by addition of 500 µL ethyl acetate. The ethyl acetate extract of the reaction mixture was then dried and re-dissolved in 100 µL methanol and analyzed with HPLC or LC-MS. For purification of **23**, the reaction described above was scaled up 500-fold and incubated overnight at 30°C. The product **23** was purified from the ethyl acetate extract of the reaction mixture by reversed-phase semi-preparative HPLC (Phenomenex Luna C18(2), 5 µm, 10 mm ID x 250 mm) using 4:1 (v:v) acetonitrile:water.

Heterologous reconstitution of *cla* pathways (partial) in *E. coli* host Coding regions of *claX1*, *claX2*, *abeX1*, *abeX2*, *claM1* and *claM3* were amplified from pYD1 or the plasmids carrying synthetic genes codon optimized for *E. coli* (synthetic genes were purchased from Bio Basic Canada, Inc.) using Q5 DNA polymerase as recommended by the manufacturer. Amplicons were cloned into a variety of commercially available *E. coli* expresssion vectors (Table S1). All the genes were placed directly under the T7 promoter. The resulting

recombinant vectors were then introduced into *E. coli* host individually or with different combinations, resulting in various engineered *E. coli* strains (Table S1).

***In vivo* biotransformation** Overnight cultures of *E. coli* strains expressing Cla enzyme(s) were inoculated into 100 mL LB medium with appropriate antibiotics and incubated with shaking (200 rpm) at 30°C to an optical density at 600 nm of 0.6. IPTG was then added to a final concentration of 0.1 mM. After 12 h of further incubation at 16°C, the cells were harvested by centrifugation and resuspended in 10 mL LB medium, and compound **21** (stock at 10 mg/mL in DMSO) was added to a final concentration of 5 μg/mL. The cultures were further incubated at 30°C and 200 rpm for 8 h or 12 h before being extracted twice with 1 volume of ethyl acetate. Combined extracts were dried under vacuum and dissolved in 500 μL DMSO, 20 μL of which was used for HPLC analysis performed as described above. In the case of heterologous production of cladoniamide B, similar assays were performed except that 3,9-dichloro-arcyriaflavin A (final concentration: 20 μg/mL) was provided to the culture and the culture extract was eventually dissolved in 100 μL DMSO.

Construction of *claY*-disrupted mutant in *S. uncialis* For the construction of *claY*-disrupted *S. uncialis* mutant, *claY* was first interrupted in pYD1 by insertion of an *aac(3)IV* gene amplified from pHY773 into its coding region, using the λ-RED-mediated PCR targeting method, generating pYD36. An 8-kb KpnI-digested fragment containing the *claY* knock-out cassette from pYD36 was cloned into pMRD400,[8] resulting in pYD17. After confirmation with restriction and PCR analysis, pYD17 was introduced into *E. coli* ET12567/pUZ8002 for *E. coli-Streptomyces* conjugation, which is based on standard protocol and performed as follows. As *S. uncialis* recipients, ~ 10^9 *S. uncialis* spores were heat-shocked in 500 μl of 2 × YT medium at 50°C for 10 min, followed by 3 h at 30°C at 225 rpm. *E. coli* donors were prepared according to standard procedure. For conjugation, the recipients and donors were mixed and distributed onto ISP4 agar supplemented with 20 mM $MgCl_2$. The plates were then incubated at 30°C for 20 h and overlaid with 500 μl of water containing 25 μg/mL nalidixic acid and 50 μg/mL apramycin to select for *S. uncialis* exconjugates. Exconjugants were inoculated onto R5 agar for two rounds of non-selective growth before selection by replica plating for thiostrepton-sensitive and apramycin-resistent colonies. Disruption of *claY* was confirmed by PCR analysis using claY-NdeI-F/claY-XbaI-R primers (Table S2).

Cytotoxicity assays Cytotoxicity against human colon tumor cell line HCT-116 (ATCC CCL-247) was analyzed by the MTT assay.[9] Cells were seeded as 100-μL aliquots into 96-well plates at a density of 10^5 cells/mL and diluted two-fold with McCoy's 5A Modified Media supplemented with 10% FBS, 50 IU/mL penicillin and 50 μg/mL streptomycin prior to overnight growth at 37°C in a 5% CO_2 chamber. Compounds were dissolved in DMSO, diluted in culture medium in concentrations and added to the cells at final concentrations of 50, 25, 12.5, 6.25, 3.13, 1.56 and 0.78 μg/mL (for **13**/**14**, **15, 16, 17**). For the blank control, 100 μL of culture medium was added. For the DMSO control, 100 μL of culture medium with 0.5% DMSO was added. After 72 h of incubation, 50 μL of MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) (2.5 mg/mL) was added to each experimental well, and plates were returned to the incubator for 3 h. An aliquot of 230 μL of medium was aspirated from each well. 100 μL of DMSO was added to lyse the cells and solubilize

the crystals. The samples were read on a DTX 880 (Beckman Coulter Inc.) plate reader at 570 nm. Cell viability at each compound concentration was expressed as a percentage of untreated controls, and IC_{50} values were determined using Prism5 (GraphPad, CA, USA).

X-Ray Crystallography Methods

Crystallization Compound **10** was crystallized from a saturated solution of **10** in 4:1 EtOAc:MeOH, which was left standing in a glass vial, with a loose cap, at 2ºC for ~1 week.

Data Collection A yellow blade crystal of $C_{20}H_{10}N_3O_4Cl$ having approximate dimensions of 0.02 x 0.10 x 0.33 mm was mounted on a glass fiber. All measurements were made on a Bruker APEX DUO diffractometer with cross-coupled multilayer optics using Cu-Kα radiation**.** The data were collected at a temperature of -183.0 ± 0.1ºC to a maximum 2θ value of 112.3º. Data were collected in a series of ϕ- and ω-scans in 1.0º oscillations using 20.0- and 60.0-s exposures. The crystal-to-detector distance was 49.85 mm.

Data Reduction The material crystallizes as a two-component 'split-crystal' with components one and two related by a 3.1º rotation about the (1.000 -0.05 -0.47) real axis. Data were integrated for both components, including both overlapped and non-overlapped reflections. In total 45628 reflections were integrated (19828 from component one only, 19674 from component two only, 6126 overlapped). Data were collected and integrated using the Bruker SAINT[10] software packages. The linear absorption coefficient, μ, for Cu-Kα radiation is 19.67 cm^{-1}. Data were corrected for absorption effects using the multi-scan technique (TWINABS[11]), with minimum and maximum transmission coefficients of 0.716 and 0.967, respectively. The data were corrected for Lorentz and polarization effects.

Structure Solution and Refinement The structure was solved by direct methods[12] using non-overlapped data from the major twin component. Additionally, the material crystallizes with disordered solvent in the lattice. The exact nature of the solvent could not be resolved (the material was crystallized from a mixture of MeOH and EtOAc). The PLATON/SQUEEZE[13] program was used to generate a 'solvent-free' data set. Subsequent refinements were carried out using an HKLF 5 format data set containing complete data from both components including overlaps. All hydrogen atoms were placed in calculated positions, and as such any hydrogen-bonding parameters are somewhat ambiguous. Finally, the absolute configuration (S and R at C2 and C3, respectively) was determined on the basis of the refined Flack x-parameter.[14] The final cycle of full-matrix least-squares refinement[15] on F^2 was based on 11300 reflections and 510 variable parameters and converged (largest parameter shift was 0.00 times its esd) with unweighted and weighted agreement factors of:

$R1 = \Sigma \, ||Fo| - |Fc|| \, / \, \Sigma \, |Fo| = 0.134$

$wR2 = [\Sigma \, (\, w \, (Fo^2 - Fc^2)^2 \,) / \, \Sigma \, w(Fo^2)^2]^{1/2} = 0.324$

The standard deviation of an observation of unit weight[16] was 1.05. The weighting scheme was based on counting statistics. The maximum and minimum peaks on the final difference Fourier map corresponded to 0.77 and –0.40 $e^-/Å^3$, respectively. Neutral atom scattering factors were taken from Cromer and Waber.[17] Anomalous dispersion effects were included in Fcalc;[18] the values for Δf' and Δf'' were those of Creagh and McAuley.[19] The

values for the mass attenuation coefficients are those of Creagh and Hubbell.[20] All refinements were performed using the SHELXL-2012[21] via the OLEX2[22] interface. This structure has been deposited at the Cambridge Crystallographic Data Centre with deposition number CCDC 1027820.

REFERENCES

(1) Williams, D. E., Davies, J., Patrick, B. O., Bottriell, H., Tarling, T., Roberge, M., and Andersen, R. J. (2008) Cladoniamides A-G, tryptophan-derived alkaloids produced in culture by Streptomyces uncialis. *Org. Lett. 10*, 3501–4.

(2) Gomez-Escribano, J. P., and Bibb, M. J. (2011) Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters. *Microb. Biotechnol. 4*, 207–15.

(3) Gust, B., Challis, G. L., Fowler, K., Kieser, T., and Chater, K. F. (2003) PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc. Natl. Acad. Sci. U. S. A. 100*, 1541–6.

(4) Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000), in *Practical Streptomyces Genetics*. The John Innes Foundation, Colney, Norwich.

(5) Du, Y.-L., Ding, T., and Ryan, K. S. (2013) Biosynthetic O-Methylation Protects Cladoniamides from Self-destruction. *Org. Lett. 15*, 2538–2541.

(6) Zhou, M., Jing, X., Xie, P., Chen, W., Wang, T., Xia, H., and Qin, Z. (2012) Sequential deletion of all the polyketide synthase and nonribosomal peptide synthetase biosynthetic gene clusters and a 900-kb subtelomeric sequence of the linear chromosome of Streptomyces coelicolor. *FEMS Microbiol. Lett. 333*, 169–179.

(7) Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

(8) Du, Y. L., Shen, X. L., Yu, P., Bai, L. Q., and Li, Y. Q. (2011) Gamma-butyrolactone regulatory system of Streptomyces chattanoogensis links nutrient utilization, metabolism, and development. *Appl. Env. Microbiol. 77*, 8415–26.

(9) Freshney, R. I. (2000) Culture of Animal Cells: A Manual of Basic Technique. Wiley-Liss, New York.

(10) Bruker. (2013) SAINT. Bruker AXS Inc., Madison, Wisconsin, USA.

(11) Bruker. (2012) TWINABS. Bruker AXS Inc., Madison, Wisconsin, USA.

(12) Altomare, A., Burla, M. C., Camalli, M., Cascarano, G. L., Giacovazzo, C., Guagliardi, A., Moliterni, A. G. G., Polidori, G., and Apagna, R. (1999) SIR97. *J. Appl. Crystallogr. 32*, 115–119.

(13) Van der Sluis, P., and Spek, A. L. (1990) BYPASS: an effective method for the refinement of crystal structures containing disordered solvent regions. *Acta Crystallogr. A 46*, 194–201.

(14) Parsons, S., Flack, H. D., and Wagner, T. (2013) Use of intensity quotients and differences in absolute structure refinement. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater. 69*, 249–259.

(15) Least Squares function minimized:

$$\Sigma w(F_o^2-F_c^2)^2$$

(16) Standard deviation of an observation of unit weight:

$$[\Sigma w(F_o^2-F_c^2)^2/(N_O-N_V)]^{1/2}$$

where: N_O = number of observations

N_V = number of variables

(17) Cromer, D. T., and Waber, J. T. (1974) International Tables for X-Ray Crystallography. The Kynoch Press, Birmingham, England. Vol. IV, Table 2.2A.

(18) Ibers, J. A., and Hamilton, W. C. (1964) Dispersion corrections and crystal structure refinements. *Acta Crystallogr. 17*, 781–782.

(19) Creagh, D. C., and McAuley, W. J. (1992) International Tables for Crystallography (Wilson, A. J. C., Ed.). Kluwer Academic Publishers, Boston. Vol. C., Table 4.2.6.8., pp. 219-222.

(20) Creagh, D. C., and Hubbell, J. H. (1992) International Tables for Crystallography (Wilson, A. J. C., Ed.). Kluwer Academic Publishers, Boston. Vol. C., Table 4.2.4.3, pp. 200-206.

(21) Sheldrick, G. M. (2008) A short history of *SHELX*. *Acta Crystallogr. A 64*, 112–122.

(22) Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K., and Puschmann, H. (2009) *OLEX2* : a complete structure solution, refinement and analysis program. *J. Appl. Crystallogr. 42*, 339–341.