Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Control of Stereoselectivity in Diverse Hapalindole Metabolites is Mediated by Cofactor-Induced Combinatorial Pairing of Stig Cyclases**

*Shasha Li, Sean A. Newmister, Andrew N. Lowell, Jiachen Zi, Callie R. Chappell, Fengan Yu, Robert M. Hohlman, Jimmy Orjala, Robert M. Williams, and David H. Sherman**

ange_201913686_sm_miscellaneous_information.pdf

Supporting Information

Table of Contents

General materials and methods.

All NMR spectra were acquired on Varian 400, 600 and 700 MHz spectrometers. Proton and carbon signals are reported in parts per million (d) using residual solvent signals as an internal standard. The LC–MS analysis was performed on Shimadzu 2010 EV APCI spectrometer. Preparative-scale HPLC was performed using an Agilent Extend C18 10 mm 10 × 250 mm column, and a mobile phase of 70–90% acetonitrile in water over 28 min. High-resolution APCIMS spectra and protein mass spectrometry were obtained from an Agilent 6520 Q-TOF mass spectrometer equipped with an Agilent 1290 HPLC system at the University of Michigan core facility in the Department of Chemistry, with MS grade solvents. Optical rotations were obtained using a Jasco P2000 polarimeter at 23°C.

Escherichia coli strain DH5a (Invitrogen) was used for plasmid manipulation, and BL21(DE3/pRARE) for protein expression. KOD Xtreme Hot Start DNA polymerase (EMD Millipore) was used for polymerase chain reactions. Restriction endonucleases (NheI, XhoI, and BamHI) and T4 DNA ligase were purchased from New England BioLabs. Primers were purchased from Integrated DNA Technologies. PureLink Quick Plasmid Miniprep Kit (Invitrogen) was used to prepare plasmid DNA. All cloned plasmids were confirmed by Sanger sequencing at the University of Michigan DNA Sequencing Core. Isopropyl-D-thiogalactopyranoside (IPTG; GoldBio) was used to induce expression; benzonase and lysozyme used in purification were purchased from Sigma-Aldrich; phenylmethane sulfonyl fluoride (PMSF) was dissolved in isopropanol and used as serine protease inhibitor during protein purification. Ni-NTA agarose from Invitrogen was used to purify 6× His-tag proteins. Terrific broth, lysogeny broth and agar (EMD Millipore) were used for all *E. coli* culturing.

Protein preparation.

Vector pET28 (Novagen) was used to build the expression construct for all cyclases with their N-terminal leading peptides truncated.[1] Except *famC1/famC2/famC3/famC4* and *hpiC1*, which were cloned from codon optimized synthetic genes (IDT gBlocks), all others were cloned from genomic DNA: *filC2/filC3/filC4* from IL-199-3-1 genomic DNA, *wepC1/wepC2* from SAG 16.93 genomic DNA, *fimC1/fimC4* from UTEX 1829 genomic DNA, *hpiC2/hpiC3/hpiC4* from ATCC 43239 genomic DNA. *Westiellopsis prolifica* SAG 16.93 was obtained from Professor Jimmy Orjala (University of Illinois at Chicago). The genomic DNA preparation was described in our previous report.[2] Site-directed mutagenesis of FamC1, FamC2, FamC3, FamC4 and HawC3 was performed using a single primer method based on QuikChange mutagenesis (Agilent Genomics). All primers used are listed in **Table S1**.

Cyclase production and purification procedures were described previously,[1-2] and the protein sequences are shown in **Figure S3**.

Chemical reagents.

The synthesis of indole isonitrile (**Figure S1**) was described previously,[2] geranyl pyrophosphate (GPP) was purchased from supplier http://Isoprenoids.com/ (purity > 95%). Authentic hapalindole H (**4a**) and 12-*epi*-hapalindole C (**3b**) were isolated from enzymatic reaction described in Ref. [1], authentic 12-*epi*-hapalindole H (**5a**) used in **Figure S8** was obtained from enzymatic reaction described in this report (FilC2-FilC3), authentic hapalindole U (**2a**) and hapalindole C (**2b**) were provided by Prof. Phil Baran lab.

***In vitro* enzymatic assays.**

All Stig cyclase activity assays were conducted in a 50 µL scale and incubated at 37 °C for 3 h. These reactions were extracted twice with ethyl acetate, the organic layers were dried under N_2 gas and re-dissolved in 50 µL of acetonitrile for LC-MS or HPLC analysis. The reaction conditions for each experiment (see figures) are described in detail below or along with each of the SI figures. For structural characterization of enzymatic products, the reactions were scaled up to 5 ml and incubated under identical conditions. The extracted products were purified by HPLC equipped with an Agilent Extend C18 10 mm 10 × 250 mm column, and a mobile phase gradient of 70–90% acetonitrile in water over 28 min. The products were dissolved in C_6D_6 or $CDCl_3$ for NMR analysis.

Figure 2 and 3. The experimental conditions are described along with Figure S4.

Figure 4. The experimental conditions are described along with Figure S12.

Table S1. Primers used in this study (5′ – 3′).

Genes	Accession No.	Forward primer	Reverse primer
filC2	*KY026488.1*	TGCTGCTAGCGCAGTTACTACTTCCATTC	GTAACTCGAGTTAGGTATCAACGGTTTCTG
filC3	*KY026488.1*	TGCTGGATCCACAGGTGCTGTTTCTATTCC	ATACCTCGAGCTAAATTACAGCCGATTCAAC
filC3ΔNSVDIG		GTTGATATTGGCAATTCAGCAGACGGCTACAAAG	TGAATTGCCAATATCAACGAGCAACGTGTAATTTTTG
wepC1	MT024692	CTTCTGAGAATCTCTACTTCCAAGGCGCTAGCAAAGGTGCTGTTTCTATTCCGATAAAC	GATCTCAGTGGTGGTGGTGGTGGTGCTCGAGTTAGGTCTCAGTTGGTTCTGCCGTT
wepC1-D214A		CCTCTTACAAGACAAATTTTCTGGTCTTGCGTTTGACAATGTACGCTTAACGGC	
wepC2	MT024693	CTTCTGAGAATCTCTACTTCCAAGGCGCTAGCGCAAATGTGATTTCAATTCC	GATCTCAGTGGTGGTGGTGGTGGTGCTCGAGTTAAATATTGGTAGGTTGTGC
wepC2-D215A		AATCTTTTACAAAGTTCATCTGCAAATATTGCGTTCGATAATGTCCGCCTAATTGC	
famC4	KX451322.1	AAAGCTAGCGCGAACGTTATCCCGATTC	AAAAAACTCGAGTTAGATGTTGGTCGGTTGC
fimC1	*KY026487.1*	AAAAGCTAGCACATCTGCTGTTTCCATTC	GTCCTCGAGTTAAGTCTCAGTGGGTTCTGTG
fimC4	*KY026487.1*	TGCTGCTAGCGCAAATGTGATTCAAATTCC	GTGACTCGAGTTAAATATTGGTAGGTTGTGCAG
filC4	*KY026488.1*	TGCTGCTAGCGCAAATTTGATTCCAATTCC	GTGACTCGAGTTAAATATTGGTAGGTTGTGC
famC1-D214A	KX451322.1	GCTGCAAGATAAGTTTAGCGGCCTG**gcg**TTTGATAATGTGCGCCTGACCACCG	
famC2-D217A	KX451322.1	GCAGGGTAAATTCAGCGGCCTG**gcg**TTTGATAACGTGCGTCTGATCACC	
famC3-D214A	KX451322.1	CTGCAAACCCTGAGCGGCAACATC**gcg**TTCGATAACGTGCGTCTGAGCGTTG	
famC4-D215A		CTGCAAAGCAGCAGCGCGAACATT**gcg**TTTGATAACGTGCGTCTGACCGCG	
famC1-EFQK$_{141-144}$		GCGGGCGAATTCCAGAAACTGAGCTTCGCGGGTTTTCCGGGCTATCGTG	GCTCAGTTTCTGGAATTCGCCCGCCAGGTTACCAACATCCACGGTCAG
hpiC1-TFKG$_{141-144}$	*KJ742064.1*	GTGGCACGTTTAAAGGTATTAGCCTGGCGGGTTTTCCGGGCTAC	CTAATACCTTTAAACGTGCCACCGAAGTTACCAACGTCCACTTTCAGGGTGTAC
hpiC2	*KJ742064.1*	GAGAATCTCTACTTCCAAGGCGCTAGCGCAGTTACTACTTCCATTCCCA	GTGGTGGTGGTGGTGGTGCTCGAGTTAGGTATCAACGGTTTCTGTAATTAAG
hpiC3	*KJ742064.1*	GACAGCAAATGGGTCGCGGATCCACAGGTGCTGTTTCTATTCCTATC	GTGGTGGTGGTGGTGGTGCTCGAGCTATTCAACACTTAAACGAACATTATC
hpiC4	*KJ742064.1*	GAGAATCTCTACTTCCAAGGCGCTAGCGCAAATTTGATTCCAATTCCCATC	GTGGTGGTGGTGGTGGTGCTCGAGTTAAATATTGGTAGGTTGTGCAGTTAG

Table S2. Summary of important combinations of Stig cyclases reported in this manuscript and their products.

Subunit 1	Subunit 2	Additive ($CaCl_2$)	Product 1 (major)	Product 2 (minor)	Product 3 (minor)
FamC2 or FilC2	–	–	n.r.		
		+	5a	3b	5b
FamC2	FamC3	–/+	4a		
FilC2	FilC3	–	n.r.		
		+	5a	3b	5b
FilC2	FamC3	+	4a (80%)	5a (20%)	
	FilC3ΔNSVDIG	+	5a	3b	
FamC2	FilC3	+	5a	3b	
	FilC3ΔNSVDIG	+	4a		
HpiC2	HpiC3	+	4a		
HpiC2	FilC3	+	5a		
WepC1	–	–	n.r.		
		+	3a	3b	
WepC1	WepC2	–	3a	3b	
		+	2a (83.3%)	2b (16.7%)	4a (trace)
FamC1	–	–/+	3a		
FamC1	FamC4	+	2a		
FilC1	FilC4	+	2a		
FimC1	FimC4	+	2a		
FamC1	FamC3	+	2a	3a	
FilC1	FilC3	+	3a	2a	
FamC1-EFQK$_{141-144}$	(–/+) FamC4	+	3a		
HpiC1	(–/+) HpiC4	+	3a	3b	
HpiC1-TFKG$_{141-144}$	–	+	3a	3b	
	HpiC4	+	2a	3a	3b
FamC3/FilC3/HpiC3/WepC2 /FamC4/FilC4/FimC4/HpiC4		–/+	n.r.		

n.r.: no reaction

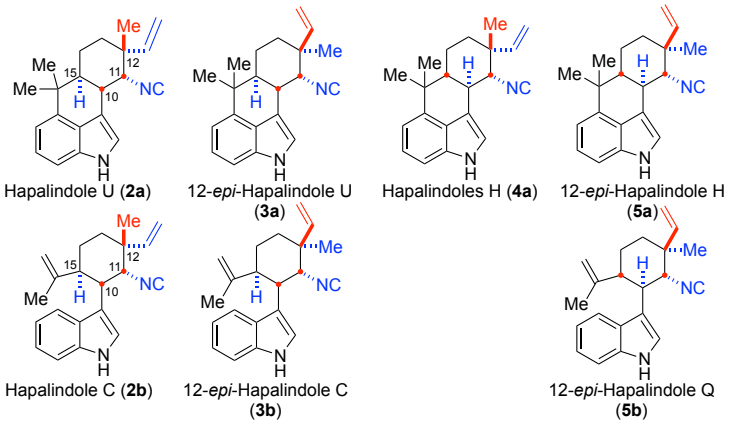


Figure S1. Proposed biogenesis of hapalindole-type alkaloids. The four subgroups, including hapalindoles, fischerindoles, ambiguines and welwitindolinones, are generated from the common units, geranyl pyrophosphate (GPP) and cis-indole isonitrile. Hapalindoles can undergo a C-2 prenylation reaction catalyzed by FamD1 to produce tetracyclic ambiguines, while fischerindoles are presumed to be converted to welwitindolinones through oxidative ring rearrangement.

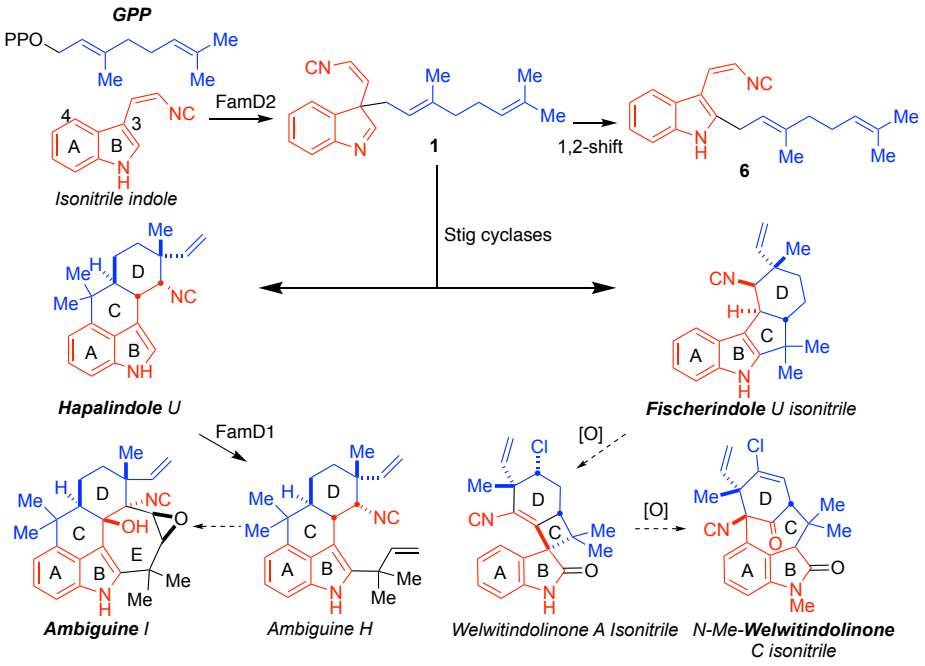
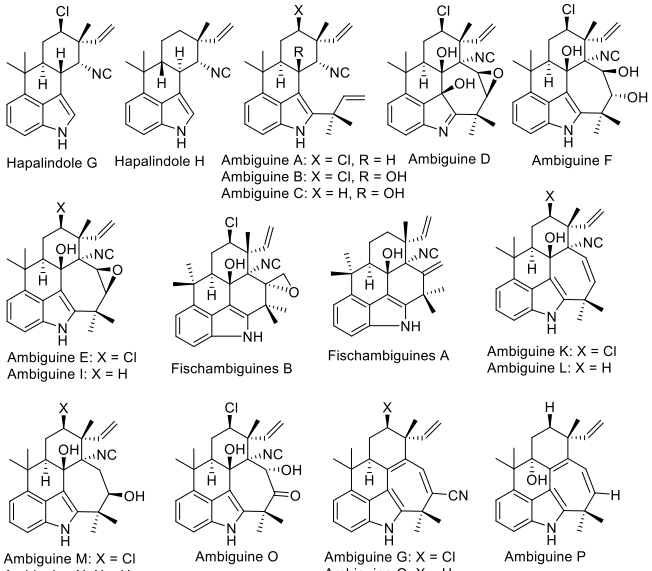
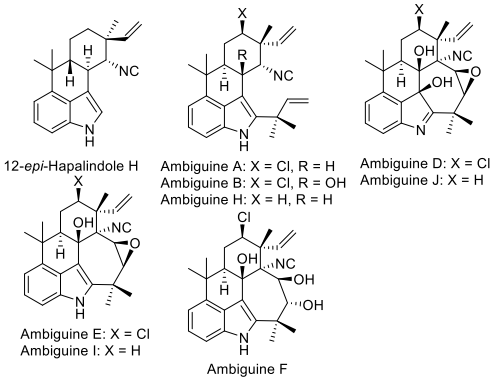


Figure S2. Isolated hapalindole-type alkaloids from *Fischerella ambigua* UTEX 1903, *Fischerella* sp.TAU IL-199-3-1 and *Westiellopsis prolifica* SAG 16.93.

***Fischerella ambigua* (UTEX 1903)**



***Fischerella sp.* (TAU IL-199-3-1)**



***Westiellopsis prolifica* SAG 16.93 (Unpublished results from Orjala Lab)**

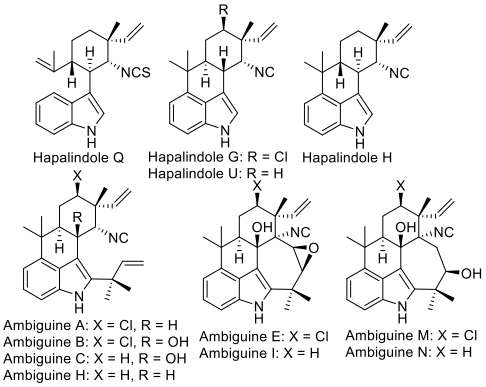


Figure S3. Protein sequence alignment of Stig cyclases analyzed in this report. All Stig cyclases including their corresponding mutants were overexpressed with the N-terminal transmembrane truncated (the red-arrow shows the position where the expression sequence starts). The Asp (D) residue shown by a blue-arrow is the 100% conserved aspartic acid across all identified Stig cyclases which was mutated to inactivate the cyclases.

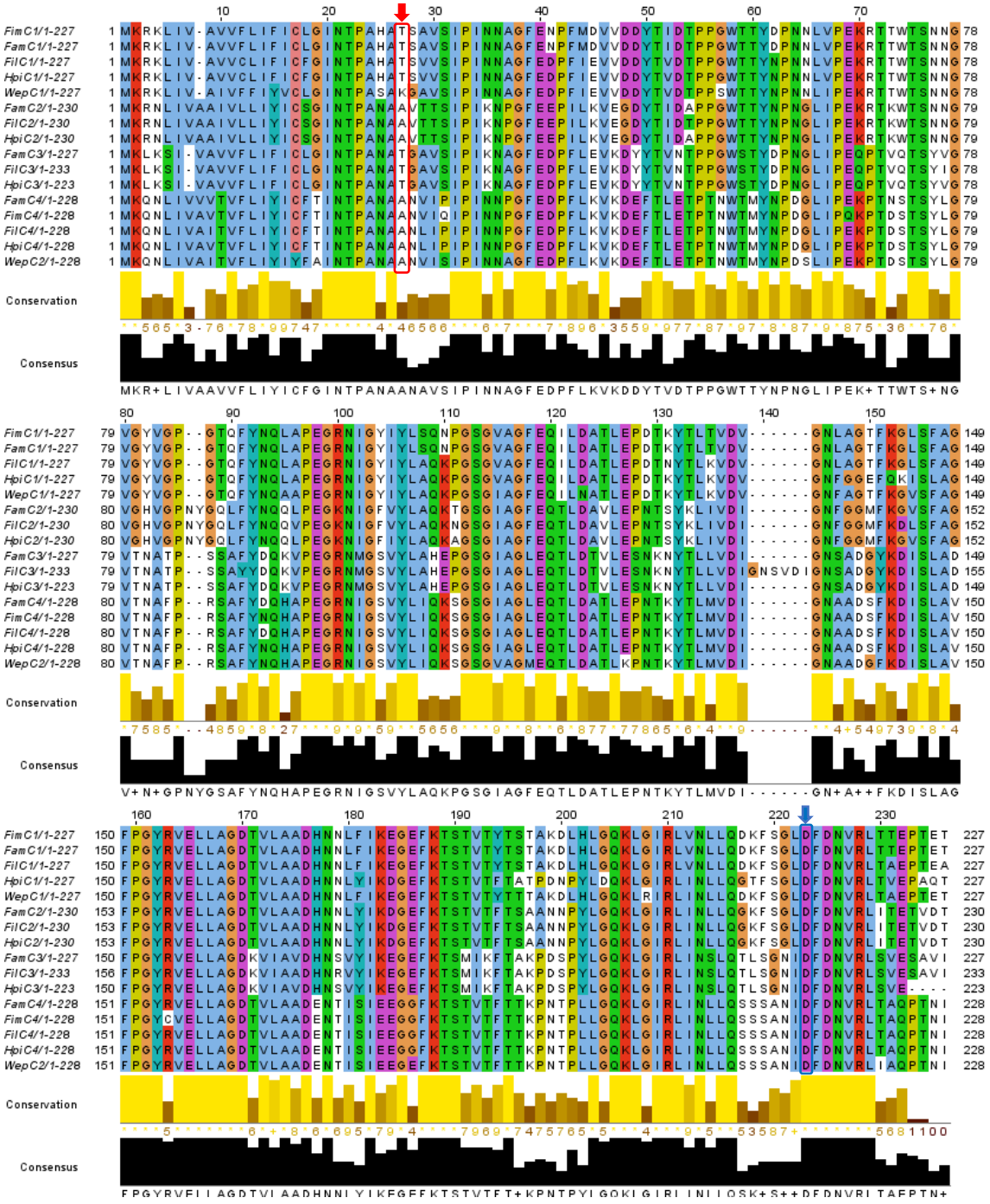


Figure S4. Complete HPLC traces of in vitro assays to identify the key cofactor for activating FilC2 and FilC3. The assay was conducted in 50 µL scale containing 50 mM Tris buffer (pH 7.8), 1 mM of **1**, 5 µM of each Stig cyclase, 5 mM of the respective metal ion cofactor or 10 µL of cell-free lysate or 10 µL of BG11 additives. The reaction was incubated at 37 °C and quenched as described in general assay conditions. The method of LC-MS analysis is: Extend C18 5 mm 4.6 × 150 mm column, and mobile phase of 70–90% gradient acetonitrile in water over 22 min with flowrate of 0.4 mL/min. Important exogenous components include: lysate, cell-free lysate of *Fischerella* sp.TAU IL 199-3-1[2]; BG-11, a metal-rich media commonly used in cyanobacterial cultivation; the metal components from the BG-11 media.

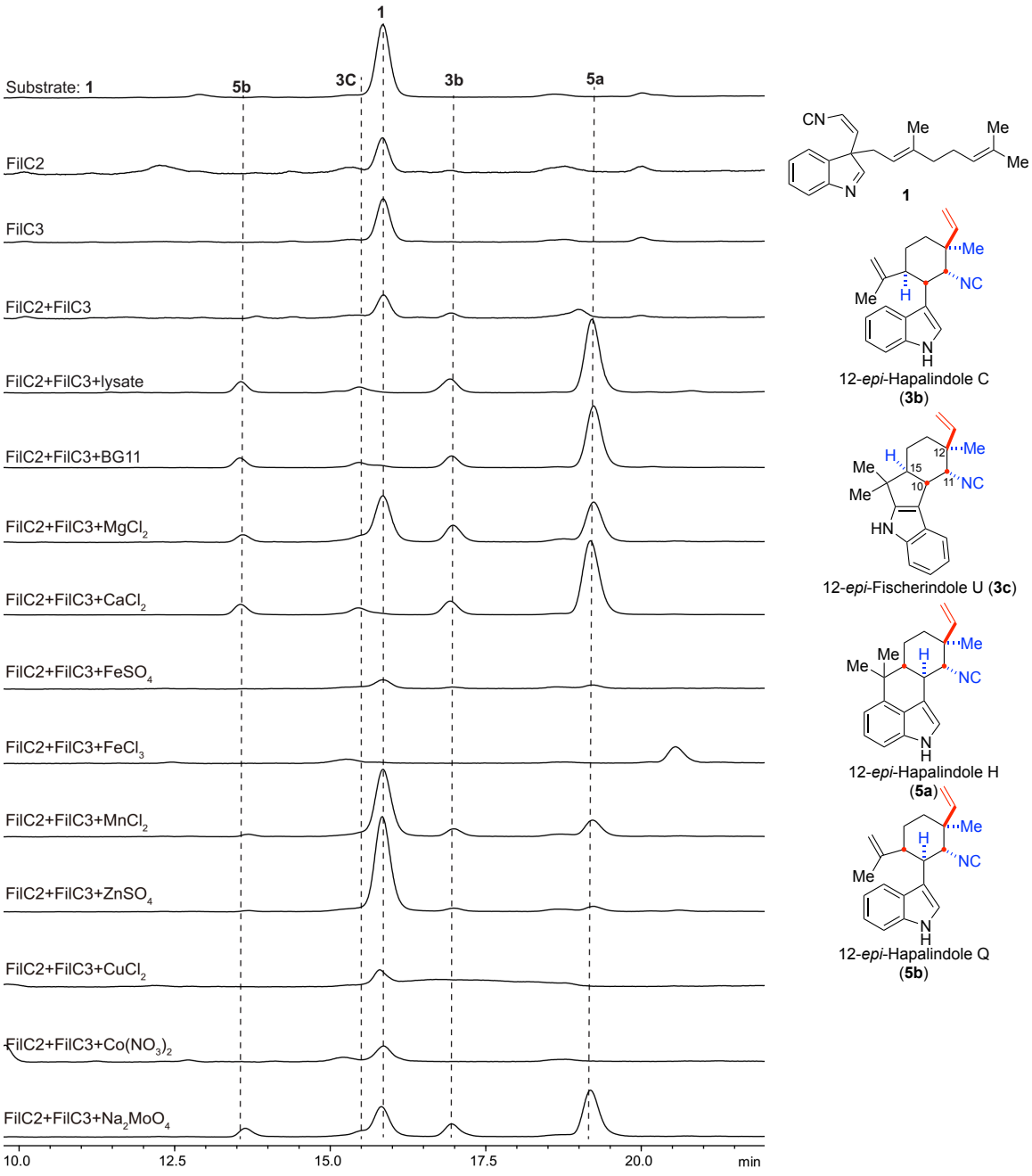


Figure S5. ^{1}H NMR spectrum overlay of 12-*epi*-hapalindole H (**5a**) with literature data in $CDCl_3$.

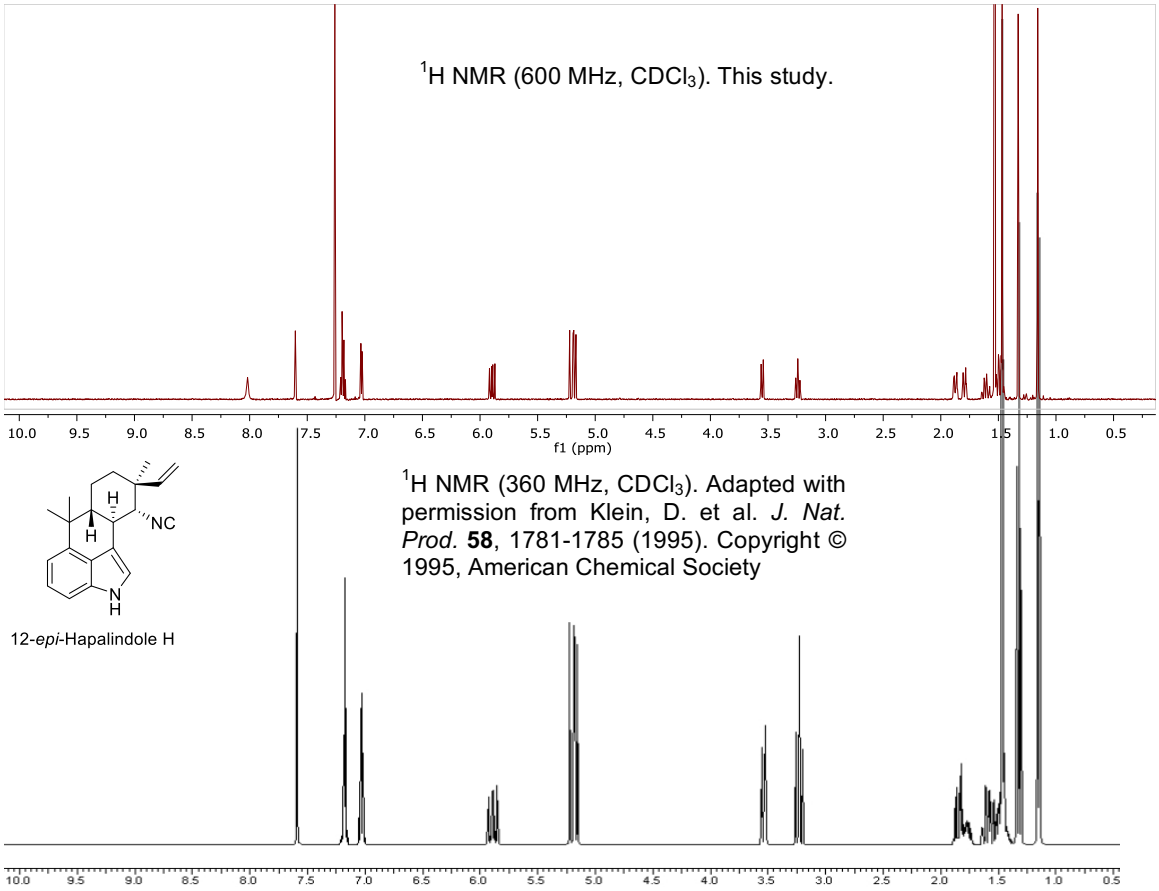


Table S3. The complete NMR spectroscopic data table for 12-*epi*-hapalindole H (**5a**) in benzene-d_6.

Position	δ_{13C}	δ_{1H} (multi, *J* value)	COSY	HMBC	NOESY
1	N	6.74 (bs)			2
2	118.7	7.56 (s)		3, 8, 9, 10	1, 10, 11
3	113.2				
4	140.9				
5	112.8	7.06 (d, *J* = 7.1 Hz)	6	6, 7, 9, 16	17, 18
6	123.2	7.27 (t, *J* = 7.3 Hz)	5,7	4, 5, 7, 8	
7	108.6	6.96 (d, *J* = 8.0 Hz)	6	5, 6, 8, 9	
8	134.0				
9	125.8				
10	36.3	3.09 (t, *J* = 11.1 Hz)	11, 15	2, 3, 9, 11, 12, 14, 15, 16	14ax, 17, 19
11	65.5	3.22 (d, *J* = 11.1 Hz)	10	3, 10, 12, 13, 19, 20, 23	13ax, 14eq, 15
12	40.2				
13	35.9	(Heq) 1.35-1.30 (1H)	14	11, 12, 14, 15, 19	
		(Hax) 0.96-0.91 (1H)			
14	20.6	(Hax) 1.35-1.30	13, 15	12, 13, 15	
		(Heq) 1.15-1.08			
15	49.6	1.04-0.98 (m)	10, 14		11
16	37.4				
17	24.9	0.97 (s)		4, 15, 16, 17	
18	24.8	1.28 (s)		4, 15, 16, 18	
19	17.0	1.10 (s)		11, 12, 13, 20	10
20	146.0	5.73 (dd, *J* = 10.9, 17.5 Hz)	21	11, 12, 13, 19	
21	113.4	(cis) 5.07 (d, *J* = 10.7 Hz)	20	11, 12, 19, 20	
		(trans) 5.09 (d, *J* = 17.3 Hz)			
23	160.8				

Figure S6. ^{1}H NMR spectrum of 12-*epi*-hapalindole Q (**5b**) produced by FilC2-FilC3 heterodimer and data overlay (identical solvents).

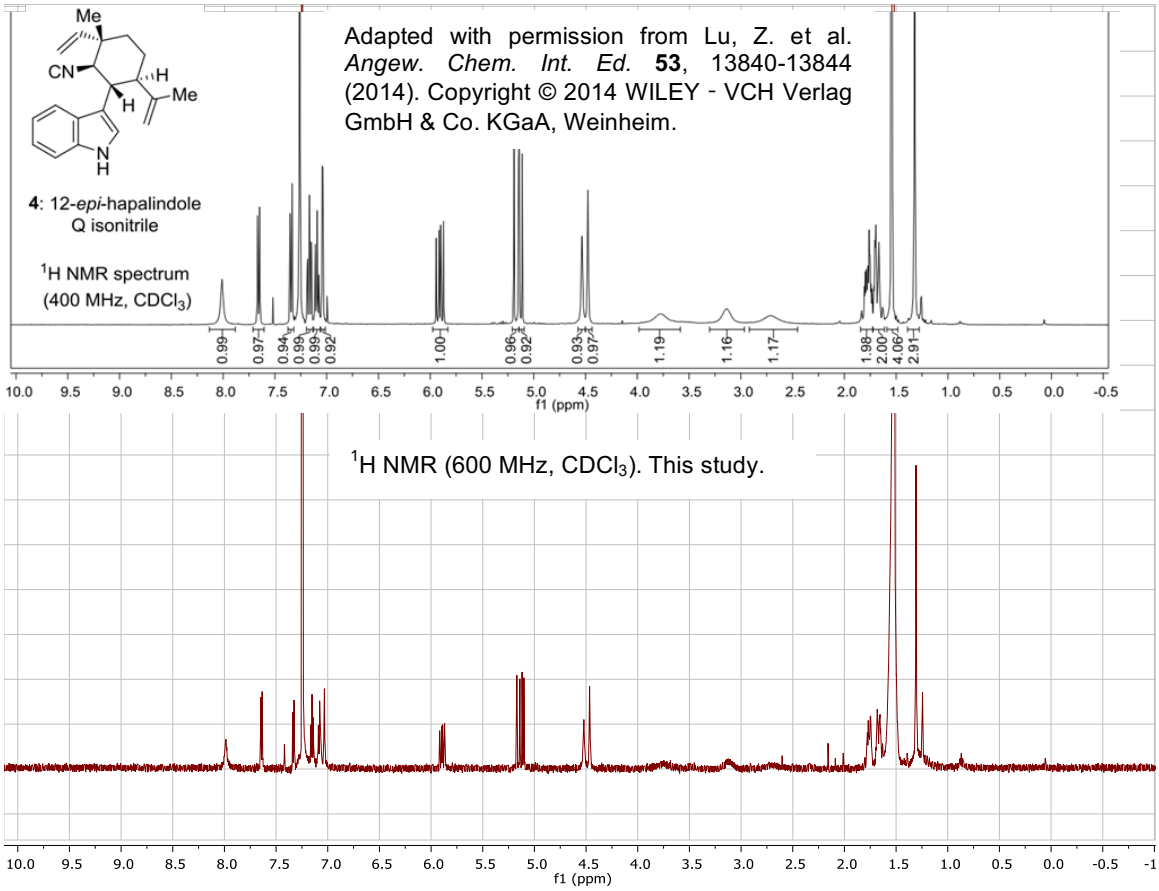


Figure S7. Complete HPLC traces for FamC2, FamC3, FilC2, and FilC3 in vitro assay to assess the effect of supplemental Ca^{2+} on metabolite stereochemical configuration. **a)**. The assay was conducted in 50 µL scale containing 50 mM Tris buffer (pH 7.8), 1 mM of **1**, 5 µM of each Stig cyclase, with or without 5 mM of $CaCl_2$, and incubated at 37 °C. The reaction was quenched as described in the general assay conditions and analyzed by HPLC with method: Extend C18 5 mm 4.6 × 150 mm column, and a mobile phase gradient of 50–72% acetonitrile in water over 23 min with flowrate 1.5 mL/min. Peak **7** was not identified due to low yield and no authentic standard. **b**). The assay was conducted as a one-pot reaction in 50 µL scale containing 50 mM Tris buffer (pH 7.8), 1 mM of cis-indole isonitrile, 1.2 mM of geranyl pyrophosphate (GPP), 5 µM of FamD2, 5 mM of $MgCl_2$, 15 µM of each Stig cyclase, with or without 5 mM of $CaCl_2$, and incubated at 37 °C. The method of LC-MS analysis is: Extend C18 5 mm 4.6 × 150 mm column, and mobile phase of 70–90% gradient acetonitrile in water over 22 min with flowrate of 0.4 mL/min.

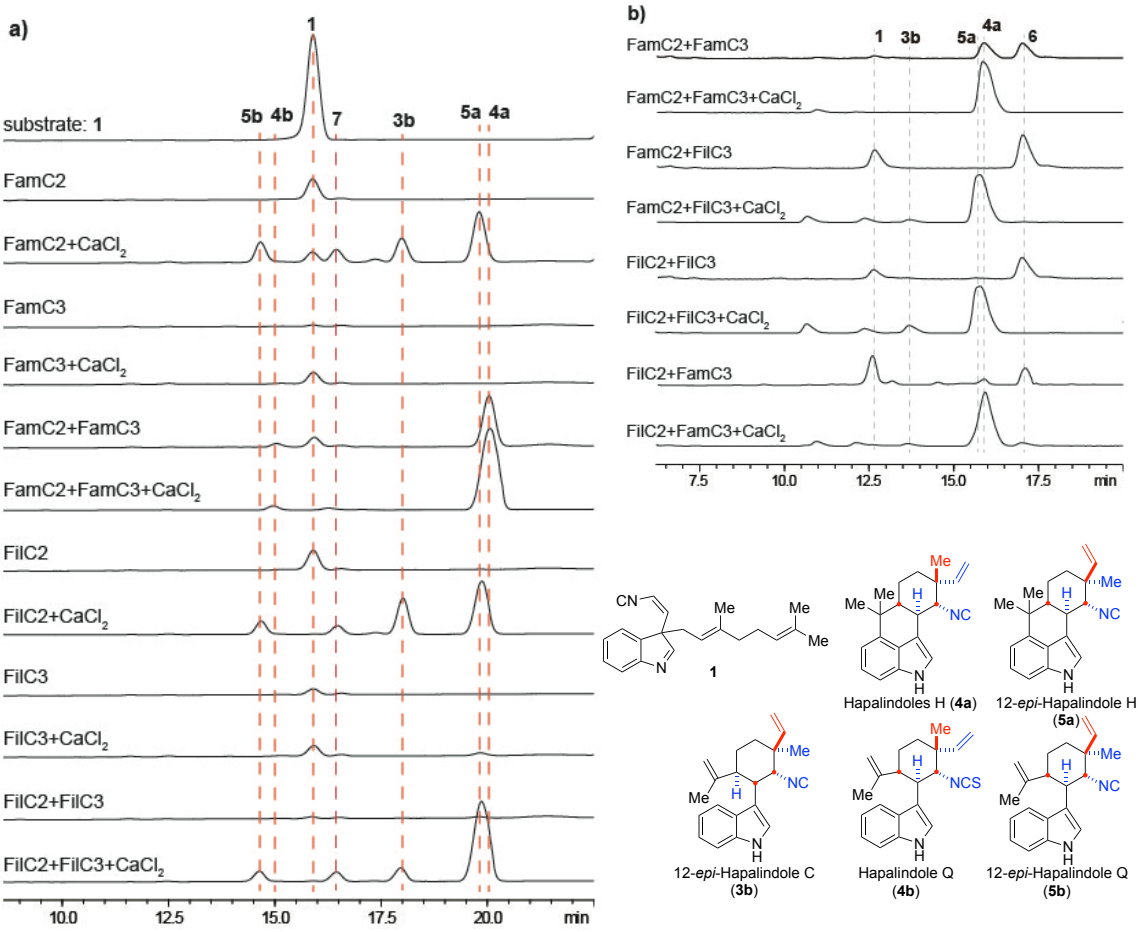


Figure S8. ^{1}H NMR spectrum of FilC2-FamC3 major product, which is a mixture of hapalindole H **(4a)**; $[\alpha]_D^{23}$= +45.8 (c=0.08, CH_2Cl_2) and 12-*epi*-hapalindole H **(5a)** with a ratio 4:1. The spectra were analyzed in benzene-d_6.

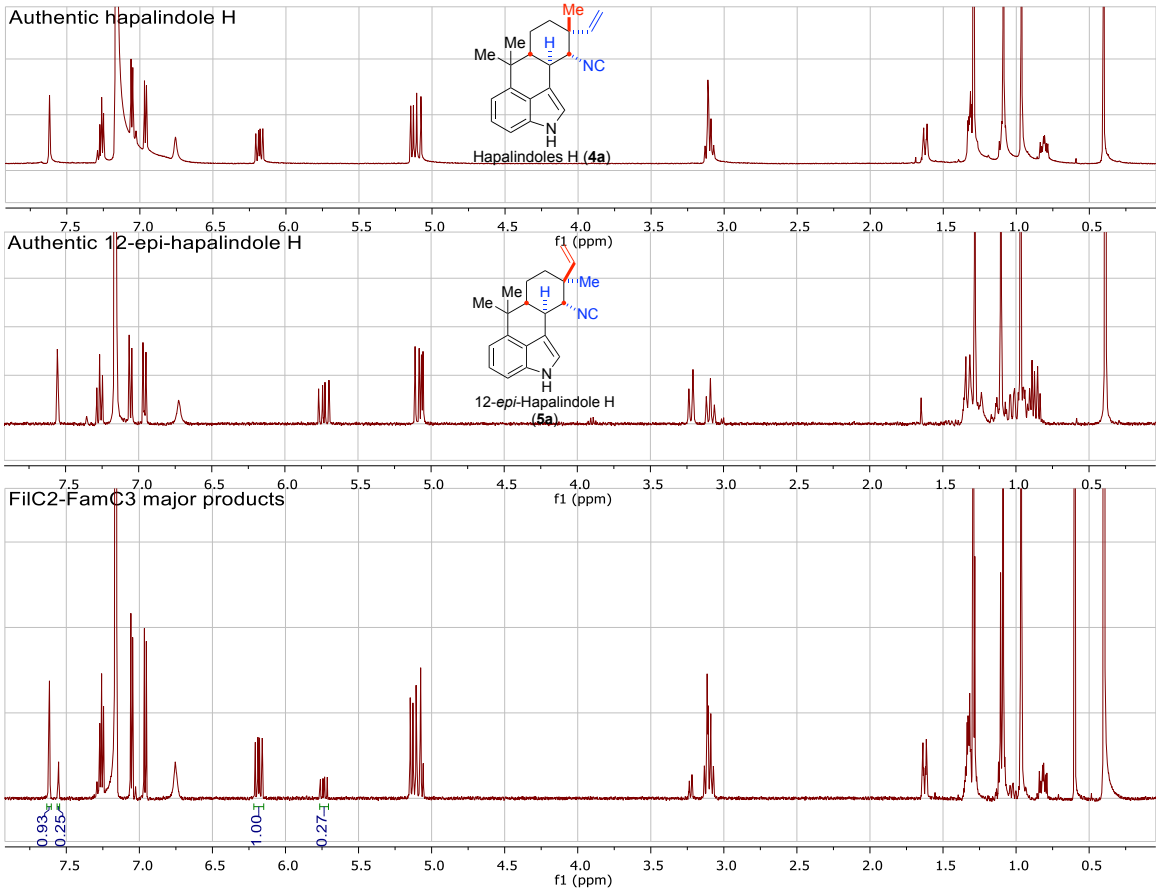


Figure S9. ^{1}H NMR spectrum of one of the FilC2-FamC3 minor product, by comparing to authentic hapalindole C (**2b**) and 12-*epi*-hapalindole C (**3b**) in benzene-d_6, this molecule was identified as hapalindole C (**2b**).

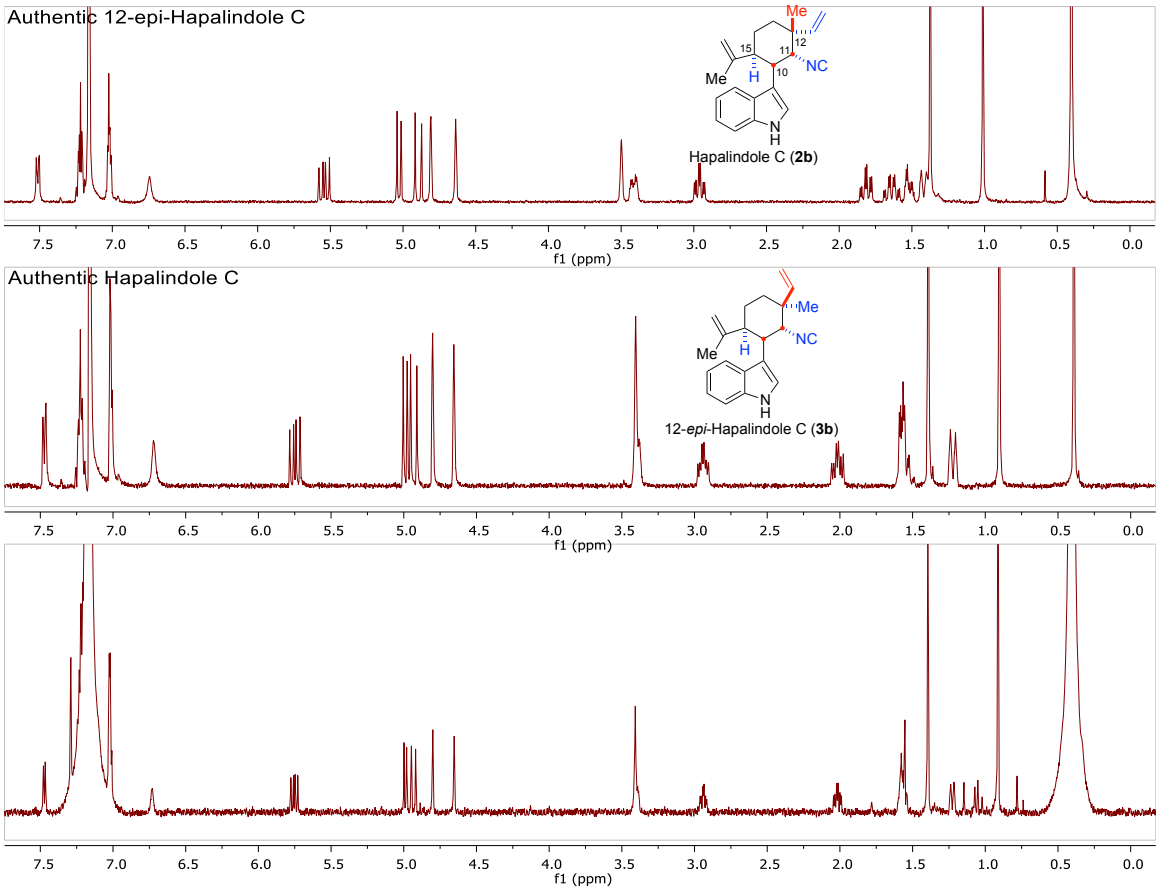


Figure S10. HPLC traces of HpiC2, HpiC3 in vitro assay. The assay was conducted as a one-pot reaction in 50 µL scale containing 50 mM MES buffer (pH 6.5), 1 mM of cis-indole isonitrile, 1.2 mM of geranyl pyrophosphate (GPP), 5 µM of FamD2, 5 mM of $MgCl_2$, 10 µM of each Stig cyclase, 5 mM of $CaCl_2$, and incubated at 30 °C. The reaction was quenched as described using the general assay conditions and analyzed on HPLC with method: XBridge Shield RP18 3.5 µm, 3.0 x 150mm column, and a mobile phase gradient of 50–72% acetonitrile in water over 16 min with flowrate 0.8 mL/min.

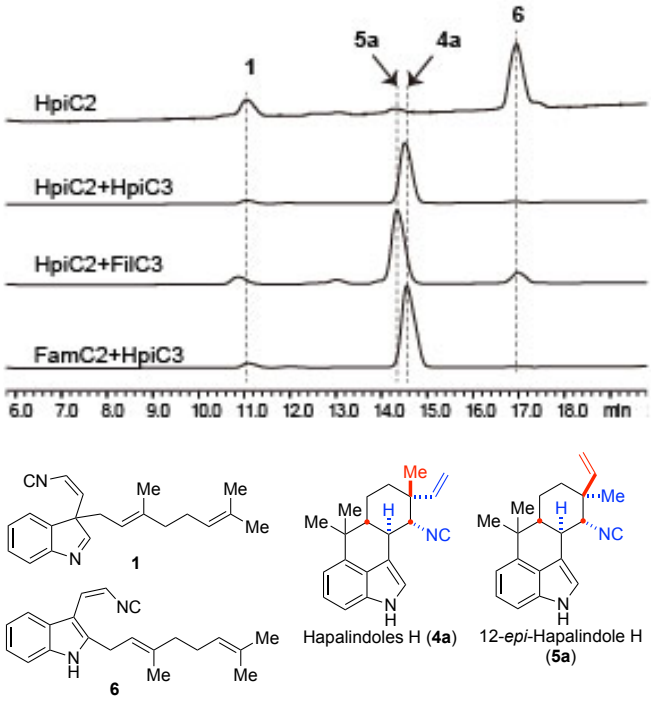


Figure S11. ^{1}H NMR spectrum of the major product from the WepC1-WepC2 reaction, which is a mixture of hapalindole U (**2a**); $[\alpha]_D^{23}$= +4.4 (c=0.15, CH_2Cl_2) and hapalindole C (**2b**) (ratio 5:1).

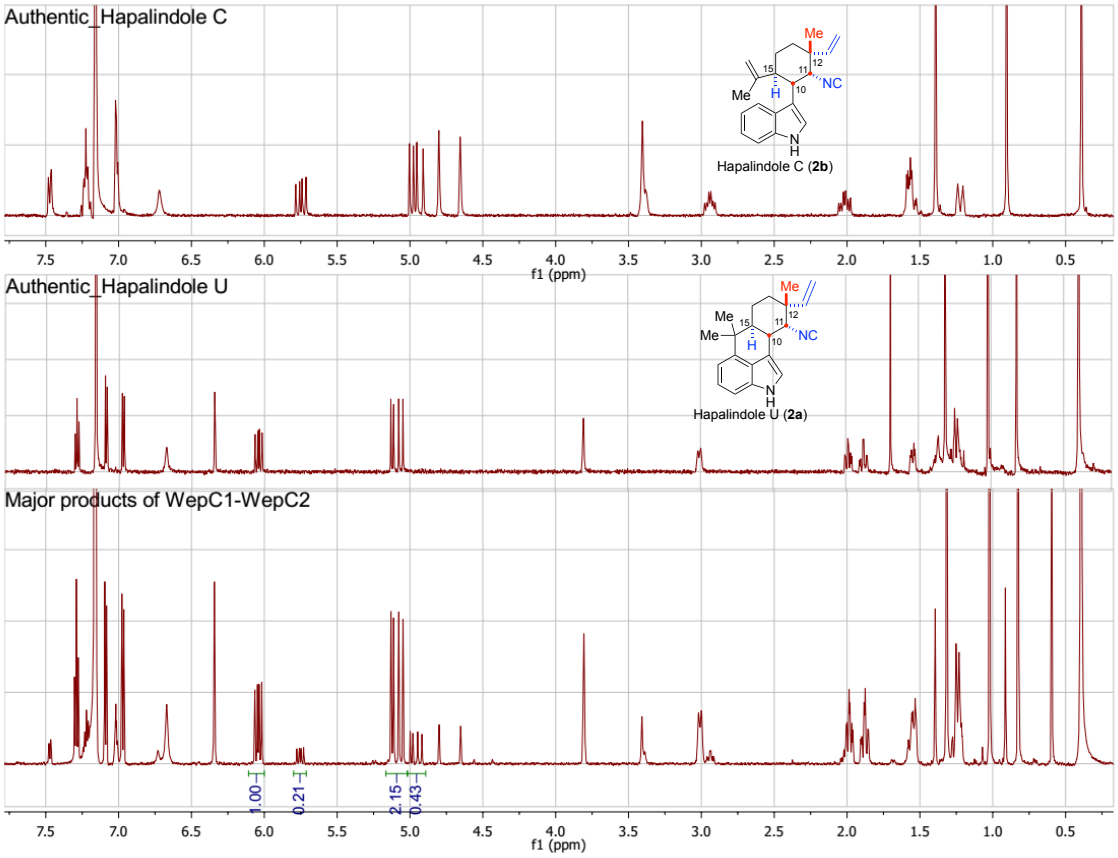


Figure S12. Complete HPLC traces of in vitro reactions using WepC1-WepC2 homologs. The products were obtained from a one-pot reaction (50 µL scale) containing 50 mM Tris buffer (pH 7.8), 1 mM of cis-indole isonitrile, 1.2 mM of geranyl pyrophosphate (GPP), 5 µM of FamD2, 5 mM of $MgCl_2$, 10 µM of each Stig cyclase, 5 mM of $CaCl_2$, and incubated at 30 °C. The reaction was quenched as described under general assay conditions and analyzed on HPLC with method: XBridge Shield RP18 3.5 µm, 3.0 x 150mm column, and a mobile phase gradient of 50–72% acetonitrile in water over 16 min with flowrate of 0.8 mL/min.

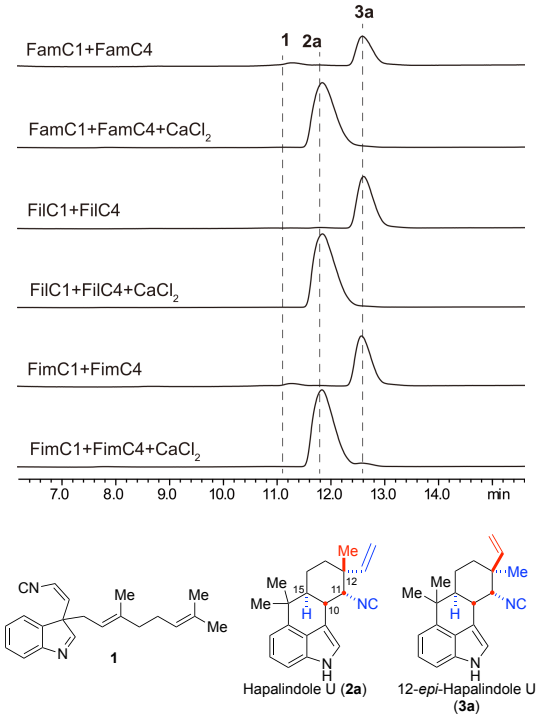


Figure S13. Calcium titration assay. The experiment was conducted as a one-pot reaction (50 µL scale) containing 50 mM MES buffer (pH 6.5), 1 mM of cis-indole isonitrile, 1.2 mM of geranyl pyrophosphate (GPP), 5 µM of FamD2, 5 mM of $MgCl_2$, 10 µM of each Stig cyclase, variant concentrations of $CaCl_2$, and incubated at 30 °C. The reaction was quenched as described under general assay conditions and analyzed on HPLC with method: XBridge Shield RP18 3.5 µm, 3.0 x 150mm column, and a mobile phase gradient of 50–72% acetonitrile in water over 16 min with a flowrate 0.8 mL/min. A) The product was gradually switched from **3a** to **2a** by increasing Ca^{2+} concentration. B) WepC1 was activated by increasing Ca^{2+} concentration, and the reaction was gradually switched to hapalindoles **3a**/**3b** instead of the side product **8**, which will be generated if no active Stig cyclase is presented. When the UV area absorption of the desired product peak reached maximum amplitude, the corresponding Ca^{2+} concentration was considered optimal for Stig cyclase activity.

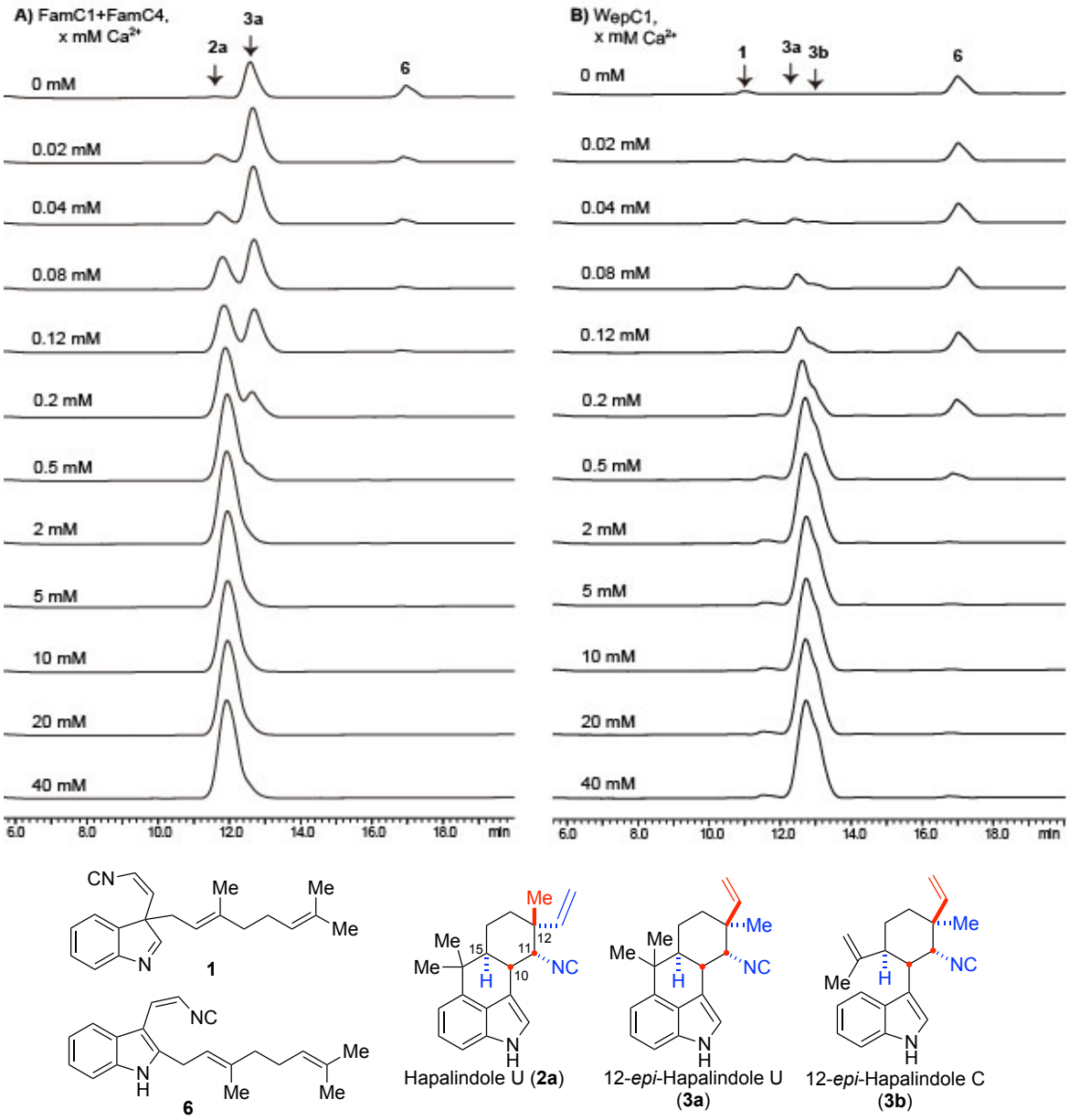
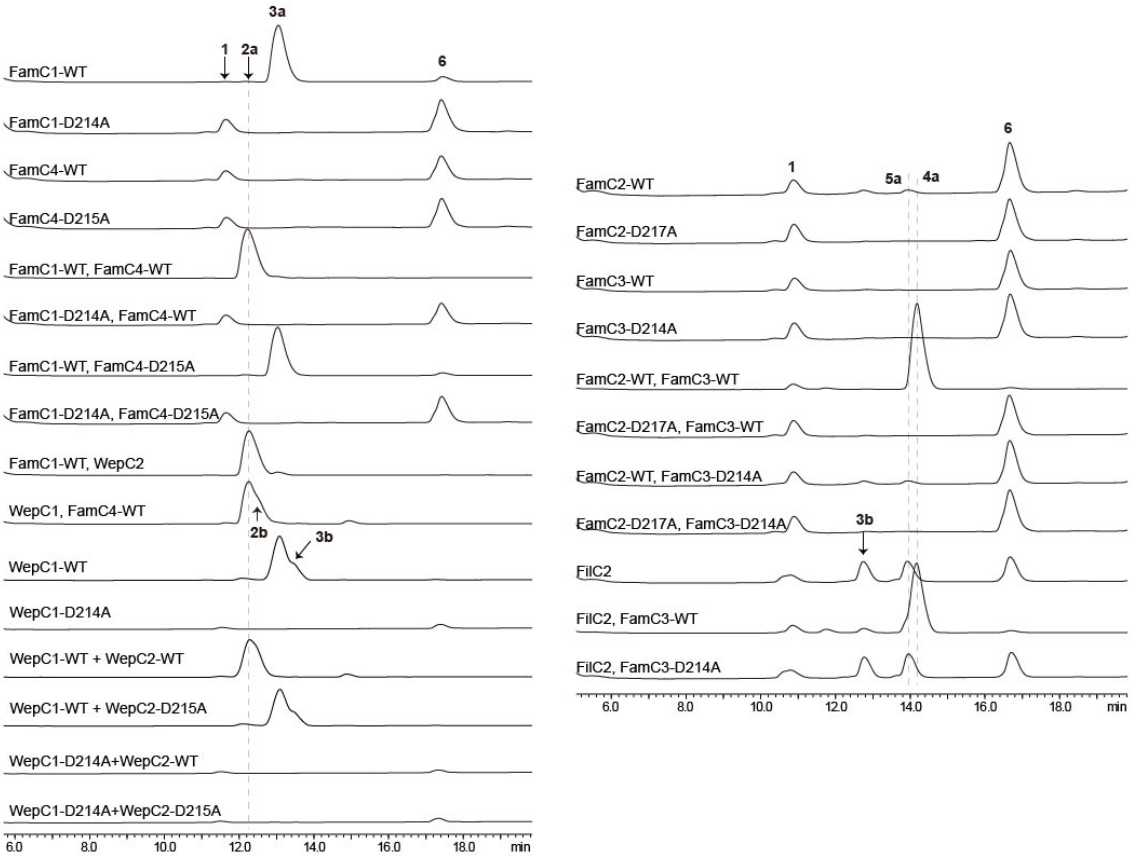


Figure S14. Complete HPLC traces of in vitro assay for FamC-type cyclase mutants. The assay was conducted as a one-pot reaction (50 µL scale) containing 50 mM MES buffer (pH 6.5), 1 mM of cis-indole isonitrile, 1.2 mM of geranyl pyrophosphate (GPP), 5 µM of FamD2, 5 mM of $MgCl_2$, 10 µM of each Stig cyclase, and 5 mM of $CaCl_2$, and incubated at 37 °C. The reaction was quenched as described in the general assay conditions and analyzed on HPLC with method: XBridge Shield RP18 3.5 µm, 3.0 x 150mm column, and a mobile phase gradient of 50–72% acetonitrile in water over 16 min with a flowrate 0.8 mL/min.



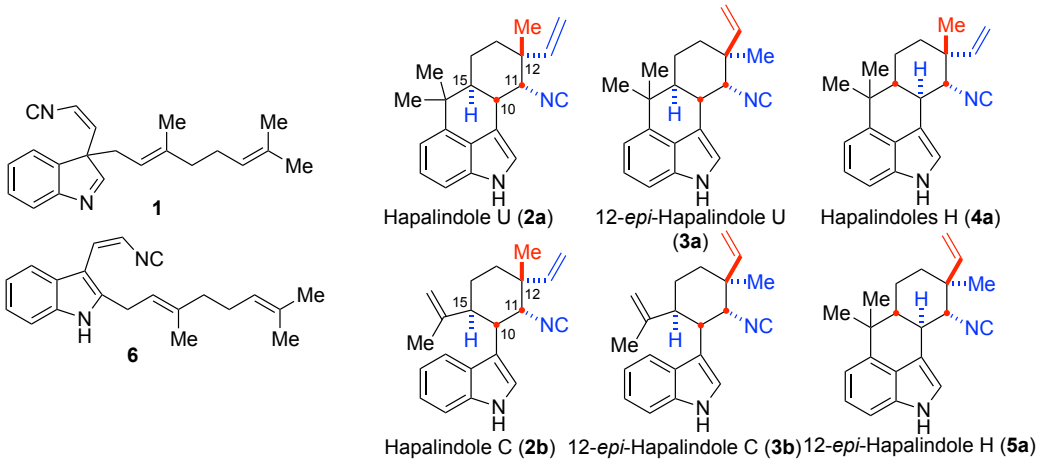
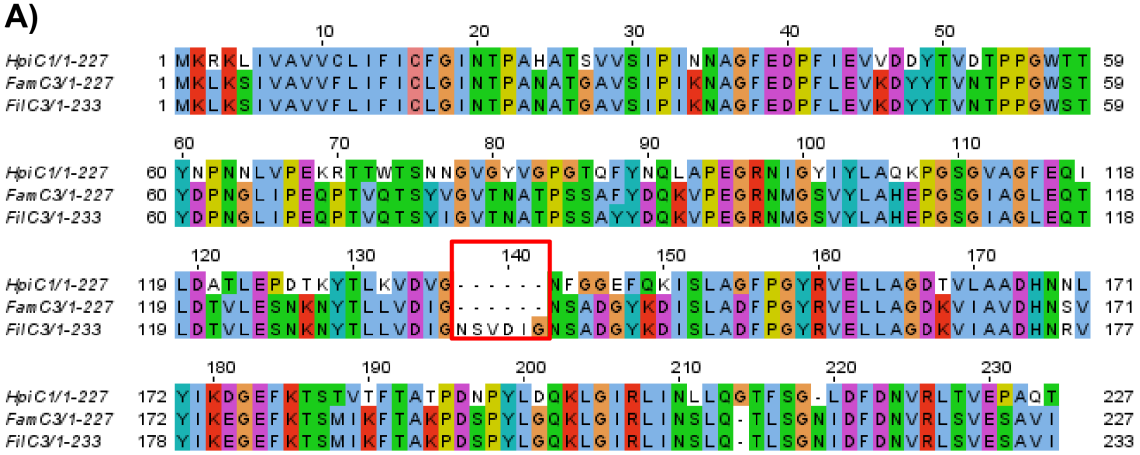


Figure S15. Protein sequence of FilC3. A) Compared to HpiC1 and FamC3, and other Stig cyclases, FilC3 contains a six amino-acid insertion (NSVDIG) highlighted in the red-box; B) Protein structure of HpiC1. The FilC3-insertion is located between HpiC1-G136/N137, which is very close to the HpiC1 active site and plays a role in reactivity and hapalindole product profile.

A)



B)

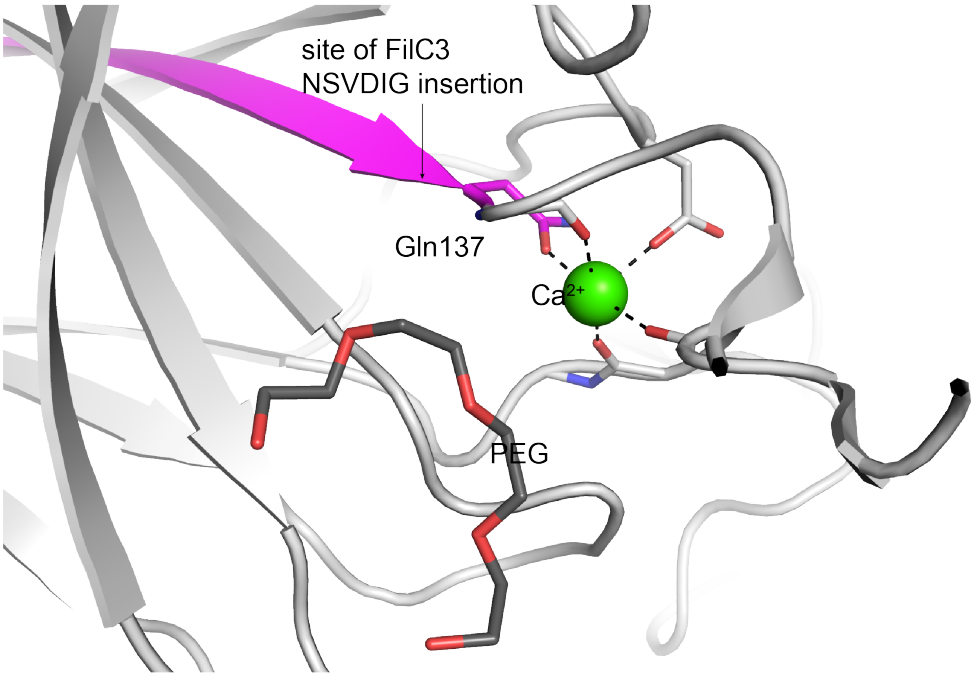
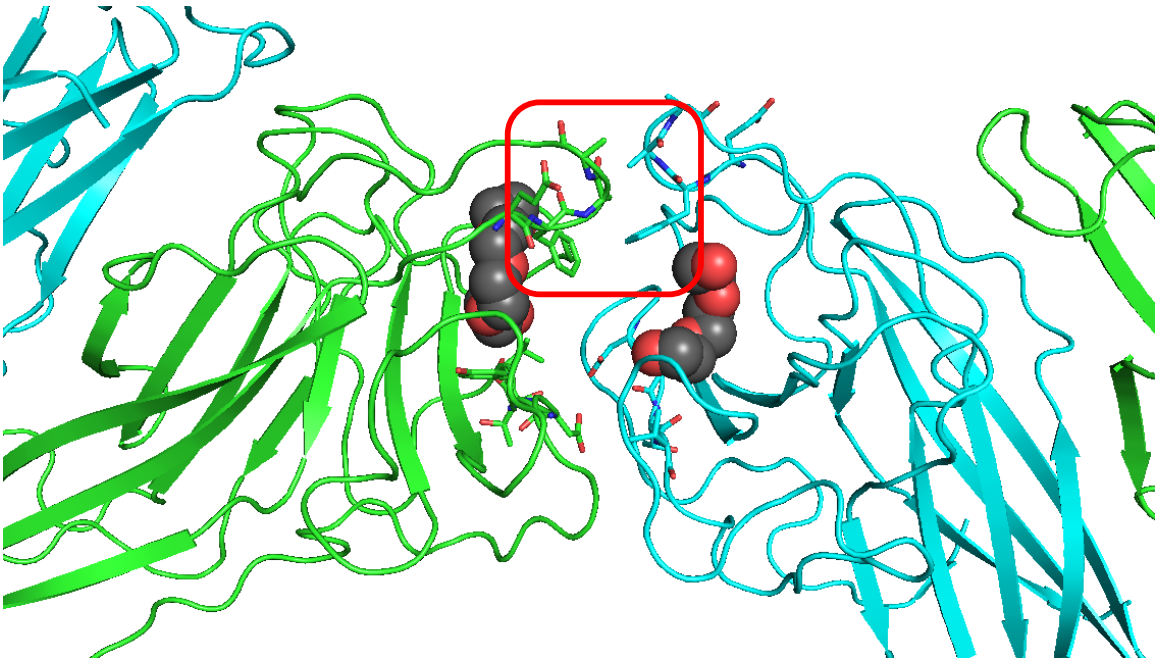


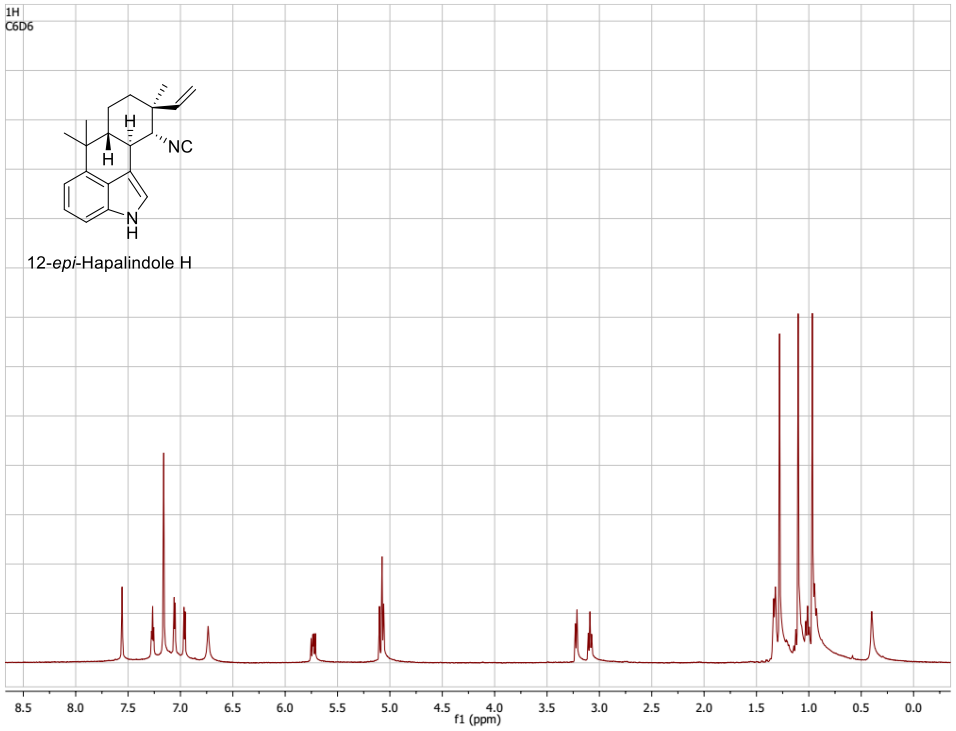
Figure S16. Oligomeric structure of HpiC1 indicating the amino acid residues involved in protein-protein interactions (in red box). Sequence alignment among the four Stig cyclases shows that HpiC4 and FamC4 have the same residues, while HpiC1 and FamC1 are different from each other. A bound polyethylene glycol is shown in space-fill representation.

Interface residues:	
HpiC1	EFQK
FamC1	TFKG
HpiC4	SKDI
FamC4	SKDI

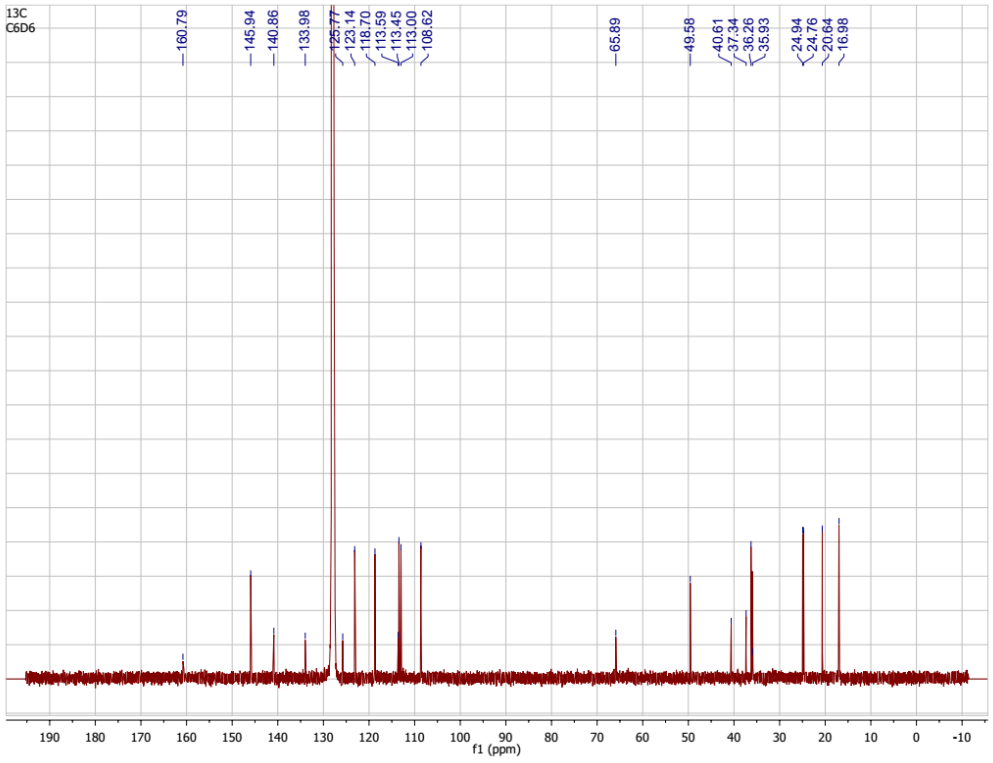


SPECTRA SECTION.

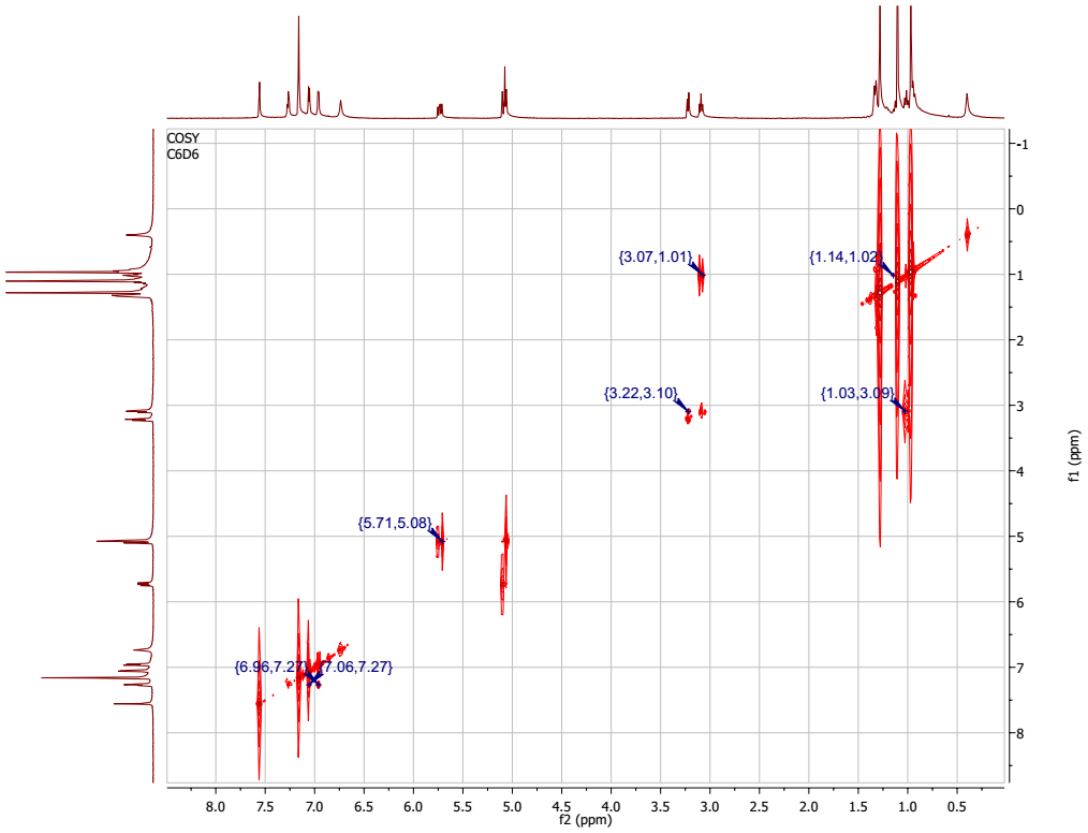
^{1}H NMR spectrum of 12-*epi*-hapalindole H (**5a**) in benzene-d_6 at 700 MHz.



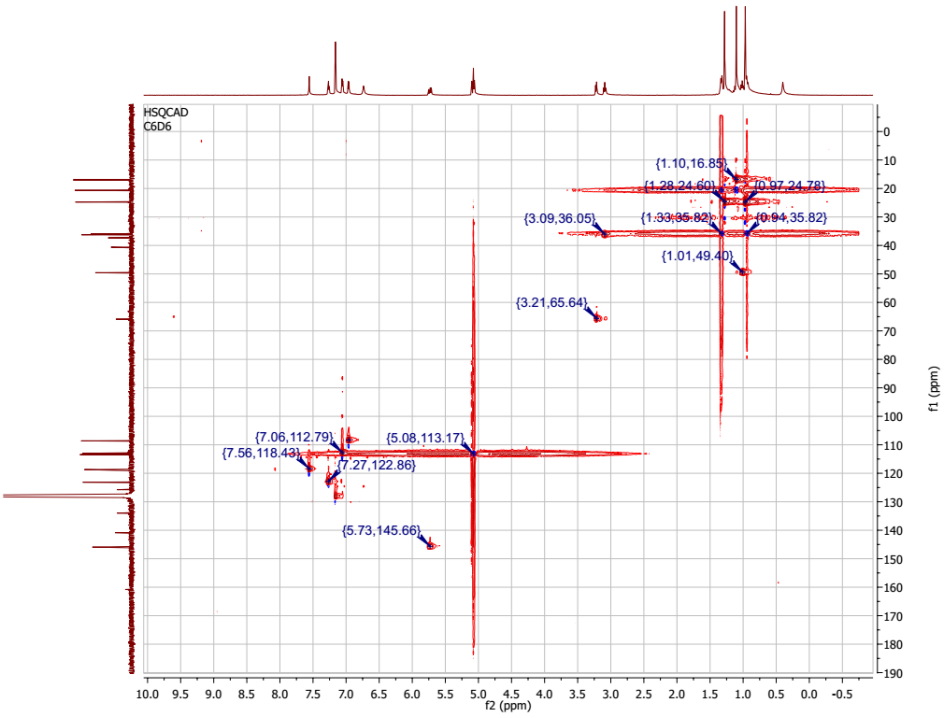
^{13}C NMR spectrum in benzene-d_6 at 176 MHz.



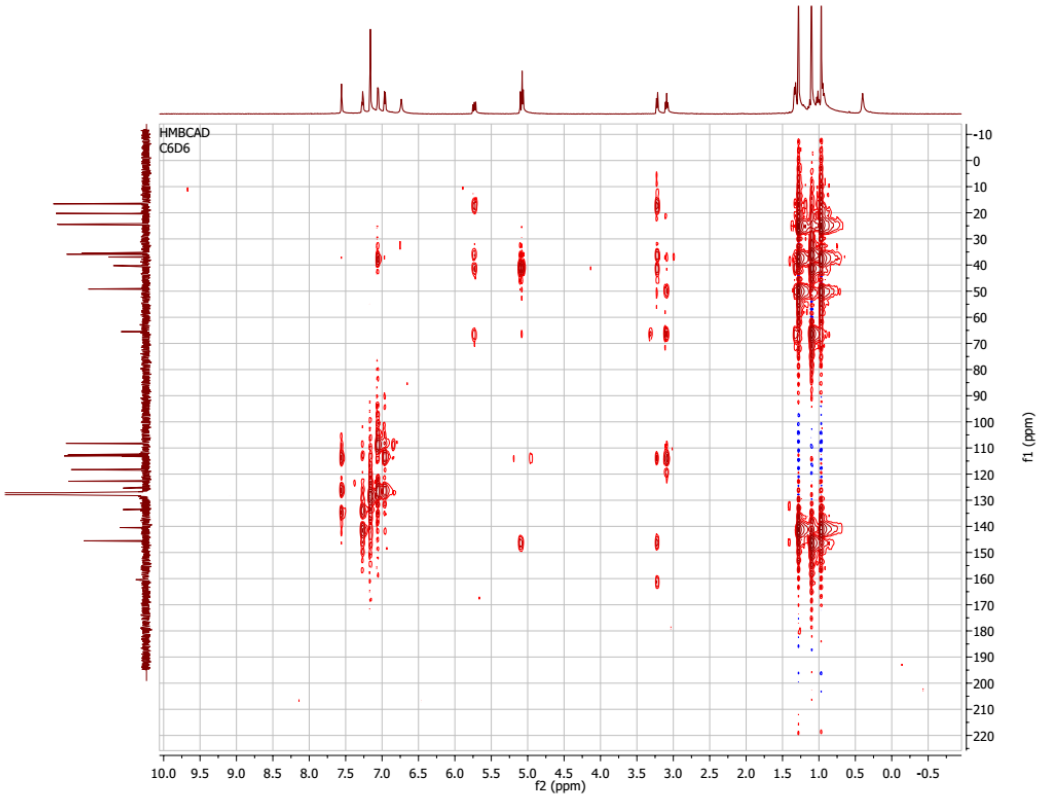
COSY spectrum in benzene-d_6 at (700, 700) MHz.



HSQC spectrum in benzene-d_6 at (700, 176) MHz.



HMBC spectrum in benzene-d_6 at (700, 176) MHz.



References.

[1] S. Li, A. N. Lowell, S. A. Newmister, F. Yu, R. M. Williams, D. H. Sherman, *Nat. Chem. Biol.* **2017**, *13*, 467-469.

[2] S. Li, A. N. Lowell, F. Yu, A. Raveh, S. A. Newmister, N. Bair, J. M. Schaub, R. M. Williams, D. H. Sherman, *J. Am. Chem. Soc.* **2015**, *137*, 15366-15369.