

Running title: *Aspergillus fumigatus* anthraquinone biosynthesis

Genome-based cluster deletion reveals endocrocin biosynthetic pathway in *Aspergillus fumigatus*

**Fang Yun Lim¹, Yanpeng Hou², Yiming Chen³, Jee-Hwan Oh⁴, Inhyung Lee⁴, Tim S Bugni²,
Nancy P Keller^{2,3,*}**

¹Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA

²Pharmaceutical Sciences Division, University of Wisconsin-Madison, Madison, WI, USA

³Department of Medical Microbiology and Immunology, University of Wisconsin-Madison, Madison, WI, USA

⁴Department of Advanced Fermentation Fusion Science and Technology, Kookmin University, Seoul 136-702, Korea

Keywords: endocrocin, anthraquinone, non-reducing polyketide synthase, LaeA, metallo- β -lactamase thioesterase, emodin, asperthecin

***Corresponding author**

Professor Nancy P. Keller

Department of Bacteriology

University of Wisconsin-Madison

1550 Linden Drive

Madison WI, U.S.A.

SUPPLEMENTAL TABLE

Table S1. Primers used in this study.

Primer Name	Primer Sequence (5'-sequence-3')	Purpose
OE Afu4g00210 5' Flank For	GCCTTGGAATAGAGCCATGTAC	<i>OE::encA</i> 5' Flank
OE Afu4g00210 5' Flank Rev	CAATTCGCCCTATAGTGAGTCGTATTAC GGTTGAATGTAGATCAGGTCAAC	
OE Afu4g00210 3' Flank For	GCTACCCCGCTTGAGCAGACATCACCAT GCAAGGCCCAAGCCAATTG	<i>OE::encA</i> 3' Flank
OE Afu4g00210 3' Flank Rev	CGTGTCGGTATGACTCTCTAC	
FY 4G00210 INTKO 5'F FOR	CCAATTGGCCCTGTTCTAC	<i>AencA</i> 5' Flank
FY 4G00210 INTKO 5'F REV	CAGCTTATCGATGATAAGCTGTCAAACA TGGGCTGGTTCGATGAAGCAGCC	
FY 4G00210 INTKO 3'F FOR	CGTGTTGATAGCACACCCTCGGAATAGT CACGCAGGCGGGCGATTCAAC	<i>AencA</i> 3' Flank
FY 4G00210 INTKO 3'F REV	GACGGTCTTTGCGTCTACC	
FY Extended T7 FOR	CGTAATACGACTCACTATAGGG	<i>gpdA(p)::pyrG</i> for OE cassette
FY <i>gpdA(p)</i> Fusion REV	CATGGTGATGTCTGCTCAAG	
FY A. para <i>pyrG</i> 2KB FOR	CATGTTTGACAGCTTATCATCG	<i>pyrG</i> for deletion cassette
FY A. para <i>pyrG</i> 2KB REV	GACTATTCCGAGGGTGTGCTAT	
FY 4G00220 5'F FOR	CCAGACCTTGCAGAGGATC	<i>AencB</i> 5' Flank
FY 4G00220 5'F REV	CTGAGATCCATAGGATCAGCTTATCGAT GCTATGGCCTGAGTTCAGCGG	
FY 4G00220 3'F FOR	CGTGTTGATAGCACACCCTCGGAATAGT CCGCCAATGTGAGCCAAGAGAA	<i>AencB</i> 3' Flank
FY 4G00220 3'F REV	TTCAGGTGCGTCGATTCC	
FY 4G00220 NEST FOR	GATGTCGTTGCTGTGCTCG	<i>AencB</i> deletion cassette
FY 4G00220 NEST REV	GCATGGGATCTGTCTGCA	
FY 4G00225 5'F FOR	ATCACAGATCGCCACTGC	<i>AencC</i> 5' Flank whole ORF deletion
FY 4G00225 5'F REV	CTGAGATCCATAGGATCAGCTTATCGAT GACATGACGAACAA CCTACGGC	
FY 4G00225 3'F FOR	CGTGTTGATAGCACACCCTCGGAATAGT CGCTAGACATCGAGGTAGTGGT	<i>AencC</i> 3' Flank whole ORF deletion
FY 4G00225 3'F REV	CCAGTATGACCAAGCCAC	
FY 4G00225 NEST FOR	CCGAGCGAGATGCAAGGA	<i>AencC</i> whole ORF deletion cassette
FY 4G00225 NEST REV	CTATGGCCATGACAGCCAAG	
FY 4G00225INT 5'F FOR	GTCTATGGCCTGAGTTCAG	<i>AencC</i> 5' Flank internal deletion
FY 4G00225INT 5'F REV	CTGAGATCCATAGGATCAGCTTATCGAT GCACAGAGACCAGGCTGTGTGCTG	
FY 4G00225INT 3'F FOR	ACGTGTTGATAGCACACCCTCGGAATAG	<i>AencC</i> 3' Flank

	TCAGCATGGTGTGCTTCCGATCGC	internal deletion
FY 4G00225INT 3'F REV	GGGATCGCTCCAAGATCG	
FY 4G00230 5'F FOR	GAGGTTCTGGAGTGGCTAT	<i>AencD</i> 5' Flank
FY 4G00230 5'F REV	CTGAGATCCATAGGATCAGCTTATCGAT GGCTAGCTTATGAGTGCCTGCA	
FY 4G00230 3'F FOR	CGTGTGATAGCACACCCTCGGAATAGT CGGTCCTTGCCTAAGTAGACC	<i>AencD</i> 3' Flank
FY 4G00230 3'F REV	ATCGATCAGCAGCCGTTT	<i>AencD</i> deletion cassette
FY 4G00230 NEST FOR	TGTTGGACTCTCGTTCGC	
FY 4G00230 NEST REV	CAGGTTGCAAGAGCCAT	
FY 4G00200B 5'F FOR	GGTTCAACCTAGTGGACT	Δ AFUA_4G00200 5' Flank
FY 4G00200B 5'F REV	CTGAGATCCATAGGATCAGCTTATCGAT GACAATGGCGACTGGATGC	
FY 4G00200B 3'F FOR	CGTGTGATAGCACACCCTCGGAATAGT CGAGTGCGGACGAGGAATTCTA	Δ AFUA_4G00200 3' Flank
FY 4G00200B 3'F REV	CGCGTAGTATGGCTCAAC	Δ AFUA_4G00200 deletion cassette
FY 4G00200B NEST FOR	CATATCCATCTCGCAAACCGC	
FY 4G00200B NEST REV	ACAATGGCGACTGGATGC	
FY 4G00240 5'F FOR	CGTCCATACAGATGCGGT	Δ AFUA_4G00240 5' Flank
FY 4G00240 5'F REV	CTGAGATCCATAGGATCAGCTTATCGAT GCATGCAAGGTTAGCCCGAAGTG	
FY 4G00240 3'F FOR	ACGTGTTGATAGCACACCCTCGGAATAG TCGGACCGGAGAGATCCAAAGATC	Δ AFUA_4G00240 3' Flank
FY 4G00240 3'F REV	CGCTCCTAAGAAATGAGCAC	Δ AFUA_4G00240 deletion cassette
FY 4G00240 NEST FOR	CGGCCTTCTAACACGCCA	
FY 4G00240 NEST REV	GTGTGTCCCGTGAGTCAG	
FY 4G00250 5'F FOR	GGTAGGTGTAGAGAGCAGG	Δ AFUA_4G00250 5' Flank
FY 4G00250 5'F REV	CTGAGATCCATAGGATCAGCTTATCGAT GAGGAGAGGATCCCGTGTCGAAC	
FY 4G00250 3'F FOR	ACGTGTTGATAGCACACCCTCGGAATAG TCGATCTTAGTCAAGCCTCCGGTA	Δ AFUA_4G00250 3' Flank
FY 4G00250 3'F REV	CGCTCCTGTTACTGACGG	Δ AFUA_4G00250 deletion cassette
FY 4G00250 NEST FOR	GCAATCATCTAGCCCGCG	
FY 4G00250 NEST REV	GCGGTGACAGGGAATAAC	
FY 4G00210 INTB FOR	CAAGAGTTGGAACCATAACCTC	Northern probe of <i>encA</i>
FY 4G00210 INTB REV	GCTGCAGATTGATGATGTCAG	Northern probe of <i>encB</i>
FY 4G00220 INT FOR	GGACCAACACATACATTCTGGGC	
FY 4G00220 INT REV	CTGCAGCGCAAGTAGCACTTG	Northern probe of <i>encC</i>
FY 4G00225 INT FOR	TACAGCACACAGCCTGGTCT	
FY 4G00225 INT REV	CTCCAGTGACTGAGCAGCTC	Northern probe of <i>encD</i>
FY 4G00230 INT FOR	GACCAAGCCACAGCCACC	
FY 4G00230 INT REV	CTCTCTCGACGGTGCCTC	
FY 4G00200 INT FOR	CGACTATGACGAGGAGGAAGTC	Northern probe of F box protein
FY 4G00200 INT REV	CCATCTCGCTCTTAAGGAAGGG	
FY 4G00240 INTB FOR	GCCTGCGTACGAGCCTTCTT	PCR screen of Δ AFUA_4G00240
FY 4G00240 INTB REV	CGGCTCTCTTCGATGCGC	

FY 4G00250 INT FOR	CGCCGATATTCTCGAGTCTG	PCR screen of Δ AFUA_4G00250
FY 4G00250 INT REV	GGATGGCAATTCCTGCTTC	
FY <i>gpdA</i> int FOR	GAAGGGTGGTGCCAAGAAG	Northern probe of <i>gpdA</i> , multiplex PCR internal control
FY <i>gpdA</i> int REV	CAACGGAGACGTTGGAGGT	
FY <i>fumi laeA</i> Int FOR	CCTGACAGATACCCTTGC	Northern probe of <i>laeA</i>
FY <i>fumi laeA</i> Int REV	GCCTCCGATTTAACATCGGC	
FY <i>encD</i> comp (NotI) FOR	TATAGCGGCCGCTCTTGAGGAACGTCTG CCTGC	EncD complementation insert
FY <i>encD</i> comp (XbaI) REV	TATATCTAGAGCTGCTCAGTCACTGGAG	
FY <i>pyrG::gpdAp::encD</i> 5'FOR	GTGGCTGTTGCAGGAATG	OE:: <i>encD</i> 5' Flank
FY <i>pyrG::gpdAp::encD</i> 5'REV	CAATTCGCCCTATAGTGAGTCGTATTAC GACTGGAATAGTATAACTATTTTCTTG	
FY <i>pyrG::gpdAp::encD</i> 3'FOR	GCTACCCCGCTTGAGCAGACATCACCAT GACCAAGCCACAGCCACCT	OE:: <i>encD</i> 5' Flank
FY <i>pyrG::gpdAp::encD</i> 3'REV	GCGCTAGACATCGAGGTAGTG	

SUPPLEMENTAL FIGURES

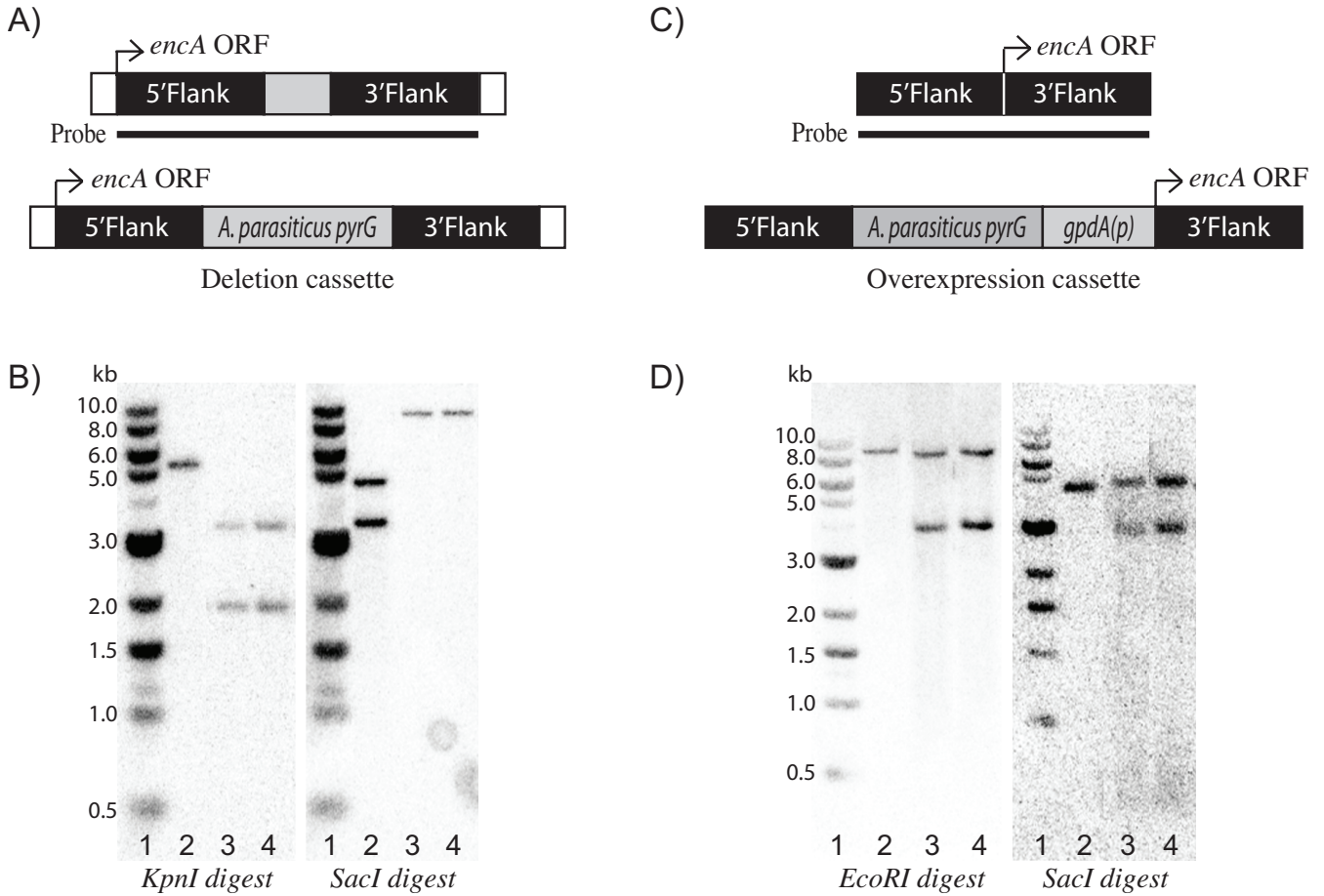


Figure S1. Construction of $\Delta encA$ and overexpression *encA* mutants. A) Deletion construct. B) Southern analysis of $\Delta encA$ transformants. Lane 1 (Ladder), Lane 2 (Parental strain), Lane 3 (Transformant 1), and Lane 4 (Transformant 2). C) Overexpression construct. D) Southern analysis of overexpression *encA* transformants Lane 1 (Ladder), Lane 2 (Parental strain), Lane 3 (Transformant 1), and Lane 4 (Transformant 2).

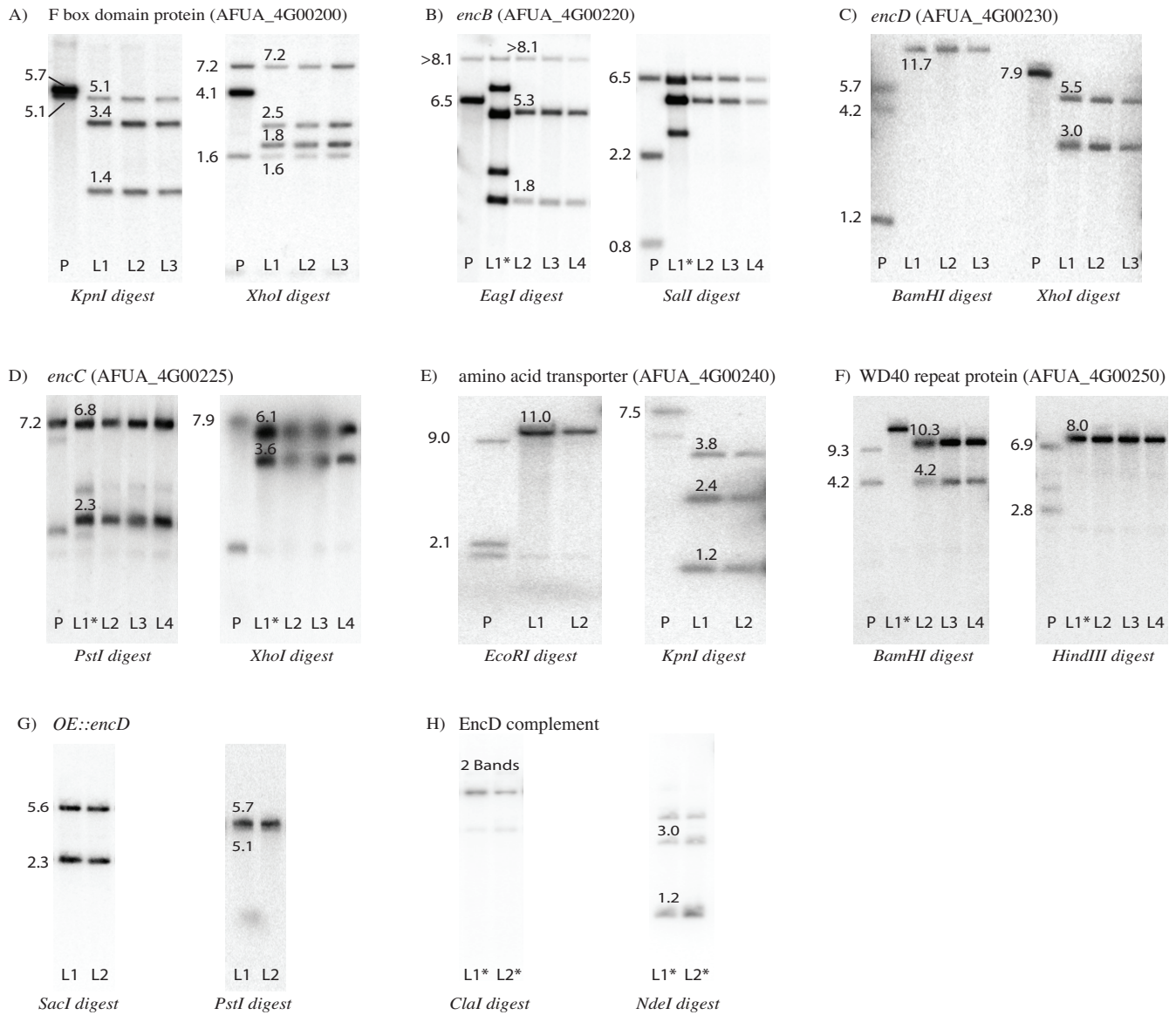
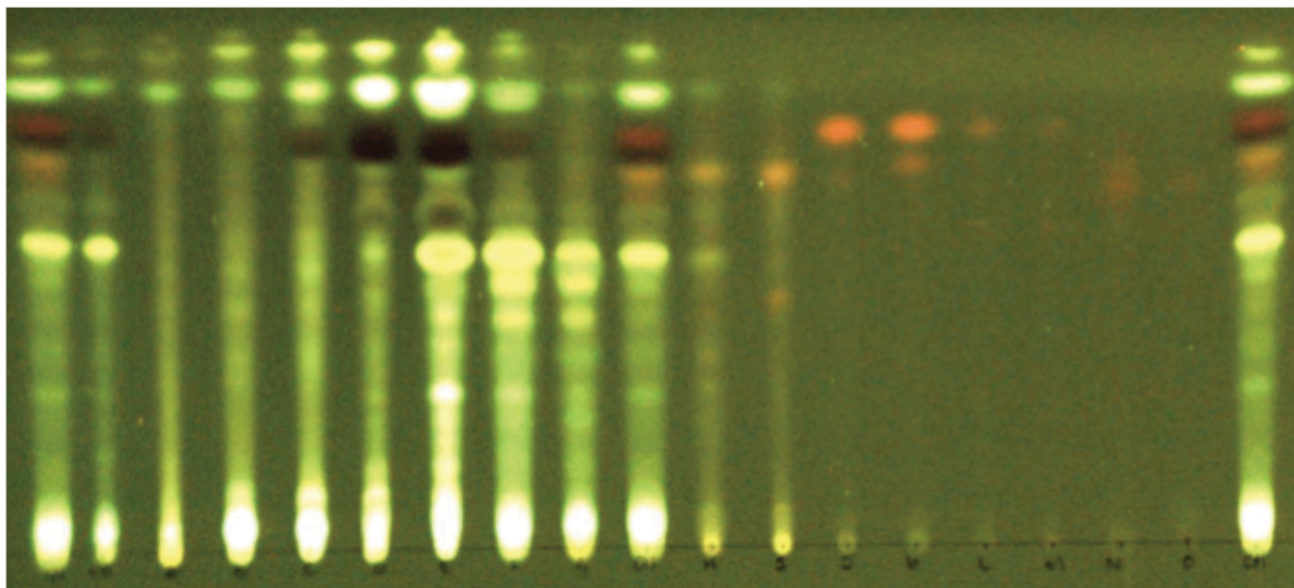
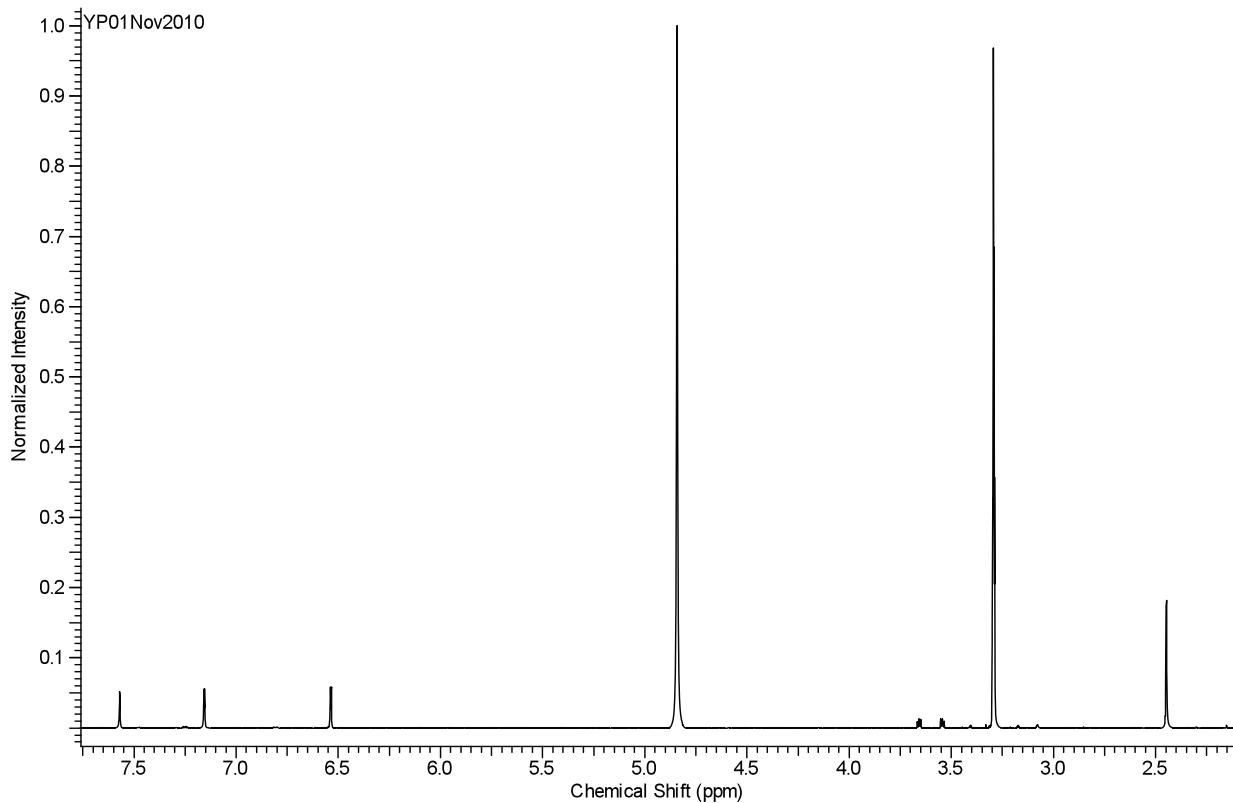


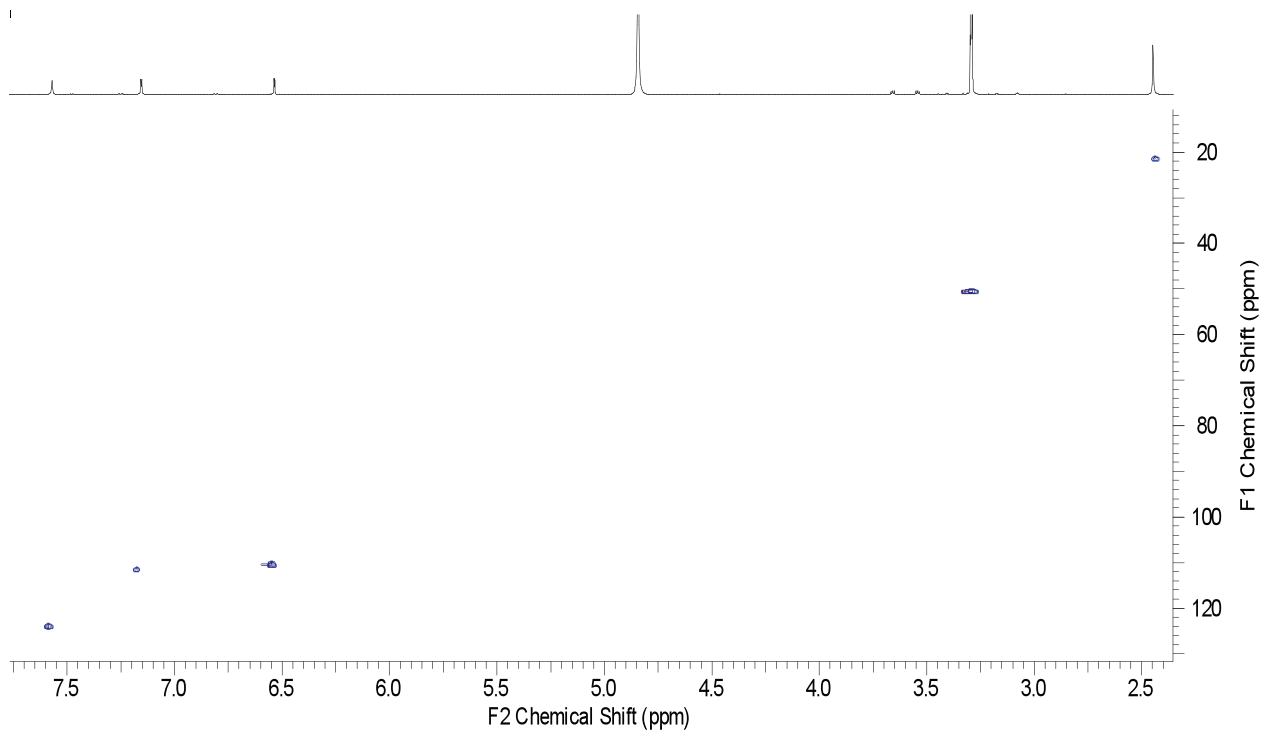
Figure S2. Southern analysis of *A. fumigatus* mutants. A) Δ AFUA_4G00200 B) Δ encB C) Δ encD D) Δ encC E) Δ AFUA_4G00240 F) Δ AFUA_4G00250 G) Overexpression EncD H) EncD complement. P lanes denote parental strain while lanes L1-L4 denote transformants that are positive via initial PCR screen. * Transformants that contain ectopic integration or multiple integrations of the transformation cassette.



1 2 A B C D E F G 1 H I J K L M N O 1
Figure S3. TLC analysis of fractions FY01A-O. Plates were developed using Toluene: Ethyl acetate: Formic acid (5:4:0.8; v/v/v). 1 corresponds to the wild type total crude extract and 2 correspond to the *ΔencA* total crude extract.



^1H NMR of endocrocin (600MHz, CD_3OD)



HSQC of endocrocin (600MHz, CD_3OD)

Figure S4. NMR spectra of endocrocin purified from crude extracts of *A. fumigatus*.

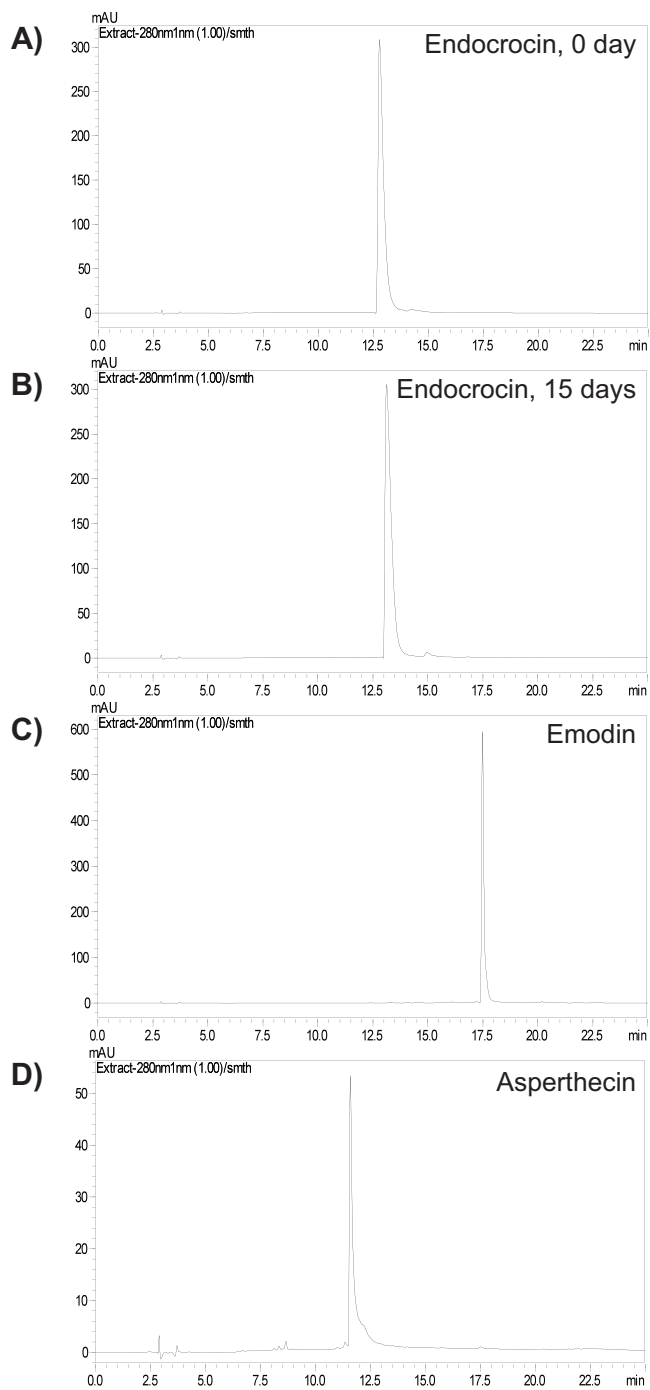


Figure S5. Endocrocin stability assay. A) HPLC chromatogram of endocrocin after incubation at 37°C for 0 days. B) HPLC chromatogram of endocrocin after incubation at 37°C for 15 days. C) HPLC chromatogram of emodin. D) HPLC chromatogram of asperthecin.