Electronic Supplementary Material (ESI) for ChemComm.

Supplementary Information

Directly utilizing an endogenous gene to dissect regulatory elements in the biosynthetic gene cluster of nosiheptide**

Min Ma, Yanjiu Xue, Weiying Liu, Hong Zhang, Lingyi Kong, Shuzhen Wang, Yijun Chen**

[*] State Key Laboratory of Natural Medicines and Laboratory of Chemical Biology

China Pharmaceutical University

24 Tongjia Street, Nanjing, Jiangsu Province, 210009 (PRC)

Fax: 86-25-83271031

E-mail: yjchen@cpu.edu.cn; shuzhenwang@cpu.edu.cn

.

Table of contents

Materials and Methods

Materials and general methods. Details of bacterial strains and plasmids used in this study are given in **Table S1**. Primers used in this study are listed in **Table S2**. Prime Star Max DNA polymerase and rTaq DNA polymerase were purchased from TaKaRa Biotechnology Co., Ltd. (Dalian, China). Primer synthesis and DNA sequencing were performed at Nanjing Springen Biotech Co., Ltd.. Restriction enzymes and other enzymes were purchased from Fermentas international INC (USA). Unless otherwise stated, biochemical materials and chemicals were purchased from Sigma Co. (St. Louis, USA). The standard DNA manipulation in *Escherichia coli* and *Streptomyces actuosus* were the same as the procedures described previously.[1]

Construction of the host-vector system pWHM7-L1101. The vector, pWHM7, was constructed as follows. Initially, a 300bp synthetic polylinker containing five unique restriction sites (*Eco*RI, *Xba*I, *Pst*I, *Sph*I, *Bam*HI, *Hin*dIII) and two fd terminators preceding *Xba*I site was digested by the restriction enzymes *Eco*RI/*Hin*dIII and introduced into pWHM3, yielding pWHM5. Subsequently, a 1266 bp fragment carrying the apramycin resistance gene (*aacC*) and conjugation gene (*oriT*) were obtained using primer pairs *oriT-aacC* F/*oriT-aacC* R and template plasmid pIJ773. A total 50 μL PCR reaction mixture contained: 25 μl Prime STAR Max Premix (2 ×), 1.0 μl of degenerate PCR primers (20 mM) each, 1 μl template plasmid pIJ773 (ca. 0.2 μg) and 2.5 μL DMSO. The amplification conditions were as follows: a) 30 cycles of 98 ℃ 10 s, 72 ℃ 1.0 min; b) 72 ℃ 5 min for 1 cycle. The purified and recovered fragments were cloned into pMD19-T simple vector after adding A at 3' terminus using rTaq DNA polymerase, generating recombinant plasmids pMD-*oriT aacC*. After digestion with *Eco*RI, the *oriT-aacC* fragment was purified and ligated with pWHM5, generating pWHM7. The introduction of this fragment conferred apramycin resistance and the ability of conjugation between *E.coli* and *Streptomyces*. All DNA inserts were confirmed by sequencing (Springen, Nanjing).

In-frame deletion of *nosA* gene in *S. actuosus* ATCC 25421 was performed according to the method described

previously.[2]

Verification of L1101-pWHM7 system. Negative control vector pWRA was constructed as follows. First, a fragment containing the promoterless *nosA* gene and its ribosomal binding site was obtained by PCR amplification using primer pairs RBS-*nosA* U/ RBS-*nosA* D. The resulting PCR products, 563bp long, were cloned into pMD19-T simple vector after adding A at 3' terminus using rTaq DNA polymerase without primers, generating recombinant plasmid pMD-RA. The negative control vector pWRA was obtained by ligating the *Xba*I/*Hin*dIII treated products of pMD-RA and vector pWHM7. To construct positive control vector pWEA, a 795bp seamless fusion of P*ermE** and the promoterless *nosA* gene was obtained by overlap PCR amplification. Two following primer pairs were used: P*ermE*-nosA* U1/P*ermE*-nosA* D1, P*ermE*-nosA* U2/P*ermE*-nosA* D2, in which there are 15 complementary base pairs between primers P*ermE*-nosA* D1 and P*ermE*-nosA* U2 at 5' end. The procedures to construct pWEA were the same as the vector for negative control.

The vectors were transformed into methylation-deficient *E. coli* ET12567/pUZ8002 and then introduced into L1101 by conjugation. The mutant strains L1101/pWRA and L1101/pWEA were screened by the phenotype of apramycin resistance.

Cloning of DNA fragments containing transcriptional activity in *S. actuosus*. The DNA fragments presumably responsible for the regulation of gene transcription were isolated from *S. actuosus* and ligated with the promoterless *nosA* gene through overlap PCR amplification. The resulting fragments were purified and cloned into pMD-19T simple vector. These recombinant plasmids were digested and ligated with pWHM7 treated by the same restriction enzymes, generating derivatives of pWHM7. All transcriptional fusion plasmids were transformed into methylation-deficient *E.coli* host ET12567/pUZ8002 and then introduced into L1101 by conjugation. The derivatives of L1101 carrying the recombinant plasmids were screened by the phenotype of apramycin resistance. All DNA inserts were confirmed by sequencing (Springen, Nanjing).

RNA isolation and 5' RACE analysis of transcriptional start point. Isolation of total RNA from *S. actuosus* ATCC 25421 and 5' RACE was performed according to the methods described previously.[3] 5' RACE was carried out with cDNA Amplification kits for 5' RACE System for Rapid Amplification of cDNA ends (Invitrogen). The primers used in the experiments are listed in Table S1.

5' Deletion analysis of the core promoter. To analyze the core promoter region in $NCR_{L\text{-}M}$, we developed series of recombinant fusions of pWHM7 carrying 5' truncated fragments of $NCR_{L\text{-}M}$ or inverse $NCR_{L\text{-}M}$ upstream the promoterless *nosA* gene. All recombinant plasmids were constructed as the method described in the section **Cloning of DNA fragments containing transcriptional activity in *S. actuosus*.** The primers used in the experiments are listed in Table S1.

Fermentation and HPLC analysis of nosiheptide in the cultures. Fermentation of *S. actuosus* wild-type, L1101 and other mutant strains were carried out according to the method described previously.[1b] HPLC analysis of nosiheptide (**2**) was performed according to the procedures described by Liu et al.[4] The concentration of nosiheptide in fermentation culture was quantified by measuring the peak area and calculated according to the standard curve made by using commercial nosiheptide.

Table S1. Bacterial strains and plasmids used and constructed in this study.

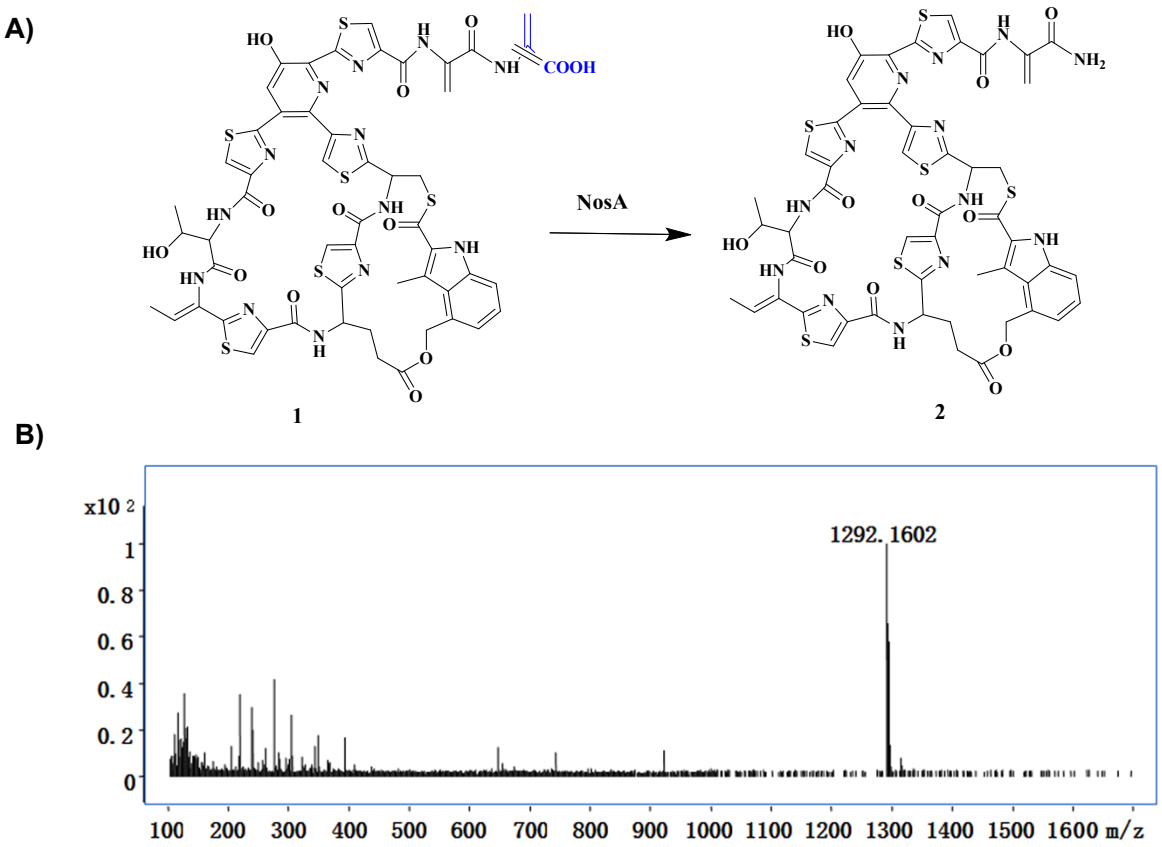
Strains/plasmids	Characteristic(s)	Source/Reference
***E. coli* strain**		
ET12567/pUZ8002	Donor strain for conjugation between *E. coli* and *Streptomyces*	Our laboratory
DH5α	Host for general cloning	TIANGEN, China
Streptomyces actuosus		
ATCC 25421	Wild type strain, nosiheptide producing	ATCC
L1101	Derivative of ATCC 25421, *nosA* in-frame deletion mutant	[5]
L1101/pWEA	Derivative of L1101, harboring positive control vector pWEA	This study
L1101/pWRA	Derivative of L1101, harboring negative control vector pWRA	This study
L1101/pWMA	Derivative of L1101, harboring pWMA	This study
L1101/pWLA	Derivative of L1101, harboring pWLA	This study
L1101/pWMX	Derivative of L1101, harboring pWMX (X=1, 2, 3, 4, 5)	This study
L1101/pWLY	Derivative of L1101 harboring pWLY (Y=6, 7, 8, 9, 10)	This study
Plasmids		
pMD19-T	*E. coli* subcloning vector	TaKaRa
pMD-*oriT aacC*	Derivative of pMD-19T carrying origin of transfer and apramycin resistance gene	This study
pMD-EA	Derivative of pMD19-T carrying a seamless fusion of promoter P*ermE** and promoterless *nosA* gene	This study
pMD-RA	Derivative of pMD19-T carrying a promoterless *nosA* gene containing ribosomal binding site	This study
pWHM3	*E. coli-Streptomyces* shuttle vector	[6]
pWHM5	Derivative of pWHM3 containing a synthesized polylinker	This study
pWHM7	Derivative of pWHM5 carrying apramycin resistance gene and origin of transfer gene	This study
pWRA	Derivative of pWHM7 carrying a promoterless *nosA* gene containing ribosomal binding site	This study
pWEA	Derivative of pWHM7 carrying a seamless combination of *PermE** promoter and promoterless *nosA* gene	This study
pSET152E	pSET152 derivative containing P*ermE** promoter	Our laboratory
pIJ773	A PCR targeting vector	[7]
pWMA	Derivative of pWHM7 carrying a seamless combination of NCR_M and *nosA* ORF	This study
pWLA	Derivative of pWHM7 carrying a seamless combination of NCR_L and *nosA* ORF	This study
pWMX	Derivative of pWHM7 carrying a seamless fusion of 5' truncated NCR_M and *nosA* ORF (X=1, 2, 3, 4, 5)	This study
pWLY	Derivative of pWHM7 carrying a seamless fusion of 5' truncated NCR_L and *nosA* ORF (Y=6, 7, 8, 9, 10)	This study

Table S2. Primers used in this study[a]

Primer Name	Primer Sequence (5'-3')
oriT-aacC F	CGGAATTCAGGAACTTATGAGCTCAGCCAATC
oriT-aacC R	ATGAATTCCGAAGTTCCCGCCAGCCTCGCAGA
RBS-*nosA* U:	TCTAGAGAGTTCACCGCGAACCTCTTCCAC
RBS-*nosA* D:	AAGCTTCGCGGAAGGCGCTGATTTTTGGTG
P*ermE**-*nosA*U1:	TCTAGACGAGGTCCAGCCCGACCCGAG
P*ermE**-*nosA* D1:	TCGGTCATATCGATACCGTCGATC
P*ermE**-*nosA* U2:	GTATCGATATGACCGAACACCCCG
P*ermE**-*nosA* D2:	AAGCTTGCGGAAGGCGCTGATTTTTGGTGG
NCR_MA U1	TCTAGACGTCATGTGAATTCCTCTGCGAAC
NCR_MA D1	TCGGTCAT CGGAAACCCCTTCCAT
NCR_MA U2	GGTTTCCG ATGACCGAACACCCCG
NCR_MA D2	AAGCTT GCGGAAGGCGCTGATTTTTGGTGG
NCR_LA U3	TCTAGAGTCCACCGGAAACCCCTTCCATGT
NCR_LA D3	GTTCGGTCAT GTGAATTCCTCTGC
NCR_LA U4	GAATTCAC ATGACCGAACACCCCG
NCR_LA D4	AAGCTTGCGGAAGGCGCTGATTTTTGGTGG
NCR_MA cut71-U	TCTAGAAAGACACCGAGGAAACGGAGCACT
NCR_MA cut144-U	TCTAGACTCCACGGCGGTTGTGCAGCAGAC
NCR_MA cut198-U	TCTAGAGCCTTTTTGTGCCGTCACAGCAGT
NCR_MA cut258-U	TCTAGA GGAATCCTTGACTCACGGAAGACC
NCR_MA cut280-U	TCTAGAGTCCACCGGAAACCCCTTCCATGT
NCR_MA D1	AAGCTT GCGGAAGGCGCTGATTTTTGGTGG
NCR_LA cut50-U	TCTAGA CTGCAGCGCGTGCCTCGGAGGCTAT
NCR_LA cut70-U	TCTAGA GCTATGCAGCGGGTCTTCCGTGAG
NCR_LAcut95-U	TCTAGA CAAGGATTCCGGCCCTTTTGCAT
NCR_LA cut120-U	TCTAGA CTGCTGATTCCCCGCTGCTGTAC
NCR_LA cut246-U	TCTAGA GTGGGGAGGGTAACGGAGAAAGAG
NCR_LA D4	AAGCTTGCGGAAGGCGCTGATTTTTGGTGG
R857-1	ACGGGTGTCCACCGT
R857-2	TCGGCGCGTACGTCCTCCA
R857-3	GGCAGGACGAAGTCACCGGT
R858-1	GACATGACCTTGGCC
R858-2	GGCTCTCGTCCAGGAACTCG
R858-3	CGTCCAGGAACTCGGAGATC

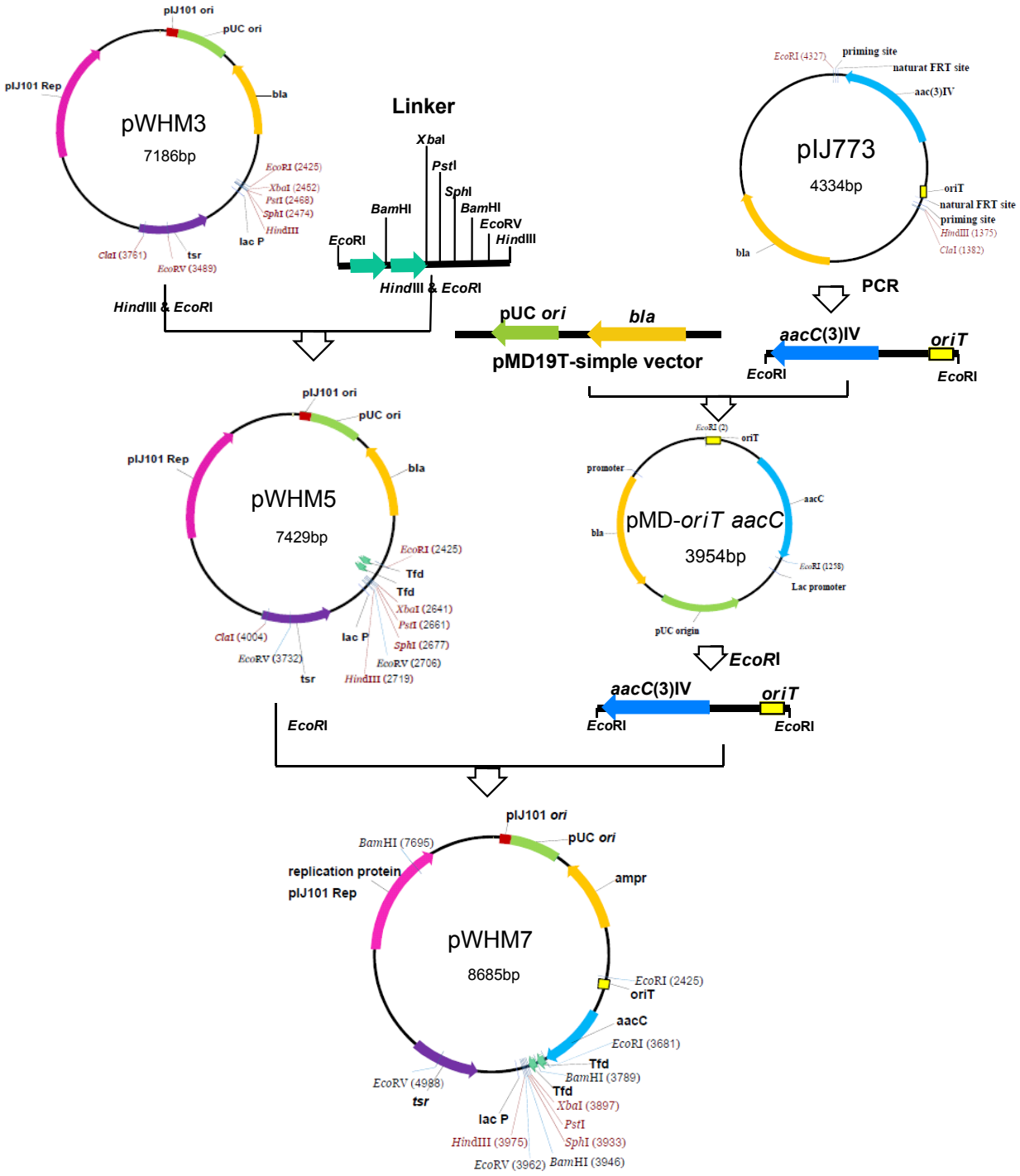
[a]Underlined letters are restriction sites for *Hind*III (AAGCTT), *Xba*I (TCTAGA) and *EcoR*I (GAATTC).

Figure S1. Confirmation of the host strain L1101 on the accumulation of **1.**



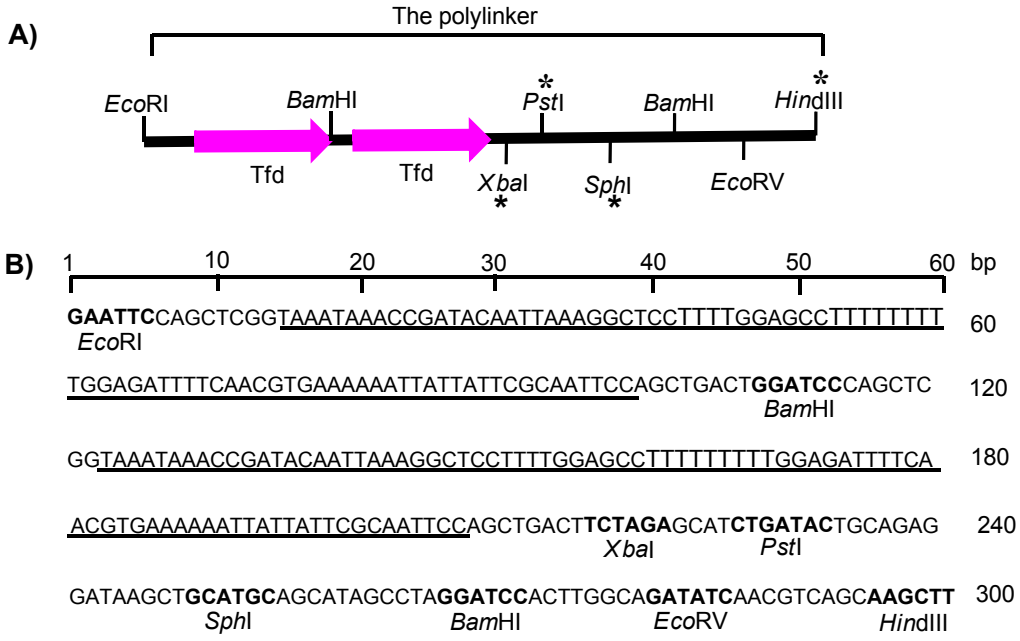
A) NosA-catalyzed enamine dealkylation; **B**) HR-ESI-MS of substrate **1** for NosA.

Figure S2. Construction of promoter-probe plasmid pWHM7.



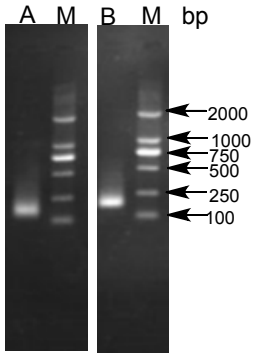
Plasmid pWHM5 was obtained by ligating the 7134-bp *Eco*RI-*Hin*dIII fragment of *E.coli-Streptomyces* shuttle vector pWHM3 and a 300-bp polylinker carrying two-copy fd terminator of *E. coli* Phage fd and five mono-clone restriction sites (*Eco*RI, *Xba*I, *Pst*I, *Sph*I, *Hin*dIII). Subsequently, a 1266-bp *EcoR*I fragment amplified from pIJ773 carrying *aacC* and *oriT* was subcloned into pMD19-T, resulting in pMD-*oriT aacC*. Introduction of this fragment into *Eco*RI site of pWHM5 yielded plasmid pWHM7, an *E. coli-Streptomyces* conjugating promoter-probe vector suitable for the cloning of promoter-active fragment and the promoterless reporter gene. *oriT,* origin of transfer; *ori*, origin of DNA replication; pIJ101 Rep, the replication protein of pIJ101; Tfd, fd terminator; *bla*, ampicillin resistance gene; *aacC*, apramycin resistance gene; *tsr* ,thiostrepton resistance gene.

Figure S3. Restriction map and nucleotide sequence of the synthetic polylinker.



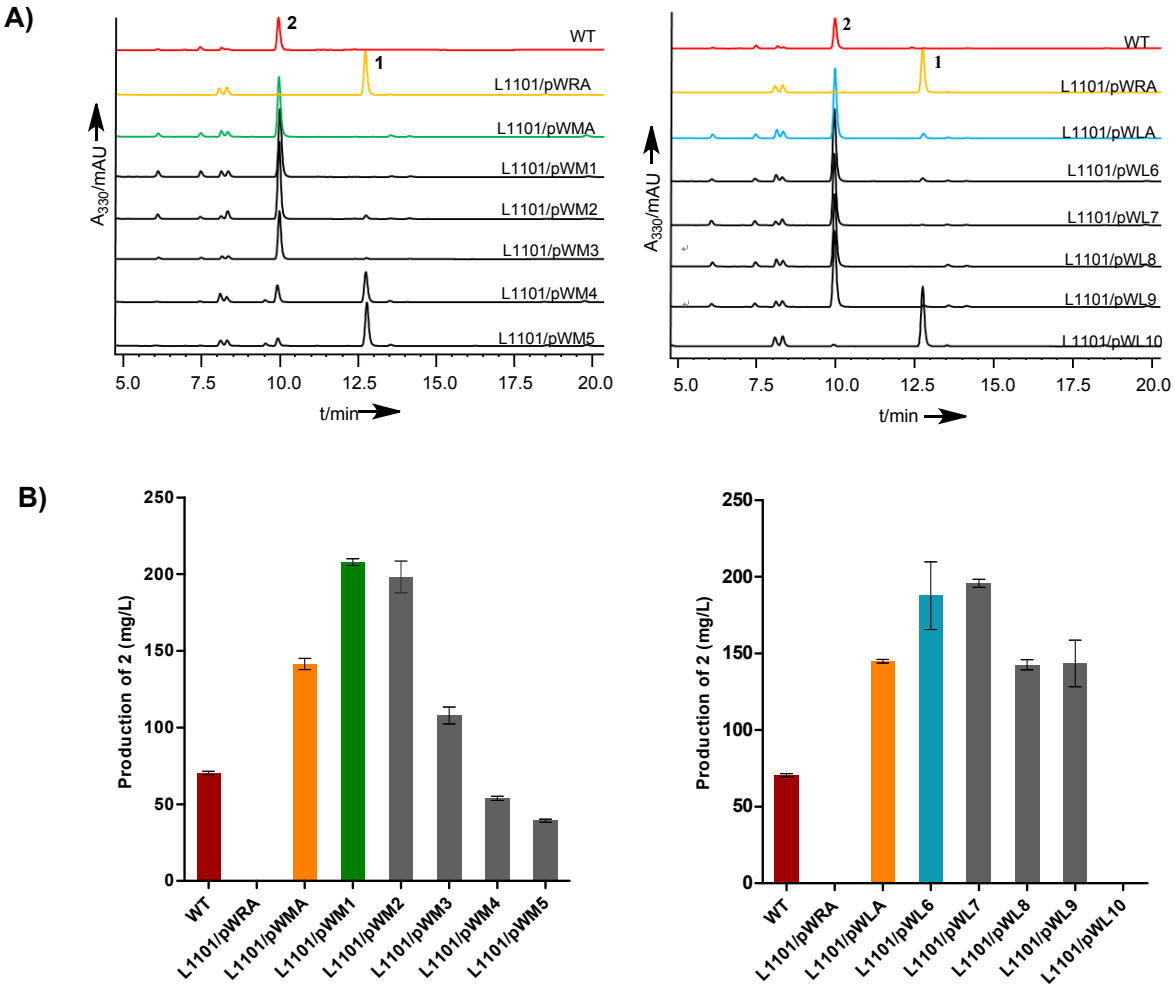
A) Restriction map of the polylinker. **B)** Nucleotide sequence of the synthetic polylinker. Underlined letters are fd terminators. The asterisk indicates the mono-clone sites for inserts in pWHM7.

Figure S4. Agarose gel-electrophoresis of 5’ RACE products.



Reaction products were analyzed by gel-electrophoresis on a 1.5% agarose gel (1 × TBE) and stained with 0.5 μg /mL ethidium bromide. The products were purified and then sequenced to determine the transcriptional start sites of *nosL* and *nosM* genes. M: DL2000 DNA ladder; lane A: 5’ RACE products of *nosL* gene; lane B: 5’ RACE products of *nosM* gene

Figure S5. Determination of 5’ boarders of the divergent promoters in both strands by 5’ deletion.



A) HPLC analysis of products in fermentation extracts of wild-type strain and its mutants L1101 carrying derivatives of pWHM7. **B)** Production of nosiheptide from wild-type and its mutants L1101 carrying derivatives of pWHM7. Error bars indicate standard deviations from three independent experiments. WT, wild-type strain; L1101, *nosA*-inactivated mutant; pWRA, derivative of pWHM7 carrying the promoterless *nosA* gene retaining its ribosomal binding site; pWMA and pWLA are derivatives of pWHM7 carrying fusions of $NCR_{L\text{-}M}$ in both orientations and the promoterless *nosA* gene, respectively; pWMX (X=1, 2, 3, 4, 5) and pWLY (Y=6, 7, 8, 9, 10) are two sets of derivatives of pWHM7 carrying 5’ truncated $NCR_{L\text{-}M}$ in both orientations and a downstream *nosA* gene. All derivatives of pWHM7 are non-integrative in the host strain of L1101.

References

[1] a) T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood, *Practical streptomyces genetics*, The John Innes Foundation Norwich, UK, **2000**; b) W. Liu, M. Ma, Y. Xue, N. Liu, S. Wang, Y. Chen, *ChemBioChem* **2013**, *14*, 573-576.

[2] Y. Yu, H. Guo, Q. Zhang, L. Duan, Y. Ding, R. Liao, C. Lei, B. Shen, W. Liu, *J. Am. Chem. Soc.* **2010**, *132*, 16324-16326.

[3] E. Scotto-Lavino, G. Du, M. A. Frohman, *Nat. Protoc.* **2007**, *1*, 2555-2562.

[4] W. Liu, Y. Xue, M. Ma, S. Wang, N. Liu, Y. Chen, *ChemBioChem* **2013**, *14*, 1544-1547.

[5] Y. Yu, H. Guo, Q. Zhang, L. Duan, Y. Ding, R. Liao, C. Lei, B. Shen, W. Liu, *J. Am. Chem. Soc.* **2010**, *132*, 16324–16326.

[6] J. Vara, M. Lewandowska-Skarbek, Y.-G. Wang, S. Donadio, C. Hutchinson, *J. Bacteriol.* **1989**, *171*, 5872-5881.

[7] B. Gust, G. L. Challis, K. Fowler, T. Kieser, K. F. Chater, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1541-1546.