# Cloning and heterologous expression of the spectinabilin biosynthetic gene cluster from *Streptomyces spectabilis*†‡

**Yoo Seong Choi,§ Tyler W. Johannes,§¶ Michael Simurdiak, Zengyi Shao, Haige Lu and Huimin Zhao***

***Received 5th November 2009, Accepted 20th November 2009***
***First published as an Advance Article on the web 22nd December 2009***
**DOI: 10.1039/b923177c**

**Spectinabilin is a rare nitrophenyl-substituted polyketide metabolite. Here we report the cloning and heterologous expression of the spectinabilin gene cluster from *Streptomyces spectabilis*. Unexpectedly, this gene cluster is evolutionarily closer to the aureothin gene cluster than to the spectinabilin gene cluster from *Streptomyces orinoci*. Moreover, the two nearly identical spectinabilin gene clusters use a distinctly different regulation mechanism.**

Nitro group-containing natural products are relatively rare in nature, with only about 150 known examples,[1] but include such well-known antibiotics as chloramphenicol and pyrrolnitrin. Spectinabilin **1** is an unusual chain-extended nitrophenyl-containing polyketide compound produced by *Streptomyces spectabilis*[2] and *Streptomyces orinoci*[3] that exhibits antiviral[2] and antimalarial[4] activities. Although the biosynthetic pathway for spectinabilin production was recently proposed,[5] heterologous production of spectinabilin has never been achieved. In this work, we report the cloning and heterologous expression of the complete spectinabilin gene cluster from *S. spectabilis* and the discovery of some unexpected features in the gene cluster.

A fosmid library prepared from *S. spectabilis* genomic DNA was extensively screened using degenerate primers and fosmid/insert-junction sequencing.[6] The fosmid 79C was isolated from the fosmid library and integrated into the chromosome of *S. lividans* 66. The exconjugant integrant was verified by PCR and subsequently checked under a wide variety of conditions for spectinabilin production. Concentrated media extract prepared from 79C-*S. lividans* was analyzed by LC-MS/MS. No spectinabilin peak was observed from the concentrated extract of wild-type *S. lividans* 66. In comparison, a clear peak with the same retention time (23.4 min), relative molecular weight value (478.3 Da), and MS/MS fragmentation pattern as authentic spectinabilin was observed from the concentrated media extract of 79C-*S. lividans* (Fig. 1).

*Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana IL 61801, USA. E-mail: zhao5@illinois.edu; Fax: (001) 217-333-5052*

† Electronic supplementary information (ESI) available: Materials and experimental. See DOI: 10.1039/b923177c

‡ This work was partially supported by the National Institutes of Health grant GM077596 and the University of Illinois at Urbana-Champaign.

§ These two authors contributed equally.

¶ Current address: The University of Tulsa, Department of Chemical Engineering, 800 South Tucker Drive, Tulsa, OK 74104, USA.

After confirming heterologous production of spectinabilin, the 45.3 kb insert of the 79C fosmid was sequenced using a transposon-based strategy.[6] Analysis of the sequenced fosmid revealed fourteen open reading frames (ORFs), which were designated *spnA-M* for spectinabilin biosynthesis (Fig. 2). The Spn proteins were assigned possible functions based on their homology to proteins of known function (Table 1). In brief, ORF*spnD* encodes a transcriptional activator, which has 87% sequence similarity with the transcriptional activator for aureothin biosynthesis, *aurD*. Four ORFs (*spnA*, *spnA′*, *spnB*, *spnC*) encode type I polyketide synthases, four ORFs (*spnE*, *spnF*, *spnG*, *spnK*) are involved in biosynthesis of the pNBA starter unit, and two ORFs (*spnH*, *spnI*) are involved in post-PKS reactions (Fig. 3). SpnA (1937 aa), like AurA,[3] is an iterative modular type I polyketide synthase with the AT domain of methylmalonyl-CoA specificity of RVDVV-7-M-1-S-1-A-2-W.[7] As a result, six PKS modules of the spectinabilin cluster catalyze seven rounds of elongation and reduction. SpnA′ (3626 aa) is a biomodular protein of the same domain architecture of SpnA (KS, AT, DH, KR and ACP), which is responsible for two successive unsaturated chain elongation steps using methylmalonyl-CoA, due to the presence of the consensus sequence of AT domain for incorporating propionate extender units.[7] SpnB (2183 aa) contains a sequence of malonyl-CoA specificity in the AT domain, QTAYT-7-E-1-A-1-YHLL near the active site GHSVG, which leads to a full reduction of the β-keto group after condensation of malonyl-coA.[7] SpnC (2242 aa) carries two modules of KS, AT, and ACP domains and a C-terminal TE domain and catalyzes the last two Claisen condensations. Similar to AurC,

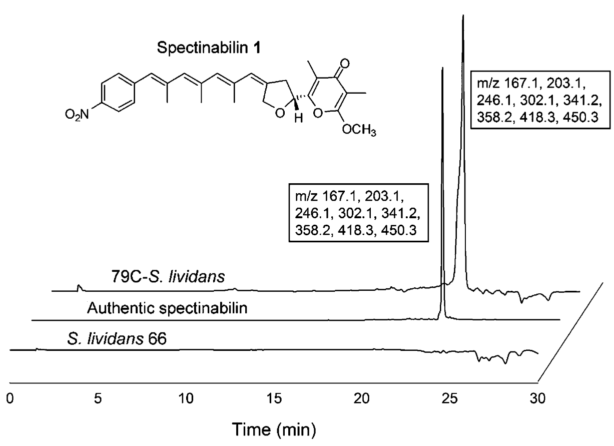


**Fig. 1** LC-MS/MS analysis of wild type *S. lividans* 66, authentic spectinabilin, and 79C-*S. lividans*.

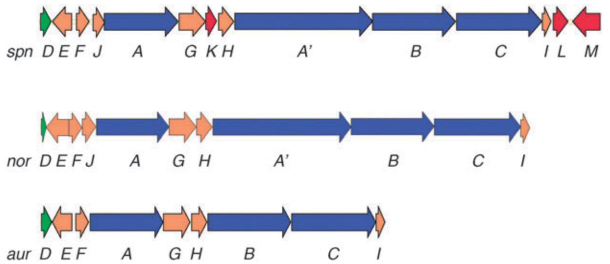


**Fig. 2** Organization of the three gene clusters including the spectinabilin biosynthetic gene cluster from *S. spectabilis* (*spn*), the proposed spectinabilin biosynthetic gene cluster from *S. orinoci* (*nor*), and the aureothin biosynthetic gene cluster from *S. thioluteus* (*aur*). Each arrow indicates the direction of transcription and relative sizes of the ORFs. Blue ORFs, type I PKS genes; orange ORFs, genes involved in either start unit biosynthesis or post-PKS reactions; green ORFs, regulatory genes; red ORFs, genes not found in the *aur* and *nor* gene clusters.

the AT domain of module 5 contains a consensus motif that confers methylmalonyl-CoA specificity and the AT domain of module 6 is greatly reduced in size and shows an altered signature motif.

The basic structure of the spectinabilin gene cluster from *S. spectabilis* is quite similar to the proposed spectinabilin cluster from *S. orinoci*.[5] However, there are some striking differences. Despite similarity between the two gene clusters, they appear to be regulated differently. Although located at the same position in their respective gene clusters, *spnD* and *norD* encode transcriptional regulators from different families. Interestingly, SpnD belongs to the AfsR family of global regulatory proteins, while NorD is a member of the ArsR family, which functions as a repressor. Furthermore, although the two spectinabilin gene clusters share significant homology, the gene cluster isolated from *S. spectabilis* appears to be evolutionarily more closely related to the aureothin biosynthetic gene cluster. Aureothin, produced by *S. thioluteus*,[11] differs from spectinabilin only in the number of diene moieties and exhibits similar antifungal, cytotoxic, and insecticidal activities.[12] The ORFs *spnD*, *spnE*, *spnF*, *spnA*, *spnG*, and *spnC* share higher (69–86%) sequence identity with their counterparts in the aureothin gene cluster than their counterparts in the spectinabilin cluster from *S. orinoci*. However, the ORFs *spnJ*, *spnH*, *spnA′*, *spnB*, and *spnI* share higher sequence identity (71–77%) with their counterparts in the spectinabilin gene cluster from *S. orinoci* than their counterparts in the aureothin cluster from *S. thioluteus*. Interestingly, the ORFs *spnK*, *spnL*, and *spnM* have no counterparts in either of the other two clusters.

The ORF *spnK* is located between *spnG* and *spnH* and encodes for a protein with homology to ADC lyases. ADC lyases are involved in the biosynthesis of *p*-aminobenzoic acid (pABA),[13] a starter unit for spectinabilin biosynthesis. Thus, SpnK (288 aa) most likely serves to up-regulate the production of pABA. SpnL (395 aa) shares high similarity (92%) with beta-ketoadipyl-CoA thiolases, which are known to catalyze the production of acetyl-CoA.[14] Acetyl-CoA carboxylases are widely distributed among bacteria including *Streptomyces* species, and catalyze irreversible carboxylation of acetyl-CoA to produce malonyl-CoA. Thus, SpnL most likely serves to increase the concentration of malonyl-CoA and methylmalonyl-CoA, which are substrates used in polyketide assembly. In addition, SpnJ shows high sequence identity with DAHP synthases, which are involved in the shikimate pathway for chorismate biosynthesis.[15] Since chorismate is used as a precursor in the biosynthesis of pABA, SpnJ possibly facilitates the biosynthesis of spectinabilin, similar to SpnL. Finally, SpnM (715 aa) shares high similarity (86%) with a membrane transport protein in *S. coelicolor* A3(2), and it may be involved in the transport of spectinabilin.

Previous phylogenetic analysis of the *nor* and *aur* polyketide synthases suggested that these two biosynthetic gene clusters share a common evolutionary origin.[5] The high degree of sequence similarity and the identical gene order of the *spn* gene cluster with the *aur* gene cluster also suggest a common ancestor. Remarkably, homologs of ADC lyase (*spnK*) and DAHP synthase (*spnJ*) were only found in the *spn* gene cluster. ADC lyases are mainly observed in bacteria, and have not been reported in eukaryotes,[16] even though eukaryotes use pABA synthases for the biosynthesis of pABA. Similarly, the evolutionary loss of the shikimate pathway including DAHP

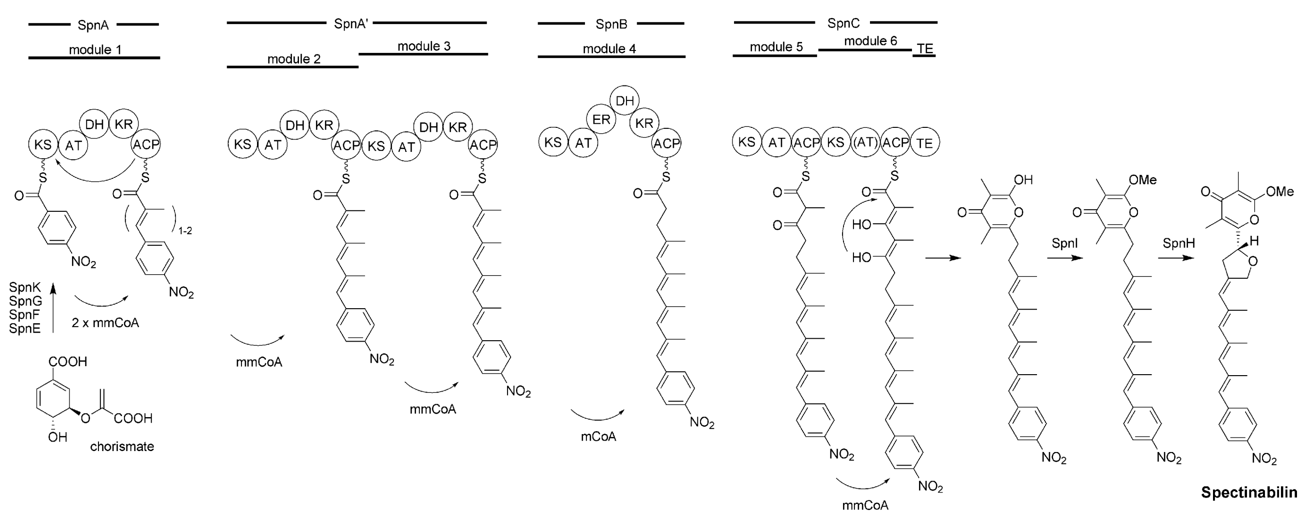


**Fig. 3** Proposed model for spectinabilin biosynthesis. mmCoA, methylmalonyl-CoA; mCoA, malonyl-CoA.

**Table 1** Proposed functions of ORFs in the spectinabilin biosynthetic gene cluster

| Protein | Amino acids | Proposed function | Sequence similarity | Identity/ similarity (%)[a] | Accession number | Reference |
|---|---|---|---|---|---|---|
| SpnD | 279 | Transcriptional activator | AurD, *S. thioluteus* | 82/87 | CAE02599 | 3 |
| SpnE | 505 | Acyl-CoA ligase | AurE, *S. thioluteus* | 86/92 | CAE02600 | 3 |
| SpnF | 341 | Amine oxidase | AurF, *S. thioluteus* | 75/84 | 3CHH_A | 8 |
| SpnJ | 298 | DAHP synthase | NorJ, *S. orinoci* | 75/83 | CAO85892 | 5 |
| SpnA Module1 | 1937 | Polyketide synthase KS AT DH KR ACP | AurA, *S. thioluteus* | 73/80 | CAE02602 | 3 |
| SpnG | 697 | PABA synthase | AurG, *S. thioluteus* | 79/86 | CAE02603 | 3 |
| SpnK | 288 | ADC lyase | ADC lyase, *Nocardiopsis dassonvillei* subsp. *dassonvillei* DSM 43111 | 46/56 | ZP04334621 | – |
| SpnH | 409 | Cyt P-450 monooxygenase | NorH, *S. orinoci* | 77/85 | CAO85895 | 5 |
| SpnA′ Module2 Module3 | 3626 | Polyketide synthase KS AT DH KR ACP KS AT DH KR ACP | NorA', *S. orinoci* | 71/78 | CAO85896 | 5 |
| SpnB Module4 | 2183 | Polyketide synthase KS AT ER DH KR ACP | AurB, *S. thioluteus* | 71/80 | CAE02605 | 3 |
| SpnC Module5 Module6 | 2242 | Polyketide synthase KS AT ACP KS AT ACP TE | AurC, *S. thioluteus* | 69/78 | CAE02606 | 3 |
| SpnI | 232 | *O*-Methyltransferase | NorI, *S. orinoci* | 73/83 | CAO85899 | 5 |
| SpnL | 395 | β-Ketoadipyl-CoA thiolase | β-Ketoadipyl-CoA thiolase, *S. ambofaciens* ATCC 23877 | 87/92 | CAJ88821 | 9 |
| SpnM | 715 | Membrane transport protein | Membrane transport protein, *S. coelicolor* A3(2) | 79/86 | NP631036 | 10 |

[a] Based on BLAST search of the NCBI database on October 25, 2009.

synthases occurred in eukaryotes.[17] Thus, it is possible that both *spnK* and *spnJ* were evolutionarily lost in the *aur* gene cluster. Moreover, the *spnA′* gene in the *spn* cluster might also have been lost in the *aur* gene cluster.

## Conclusions

In conclusion, this study successfully demonstrates the heterologous production of a rare nitrophenyl-substituted polyketide spectinabilin. Detailed sequence analysis revealed a close evolutionary relationship between the spectinabilin gene cluster and the aureothin gene cluster. The production of spectinabilin seems to be regulated differently in two very closely related strains of *Streptomyces*, which represents a rare example in natural product biosynthesis. This study has laid out the foundation for further metabolic engineering to increase the production of spectinabilin and combinatorial biosynthesis to generate new spectinabilin analogs with improved biological properties.

## Notes and references


1 H. Laatsch, *Antibase 2003*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2003.
2 K. Kakinuma, C. A. Hanson, J. Rinehart and K. L., *Tetrahedron*, 1976, **32**, 217–222.
3 J. He and C. Hertweck, *Chem. Biol.*, 2003, **10**, 1225–1232.
4 M. Isaka, A. Jaturapat, J. Kramyu, M. Tanticharoen and Y. Thebtaranonth, *Antimicrob. Agents Chemother.*, 2002, **46**, 1112–1113.
5 N. Traitcheva, H. Jenke-Kodama, J. He, E. Dittmann and C. Hertweck, *ChemBioChem*, 2007, **8**, 1841–1849.
6 R. D. Woodyer, Z. Shao, P. M. Thomas, N. L. Kelleher, J. A. Blodgett, W. W. Metcalf, W. A. van der Donk and H. Zhao, *Chem. Biol.*, 2006, **13**, 1171–1182.
7 S. F. Haydock, J. F. Aparicio, I. Molnar, T. Schwecke, L. E. Khaw, A. Konig, A. F. Marsden, I. S. Galloway, J. Staunton and P. F. Leadlay, *FEBS Lett.*, 1995, **374**, 246–248.
8 Y. S. Choi, H. Zhang, J. S. Brunzelle, S. K. Nair and H. Zhao, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 6858–6863.
9 F. Barona-Gomez, S. Lautru, F. X. Francou, P. Leblond, J. L. Pernodet and G. L. Challis, *Microbiology*, 2006, **152**, 3355–3366.
10 S. D. Bentley, K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill and D. A. Hopwood, *Nature*, 2002, **417**, 141–147.
11 K. Maeda, *J. Antibiot. (Tokyo)*, 1953, **6**, 137–138.
12 J. L. Schwartz, M. Tishler, B. H. Arison, H. M. Shafer and S. Omura, *J. Antibiot. (Tokyo)*, 1976, **29**, 236–241.
13 J. M. Green, W. K. Merkel and B. P. Nichols, *J. Bacteriol.*, 1992, **174**, 5317–5323.
14 C. S. Harwood and R. E. Parales, *Annu. Rev. Microbiol.*, 1996, **50**, 553–590.
15 R. Bentley, *Crit. Rev. Biochem. Mol. Biol.*, 1990, **25**, 307–384.
16 G. J. Basset, S. Ravanel, E. P. Quinlivan, R. White, J. J. Giovannoni, F. Rebeille, B. P. Nichols, K. Shinozaki, M. Seki, J. F. Gregory, 3rd and A. D. Hanson, *Plant J.*, 2004, **40**, 453–461.
17 T. A. Richards, J. B. Dacks, S. A. Campbell, J. L. Blanchard, P. G. Foster, R. McLeod and C. W. Roberts, *Eukaryotic Cell*, 2006, **5**, 1517–1531.


 