Biosynthesis

DOI: 10.1002/ange.200501941

Cloning and Heterologous Expression of Isocyanide Biosynthetic Genes from Environmental DNA**

Sean F. Brady and Jon Clardy*

Cultured soil bacteria have been an important source of biologically active, naturally occurring small molecules. Uncultured soil bacteria, which outnumber their cultured counterparts by at least two orders of magnitude, are likely to be an equally important source of such molecules. In an attempt to access the biosynthetic potential of uncultured soil bacteria, we have explored an approach that involves the cloning and heterologous expression of DNA extracted directly from environmental samples (environmental DNA, eDNA) into readily cultured bacteria. Earlier, we described the identification and characterization of small-molecule antibiotics from antibacterially active eDNA clones found using a high-throughput phenotypic screen.[1] This approach directly couples the biosynthesis of each natural product that is found to a relatively small piece of cloned eDNA and therefore permits the characterization of both new natural products and their biosynthetic gene clusters simultaneously. Herein, we report the isolation and characterization of the isocyanide-containing eDNA-derived antibiotic **1**, its biosynthetic genes (*isnA* and *isnB*), the origin of the isocyanide nitrogen atom, and the general outline of the biosynthesis of **1**.

CSLG18, the eDNA clone that produces **1**, was found in a cosmid library constructed from blunt-ended gel-purified high-molecular-weight eDNA extracted directly from soil collected in Boston (MA, USA). The cosmid library was screened for antibacterially active clones by using a top agar overlay which contained *Bacillus subtilis*.[1] Clones that produced a zone of growth inhibition in the overlayed *B. subtilis* lawn were recovered from the assay plates and tested for the production of extractable organic antibacterial activities.[1]

Bioassay-guided fractionation of the antibacterially active extract in ethyl acetate from cultures of CSLG18 led to the isolation of **1**, an isocyanide-functionalized C3-substituted indole. The indole and *trans* olefin were readily inferred from standard 1D and 2D NMR spectroscopic experiments.[2] The triplet resonance observed at $\Delta\delta = 107.4$ ppm in the ^{13}C NMR spectrum suggested the presence of the isocyanide functionality as ^{13}C–^{14}N coupling between the nitrogen atom and the adjacent carbon atom of the isocyanide group results in a triplet resonance for the carbon atoms, which is a hallmark of this functional group. The proposed structure of **1** was subsequently confirmed by single-crystal X-ray diffraction (Figure 1).[3] Although the *cis* isomer of **1** has been obtained from cultured bacteria, to the best of our knowledge **1** has never been reported as a natural product.[4]

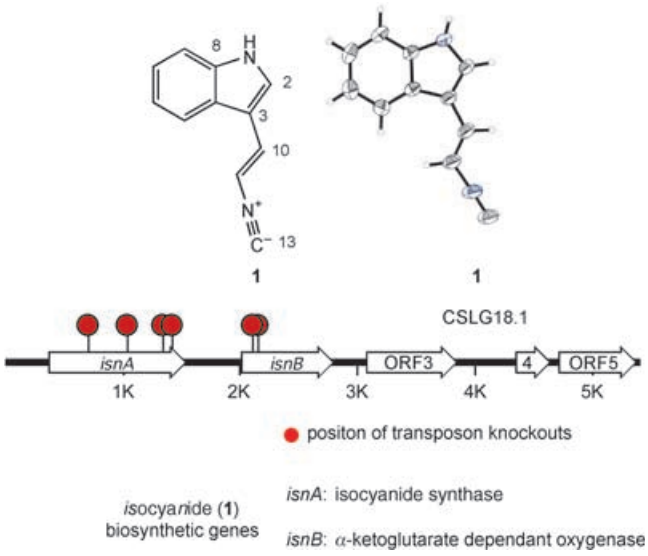


Figure 1. The natural-product antibiotic **1** was isolated from the culture broth of a 5.5-kB antibacterially active subclone of the eDNA cosmid, pCSLG18. The genes *isnA*/*B* are necessary and sufficient to confer the production of **1** to *E. coli*.

The genes responsible for the biosynthesis of **1** were sought in pCSLG18.1, a 5.5-kB antibacterially active EcoRI subclone (58.8% GC content) of the original eDNA cosmid, pCSLG18.[5] Two predicted open-reading frames (ORFs) contained in pCSLG18.1 (Figure 1) were identified by transposon mutagenesis as necessary for the production of **1** and have been given the names *isnA* and *isnB* (*is*ocya*n*ide, *iso*nitrile), respectively. The predicted translation product of *isnA* shows the highest-sequence identity to predicted proteins related to PvcA of pyoverdine chromophore biosynthesis (33%) and Dit1 of yeast-spore wall biosynthesis (21%). Neither PvcA nor Dit1 has been functionally characterized, although both are thought to be involved in the biosynthesis of C–N bonds.[6,7] The predicted translation product of *isnB* shows the highest-sequence identity to non-heme iron α-ketoglutarate dependent oxygenases, including the oxygenase from clavaminate biosynthesis and PvcB, a second enzyme thought to be involved in pyoverdine chromophore biosynthesis.[8] Three additional ORFs (>100 amino acids) were found in CSLG18.1; ORFs 4 and 5 are related to hypothetical proteins of unknown function and ORF 3 is related to prenyltransferases.

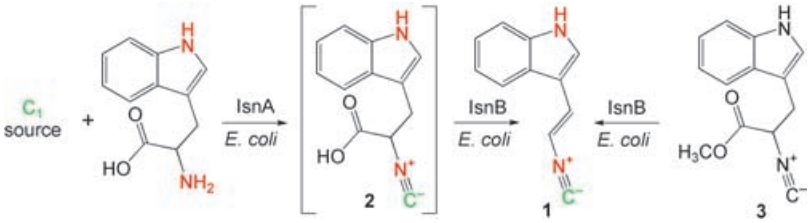
The successful expression of *isnA* and *isnB* (*isnA*/*B*), the biosynthetic genes for **1**, in *E. coli* allowed the biosynthesis of

[*] Dr. S. F. Brady, Prof. J. Clardy
Department of Biological Chemistry and Molecular Pharmacology
Harvard Medical School
240 Longwood Avenue, Boston, MA 02115 (USA)
Fax: (+1) 617-432-3702
E-mail: sean_brady@hms.harvard.edu

[**] We thank NIH CA24487 (J.C.) and the Ellison Medical Foundation (J.C.) for support of this work. We thank the *E. coli* Genetic Stock Center (CGSC) for providing the *E. coli* strains used in the feeding experiments.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

1 and, more specifically, the isocyanide functional group to be studied in detail. As with many isocyanide-containing natural products isolated from cultured bacteria and fungi, the structure of **1** suggests that the isocyanide nitrogen atom is likely to be derived from the amine of an amino acid. Feeding studies that used native-isocyanide-producing organisms have not, however, always supported this hypothesis.[9] The cloning and heterologous expression of isocyanide biosynthetic enzymes in *E. coli* allowed us to perform controlled feeding experiments which could not be carried out in less well-understood microbes and plants.



Scheme 1. Feeding experiments in *E. coli* indicate IsnA uses tryptophan in the biosynthesis of **1**. The proposed intermediate **2** is suggested by feeding experiments that use the methyl ester **3**.

The role of tryptophan in the biosynthesis of **1**, and its connection with the isocyanide nitrogen atom especially, was investigated by using *isnA*/*B* expressed in an *E. coli* host with a genetic background selected to eliminate the ambiguity seen in feeding experiments carried out in wild-type-isocyanide-producing organisms. To assess the role of tryptophan in the biosynthesis of **1**, doubly labeled ^{15}N-tryptophan was fed to a culture of a tryptophan auxotroph (*trpC*) and transaminase-deficient (*aspC*, *ilvE*, and *tyrB*) strain of *E. coli* (DL39W) transformed with *isnA*/*B*.[10] Incorporation of the ^{15}N label into **1** was determined by LCMS. The tryptophan auxotrophy ensures that the doubly labeled ^{15}N-tryptophan is the only tryptophan present in the culture, and the transaminase deficiencies ensure that the ^{15}N-labeled amine group of tryptophan is not lost during the feeding experiment. If both nitrogen atoms are incorporated into **1**, the observed *m*/*z* value $[M+H]^+$ for **1** would be 171; whereas if tryptophan is not used in the biosynthesis of **1**, no ^{15}N label will be incorporated into **1** which would result in an observed *m*/*z* value of 169. If only one of the two ^{15}N-labeled nitrogen atoms, either from the indole unit or the free amine group, is incorporated into **1** the observed *m*/*z* value for **1** would be 170.

In the presence of the doubly labeled ^{15}N-tryptophan, the general transaminase- and tryptophan-deficient strain of *E. coli* transformed with *isnA*/*B* produces **1** with *m*/*z* 171 being observed. Both the nitrogen atoms in **1**, including the isocyanide nitrogen atom, are therefore derived from tryptophan (Figure 2 and Scheme 1). Almost all of **1** obtained from a wild-type (transaminase-proficient) *E. coli*, contained a single ^{15}N label (*m*/*z* 170). Labeling studies performed in wild-type-isocyanide-producing organisms would be likely to give the same result and could explain the ambiguous feeding results obtained from these organisms previously. Many of the drawbacks traditionally associated with feeding studies can be significantly reduced or completely eliminated in model microbial systems by using mutants that create an appropriately controlled metabolic background.

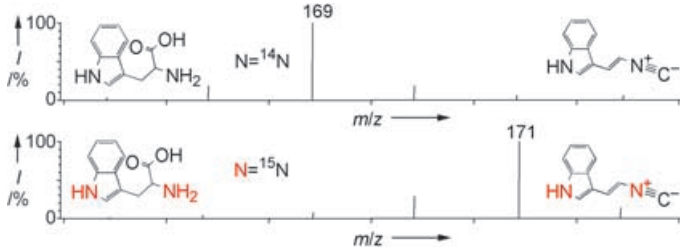


Figure 2. The *m*/*z* values observed for **1** obtained from cultures of *E. coli* transformed with *isnA*/*B* and grown in either ^{14}N-tryptophan or ^{15}N-tryptophan. A transaminase-deficient (*aspC*, *ilvE*, and *tyrB*) *E. coli* strain was used in these feeding experiments to prevent the loss of the labeled amine group.

IsnA and IsnB individually expressed as glutathione-*S*-transferase fusion proteins (IsnA-GST and IsnB-GST, respectively) were used to elucidate the overall scheme for the biosynthesis of **1** (Table 1). Cultures of *E. coli* that overexpressed either IsnA-GST or IsnB-GST individually did not accumulate **1**. However, **1** accumulated in the culture broth when these two strains were cocultured in the same flask. IsnA and IsnB are therefore necessary and sufficient to confer the production of **1** to *E. coli*, and one of the two enzymes must produce a diffusible intermediate that is converted into **1** by the other. *E. coli* that express IsnA-GST grown in media preconditioned with the growth of *E. coli* that express IsnB-GST does not lead to the accumulation of **1** in the culture broth; however, cultures of *E. coli* that express IsnB-GST grown in media preconditioned with the growth of *E. coli* that express IsnA-GST do accumulate **1** in the culture broth. The diffusible intermediate is therefore produced by IsnA and converted into **1** by IsnB (Scheme 1). All attempts to isolate this intermediate from cultures that either overexpress IsnA-GST or are transformed with *isnA*/*B* were unsuccessful.

Table 1: Results from feeding experiments used to decipher the overall biosynthetic approach to **1**.

E. coli (strain 1)	*E. coli* (strain 2)	Culturing treatment	**1**[a]
isnA + *isnB*			+
isnA–GST			–
isnB–GST			–
isnA–GST	*isnB*–GST	cocultured[b]	+
isnA–GST	*isnB*–GST	preconditioned[c]	+
isnB–GST	*isnA*–GST	preconditioned[c]	–
isnB–GST		*N*-methyl tryptophan[d]	–
isnB–GST		*N*-formyl tryptophan[d]	–
isnB–GST		*N*-methyl tryptamine[d]	–
isnB–GST		*N*-formyl tryptamine[d]	–

N-formyl and *N*-methyl compounds have been proposed as possible isocyanide intermediates in what could be a two- or three-step isocyanide biosynthetic scheme.[11] However, *N*-formyl and *N*-methyl derivatives of tryptophan and tryptamine (0.1 mg mL^{-1}) did not serve as substrates for the

production of **1** when added to cultures that overexpress IsnB-GST. It seemed plausible that IsnA might form the isocyanide function in a single enzymatic step and that the role of IsnB would be the oxidative decarboxylation of the proposed isocyanide-functionalized intermediate **2**. In our hands, the synthetic intermediate **2** was not stable and, therefore, could not be used in feeding experiments to test this hypothesis. However, we were able to synthesize the methyl ester of **2** (**3**) and thought that the slow cleavage of the methyl ester by promiscuous bacterial esterases might circumvent the instability problem.[12] The addition of **3** ($0.1\ mg\,mL^{-1}$) to cultures that express IsnB-GST led to the accumulation of **1** in the culture broth, and when the cultures of CSLG18.1 were spiked with **3**, they showed increased production of **1** relative to the control cultures of CSLG18.1. Compound **1** was not detected in the extracts in ethyl acetate from the vector control cultures spiked with **3** or from IsnA-GST cultures spiked with **3**. These feeding studies suggest that the isocyanide group seen in **1** is formed in a single enzymatic step on IsnA and that the proposed isocyanide-containing intermediate **2** is oxidatively decarboxylated by IsnB to give **1** (Scheme 1). The lack of stability of **2** is likely to be responsible for our inability to isolate the proposed intermediate.

The eDNA approach, used in this study to discover antibiotic **1**, directly couples natural products with their biosynthetic gene clusters, thus providing both a means for the isolation of previously inaccessible natural products from uncultured bacteria and a direct link to their biosynthetic enzymes. Characterization of the antibacterially active eDNA clone CLSG18 led to the identification of the isocyanide-containing natural-product antibiotic **1**, and in turn to the first isocyanide synthase, IsnA. The origin of the isocyanide carbon atom is an intriguing, but unanswered question, for all naturally occurring isocyanides. This question will be addressed in due course.

Received: June 4, 2005
Published online: October 5, 2005

Keywords: biosynthesis · DNA · heterologous expression · isocyanides · natural products

[1] a) S. F. Brady, J. Clardy, *J. Am. Chem. Soc.* **2000**, *122*, 12903; b) S. F. Brady, C. J. Chao, J. Handelsman, J. Clardy, *Org. Lett.* **2001**, *3*, 1981; c) S. F. Brady, C. J. Chao, J. Clardy, *J. Am. Chem. Soc.* **2002**, *124*, 9968; d) S. F. Brady, C. J. Chao, J. Clardy, *Appl. Environ. Microbiol.* **2004**, *70*, 6865; e) S. F. Brady, J. Clardy, *J. Nat. Prod.* **2004**, *67*, 1283; see the Supporting Information for library construction and screening methods.

[2] ^{13}C NMR (125 MHz, CD_2Cl_2): $\delta = 163.9$, 137.5, 130.7, 127.3, 125.1, 123.8, 121.8, 120.3, 112.4, 111.24, 107.4 ppm (t); ^1H NMR (500 MHz, CD_3OD): $\delta = 8.58$ (1H, s), 7.68 (1H, d, 7.5), 7.45 (1H, d, 8), 7.39 (1H, d, 3), 7.28 (1H, m), 7.24 (1H, m), 7.16 (1H, d, 14.5), 6.37 ppm (1H, d, 14); HRMS (ESI-TOF): *m*/*z* calcd for $C_{11}H_9N_2$: 169.0766; found 169.0774 $[M+H]^+$; see the Supporting Information for a description of the isolation procedure, structure-elucidation methods, and the ^1H and ^{13}C NMR spectra of **1**.

[3] CCDC-273844 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from the Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

[4] a) J. R. Evans, E. J. Napier, P. Yates, *J. Antibiot.* **1976**, *29*, 850; b) I. Hoppe, U. Schollkopf, *Liebigs Ann. Chem.* **1984**, 600.

[5] GenBank accession number DQ084328.

[6] a) P. Briza, M. Eckerstorfer, M. Breitenbach, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 4524; b) P. Briza, H. Kalchhauser, E. Pittenauer, G. Allmaier, M. Breitenbach, *Eur. J. Biochem.* **1996**, *239*, 124.

[7] A. Stintzi, Z. Johnson, M. Stonehouse, U. Ochsner, J. M. Meyer, M. L. Vasil, K. Poole, *J. Bacteriol.* **1999**, *181*, 4118.

[8] a) E. N. Marsh, M. D. Chang, C. A. Townsend, *Biochemistry* **1992**, *31*, 12648; b) J. E. Hodgson, A. P. Fosberry, N. S. Rawlinson, H. N. Ross, R. J. Neal, J. C. Arnell, A. J. Earl, E. J. Lawlor, *Gene* **1995**, *166*, 49.

[9] a) H. Achenbach, H. Grisebach, *Z. Naturforsch B* **1965**, *20*, 137; b) V. Bornemann, G. M. L. Patterson, R. E. Moore, *J. Am. Chem. Soc.* **1988**, *110*, 2339; c) P. J. Scheuer, *Acc. Chem. Res.* **1992**, *25*, 433.

[10] D. M. LeMaster, F. M. Richards, *Biochemistry* **1988**, *27*, 142; see the Supporting Information for a detailed description of the methods used in these feeding experiments.

[11] M. S. Puar, H. Munayyer, V. Hegde, B. K. Lee, J. A. Waitz, *J. Antibiot.* **1985**, *38*, 530.

[12] a) D. H. R. Barton, G. Bringmann, W. B. Motherwell, *Synthesis* **1980**, 68; b) N. Hashimoto, T. Aoyama, T. Shioiri, *Chem. Pharm. Bull.* **1981**, *29*, 1475.

