J|A|C|S
COMMUNICATIONS

Published on Web 09/19/2007

Natural Products from *isnA*-Containing Biosynthetic Gene Clusters Recovered from the Genomes of Cultured and Uncultured Bacteria

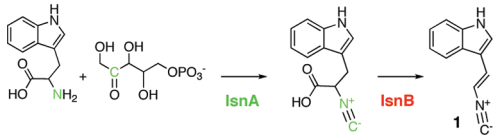
Sean F. Brady,* John D. Bauer, Michael F. Clarke-Pearson, and Rachel Daniels

Laboratory of Genetically Encoded Small Molecules, The Rockefeller University, New York, New York 10021

Received July 23, 2007; E-mail: sbrady@rockefeller.edu

One of the key insights to arise from the large-scale sequencing of bacterial genomic DNA is that traditional culture-based strategies used to discover natural products have only provided access to a small fraction of the biosynthetic diversity encoded in bacterial genomes. The complete sequencing of many bacterial genomes indicates that often only a subset of the biosynthetic gene clusters present in a genome is expressed under laboratory fermentation conditions, and the sequencing of DNA extracted directly from environmental samples suggests that the majority of bacteria present in nature have not been cultured in the laboratory.[1−5] To circumvent these limitations and more thoroughly explore the biosynthetic potential of bacteria, we used the sequence of a recently characterized isonitrile synthase, *isnA*, to search fully sequenced bacterial genomes as well as DNA extracted directly from environmental samples (environmental DNA, eDNA) for operons that use *isnA* homologues in the biosynthesis of structurally diverse natural products (Scheme 1, Figure 1). This report describes the *isnA*-containing biosynthetic operons that we found and the compounds that were produced from the expression of these operons in model bacterial hosts.

Scheme 1



The first isonitrile synthase, *isnA*, was recently cloned from DNA extracted from soil and was characterized in an *Escherichia coli*-based heterologous expression system.[7] In this system, compound **1** is produced from the heterologous expression of two eDNA-derived enzymes, IsnA and IsnB (Scheme 1). In a BLAST search of fully sequenced bacterial genomes, eight potential *isnA* homologues were initially discovered, and an examination of the DNA surrounding these predicted *isnA* homologues suggested that each was part of a small molecule biosynthetic operon.[8,9] The identification of multiple, unique *isnA*-containing operons in cultured bacteria suggested that additional *isnA*-associated biosynthetic operons were also likely to be present in the genomes of uncultured bacteria. To recover *isnA*-containing biosynthetic gene clusters from uncultured bacteria, *isnA*-specific degenerate primers were used to PCR-amplify *isnA*-related sequences from the DNA cloned in four eDNA cosmid libraries.[10] In total, twelve unique *isnA*-related sequences were amplified from the approximately 400 000 cosmid clones that were screened. These new *isnA*-related sequences were then used to recover *isnA*-containing eDNA cosmid clones, and the DNA surrounding the cloned *isnA* homologues was sequenced to reveal potential small molecule biosynthetic operons.

All of the gene clusters we identified contain an *isnA* homologue and a relative of *isnB*, the oxidative enzyme originally found with *isnA* (Figure 2). In two cases, these domains are fused into a single

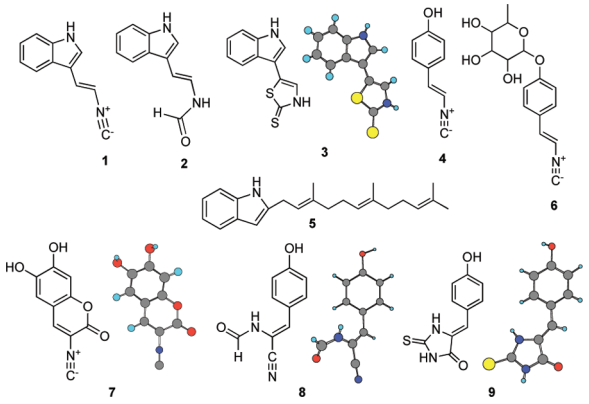


Figure 1. Clone-specific metabolites isolated from the culture broths of bacteria transformed with *isnA*-containing biosynthetic operons. Computer-generated perspective drawings are included for those metabolites that were characterized by X-ray crystallography.

predicted open reading frame (ORF). IsnA-containing operons range from simple systems with just *isnA* and *isnB* domains to larger operons that are predicted to encode the biosynthesis of more highly functionalized small molecules. The additional biosynthetic enzymes found in this collection of operons include prenyltransferases, monooxygenases, dioxygenases, carboxylases, amidotransferases, reductases, a dehydrogenase, a glycosyltransferase, and a halogenase.

To begin to access the molecules encoded by these gene clusters, each predicted *isnA*-containing operon was PCR amplified, cloned into a variety of bacterial expression vectors (pGEX-3X, pMAL-C2 or pMMB67) and transformed into either *E. coli* or *Pseudomonas aeruginosa* for expression studies. Clone-specific metabolites were detected by silica gel thin layer chromatography of the organic extracts derived from cultures of *E. coli* transformed with pathways from eDNA clone CSLG18 (**1** and **3**), eDNA clone CLRW2−34 (**5**), *Erwinia carotovora* (**4** and **6**), and *Burkholderia mallei* (**8** and **9**), as well as cultures of *P. aeruginosa* that overexpress the *isnA*-containing operon from *P. aeruginosa* (**7**) and the pathway from eDNA clone CLOLII-7 (**2**).[11,12] The structures of compounds **4**, **5**, and **6** were inferred from a combination of NMR and MS experiments, while the structures of compounds **3**, **7**, **8**, and **9** were determined by X-ray crystallography.[13] As predicted from examining the sequenced biosynthetic gene clusters, clone-specific compounds modified with an isoprene group (**5**), a sugar (**6**), an oxidized aromatic ring (**7**), and a new carbon nitrogen bond (**8**) were found in these heterologous expression studies. Some metabolites have undergone additional transformations that result in the creation of unexpected functional groups, including the nitrile in compound **8** and the heterocyclic ring systems in compounds **3**, **7**, and **9**. The identification of new metabolites from the expression of this family of operons suggests that the activation of even simple, sequenced natural product biosynthetic gene clusters using standard *E. coli*

10.1021/ja075492v CCC: $37.00

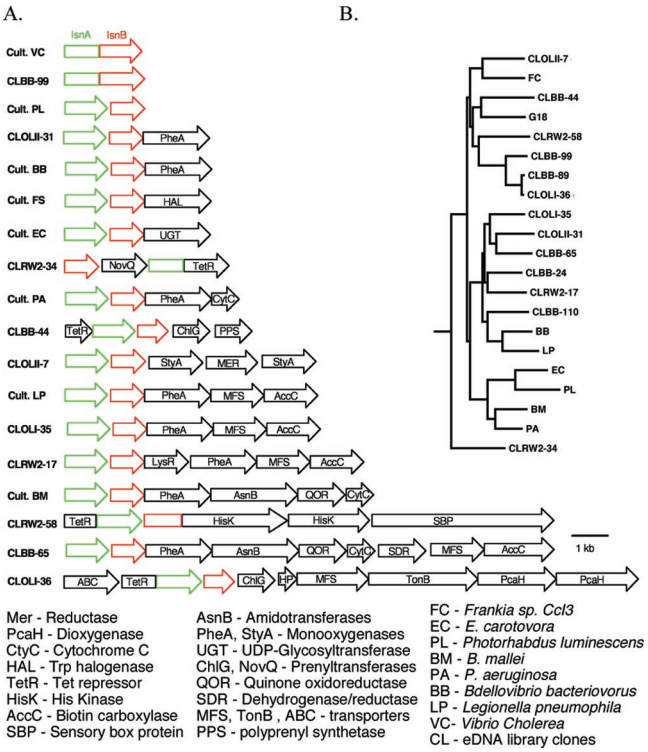


Figure 2. (A) *IsnA*-containing operons range from simple systems with just *isnA* and *isnB* domains to more complex operons that contain an array of additional biosynthetic enzymes. Genes are annotated with either an abbreviation for the general function prediction of an ORF or a commonly used gene symbol for the family of genes that is most closely related to the ORF. (B) ClustalW phylogenetic tree of *isnA* sequences.[6]

protein expression vectors should be a rewarding and straightforward strategy for the discovery of structurally novel metabolites.

While it is easy to infer the biosynthesis of some metabolites (**1**, **4**, **6**, and **7**) from the sequenced biosynthetic gene clusters, the origin of other metabolites is less clear. For example, the two sulfur-containing metabolites **3** and **9** were isolated from the culture broth of clones that also produce compounds **1** and **8**, respectively. Isothiocyanates are often isolated from the same organisms that produce isonitrile-functionalized metabolites.[14,15] It is not known if these sulfur-containing isonitrile derivatives arise enzymatically or if they arise nonenzymatically as a consequence of the inherent reactivity of the isonitrile functional group. Neither the CSLG18 (**1** and **3**) nor the *B. mallei* (**8** and **9**) derived operons contain ORFs that are predicted to add a sulfur to the heterologously produced metabolite suggesting that these and other sulfur functionalized isonitrile derivatives likely arise nonenzymatically from reactive intermediates during the fermentation process.

The genesis of many new bacterial natural products likely results from a process in which existing metabolites are modified by enzymes present in a bacterial proteome. Through successive rounds of natural selection, enzymes responsible for beneficial modifications are optimized for the biosynthesis of new metabolites.[16] For any one metabolite, or natural product substructure, this process has likely occurred numerous times, leading to a collection of evolutionarily related biosynthetic gene clusters that encode the biosynthesis of structurally distinct metabolites (Figure 2). While the recruitment of enzymes with new functions to existing gene clusters appears to be an important step in the creation of this family of *isnA*-containing operons, the ability of IsnA/IsnB homologues to produce related but not identical intermediates also appears to be an important contributor to the structural diversity seen in Figure 1. The structures derived from this family of gene clusters suggest that IsnA enzymes can biosynthesize either isonitrile (**1**, **4**, **6**, and **7**) and *N*-formyl functional groups (**2** and **8**) and that IsnB enzymes either eliminate (**1**, **4**, and **6**) or retain the carboxylic acid group (**7** and **8**).

Essentially all bacteria, from those that have not yet been cultured to those whose genomes have been completely sequenced, are potentially rich sources of, as yet, unstudied natural product biosynthetic gene clusters. The large pool of new bacterial biosynthetic clusters that can now be accessed using DNA-based screening strategies should ensure that natural products remain a productive source of structurally unique and biologically active small molecules.

Acknowledgment. This work was supported by NIH GM077516 and The Initiative for Chemical Genetics.

Supporting Information Available: Complete ref 5, methods for operon recovery and cloning, GenBank and CDCC accession numbers, and analytical data for the compounds reported here. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Hugenholtz, P.; Goebel, B. M.; Pace, N. R. *J. Bacteriol.* **1998**, *180*, 4765−74.
(2) Seow, K. T.; Meurer, G.; Gerlitz, M.; Wendt-Pienkowski, E.; Hutchinson, C. R.; Davies, J. *J. Bacteriol.* **1997**, *179*, 7360−8.
(3) Peric-Concha, N.; Long, P. F. *Drug Discovery Today* **2003**, *8*, 1078−84.
(4) Rappe, M. S.; Giovannoni, S. J. *Annu. Rev. Microbiol.* **2003**, *57*, 369−94.
(5) Venter, J. C.; et al. *Science* **2004**, *304*, 66−74.
(6) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, 4673−80.
(7) Brady, S. F.; Clardy, J. *Angew. Chem., Int. Ed.* **2005**, *44*, 7045−8.
(8) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403−10.
(9) Many *isnA* homologues have been annotated as *pvcA* after the gene from the *P. aeruginosa* pathway (Figure 2) that was thought to be involved in the biosynthesis of the pyoverdine chromophore and pseudopyoverdine, the *N*-formyl derivative of **7** (Stintzi, A.; Johnson, Z.; Stonehouse, M.; Ochsner, U.; Meyer, J.-M.; Vasil, M. L; Poole, K. *J. Bacteriol.* **1999**, *181*, 4118−4124 and Longerich, I.; Taraz, K.; Budzikiewicz, H.; Tsai, L.; Meyer, J. M. *Z. Naturforsch., C: Biosci.* **1993**, *48*, 425). Since beginning this work additional *isnA*-containing biosynthetic gene clusters have appeared in the databases. These additional sequences were not included in our analysis.
(10) Brady, S. F. *Nat. Protoc.* **2007**, *2*, 1297−305.
(11) The structure of compound **2** was inferred by the observation of a peak with a similar UV trace to that of compound **1** but with a mass of 186, 18 mass units higher than **1**. This hydrated version of compound **1** and other minor clone-specific compounds that appeared to be simple derivatives of compounds found in this study were not characterized in detail.
(12) Both the *L. pneumophila* operon and the eDNA operon from clone CLBB-65 produce colored culture broths when expressed in *E. coli* and *P. aeruginosa*, respectively. In neither case did this color partition into the organic phase. The number and diversity of molecules characterized from this collection of biosynthetic pathways would likely increase with the use of additional model hosts or by examining the aqueous culture broth.
(13) Compound **6** has been reported as a natural product from a gram-negative *Bacilli (Enterobacter sp. B20)*. Nakagawa, A.; Takahashi, S.; Miyazaki, T.; Osanai, Y.; Kosaka, K.; Nagai, K.; Arao, N.; Tanaka, K. Novel microorganism, its product and utilization thereof. U.S. Patent 7115721, 2003. A synthetic version of **5** was made in an attempt to identify new antimycobacterial agents. Kieć-Kononowicz, K.; Szymańska, E. *Il Farmaco* **2002**, *57*, 909−916.
(14) Scheuer, P. J. *Acc. Chem. Res.* **1992**, *25*, 433−439.
(15) Garson, M. J.; Simpson, J. S. *Nat. Prod. Rep.* **2004**, *21*, 164−79.
(16) Stone, M. J.; Williams, D. H. *Mol. Microbiol.* **1992**, *6*, 29−34.

JA075492V