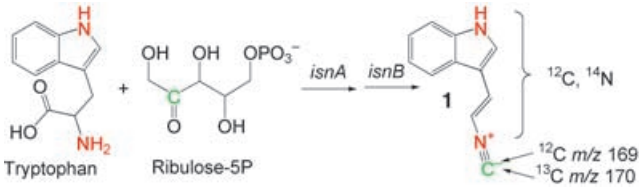
Biosynthesis

DOI: 10.1002/ange.200501942

Systematic Investigation of the *Escherichia coli* Metabolome for the Biosynthetic Origin of an Isocyanide Carbon Atom**

Sean F. Brady and Jon Clardy*

In vivo studies on the biosynthesis of natural products are often hampered by the inability to predictably control the metabolome of the producing organism. In contrast, when biosynthetic pathways are expressed in well-understood model organisms, the ability to predictably control metabolism through the use of mutant strains allows pathways to be examined with a rigor that cannot be achieved in the wild-type-producing organism. Recently, we have described the isolation and characterization of the isocyanide-containing antibiotic **1** along with its biosynthetic genes (*isnA* and *isnB*) from an environmental DNA (eDNA) clone.[1] The ability to control *Escherichia coli* metabolism systematically through the mutagenesis of genes that code for primary metabolic enzymes has enabled us to elucidate the biosynthetic origin of the isocyanide functional group found in **1** (Scheme 1).



Scheme 1. Compound **1** is derived from tryptophan and C2 of ribulose-5P. Predicted *m/z* $[M+H]^+$ values for isotopically labeled **1** are shown.

Almost fifty years after the discovery of the first isocyanide-functionalized natural product, xanthocillin, no consensus view of isocyanide biosynthesis has emerged from the extensive feeding studies that use isocyanide-producing organisms.[2,3] The origin of the carbon atom of the isocyanide group found in **1** is presented herein.

In most biosynthetic studies, a hint as to the origin of a fragment can be seen in its structure, but the lone carbon atom of an isocyanide group contains no such clue. In a normal feeding study, labeled precursors are added to an unlabeled background to decipher the origin of the atoms in a molecule; however, because of the abundance of possible isocyanide precursors, we chose to use an "inverse-labeling" strategy to study the origin of the carbon atom of the isocyanide group. In this approach, ^{12}C precursors are added to a ^{13}C background, thus eliminating the need to synthesize a ^{13}C (or ^{14}C) sample of every precursor.

Feeding studies that used native-isocyanide-producing organisms have suggested that the isocyanide carbon atom might be derived from an amino acid.[2,3] To test this hypothesis, we carried out feeding studies using *isnA*/*B*, the biosynthetic pathway for **1**, expressed in *E. coli* amino acid auxotrophs.[4] Each auxotroph was grown in minimal media that contained ^{13}C-glucose, ^{12}C-tryptophan, and the appropriate ^{12}C-amino acids needed to compensate for the auxotrophies. Compound **1** isolated from these cultures always has an observed *m/z* value of 170, which is consistent with its containing ten carbon atoms from ^{12}C-tryptophan,[5] the isocyanide carbon atom from ^{13}C-glucose, and no carbon atoms from the other ^{12}C-amino acids added to the media. By using this strategy, the amino acids colored dark blue in the overview of the *E. coli* metabolome shown in Figure 1 were ruled out as precursors in the biosynthesis of the isocyanide functional group.

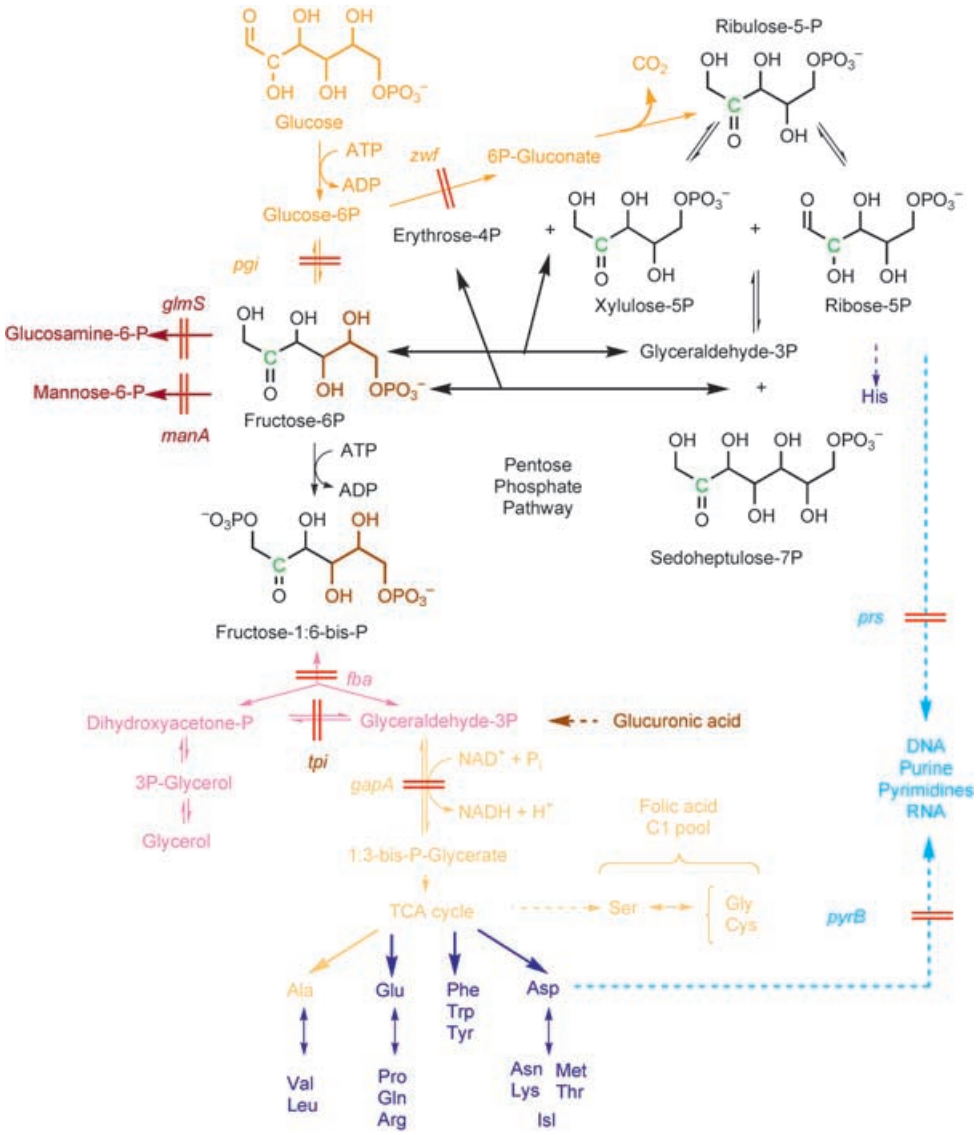
None of the amino acids tested labeled the isocyanide carbon atom, so we chose to examine the remainder of the *E. coli* metabolome systematically for the source of the isocyanide carbon atom. Figure 1 summarizes the differential labeling of **1** in a variety of *E. coli* strains deficient in key primary metabolic enzymes. Each strain was transformed with *isnA*/*B* and grown in an isotopically defined mixture of carbon sources such that unique regions of the metabolome were labeled in isotopically distinct manners. This approach allowed us to rapidly investigate the entire *E. coli* metabolome for the source of the isocyanide carbon atom. When the carbon flow between early and late glycolysis is blocked by a mutation in *gapA*, the *E. coli* metabolome is dissected into two distinct carbon pools (Figure 1). Compound **1** isolated from a *gapA* mutant transformed with *isnA*/*B* and grown in ^{13}C-glucose, ^{12}C-malate from the tricarboxylic acid cycle, and ^{12}C-tryptophan has an observed *m/z* value of 170. The isocyanide carbon atom is, therefore, derived from a descendant of ^{13}C-glucose before the conversion of glyceraldehyde-3P to 1,3-bis-phosphoglycerate by GAPDH. In addition to the dark-blue regions in Figure 1, the peach-colored area of the *E. coli* metabolome could now also be eliminated as a source of the isocyanide carbon atom.

The pentose phosphate pathway is central to the remaining portion of the *E. coli* metabolome from which the isocyanide carbon atom must be derived (Figure 1). We tested the major metabolic branches that arise from the pentose phosphate pathway as possible sources of the isocyanide carbon atom.[6] *E. coli* strains that contain mutations in genes that represent the metabolic entrance points into the following pentose pathway branch points were transformed with *isnA/B* (Figure 1):[6] glycerol-derived metabolites (*fba*), sugar phosphates (*manA* and *glmS*),

[*] 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 provision of *E. coli* strains used in the feeding experiments and Bjarne Hove-Jensen (University of Copenhagen) for the *prs* mutant.

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



Mutation(s)	^{13}C source	^{12}C source(s)	*m/z*	^{13}C/^{12}C
a.a. auxotrophs	^{13}C-glucose	^{12}C-amino acid(s)	170	^{13}C
gapA	^{13}C-glucose	^{12}C-malate	170	^{13}C
fba	^{13}C-glucose	^{12}C-3P-glycerol	170	^{13}C
manA	^{13}C-glucose	^{12}C-mannose	170/169	^{13}C/^{12}C
glmS	^{13}C-glucose	^{12}C-glucosamine	170/169	^{13}C/^{12}C
pgi/zwf	^{13}C-glucose	^{12}C-fructose	169	^{12}C
prs or *pyrB*	^{13}C-glucose	^{12}C-multiple	170	^{13}C
tpi	^{13}C-glycerol	^{12}C-glucuronic acid	170	^{13}C
zwf	^{13}C-2C-glucose	^{12}C-1,2,4,5,6C-glucose	170	^{13}C

Figure 1. The systematic differential labeling of **1** produced by *E. coli* strains deficient in specific primary metabolic enzymes (=) indicates that the isocyanide carbon atom is derived from the C2 atom (C) of an intermediate in the pentose phosphate pathway. Color is used to designate the region of metabolism that was ruled out with feeding experiments in different *E. coli* mutants. a.a. = amino acids, *gapA* = glyceraldehyde-3P dehydrogenase (GAPDH), *fba* = fructose-1,6-bisphosphate aldolase, *manA* = mannose-6-phosphate isomerase, *glmS* = L-glutamine:D-fructose-6-phosphate aminotransferase, *prs* = ribose-phosphate diphosphokinase, *pyrB* = aspartate transcarbamylase, *pgi* = phosphoglucose isomerase, *zwf* = glucose-6-phosphate-1-dehydrogenase, ATP = adenosine triphosphate, ADP = adenosine diphosphate.

nucleic acids (*prs* and *pyrB*), and glucose-derived metabolites (*pgi* and *zwf*).[7] Each strain was grown in ^{12}C-tryptophan, ^{13}C-glucose, and a ^{12}C-carbon source(s) that would supply an isotopically distinct carbon pool to the blocked region of metabolism (Figure 1).[7] In each feeding experiment, we observed that the isocyanide carbon atom was derived from the carbon source that is able to flux into the pentose phosphate pathway and not the blocked region of metabolism. In the case of *manA* and *glmS*, each of which block the exit of metabolites from the pentose phosphate pathway but not the entrance, we detected a mixture of labeled and unlabeled isocyanide groups (Figure 1). As with the other feeding studies, this suggests that the isocyanide carbon atom must be derived from the carbon source that is able to flux into the pentose phosphate pathway and not the blocked region of metabolism. The regions colored pink (glycerol metabolites, *fba*), purple (sugar phosphates, *manA* or *glmS*), orange (glucose metabolites, *pgi*/*zwf*), and light blue (nucleic acids, *prs* or *pyrB*) in Figure 1 could, therefore, now also be excluded as sources of the isocyanide carbon atom. Taken together, these feeding studies indicate that the isocyanide carbon atom is likely to be derived from one of the eight sugars directly involved in the pentose phosphate shunt (Figure 1, black), the major region of metabolism that was not eliminated by our feeding experiments.[8]

Additional feeding experiments were performed to determine which of the carbon atoms present in early glycolytic hexose intermediates feeds directly into the isocyanide carbon atom. In a triose-phosphate isomerase (*tpi*) mutant, the C1-C2-C3 and C4-C5-C6 units of early glycolytic intermediates can be differentially labeled using glycerol and glucuronic acid. Compound **1** that has been isolated from cultures of a *tpi* knockout, transformed with *isnA*/*B*, and grown in media that contain ^{13}C-glycerol and ^{12}C-glucuronic acid (Figure 1, brown) has an observed *m*/*z* value of 170. Therefore, the isocyanide carbon atom must arise from ^{13}C-glycerol,

which corresponds to the C1-C2-C3 portion of early glycolytic hexose intermediates (Figure 1, black). The carbon atoms derived from glucuronic acid (Figure 1, brown) could, therefore, also be eliminated as possible sources of the isocyanide carbon atom.

A *zwf*-deficient strain of *E. coli,* which helps to prevent the scrambling of labeled sugars in the pentose phosphate shunt by blocking the decarboxylation of 6P-gluconate, was used to determine which of the three remaining carbon atoms is incorporated into the isocyanide functionality. The *zwf*-knockout strain grown on C1-labeled ^{13}C-glucose did not produce ^{13}C-labeled **1**, whereas cultures grown on either universally labeled or C2-labeled ^{13}C-glucose almost exclusively produced ^{13}C-labeled **1** (Figure 2). In this system, the C2 atom in glucose specifically labels only six major sugars that were not excluded as a source of the isocyanide carbon atom by our other feeding studies. The six remaining metabolites (fructose-6P, fructose-1,6-bis-P, ribulose-5P, ribose-5P, xylulose-5P, and sedoheptulose-7P) are shown in Figure 1 and the C2 position of each is highlighted in green. Five of these six sugar phosphates are commercially available and were therefore tested in in vitro reconstitution experiments.[9] The in vitro reconstitution experiments using purified IsnA, IsnB, tryptophan, α-ketoglutarate, and both ribulose-5P and ribose-5P were found to produce a compound with the same retention time and mass as the naturally occurring **1**. Other commercially available C4–C7 sugars and sugar phosphates were tested in the same reconstitution system, and only arabinose-5P was found to result in isocyanide production.[10]

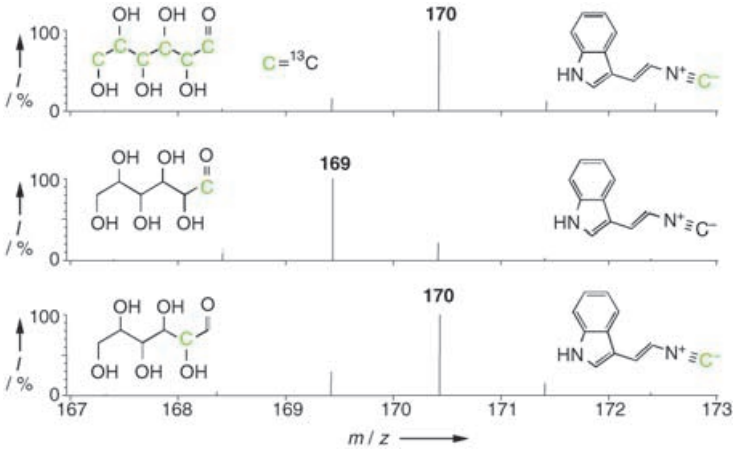
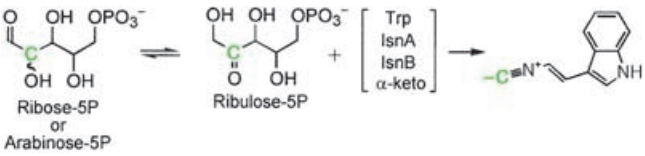


Figure 2. The *m*/*z* values observed for **1** obtained from cultures of *E. coli* transformed with *isnA*/*B* and grown in uniformly C1- and C2-labeled ^{13}C-glucose.

To investigate the source of the isocyanide carbon atom in vitro, C1- and C2-labeled ^{13}C-ribose were phosphorylated with recombinant *E. coli* ribokinase (rbsK), and each of the products was used for the in vitro biosynthesis of **1**.[11] As suggested by the in vivo feeding experiments, the C2 atom of the sugar phosphate is used in the enzymatic synthesis of **1**. C1-labeled ribose leads to the production of unlabeled **1** (*m*/*z* 169), whereas C2-labeled ribose leads to a ^{13}C-labeled isocyanide group (*m*/*z* 170). The tautomerization of ribose-5P and arabinose-5P to ribulose-5P either in solution or in an enzyme-assisted fashion could explain the use of all three sugars in the in vitro system (Scheme 2). 2-Deoxyriobose-5P, which cannot tautomerize to the equivalent keto sugar, does not serve as a substrate for the enzymatic synthesis of **1** by using this system.



Scheme 2. The tautomerization of aldo and keto sugars may explain the use of different regio- and stereochemical sugar isomers in the enzymatic synthesis of **1**.

The ability to control *E. coli* metabolism through the systematic use of strains that carry mutations in primary metabolic pathways allowed us to identify the origin of the isocyanide carbon atom in **1** as the C2 atom of ribulose-5P, or a tautomerically equivalent sugar. As shown previously, the isocyanide nitrogen atom can be traced to the amine functionality in tryptophan.[1] The isocyanide group present in many known microbial natural products can also theoretically be traced to the free amine moiety of an amino acid.[2] The proposed biosynthetic scheme for **1** may, therefore, be general for the biosynthesis of microbial-derived isocyanides. The differential labeling of natural products produced by microorganisms that carry mutations in primary metabolic genes that lead to predictable changes in the microbial metabolome can now be carried out in a systematic fashion in many model microorganisms and could be a generally useful strategy for the study of the biosynthetic origin of individual atoms found in microbial secondary metabolites.

Received: June 4, 2005
Revised: August 4, 2005
Published online: October 11, 2005

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

[1] S. F. Brady, J. Clardy, *Angew. Chem.* **2005**, *115*, 7225; *Angew. Chem. Int. Ed.* **2005**, *44*, 7063.
[2] I. Hagendorn, H. Tonjes, *Pharmazie* **1957**, *5*, 567.
[3] a) R. B. Herbert, J. Mann, *Tetrahedron Lett.* **1984**, *25*, 4263; b) R. B. Herbert, J. Mann, *J. Chem. Soc. Chem. Commun.* **1984**, 1474; c) M. S. Puar, H. Munayyer, V. Hegde, B. K. Lee, J. A. Waitz, *J. Antibiot.* **1985**, *38*, 530; d) J. E. Baldwin, H. S. Bansal, J. Chondrogianni, L. D. Field, A. A. Taha, V. Thaller, *Tetrahedron* **1985**, *41*, 1931; e) V. Bornemann, G. M. L. Patterson, R. E. Moore, *J. Am. Chem. Soc.* **1988**, *110*, 2339; f) C. W. J. Chang, P. J. Scheuer, *Comp. Biochem. Physiol. Part B* **1990**, *97*, 227; g) P. J. Scheuer, *Acc. Chem. Res.* **1992**, *25*, 433; h) K. M. Cable, R. B.

Herbert, A. R. Knaggs, J. Mann, *J. Chem. Soc. Perkin Trans. 1* **1991**, 595.

[4] Strains used in the amino acid feeding studies are referenced as follows: strain name (*E. coli* genetic stock center (CGSC) number) reference: a) Strain AB1359 (CGSC 1359) A. L. Taylor, M. S. Thoman, *Genetics* **1964**, *50*, 659; b) Strain S1228 (CGSC 6432) A. del Campillo-Campbell, A. Campbell, *J. Bacteriol.* **1982**, *149*, 469; c) Strain DG30 (CGSC 5799) D. H. Gelfand, R. A. Steinberg, *J. Bacteriol.* **1977**, *130*, 429; d) Strain KL285 (CGSC 4310) S. J. Clarke, B. Low, W. Konigsberg, *J. Bacteriol.* **1973**, *113*, 1096.

[5] All feeding experiments were carried out in the presence of ^{12}C-tryptophan, the tryptophan portion of **1** is, therefore, always derived from a ^{12}C source; see the Supporting Information for a more detailed description of the feeding-experiment methods.

[6] G. A. Sprenger, *Arch. Microbiol.* **1995**, *164*, 324.

[7] Strains used in the differential labeling studies are referenced as follows: relevant mutation(s), strain name (CGSC number) reference: a) *gapA*, DS112 (CGSC 7563) F. D. Seta, S. Boschi-Muller, M. L. Vignais, G. Branlant, *J. Bacteriol.* **1997**, *179*, 5218; b) *fba*, JM2087 (CGSC 6806) E. O. Davis, M. C. Jones-Mortimer, P. J. Henderson, *J. Biol. Chem.* **1984**, *259*, 1520; c) *manA*, JE5511 (CGSC 5505) Y. Hirota, H. Suzuki, Y. Nishimura, S. Yasuda, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 1417; d) *manA*, F500/GMS724 (CGSC 6673) M. Novel, G. Novel, *J. Bacteriol.* **1976**, *127*, 406; e) *glmS*, E111 (CGSC 5393) H. C. Wu, T. C. Wu, *J. Bacteriol.* **1971**, *105*, 455; f) *zwf* and *pgi*, DF2000 (CGSC 4873) D. G. Fraenkel, S. Banerjee, *Genetics* **1972**, *71*, 481; g) *prs*, HO733; B. Hove-Jensen, *J. Bacteriol.* **1996**, *178*, 714; h) *pyrB*, Hfr 3000 pyr (CGSC 6851) J. R. Beckwith, A. B. Pardee, R. Austrian, F. Jacob, *J. Mol. Biol.* **1962**, *5*, 618; i) *tpi*, AA200 (CGSC 5570) A. Anderson, R. A. Cooper, *J. Gen. Microbiol.* **1970**, *62*, 329; j) *zwf*, K10-15-16 (CGSC 4848).

[8] Metabolites downstream of xylulose-5P and ribulose-5P were not ruled out in these experiments.

[9] Xylulose-5P was not tested because it was not commercially available.

[10] The following sugars were tested in reconstitution experiments: ribulose-5P, ribose-5P, arabinose-5P, erythrose-4P, ribulose, ribose, arabinose, xylose, xylulose, glucose-6P, glucosamine-6P, 6P-gluconate, fructose, fructose-6P, fructose-1:6-bis-P, mannose, mannose-6P, allose, sedoheptulose-7P, and 2-deoxyribose-5P; only a small quantity of **1** is produced in the reconstitution system; experiments are being performed to look for possible cofactor and Fe requirements (see the Supporting Information for cloning, protein-expression, and reconstitution procedures).

[11] a) J. N. Hope, A. W. Bell, M. A. Hermodson, J. M. Groarke, *J. Biol. Chem.* **1986**, *261*, 7663; b) C. E. Andersson, S. L. Mowbray, *J. Mol. Biol.* **2002**, *315*, 409.

