Published in final edited form as:

Org Lett. 2017 March 17; 19(6): 1426–1429. doi:10.1021/acs.orglett.7b00355.

Identification and Characterization of Enzymes Catalyzing Pyrazolopyrimidine Formation in the Biosynthesis of Formycin A

Yeonjin Ko^{†,§}, Shao-An Wang^{†,§}, Yasushi Ogasawara^{‡,§}, Mark W. Ruszczycky[‡], and Hung-wen Liu^{*,†,‡}

[†]Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States

[‡]Division of Chemical Biology and Medicinal Chemistry, College of Pharmacy, University of Texas at Austin, Austin, Texas 78712, United States

Abstract

Genome scanning of *Streptomyces kaniharaensis*, the producer of formycin A, reveals two sets of *purA*, *purB*, *purC*, and *purH* genes. The Pur enzymes catalyze pyrimidine assembly of purine nucleobases. To test whether enzymes encoded by the second set of *pur* genes catalyze analogous transformations in formycin biosynthesis, formycin B 5'-phosphate was synthesized and shown to be converted by ForA and ForB to formycin A 5'-phosphate. These results support that For enzymes are responsible for formycin formation.

Graphical abstract

Authors are required to submit a graphic entry for the Table of Contents (TOC) that, in conjunction with the manuscript title, should give the reader a representative idea of one of the following: A key structure, reaction, equation, concept, or theorem, etc., that is discussed in the manuscript. Consult the journal's Instructions for Authors for TOC graphic specifications.

Formycin A (i.e., 8-aza-9-deazaadenosine (**1**)) is an antibiotic isolated from *Nocardia interforma*¹ and *Streptomyces kaniharaensis* SF-557.² It has been reported to demonstrate antiproliferative properties with respect to Ehrlich carcinoma, mouse leukemia L-1210, Yoshida sarcoma and HeLa cells.³ It also has antiviral activity against influenza virus A1⁴ and human immunodeficiency virus type 1.⁵ Being a C-nucleoside analog of adenosine (**2**),

*Corresponding Author: h.w.liu@mail.utexas.edu.

#Present Addresses: Laboratory of Applied Biochemistry, Graduate School of Engineering, Hokkaido University, Sapporo, Hokkaido 060-8628, Japan.

§Author Contributions: Y.K. and S.-A.W. contributed equally to this work.

Notes: The authors declare no competing financial interests.

Experimental details including synthesis of compounds, genomic analysis, gene cloning and expression, protein isolation and analytical methodologies (PDF)

Supporting Information: The Supporting Information is available free of charge on the ACS Publications website.

formycin A has been shown to be a potent inhibitor of various enzymes that accept adenosine as a substrate such as adenosine kinase from *Mycobacterium tuberculosis*, which is involved in the purine salvage pathway.⁶ Likewise, formycin A and its analogs are also known to inhibit bacterial purine nucleoside phosphorylase, which has pharmaceutical significance.⁷

While the biological activities of formycin A and its derivatives have been well documented, how they are biosynthesized in nature has remained obscure. Early *in vitro* studies revealed that formycin B 5'-phosphate (**12b**) is a likely precursor to formycin A (**1**).⁸ Feeding experiments in *N. interforma* showed that the ribosyl moiety of formycin A is derived directly from ribose likely via an intermediary produced from phosphoribosyl pyrophosphate (PRPP, **3**).⁹ However, the N3 and N8 nitrogen centers of formycin A originate from the ϵ -amino group of L-lysine (**5**).¹⁰ Importantly, L-glutamate (**4**) was found to be asymmetrically incorporated into formycin A such that the C6, C5, C4, and C9 carbons of formycin A are derived from the C1, C2, C3, and C4 carbons of glutamate (see Figure 1).^{10b,11a} Similar incorporation patterns were also observed in the formation of other C-nucleosides, such as pyrazomycin, showdomycin, and minimycin,¹¹ making L-glutamate a common precursor of this class of nucleoside antibiotics. These incorporation patterns differ from those observed for adenosine (**2**), where the C4, C5 and N7 are from glycine (**6**), N3 and N9 are from glutamine (**7**), and C6 is from bicarbonate, while C2 and C8 are from *N*¹⁰-formyl-tetrahydrofolate (*N*¹⁰-formyl-THF, see Figure 1).¹² Thus, despite the marked structural resemblance between formycin A (**1**) and adenosine (**2**), they are expected to be biosynthesized via distinct pathways. Although early isotope tracer experiments have shed some light on the possible precursors of formycin A, the link between these precursors and the final product remains a mystery.

To investigate the biosynthesis of formycin A, the genome of *S. kaniharaensis* was sequenced to generate 3011 contiguous sequences (contigs) with an average size of 2641 bp. Although the relatively short size of the contigs hampered gene assembly, preliminary genomic analysis revealed the presence of two sets of *purA*, *purB*, *purC*, and *purH* genes. The *pur* genes are known to encode enzymes catalyzing the transformation of carboxyaminoimidazole ribonucleotide (CAIR) to adenosine 5'-monophosphate (**8a**→**14a**) during the biosynthesis of purine nucleobases (see Scheme 1).¹³ Given the high sequence similarity (ca. 90%) among the conserved adenosine biosynthetic genes (i.e., the *pur* genes) present in all *Streptomyces* species, the genuine *pur* genes responsible for purine nucleotide biosynthesis in *S. kaniharaensis* could be readily identified. In contrast, the second set of *pur*-like genes other than the PurB homolog exhibit only ca. 50% sequence similarity to the genuine *pur* genes found in *S. kaniharaensis* (see Table S1). This distinction suggests that the second set of *pur*-genes, which will be referred to as the *for* genes, are more likely involved in a secondary rather than a primary metabolic pathway.

Even though formycin A (**1**) contains a pyrazolopyrimidine nucleobase instead of the imidazolopyrimidine base observed in adenosine (**2**), the presence of a seemingly second copy of *pur* genes in *S. kaniharaensis* strongly hinted that construction of the nucleobases in both compounds may follow similar biosynthetic routes. Namely, the *for* genes in *S. kaniharaensis* may be involved in formycin A biosynthesis. In particular, it was hypothesized

that ForA, B, C & H are responsible for catalyzing the transformation of **8b** to formycin A (**1**) as shown in Scheme 1. Formycin B 5'-phosphate (**12b**) represents a late intermediate in this putative pathway, and the gene products of *forA* (*purA* homolog) and *forB* (*purB* homolog) are proposed to be responsible for its conversion to formycin A 5'-phosphate (**12b**→**13b**→**14b**).

To test this hypothesis, formycin B 5'-phosphate (**12b**), the putative substrate for ForA, was synthesized according to reported procedures with some modifications as shown in Scheme 2 and described in the Supporting Information.¹⁸⁻²⁴ Meanwhile, the *forA* gene (GenBank accession no. KY629733) was heterologously overexpressed and isolated as an *N*-terminal His₆-tagged protein in *Escherichia coli*. When 1 mM **12b** was incubated with 10 μM ForA in the presence of 2.5 mM L-aspartate and 0.125 mM guanosine triphosphate (GTP) in 25 mM potassium phosphate buffer (pH 8) containing 16 mM magnesium acetate, 1 mM phosphoenolpyruvate (PEP), and 2 units pyruvate kinase at room temperature, the consumption of **12b** could be observed by anion exchange HPLC (see Figure 2). Pyruvate kinase and PEP are used to regenerate GTP from GDP in this assay. While no new peak was discernible by HPLC likely due to prolonged retention on the anion exchange HPLC column, analysis by reversed phase HPLC did reveal a new peak eluting near the void volume. High resolution LC-MS analysis of the reaction mixture demonstrated a signal at 462.0658 *m/z* (negative ion mode), which is consistent with the mass calculated for the [M – H][–] ion of **13b** (462.0668 Da). Steady state kinetic parameters for ForA were determined by coupling pyruvate formation to oxidation of NADH in the presence of lactate dehydrogenase, which can be monitored by UV absorbance. At 2.5 mM L-aspartate and 0.125 mM GTP, the *k_{cat}* for saturating **12b** is 18 ± 2 min^{–1} and the *K_M* for **12b** is 74 ± 25 μM. These are comparable to the reported kinetic constants of *E. coli* PurA (*k_{cat}* of 82 ± 7 min^{–1} and *K_M* for **12a** of 21 ± 2 μM at 5 mM L-aspartate and 0.2 mM GTP).²⁵

The *forB* gene (GenBank accession no. KY629734) was also heterologously expressed in *E. coli* and purified as an *N*-His₆-tagged protein in order to further investigate the reactions catalyzed by the *for* gene cluster. When both ForA and ForB were incubated with **12b**, a new peak appeared with a retention time of 14 min by anion exchange HPLC (see Figure 2, trace C). High resolution ESI-MS analysis (negative ion mode) of the isolated peak demonstrated a signal at 346.0556 *m/z*, which matches the predicted monoisotopic mass of the [M – H][–] ion of formycin A 5'-phosphate (**14b**) (346.0558 Da). Treatment of the ForA/ForB reaction product with calf intestinal alkaline phosphatase (CIP) resulted in a new species observed by reversed phase HPLC (Figure 2, trace E) that co-eluted with a formycin A standard (Figure 2, trace F). Isolation and MS analysis of this reaction product revealed a signal at 266.0893 *m/z* (negative ion mode) consistent with the [M – H][–] ion of formycin A (266.0895 Da). These observations demonstrate that ForA and ForB catalyze the introduction of the C6 amino group to generate formycin A 5'-phosphate (**14b**) from formycin B 5'-phosphate (**12b**) with L-aspartate serving as the nitrogen donor. These activities are analogous to those of PurA and PurB during the biosynthesis of adenosine (see Scheme 1).

Complementary to the genomic analyses, these observations provide experimental evidence that ForA and ForB represent components of the formycin A biosynthetic pathway. In order to establish that the *for* and *pur* gene clusters encode enzymes for two distinct biosynthetic

pathways, we sought to determine the substrate specificity of the ForA and PurA enzymes. The *purA* gene (GenBank accession no. KY629735) was heterologously overexpressed and purified as a C-terminal His₆-tagged protein in *E. coli*. Despite the sequence similarity of these enzymes (41% identity, 58% similarity) and the structural resemblance of the ForA substrate **12b** to the PurA substrate inosine 5'-monophosphate (**12a**, IMP), no turnover was observed upon incubation of **12b** with PurA (Figure S2). This result implies that *S. kaniharaensis* cannot produce formycin A using the Pur enzymes alone, and therefore a second biosynthetic gene cluster is indeed necessary.

In contrast to PurA, ForA can accept both **12b** and **12a** as substrates though **12b** is preferred. The relative specificity of ForA for the two substrates can be assessed quantitatively from the ratio of the corresponding specificity constants (i.e., k_{cat}/K_M) determined from a competition experiment.²⁶ This was performed by adding ForA to a mixture of approximately 0.15 mM **12b** and a 3–30-fold excess of **12a** in the presence of 5 mM L-aspartate, 0.5 mM GTP, 0.2 mM 5-aminoimidazole-4-carboxamide 1- β -D-ribofuranoside (AICAR, internal standard), 1 mM PEP, 1 mM MgCl₂, and 1 unit pyruvate kinase in 20 mM HEPES buffer at pH 7.7. As shown in Figure 3, the HPLC peaks for **12a** and **12b** from reaction aliquots at some time t could be separately integrated and standardized based on their extinction coefficients at the detector wavelength (278 nm) and the AICAR peak integration to yield the observables $y(t)$ and $x(t)$, respectively. These observables are related according to

$$\log y(t) = \beta_2 \log x(t) + \beta_1 \quad (1)$$

where β_2 is the reciprocal of the desired specificity constant ratio and β_1 is a nuisance parameter. Approximately 7 time points were taken per experiment and the entire competition experiment was repeated 5 times. The parameter β_2 was then estimated from the 5 data sets using a linear regression model developed from Eqn. (1) indicating that ForA is approximately 8-fold (95% conf. int.: 6.9–9.2) more specific for **12b** than **12a** (see Supporting Information for details). This supports the hypothesis that the *for* gene cluster is more than a redundant copy of the *pur* cluster and is instead specific for formycin A biosynthesis.

In conclusion, the discovery of two sets of homologous genes (i.e., *pur* and *for*) in the genome of *S. kaniharaensis* led to the hypothesis that the pyrazolopyrimidine of formycin A is biosynthesized in a manner analogous to the formation of the imidazolopyrimidine moiety in adenosine. This hypothesis was supported by *in vitro* demonstration of the catalytic activities of two enzymes (ForA and ForB) encoded by the *for* genes. Whereas enzymes of the *pur* gene cluster are specific for pyrimidine nucleoside biosynthesis, the *for*-encoded enzymes are selective for pyrazolopyrimidine nucleoside biosynthesis. This implies that the two sets of enzymes are indeed responsible for the production of different metabolites with the *for* genes encoding enzymes for the biosynthesis of formycin A. Furthermore, ForA together with ForB catalyze the conversion of formycin B 5'-phosphate (**12b**) to formycin A 5'-phosphate (**14b**) analogous to the reactions catalyzed by PurA and PurB during the

biosynthesis of adenosine. This study shows how sequence information garnered by genome scanning can be directly translated into the discovery of an uncharted biosynthetic pathway. Further investigation of formycin biosynthesis is being pursued with an emphasis on identifying the full gene cluster and reconstructing the biosynthetic pathway *in vitro*. Detailed knowledge of the biosynthesis of *C*-nucleoside antibiotics is essential for the repurposing and modification of these pathways to produce new antibiotics with improved biomedical properties.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by grants from the National Institutes of Health (GM035906 and GM040541) and the Welch Foundation (F-1511).

References

1. Hori M, Ito E, Takita T, Koyama G, Takeuchi T, Umezawa H. J Antibiotics Ser A. 1964; 17:96–99.
2. Suhadolnik, RJ. Nucleoside Antibiotics. Wiley-Interscience; New York, N Y: 1970.
3. a) Ishizuka M, Takeuchi T, Nitta K, Koyama G, Hori M, Umezawa H. J Antibiotics Ser A. 1964; 17:124–126. b) Ishizuka M, Sawa T, Hori S, Takayama H, Takeuchi T, Umezawa H. J Antibiot. 1968; 21:5–12. [PubMed: 5673295]
4. Takeuchi T, Iwanaga J, Aoyagi T, Umezawa H. J Antibiotics Ser A. 1966; 19:286–287.
5. Dapp MJ, Bonnac L, Patterson SE, Mansky LM. J Virol. 2014; 88:354–363. [PubMed: 24155391]
6. Long MC, Parker WB. Biochem Pharmacol. 2006; 71:1671–1682. [PubMed: 16620788]
7. a) Bzowska A, Kulikowska E, Shugar D. Biochim Biophys Act. 1992; 1120:239–247. b) Kierdaszuk B, Modrak-Wójcik A, Wierzchowski J, Shugar D. Biochim Biophys Act. 2000; 1476:109–128.
8. Sawa T, Fukagawa Y, Homma I, Wakashiro T, Takeuchi T, Hori M, Komai T. J Antibiot. 1968; 21:334–339. [PubMed: 4973121]
9. Kunimoto T, Sawa T, Wakashiro T, Hori M, Umezawa H. J Antibiot. 1971; 24:253–258. [PubMed: 5572753]
10. a) Ochi K, Iwamoto S, Hayase E, Yashima S, Okami Y. J Antibiot. 1974; 27:909–916. [PubMed: 4468276] b) Ochi K, Yashima S, Eguchi Y, Matsushita K. J Biol Chem. 1979; 254:8819–8824. [PubMed: 479162]
11. a) Buchanan JG, Hamblin MR, Sood GR, Wightman RH. JCS Chem Comm. 1980:917–918. b) Elstner EF, Suhadolnik RJ. Biochem. 1972; 11:2578–2584. [PubMed: 5045517] c) Isono K, Suhadolnik RJ. J Antibiot. 1977; 30:272–273. [PubMed: 863787]
12. Buchanan JM, Hartman SC. Adv Enzymol Relat Areas Mol Biol. 1959; 21:199–261.
13. a) Zalkin H, Dixon JE. Prog Nucl Acid Res Mol Biol. 1992; 42:259–287. b) Zhang Y, Morar M, Ealick SE. Cell Mol Life Sci. 2008; 65:3699–3724. [PubMed: 18712276]
14. a) Ju KS, Gao J, Doroghazi JR, Wang KKA, Thibodeaux CJ, Li S, Metzger E, Fudala J, Su J, Zhang JK, Lee J, Cioni JP, Evans BS, Hirota R, Labeda DP, van der Donk WA, Metcalf WW. Proc Natl Acad Sci USA. 2015; 112:12175–12180. [PubMed: 26324907] b) Penn K, Jenkins C, Nett M, Udvary DW, Gontang EA, McGlinchey RP, Foster B, Lapidus A, Podell S, Allen EE, Moore BS, Jensen PR. ISME J. 2009; 3:1193–1203. [PubMed: 19474814]
15. Madduri K, Stuttard C, Vining LC. J Bacteriol. 1989:299–302. [PubMed: 2492500]
16. a) Dumitru RV, Ragsdale SW. J Biol Chem. 2004; 279:39389–39395. [PubMed: 15262968] b) White RH. Biochem. 2011; 50:6041–6052. [PubMed: 21634403]
17. Aravind L, de Souza RF, Iyer LM. Biol Direct. 2010; 5:48. [PubMed: 20678224]

18. (a) Xu G, Moeller KD. *Org Lett.* 2010; 12:2590–2593. [PubMed: 20462275] (b) Zeng J, Vedachalam S, Xiang S, Liu X-w. *Org Lett.* 2011; 13:42–45. [PubMed: 21114334]
19. Le Coq A, Gorgues A. *Org Synth.* 1979; 59:10.
20. Buchanan JG, Edgar AR, Hutchison RJ, Stobie A, Wightman RH. *J Chem Soc Perkin Trans 1.* 1980:2567–2571.
21. Harusawa S, Matsuda C, Araki L, Kurihara T. *Synthesis.* 2006:793–798.
22. Zhou J, Yang M, Akdag A, Schneller SW. *Tetrahedron.* 2006; 62:7009–7013.
23. Buchanan JG, Stobie A, Wightman RH. *J Chem Soc Perkin Trans 1.* 1981:2374–2378.
24. Buchanan JG, Stobie A, Wightman RH. *Can J Chem.* 1980; 58:2624–2627.
25. Kang C, Sun N, Poland BW, Gorrell A, Honzatko RB, Fromm HJ. *J Biol Chem.* 1997; 272:11881–11885. [PubMed: 9115248]
26. Cornish-Bowden, A. *Fundamentals of Enzyme Kinetics.* 4. Wiley-Blackwell; Weinheim, Germany: 2012.

Figure 1.
Biosynthetic origins of formycin A and adenosine.

Figure 2.

In vitro assay of ForA and ForB. (A) Reaction mixture containing formycin B 5'-phosphate (12b) without enzymes; (B) 12b + ForA; (C) 12b + ForA + ForB. Reaction mixtures for traces (A) to (C) were analyzed after a 1.5 h incubation by HPLC using a Dionex CarboPac PA1 anion exchange column in a two-solvent system of H₂O versus 1.5 M ammonium acetate in H₂O (pH 7.0). (D) Isolated ForA/ForB product; (E) isolated ForA/ForB product + CIP; (F) coinjection of isolated ForA/ForB/CIP product and formycin A (1) standard. HPLC traces (D) to (F) were obtained using a Microsorb-MV 100-5 C18 column in a two-solvent system of 1% ammonium acetate in H₂O (pH 5.1) versus acetonitrile (see Supporting Information and text for details).

Figure 3.

Competition assay of ForA with **12a** and **12b**. (A) Sample HPLC traces at three time points from a single trial showing relative consumption of **12b** versus **12a**. HPLC traces utilized a Microsorb-MV 100-5 C18 column with a two-solvent system of 1% ammonium acetate in H₂O (pH 5.1) versus acetonitrile. (B) Plots of $\log([12a]/b_i)$ versus $\log([12b]/a)$ for the five separate trials where b_i and a are scaling factors relating the concentrations to the standardized HPLC peak integrations and shift the intercepts to facilitate visualization (see text and Supporting Information for details).

Scheme 1.

The later steps of the established biosynthetic pathway for adenosine (2) and the proposed analogous pathway for formycin A (1) formation.

Scheme 2.

Chemical synthesis of formycin B 5'-phosphate (12b).