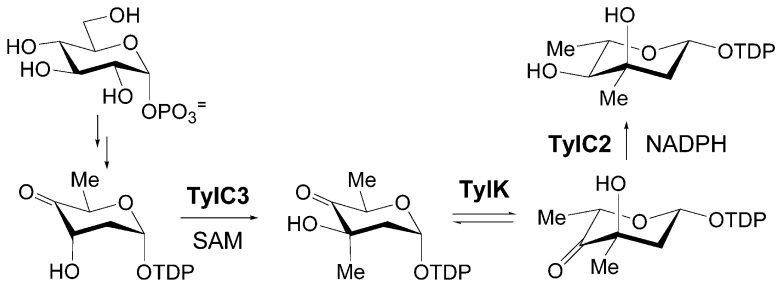
Communication

Biosynthesis of TDP-l-Mycarose: The Specificity of a Single Enzyme Governs the Outcome of the Pathway

Haruko Takahashi, Yung-nan Liu, Huawei Chen, and Hung-wen Liu

J. Am. Chem. Soc., **2005**, 127 (26), 9340-9341• DOI: 10.1021/ja051409x • Publication Date (Web): 14 June 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 4 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

Journal of the American Chemical Society is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

J|A|C|S
COMMUNICATIONS

Published on Web 06/14/2005

Biosynthesis of TDP-L-Mycarose: The Specificity of a Single Enzyme Governs the Outcome of the Pathway

Haruko Takahashi, Yung-nan Liu, Huawei Chen, and Hung-wen Liu*

Division of Medicinal Chemistry, College of Pharmacy and Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, Texas 78712

Received March 4, 2005; E-mail: h.w.liu@mail.utexas.edu

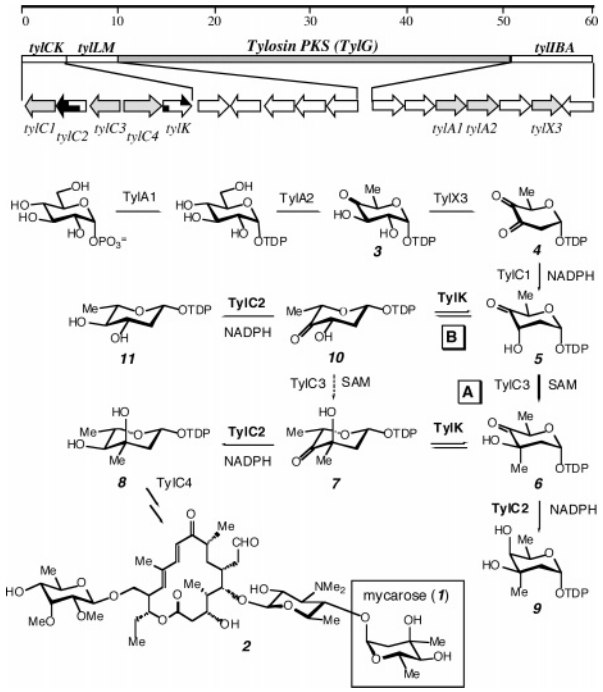
Our knowledge of the genetics for the formation of microbial secondary metabolites, including the appended unusual sugars in many cases, has grown considerably in recent years.[1] Sequence analysis of genes found in the biosynthetic gene clusters has resulted in tentative assignments for the encoded proteins and, consequently, their putative functions in the biosynthetic pathways. However, a lack of biochemical characterization of these gene products leaves many of these assignments uncertain. Mycarose (**1**), a 2,6-dideoxy-L-hexose with a C-3 methyl branch, is one of the three sugar components of the macrolide antibiotic tylosin (**2**) produced by *Streptomyces fradiae*.[2] This sugar and its derivatives are also found in olivomycin, chromomycin, erythromycin, and a few other secondary metabolites.[3] Since modification of the sugar component(s) often alters the biological activities of the parent glycoconjugates,[1c,4] a detailed understanding of mycarose biosynthesis is essential for developing strategies to control its formation so that the biological activities of the parent molecules can be varied.

Early genetic studies identified the entire gene cluster for the biosynthesis of tylosin (Scheme 1)[5] and revealed that the genes involved in the biosynthesis of mycarose are in the *tylCK* and *tylIBA* regions. These two regions were later sequenced by us and others from which 12 open reading frames (ORFs) were identified.[6,7] The genes believed to participate in mycarose formation, as shown in Scheme 1, are assigned based on sequence similarities to other sugar biosynthetic genes.[8] Among them, the proteins encoded by *tylX3*, *tylC1*, and *tylC3* have been studied,[9,10] and those of *tylK* and *tylC2* remain to be characterized.

The TylX3 and TylC1 proteins had been determined to catalyze C-2 deoxygenation (**3** → **4**) and 3-keto reduction (**4** → **5**), respectively.[9] Beyond that, two pathways are possible, which differ in the order of C-3 methylation and C-5 epimerization (route A: **5** → **6** → **7**, route B: **5** → **10** → **7**). Although previous studies had shown that TylC3 is a SAM-dependent C-methyltransferase and is capable of catalyzing the conversion of **5** to **6**,[10] it is not known whether TylC3 can also process **10** to **7**. This ambiguity can be resolved by determining the substrate specificity of TylC3, or of TylK instead, which may be responsible for C-5 epimerization. As part of our efforts to establish the sequence of events in this pathway, we carried out functional analyses of the proteins encoded by *tylK* and *tylC2*, the latter of which is predicted to convert **7**, produced in either route, to **8**. The results reported herein clearly show that TylK functions as the 5-epimerase and TylC2 functions as the 4-ketoreductase. However, both enzymes exhibit relaxed substrate specificity such that the production of L-mycarose as the sole product of the pathway depends on the absolute specificity of TylC3 for **5**.

To obtain the TylK and TylC2 proteins, their respective coding sequences were amplified from the *tylCK* region of the tylosin cluster of *S. fradiae* by the polymerase chain reaction. The *tylK* gene was cloned into two expression vectors, pET-28b(+) and pET-24b(+), to be used to transform *Escherichia coli* BL21(DE3) to generate TylK with a His_6 tag at the N- and C-terminus, respectively. Analogously, the N-terminus His_6-tagged TylC2 was obtained by cloning the *tylC2* gene into pET-28b(+) and was expressed in *E. coli* BL21(DE3). These proteins were purified to near homogeneity using Ni−NTA resin (Qiagen). The analysis of each protein by size exclusion chromatography revealed that TylK exists as a homodimer with a molecular mass of 40.2 kDa, and TylC2 consists of a monomer with a mass of 37.3 kDa.[11]

Scheme 1



When the N-terminal His_6-tagged TylK was incubated with **6**, which was generated in situ from **3** in the presence of TylX3, TylC1, TylC3, and the necessary coenzymes,[10] a new product with a retention time of 19.2 min was detected by HPLC (Figure 1B, peak b).[12] Unfortunately, poor separation of this compound from **6**, which has a retention time of 17.7 min under the HPLC conditions (Figure 1A,B, peak a), hampered isolation efforts. Nevertheless, the appearance of a new set of methyl signals at δ 1.08 (d, 5-Me) and 1.12 (s, 3-Me), in addition to those assigned to **6** at δ 1.11 (d, 5-Me) and 1.37 (s, 3-Me), in the ^{1}H NMR spectrum of the incubation mixture clearly revealed the existence of a new compound, which is likely the C-5 epimerized product **7**. When this mixture was incubated with TylC2 and NADPH, a new HPLC peak with a retention time of 13.7 min appeared (Figure 1C, peak c). The corresponding product was isolated by FPLC (MonoQ) and desalted using a Sephadex G-10 column (2.5 × 50 cm, elution with H_2O at

10.1021/ja051409x CCC: $30.25

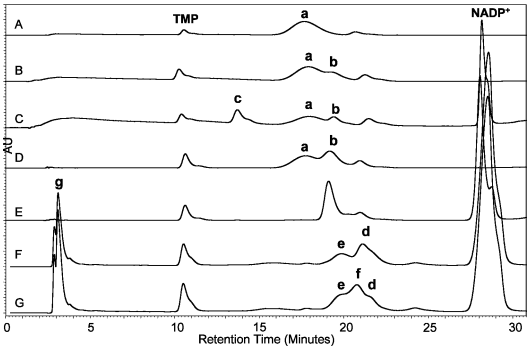


Figure 1. HPLC of various incubation mixtures. (A) Compound **6**; products resulting from the incubation of (B) N-terminal His_6-tagged TylK (260 μM) with **6** (1.5 mM) at 37 °C for 3 h; (C) TylC2 (15 μM) with NADPH (2.5 mM) along with the product mixture from B at room temperature for 3 h; (D) C-terminal His_6-tagged TylK (70 μM) with **6** (1.5 mM) at room temperature for 30 min; (E) TylC2 (7 μM) with NADPH (1.5 mM) along with the product mixture from D at room temperature for 30 min; (F) C-terminal His_6-tagged TylK (36 μM) with **5** (13 mM) at room temperature for 4 h; (G) TylC2 (8.6 μM) with NADPH (16 mM) along with the product mixture from F at room temperature for 4 h. [Peak g in (F) and (G) arises from the decomposition of **4**.]

24 mL/h). To our surprise, NMR analysis showed that this product is not the expected TDP-L-mycarose (**8**), but its D-epimer (**9**).[13] The assignment is based on the coupling constant between 4-H and 5-H (J = 1 Hz) and the nuclear Overhauser effect (NOE) between the proton at C-5 and the methyl protons at C-3. Clearly, TylC2 is a 4-ketoreductase, which appears to favor a D-sugar (such as **6**) rather than an L-sugar (such as **7**). These results initially suggested that route B might be the preferred pathway, in which the substrate available to TylC2 is **7**, and not an equilibrium mixture of **6** and **7** (via route A). This pathway would avoid the coproduction of the D-sugar **9** during the biosynthesis of L-mycarose (**8**).

Interestingly, when the C-terminal His_6-tagged TylK was incubated with **6**, the peak corresponding to the epimerized product **7** (Figure 1D, peak b) at 19.2 min was more pronounced, and its rate of formation was markedly increased.[14] This result could be due to the higher activity of the C-terminal-tagged TylK, compared with that of its N-terminal-tagged counterpart. Subsequent treatment of the incubation mixture with TylC2 and NADPH led to the complete consumption of **6**, as shown by the loss of the broad peak at 17.7 min in Figure 1E. However, no peak corresponding to **9** (at 13.7 min) could be detected. Instead, the intensity of the signal at 19.2 min increased significantly.

Isolation and characterization of species corresponding to this signal showed that it is TDP-L-mycarose (**8**) ($J_{4,5}$ = 9.8 Hz).[13] Apparently, the product of TylK (**7**) and that of TylC2 (**8**) have the same retention time under the HPLC conditions. This overlap led to the mistaken assumption that **8** was not present in the previous experiment (Figure 1C, peak b contains both **7** and **8**).[15] Thus, this observation suggests that route B is not necessarily the preferred pathway for the biosynthesis of L-mycarose as previously surmised.

To test whether TylK can also process **5** to **10**, the C-terminal His_6-tagged TylK was incubated with **5**, which was generated in situ from **3** via the reaction of TylX3 and TylC1,[9] and the reaction was monitored by HPLC. As shown in Figure 1F, a new product with a retention time of 21.1 min (peak d), in addition to **5** which appeared at 19.8 min (peak e), was observed. Its identity was deduced to be **10** based on the fact that it could be converted to **11** by the action of TylC2/NADPH (Figure 1G, peak f). The structure of **11** was confirmed by spectral analyses.[13] Interestingly, when the mixture of **5**/**10**, generated by the action of TylK followed by ultrafiltration to remove TylK, was incubated with TylC3/SAM, only compound **6** (and no **7**) was produced.[13] These results clearly demonstrate that compound **10** is not a substrate of TylC3 and that route B is not a part of the biosynthetic pathway.

The above experiments provide critical evidence assigning the roles of the key enzymes involved in TDP-L-mycarose biosynthesis and complete the reconstitution of the entire mycarose pathway in vitro. TylK can epimerize C-5 of **5**/**10** and **6**/**7**, and TylC2 can reduce the 4-keto group of **6**, **7**, and **10**.[16] The relaxed substrate specificity exhibited by these two enzymes could result in a mixture of products and prematurely terminate the pathway. However, the stringent substrate specificity of TylC3 for **5** maintains progression of the pathway toward the formation of L-mycarose via steps **6** → **7** → **8**.[17] Thus far, only a handful of unusual sugar biosynthetic pathways have been biochemically verified. An appreciation of the complexity of a multistep biosynthesis in vivo and in vitro is essential for utilizing the biosynthetic machinery in a combinatorial biology approach to produce "tailor-made" molecules or diverse structure entities. The promiscuous substrate specificity observed in the late steps of this biosynthetic pathway, in particular, may be exploited to promote structural diversity.

Acknowledgment. This work was supported by the National Institutes of Health Grants GM35906 and GM54346.

Supporting Information Available: Spectral data of compounds **8**, **9**, and **11**. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) McDaniel, R.; Welch, M.; Hutchinson, C. R. *Chem. Rev.* **2005**, *105*, 543−558. (b) Katz, L.; Ashley, G. W. *Chem. Rev.* **2005**, *105*, 499−527. (c) He, X. M.; Liu, H.-w. *Annu. Rev. Biochem.* **2002**, *71*, 701−754.
(2) Achenbach, H.; Regel, W.; Karl, W. *Chem. Ber.* **1975**, *108*, 2481−2483.
(3) Johnson, D. A.; Liu, H.-w. In *Comprehensive Chemistry of Natural Products Chemistry*; Barton, D., Nakanishi, K., Meth-Cohn, O., Eds.; Pergamon: New York; 1999; Vol. 3, pp 311−365.
(4) For reviews, see: (a) Weymouth-Wilson, A. C. *Nat. Prod. Rep.* **1997**, *14*, 99−110. (b) Thorson, J. S.; Hosted, T. J., Jr.; Jiang, J.; Biggins, J. B.; Ahlert, J. *Curr. Org. Chem.* **2001**, *5*, 139−167.
(5) (a) Fisherman, S. E.; Cox, K.; Larson, J. L.; Reynolds, P. A.; Seno, E. T.; Yeh, W.-K.; Van Frank, R.; Hershberger, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8248−8252. (b) Baltz, R. H.; Seno, E. T. *Annu. Rev. Microbiology* **1988**, *42*, 547−574.
(6) (a) Merson-Davies, L. A.; Cundliffe, E. *Mol. Microbiology* **1994**, *13*, 349−355. (b) Gandecha, A. R.; Large, S. L.; Cundliffe, E. *Gene* **1997**, *184*, 197−203. (c) Fish, S. A.; Cundliffe, E. *Microbiology* **1997**, *143*, 3871−3876.
(7) Chen, H.; Guo, Z.; Liu, H.-w. *J. Am. Chem. Soc.* **1998**, *120*, 9951−9952.
(8) (a) Summers, R. G.; Donadio, S.; Staver, M. J.; Wendt-Pienkowski, E.; Hutchinson. C. R.; Katz, L. *Microbiology* **1997**, *143*, 3251−3262. (c) Salah-Bey, K.; Doumith, M.; Michel, J.-M.; Haydock, S.; Cortés, J.; Leadlay, P. F.; Raynal, M.-C. *Mol. Gen. Genet.* **1998**, *257*, 542−553.
(9) Chen, H.; Agnihotri, G.; Guo, Z.; Que, N. L. S.; Chen, X. H.; Liu, H.-w. *J. Am. Chem. Soc.* **1999**, *121*, 8124−8125.
(10) Chen, H.; Zhao, Z.; Hallis, T. M.; Guo, Z.; Liu, H.-w. *Angew Chem., Int. Ed.* **2001**, *40*, 607−610.
(11) The calculated molecular mass, including the His_6-tag, for TylK is 22 165 Da, and that for TylC2 is 36 258 Da.
(12) The incubation mixture was loaded on a Dionex PA1 column (5 μm, 4.6 × 250 mm) after ultrafiltration through a YM-10 membrane to remove protein. A linear gradient from 200 to 350 mM ammonium acetate buffer over 30 min and 0.6 mL/min of flow rate was used to elute the reaction products (monitored at 267 nm).
(13) See Supporting Information for spectral data.
(14) It appears that the equilibrium between **6** and **7** was not reached when less reactive TylK was used.
(15) The fact that no **9** was detected in this incubation mixture may be explained by the preference of **7** over **6** as the substrate for TylC2 and the rapid rate of reaching equilibrium between **6** and **7** using the more reactive TylK.
(16) Interestingly, no reduced product of **5** by TylC2 was observed. Perhaps, the conversion is beyond the detection limit of HPLC analysis.
(17) Compound **5** must be a better substrate for TylC3 than **10** for TylC2.

JA051409X