Expression, Purification, and Characterization of TylM1, an *N*,*N*-Dimethyltransferase Involved in the Biosynthesis of Mycaminose

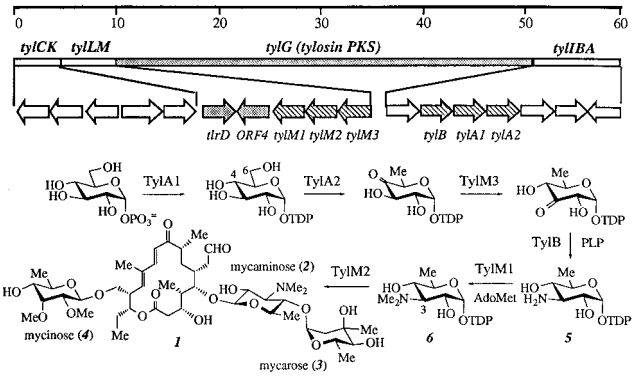
Huawei Chen, Zhihong Guo,† and Hung-wen Liu*

Department of Chemistry, University of Minnesota Minneapolis, Minnesota 55455

Received May 7, 1998

Tylosin (**1**), produced by *Streptomyces fradiae*, is an important drug used to treat veterinary Gram-positive and mycoplasma infections as well as to promote livestock growth.[1] This macrolide antibiotic is composed of a polyketide aglycon (tylactone) and three unusual sugars—mycaminose (**2**), mycarose (**3**), and mycinose (**4**). Extensive genetic and phenotypic complementation studies have led to several classes of blocked mutants, revealing the genetic organization of the tylosin biosynthetic gene (*tyl*) cluster.[2] As illustrated in Scheme 1, the *tylG* region harbors the polyketide synthase (PKS) genes, while the *tylLM*, *tylIBA*, and *tylCK* regions contain genes for mycaminose and mycarose formation.[2] In an effort to study the biosynthesis of unusual sugars found in antibiotics, we have sequenced the extended segments flanking the PKS genes from which 17 open reading frames (ORFs) in the *tylCK*, *tylLM*, and *tylIBA* regions have been identified.[3] Further analysis based on sequence similarities to other sugar biosynthetic genes,[4,5] especially those reported by Cundliffe and co-workers, who have also sequenced the *tylIBA* and *tylLM* segments of the *tyl* cluster,[6] has allowed most of these ORFs to be tentatively assigned. Three genes in the *tylIBA* region are believed to be involved in mycaminose biosynthesis[6a]—*tylA1* is for α-D-glucose 1-phosphate thymidylyltransferase, *tylA2* is for TDP-D-glucose 4,6-dehydratase, and *tylB* is likely the gene for a pyridoxal 5′-phosphate-dependent aminotransferase.[7] Other genes that have been assigned to encode enzymes involved in the mycaminose pathway reside in the *tylLM* region—*tylM1* is likely for a *S*-adenosylmethionine (AdoMet)-dependent methyltransferase, *tylM2* is for the glycosyltransferase, and *tylM3* may encode a tautomerase which displays sequence similarity to some P-450 enzymes but lacks the conserved cysteine residue that coordinates the heme iron.[6b,8] While all of these assignments are based solely on sequence analysis, the tentative identification of these genes has allowed initial speculation of their roles, leading to a possible route for mycaminose biosynthesis shown in Scheme 1.[6] To verify the proposed pathway, it is important to experimentally determine the function of each gene product. Thus, we have expressed several of the aforementioned genes and examined the catalytic properties of the purified enzymes. In this paper, we report the initial characterization of TylM1, and our results provide, for the first time, biochemical evidence establishing the role of TylM1 as the methyltransferase catalyzing *N*,*N*-dimethylation of the 3-amino group.

Scheme 1



To study the function of TylM1, the *tylM1* gene was amplified by polymerase chain reaction (PCR) and cloned into the pET-17b expression vector, and the ensuing plasmid was used to transform *Escherichia coli* BL21(DE3)pLysS host cells. Growth of the resulting construct at 30 °C and induction using isopropyl β-D-thiogalactoside (IPTG) allowed highly efficient expression of *tylM1*, in which greater than 20% of the soluble protein in the crude extract was found to be TylM1. This enzyme was purified to near homogeneity by a protocol consisting of ammonium sulfate fractionation and DEAE Sepharose and FPLC MonoQ chromatographic steps. AdoMet (0.1 mM) was included in all buffers throughout the purification to prevent the precipitation of TylM1 from the solution. Judging from a M_r of 55.2 K estimated by gel filtration and a calculated mass of 27 427 Da based on the translated sequence for each subunit, TylM1 exists as a homodimer.[9] The electronic absorption spectrum of the purified enzyme shows no absorbance above 300 nm.

The predicted substrate **5** was synthesized by the reactions delineated in Scheme 2.[10] To test whether TylM1 is the desired methyltransferase, a mixture of TylM1 (0.05 μmol), **5** (7.0 μmol), and AdoMet (40 μmol) in 1.5 mL of 50 mM potassium phosphate buffer (KP_i, pH 7.5) was incubated at 23 °C for 4 h. The enzyme

† Current address: Department of Chemistry, University of California, Berkeley, CA 94720.

(1) (a) McGuire, J. M.; Boniece, W. S.; Higgens, C. E.; Hoehn, M. M.; Stark, W. M.; Westhead, J.; Wolfe, R. N. *Antibiot. Chemother*. **1961**, *11*, 320−327. (b) Corcoran, J. W.; Huber, M. L. B.; Huber, F. M. *J. Antibiot*. **1977**, *30*, 1012−1014.

(2) (a) Baltz, R. H.; Seno, E. T. *Antimicrob. Agents Chemother*. **1981**, *20*, 214−225. (b) 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. (c) Baltz, R. H.; Seno, E. T. *Annu. Rev. Microbiol*. **1988**, *42*, 547−574.

(3) Nine other ORFs have also been identified downstream from the *tylCK* region.

(4) (a) Liu, H.-w.; Thorson, J. S. *Annu. Rev. Microbiol*. **1994**, *48*, 223−256. (b) Kirschning, A.; Bechthold, A. F.-W.; Rohr, J. In *Bioorganic Chemistry Deoxysugars, Polyketides & Related Classes: Synthesis, Biosynthesis, Enzymes*; Rohr, J., Ed.; Springer: Berlin, 1997; pp 1−84. (c) Johnson, D. A.; Liu, H.-w. In *Comprehensive Chemistry of Natural Products Chemistry*; Barton, D., Nakanishi, K., Meth-Cohn, O., Eds.; Pergamon: New York, in press.

(5) (a) Gaisser, S.; Bohm, G. A.; Cortés, J.; Leadlay, P. F. *Mol. Gen. Genet*. **1997**, *256*, 239−251. (b) Summers, R. G.; Donadio, S.; Staver, M. J.; Wendt-Pienkowski, E.; Hutchinson. C. R.; Katz, L. *Microbiology* **1997**, *143*, 3251−3262.

(6) (a) Merson-Davies, L. A.; Cundliffe, E. *Mol. Microbiol*. **1994**, *13*, 349−355. (b) Gandecha, A. R.; Large, S. L.; Cundliffe, E. *Gene* **1997**, *184*, 197−203. (c) Fish, S. A.; Cundliffe, E. *Microbiol*. **1997**, *143*, 3871−3876.

(7) Thorson, J. S.; Lo, S. F.; Liu, H.-w.; Hutchinson, R. C. *J. Am. Chem. Soc*. **1993**, *115*, 6993−6994.

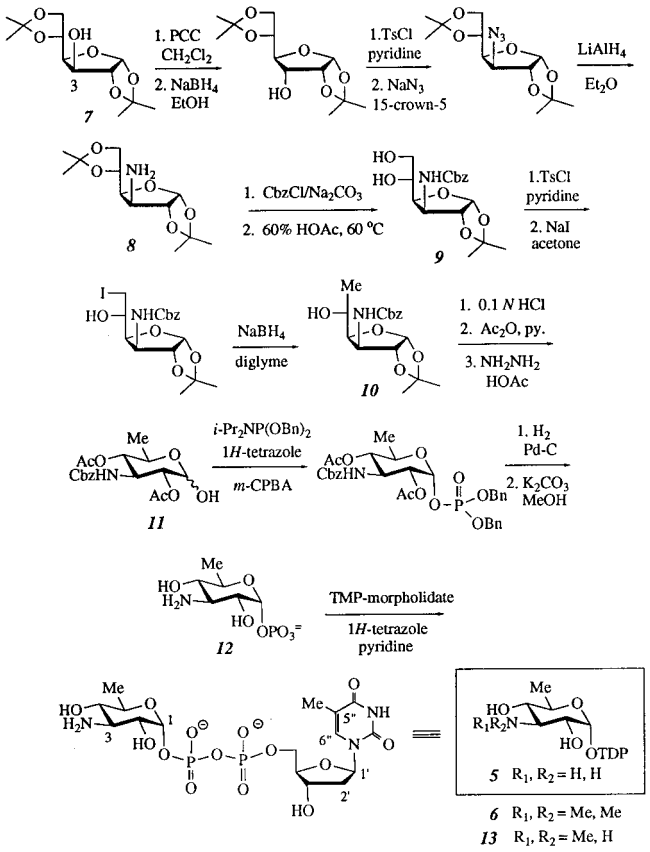
(8) The other two genes found in the *tylLM* region are[6b] *tlrD*, a resistant gene that ecodes a rRNA methyltransferase (Gandecha, A. R.; Cundliffe, E. *Gene* **1996**, *180*, 173−176), and *ORF4*, which is highly homologous to the crotonyl-CoA reductase gene (Wallace, K. K.; Bao, Z.; Hong, D.; Digate, R.; Schuler, G.; Speedie, M. K.; Reynolds, K. A. *Eur. J. Biochem*. **1995**, *233*, 954−962).

(9) N-terminal amino acid sequencing confirmed that the first 10 residues (AHSSATAGPQ) of this protein are identical to the translated *tylM1* sequence except for the deletion of the first methionine residue.

(10) Starting from 1,2:5,6-di-*O*-isopropylidene-α-D-glucose (**7**), the key intermediate **8** was prepared in five steps, which involved inversion of the C-3 hydroxyl group, nucleophilic displacement with sodium azide, and hydride reduction of the azido group, with an overall yield of 30%. After carbobenzyloxy chloride treatment to protect the C-3 amino functionality, the 5,6-isopropylidene group was selectively cleaved to give **9** (83% yield for two steps).[11] Subsequent C-6 iodination followed by $NaBH_4$ reduction gave compound **10**, which was converted to **11** via acid treatment, peracetylation, and selective removal of the 1-*O*-acetyl group with hydrazine acetate in DMF[12] (45% overall yield from **9**). Phosphate **12** was obtained from **11** in pure α-form by *m*-CPBA oxidation of the corresponding phosphite[13] followed by hydrogenation and treatment with K_2CO_3 in methanol (50% yield from **11**). After passing through an Amberlite IR-120 cation exchange column (Et_3NH^+, 1 × 30 cm), the triethylammonium salt of **12** was reacted with thymidine 5′-monophosphomorpholidate in the presence of 1*H*-tetrazole in pyridine to give **5**,[13] which was purified by Bio-Gel P-2 chromatography with 25 mM NH_4HCO_3 (60% yield).[14]

S0002-7863(98)01588-1 CCC: $15.00
Published on Web 09/11/1998

Scheme 2



was removed using a Centricon-10, and the product was isolated by FPLC MonoQ using a linear gradient of 0−0.4 M ammonium bicarbonate buffer over 25 min. Spectral characterization of the purified product confirmed that it is indeed the dimethylated hexose **6**.[15] Interestingly, when the reaction was quenched at an earlier time, two products could be detected by HPLC using an Adsorbosphere SAX column (5 μm, 4.6 × 250 mm) and 50 mM KP_i buffer (pH 3.6). The product with a retention time of 8.79 min is the dimethylated hexose **6**, and the other, at 7.16 min, is the monomethylated species **13**.[16] The time courses of mono- and dimethylation were determined by HPLC (at 279 nm) following the consumption of substrate **5** and the formation of **13** and **6**. A typical assay mixture consisted of 10 mM AdoMet,[17] 0.5 mM **5**, 10.9 μM TylM1, and 200 μL of 50 mM KP_i (pH 7.5). Aliquots of 5 μL were withdrawn at appropriate time intervals, boiled, and subjected to the aforementioned HPLC analysis using the SAX anion exchange column. The percent conversion of **5** to mono- and dimethylated products was estimated on the basis of the integration of the corresponding peaks. By fitting the data to the rate laws for an irreversible unimolecular consecutive reaction (**5** → **13** → **6**),[18] rate constants of 0.020 and 0.062 min^{-1} for the mono- and dimethylation step, respectively, were deduced with good correlation coefficients.[19] Our data clearly demonstrate that TylM1 is indeed the required methyltransferase in the biosynthesis of mycaminose, and it alone catalyzes the *N*,*N*-dimethylation, via a monomethylated intermediate, in a stepwise manner.

The above results introduce TylM1 as a new member of a small family of enzymes that are capable of catalyzing *N*,*N*-dimethylation. A few representatives of this class include: PEM-2, a phospholipid methyltransferase that catalyzes the methylation of phosphatidylethanolamine, preferentially the second and third methyl transfer steps, to form phosphatidylcholine;[20] RMT1, a protein-arginine methyltransferase that catalyzes both the N^G-mono- and N^G,N^G-asymmetric dimethylation of protein arginine residues, an important process in modulating mRNA splicing as well as growth factor localization and function;[21] and TlrA, the product of a resistance gene isolated from the tylosin producer *S. fradiae*, that catalyzes the dimethylation of a single base (A-2058) to N^6,N^6-dimethyladenine within 23S rRNA rendering the bacterial strain resistant to macrolide, lincosamide, and streptogramin B-type (MLS) antibiotics.[22] Thus, TylM1, acting on the amino group of a sugar substrate, is a distinct member of this AdoMet-dependent *N*,*N*-dimethyltransferase family.

It should be pointed out that the deduced sequence of *tylM1* reveals significant similarity to those of *eryCVI* from the erythromycin cluster of *Saccharopolyspora erythraea* (60% identity),[5] *desVI* from the methymycin cluster of *Streptomyces venezuelae* (60% identity),[23] *snoX* from the nogalamycin cluster of *Streptomyces nogalater* (54% identity),[24] *rdmD* from the rhodomycin cluster of *Streptomyces purpurascens* (50% identity),[25] and *srmX* from the spiramycin cluster of *Streptomyces ambofaciens* (47% identity).[26] All of them contain a short consensus sequence near the N-terminus, LLDV(I)ACGTG, a conserved motif for many AdoMet binding proteins.[5a,27] Although such a sequence analysis has allowed speculation of their catalytic roles as methyltransfereases, their actual functions have never been verified biochemically. The fact that TylM1 has now been fully established as an AdoMet-dependent *N*,*N*-dimethyltransferase furnishes compelling evidence suggesting a similar role for these proteins in their respective biosynthetic pathways.

Acknowledgment. We express our great appreciation to Dr. Eugene Seno and the Lilly Research Laboratories for their generous gifts of the plasmids pSET552, pHJL309, pHJL311, and POJ190. This work is supported in part by the National Institutes of Health Grant GM54346.

JA9815881

(11) Alais, J.; David, S. *Carbohydr. Res.* **1992**, *230*, 79−87.
(12) Excoffier, G.; Gagnaire, D.; Utille, J.-P. *Carbohydr. Res.* **1975**, *39*, 368−373.
(13) Wittmann, V.; Wong, C.-H. *J. Org. Chem.* **1997**, *62*, 2144−2147.
(14) Spectral data of **5**: ^{1}H NMR (D_2O) δ 1.10 (3H, d, J = 6.3 Hz, 5-Me), 1.74 (3H, s, 5″-Me), 2.17−2.22 (2H, m, 2′-Hs), 3.18 (1H, t, J = 10.2, 4-H), 3.24 (1H, t, J = 10.2, 3-H), 3.61 (1H, ddd, J = 10.2, 3.3, 3.1, 2-H), 3.77−3.87 (1H, m, 5-H), 3.99−4.01 (3H, m, 4′-H, 5′- Hs), 4.41−4.44 (1H, m, 3′-H), 5.40 (1H, dd, J = 7.2, 3.3, 1-H), 6.17 (1H, dd, J = 7.2, 6.9, 1′-H), 7.55 (1H, s, 6″-H); ^{13}C NMR (D_2O) δ 12.8, 17.6, 39.7, 55.9, 66.6 (d, J = 6.0 Hz), 70.1, 70.6 (d, J = 8.4), 72.1, 73.3, 86.2, 86.4 (d, J = 9.2), 95.6 (d, J = 6.1), 112.8, 138.5, 153.0, 167.9; ^{31}P NMR (D_2O) δ −13.47 (d, J = 20.7), −11.75 (d, J = 20.7); high-resolution MALDI-MS calcd for $C_{16}H_{25}N_3O_{14}P_2$ ($M + H)^+$ 546.0890, found 546.0894.
(15) Spectral data of **6**: ^{1}H NMR (D_2O) δ 1.11 (3H, d, J = 6.3 Hz, 5-Me), 1.75 (3H, d, J = 1.2, 5″-Me), 2.18−2.22 (2H, m, 2′-Hs), 2.85 (6H, s, NMe_2), 3.36−3.46 (2H, m, 3-H, 4-H), 3.82−3.87 (2H, m, 2-H, 5-H), 4.01−4.02 (3H, m, 4′-H, 5′-Hs), 4.42−4.46 (1H, m, 3′-H), 5.42 (1H, dd, J = 7.2, 3.3, 1-H), 6.18 (1H, dd, J = 6.9, 6.7, 1′-H), 7.56 (1H, d, J = 1.2, 6″-H); high-resolution MALDI-MS calcd for $C_{18}H_{29}N_3O_{14}P_2$ ($M + H)^+$ 574.1203, found 574.1196.
(16) The identity of **13** was confirmed by its incorporation of radioactivity when [^{3}H-CH_3]AdoMet was used as the methyl donor and by determining its exact mass using high-resolution MALDI-MS: calcd for $C_{17}H_{27}N_3O_{14}P_2$ ($M + H)^+$ 560.1047, found 560.1023.
(17) Under this condition, the reaction rate is no longer dependent on the concentration of AdoMet.
(18) The rate laws for an irreversible unimolecular consecutive reaction of (**A** → **B** → **C**) are $v_A = -dx/dt = k_1x$, $v_B = k_1x - k_2y$, and $v_C = k_2y$, in which x, y, and z are the respective concentrations of **A**, **B**, and **C** at time t. Assuming the initial concentration (at time 0) of **A** is a and those of **B** and **C** are nil, data fitting to any two of the following three equations could solve k_1 and k_2: $x = a\exp(-k_1t)$, $y = ak_1[\exp(-k_1t) - \exp(-k_2t)]/(k_2 - k_1)$, and $z = a[1 - k_2\exp(-k_1t)/(k_2 - k_1) + k_1\exp(-k_2t)/(k_2 - k_1)]$. Since such a treatment holds only under subsaturation conditions, the fact that our kinetic data fit well to the above equations strongly suggests a large K_m for **5** (much greater than 0.5 mM).
(19) The correlation coefficients are greater than 0.995 for all data fittings.
(20) Kodaki, T.; Yamashita, S. *Eur. J. Biochem.* **1989**, *185*, 243−251.
(21) (a) Gary, J. D.; Lin, W.-J.; Yang, M. C.; Herschman, H. R.; Clarke, S. *J. Biol. Chem.* **1996**, *271*, 12585−12594. (b) Lin, W.-J.; Gary, J. D.; Yang, M. E.; Clarke, S.; Herschman, H. R. *J. Biol. Chem.* **1996**, *271*, 15034−15044.
(22) Zalacain, M.; Cundliffe, E. *J. Bacteriol.* **1989**, *171*, 4254−4260.
(23) Xue, Y.; Zhao, L.; Liu, H.-w.; Sherman, D. H. *Proc. Natl. Acad. Sci. U.S.A.*, in press.
(24) Ylihonko, K. Tuikkanen, J.; Jussila, S.; Cong, L.; Mäntsälä, P. *Mol. Gen. Genet.* **1996**, *251*, 113−120.
(25) Niemi, J.; Mäntsälä, P. *J. Bacteriol.* **1995**, *177*, 2942−2945.
(26) Geistlich, M.; Losick, R.; Turner, J. R.; Rao, R. N. *Mol. Microbiol.* **1992**, *6*, 2019−2029.
(27) (a) Kagan, R. M.; Clarke, S. *Arch. Biochem. Biophys.* **1994**, *310*, 417−427. (b) Schluckebier, G.; O'Gara, M.; Saenger, W.; Cheng, X. *J. Mol. Biol.* **1995**, *247*, 16−20.