Communications to the Editor

Biosynthesis of Yersiniose: Attachment of the Two-Carbon Branched-Chain Is Catalyzed by a Thiamine Pyrophosphate-Dependent Flavoprotein

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

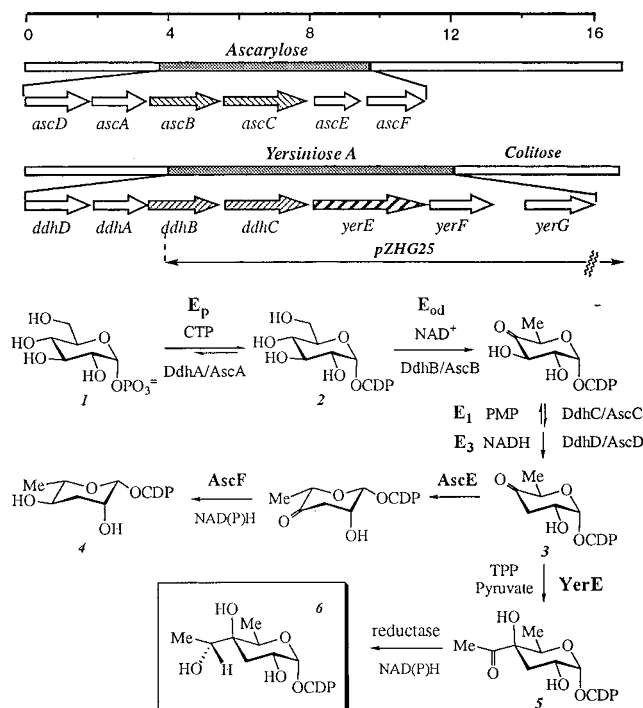
Department of Chemistry
University of Minnesota
Minneapolis, Minnesota 55455

Received August 26, 1998

The branched-chain sugars are an important class of carbohydrates present in a wide variety of natural sources.[1] They are derived from appropriate sugar precursors with one of the secondary hydrogen or hydroxyl groups replaced by an alkyl chain. Most naturally occurring branched-chain sugars contain a methyl branch, but examples that carry a two-carbon side chain or even an longer branched chain also exist. Previous feeding experiments have shown that the methyl branch is derived from *S*-adenosylmethionine (*S*-AdoMet),[1] and the two-carbon branch is from pyruvate.[1,2] While the C-methylation is believed to be simply a result of α-deprotonation of a ketosugar precursor to generate an enolate, followed by a nucleophilic attack on *S*-AdoMet,[1a,b,3] the mode of two-carbon branch attachment remains obscure. It has been speculated that the reaction may be thiamine pyrophosphate (TPP)-dependent, involving the formation of hydroxyethyl-TPP as the nucleophilic two-carbon donor.[1,2] However, the inability of well-known TPP inhibitors to affect the production of these two-carbon branched-chain sugars[2] and the failure to detect radioactive products in the incubation of 1-([1-^{14}C]-hydroxyethyl)-TPP with cell-free extracts of the producing strain[2] have cast doubts on the proposed mechanism. In an attempt to decipher this mechanistic ambiguity, we decided to study the biosynthetic formation of yersiniose A (**6**), a prototypical hydroxyethyl-branched 3,6-dideoxyhexose found in the *O*-antigen of *Yersinia pseudotuberculosis* VI.[4]

Early studies on 3,6-dideoxyhexoses have identified 3,6-dideoxy-4-hexulose (**3**) as a common intermediate for the biosynthesis of this class of unusual sugars, from which different isomers are formed by epimerization and reduction.[3] As illustrated by the formation of CDP-L-ascarylose (**4**) found in the *O*-antigen of *Y. pseudotuberculosis* V, this key intermediate is produced from α-D-glucose 1-phosphate (**1**) in four steps catalyzed by α-D-glucose-1-phosphate cytidylyltransferase (E_p),[5] CDP-D-glucose 4,6-dehydratase (E_{od}),[6] CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose 3-dehydrase (E_1),[7] and CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose 3-dehydrase reductase (E_3)[8] in sequence (Scheme 1). Beyond this point, epimerization at C-5 by an epimerase (AscE), followed by reduction at C-4 by a reductase (AscF) will give CDP-L-ascarylose (**4**).[9] Since yersiniose A (**6**) is a C-4-branched 3,6-dideoxyhexose, its formation must follow an analogous route that includes the same set of transformations to generate **3** as an intermediate.[3] It is conceivable that CDP-D-yersiniose A is formed by coupling of **3** with a two-carbon acyl anion equivalent to give **5**, which then undergoes a stereospecific reduction of the side chain. Since the two strains producing ascarylose and yersiniose A are intimately related and the proposed biosynthetic pathways of these two sugars are very similar, the gene sequence and organization of the yersiniose cluster should be highly homologous to those of the *asc* cluster.[10] Thus, the *ascC* gene,[11] which encodes E_1 from the *asc* cluster, was used as a probe to screen a genomic DNA library of *Y. pseudotuberculosis* VI for the *yer* cluster to gain access to the gene encoding the branched-chain coupling enzyme.

Scheme 1



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

(1) (a) Grisebach, H.; Schmid, R. *Angew. Chem., Int. Ed. Engl.* **1972**, *11*, 159−173. (b) Grisebach, H. *Adv. Carbohydr. Chem. Biochem.* **1978**, *35*, 81−126. (c) Williams, N.; Wander, J. In *The Carbohydrates: Chemistry and Biochemistry*; Pigman, W., Horton, D., Eds.; Academic Press: New York, 1980; Vol. 1B, pp 761−798. (d) Yoshimura, J. *Adv. Carbohydr. Chem. Biochem.* **1984**, *42*, 69−134. (e) Lindberg, B. *Adv. Carbohydr. Chem. Biochem.* **1990**, *48*, 279−318.

(2) (a) Schmid, R.; Grisebach, H.; Karl, W. *Eur. J. Biochem.* **1970**, *14*, 243−252). (b) Matern, U.; Grisebach, H. *Eur. J. Biochem.* **1972**, *29*, 5−11. (c) Matern, U.; Grisebach, H. *Z. Naturforsch.* **1974**, *29*, 407.

(3) (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; 1−84.

(4) (a) Gorshkova, R. P.; Zubkov, V. A.; Isakova, V. V.; Ovodov, Yu. S. *Carbohydr. Res.* **1984**, *126*, 308−312. (b) Gorshkova, R. P.; Zubkov, V. A.; Isakov, V. V.; Ovodov, Yu. S. *Bioorg. Khim.* **1987**, *13*, 1146−1147.

(5) Thorson, J. S.; Kelly, T. M.; Liu, H.-w. *J. Bacteriol.* **1994**, *176*, 1840−1849.

(6) (a) Yu, Y.; Russell, R. N.; Thorson, J. S.; Liu, L.-d.; Liu, H.-w. *J. Biol. Chem.* **1992**, *267*, 5868−5875. (b) He, X.; Thorson, J. S.; Liu, H.-w. *Biochemistry* **1996**, *35*, 4721−4731.

(7) (a) Thorson, J. S.; Liu, H.-W. *J. Am. Chem. Soc.* **1993**, *115*, 7539−7540. (b) Lei, Y.; Ploux, O.; Liu, H.-w. *Biochemistry* **1995**, *34*, 4643−4654.

(8) (a) Miller, V. P.; Thorson, J. S.; Ploux, O.; Lo, S. F.; Liu, H.-w. *Biochemistry* **1993**, *32*, 11934−11942. (b) Johnson, D. A.; Gassner, G. T.; Bandarian, V.; Ruzicka, F. J.; Ballou, D. P.; Reed, G. H.; Liu, H.-w. *Biochemistry* **1996**, *35*, 15846−15856.

(9) Thorson, J. S.; Lo, S. F.; Liu, H.-w. *J. Am. Chem. Soc.* **1993**, *115*, 5827−5828.

(10) Thorson, J. S.; Lo, S. F.; Ploux, O.; Liu, H.-w. *J. Bacteriol.* **1994**, *176*, 5483−5493.

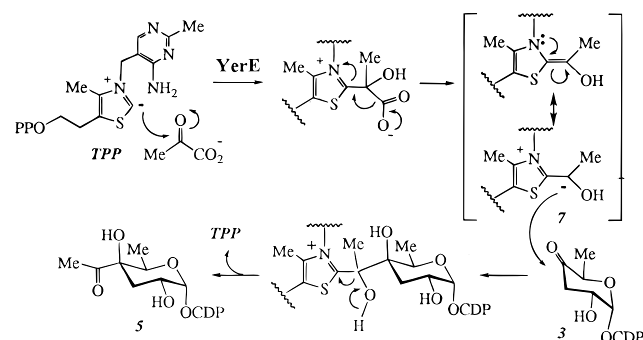
(11) According to a new nomenclature for genes involved in the biosynthesis of sugars found in lipopolysaccharides, the *ascC* gene should be designated as *ddhC* (Reeves, P. R.; Hobbs, M.; Valvano, M. A.; Skurnik, M.; Whitfield, C.; Coplin, D.; Kido, N.; Klena, J.; Maskell, D.; Raetz, C. R. H.; Rick, P. D. *Trends Microbiol.* **1996**, *4*, 495−503).

10.1021/ja983071v CCC: $15.00 © 1998 American Chemical Society
Published on Web 10/30/1998

Three positive and identical plaques, each bearing an insert of about 10 kb,[12] were isolated from the library screening. It was found that the cloned insert (*p*ZHG25) contains an 841 bp stretch (*ddhB*) identical to the 3′-terminal portion of the *ascB* gene in the *asc* cluster, and the open reading frame (ORF) immediately downstream from the truncated *ddhB* gene (*ddhC*) has a sequence identical to that of the *ascC* gene (Scheme 1).[10] These genetic results firmly substantiated the close resemblance of the yersiniose and ascarylose biosynthetic machinery.[13] Further analysis of this cluster led to the identification of three additional ORFs (Scheme 1).[14] Of particular interest is the strong similarity of the translated sequence of one of the ORFs (*yerE*) with that of the large subunit of FAD-containing acetolactate synthases (32% identity),[15] whose activity is TPP-dependent. Studies of this class of acetolactate synthases have shown that the flavin coenzyme plays no obvious role in the catalysis, whereas TPP is directly involved in the generation of an "acetyl carbanion" from pyruvate and its subsequent condensation with another molecule of pyruvate or 2-ketobutyrate.[15b] In view of the high sequence homology as well as the chemical relevance of the catalysis, the *yerE* encoded protein (YerE) must be akin to acetolactate synthases as a TPP-dependent flavoenzyme and must catalyze the branched-chain attachment in the biosynthesis of yersiniose A.

To verify the catalytic role of YerE, the *yerE* gene was amplified by polymerase chain reaction (PCR), cloned into a pET 24(+) expression vector, and expressed in *Esherichia coli* BL21-(DE3) cells. The expressed C-terminal His-tagged YerE protein was purified to near homogeneity by a Ni-NTA column (Qiagen), and its identity was confirmed by N-terminal sequencing. The purified YerE contains bound FAD, giving a typical flavin spectrum with two absorption maxima at 370 and 450 nm. Judging from a M_r of 117K estimated by gel filtration and a calculated mass of 63372 Da for each subunit, YerE must exist as a homodimer. Both TPP and Mg^{2+} are required for full activity, and the optimal pH was determined to be 8.0. Upon incubation with TPP (6 μmol) and pyruvate (250 μmol) at 24 °C for 1 h in 100 mM potassium phosphate buffer (pH 8.0, 1.0 mL) containing 3 μmol Mg^{2+}, YerE (0.75 nmol) quantitatively converted the predicted substrate **3** (12.0 μmol), which was prepared enzymatically from CDP-D-glucose (**2**) using purified E_{od}, E_1, and E_3, to **5**. This product was isolated from the incubation mixture, after removal of the proteins using Centricon 10, by FPLC MonoQ using a linear gradient of 25−90 mM ammonium bicarbonate buffer (pH 8) over a 25-min elution cycle. The spectral data of the purified product are consistent with **5**,[16] and the assigned role of YerE as the branched-chain coupling enzyme is thus confirmed. Since no loss of catalytic efficiency was discernible when the reaction was performed under anaerobic conditions with YerE prereduced with dithionite, this observation strongly suggests that the flavin coenzyme is not directly involved in catalysis.[17]

Scheme 2



Thus, cloning and sequencing of the *yer* cluster have allowed the gene, *yerE*, encoding the branched-chain coupling enzyme to be identified. Purification and characterization of YerE protein have provided, for the first time, unambiguous evidence confirming that the side-chain attachment in the biosynthesis of yersiniose and perhaps other two-carbon branched-chain sugars as well is TPP-dependent. As depicted in Scheme 2, coupling the TPP-bound acyl intermediate **7** to the 4-keto group of **3** to give **5** can be readily accomplished by a well-established mechanism found for other TPP-dependent enzymes.[18] Clearly, this study has infused significant mechanistic insights into the general biosynthetic routes leading to the production of branched-chain sugars and also allowed the integration of the yersiniose biosynthetic pathway into a unified model for the construction of all 3,6-dideoxyhexoses.[3a,10,13]

JA983071V

(12) The radiolabeled probe was prepared with the Multiprime DNA Labeling kit (Amersham) using *asc*C as the template.[10] The genomic DNA library of *Y. pseudotuberculosis* VI was constructed by digesting its DNA with *Eco*R I and ligating the fragments into a λZAPII vector. After packaging the ligation product with a phage extract as per the procedures specified by the manufacturer (Stratagene), the genomic library had a titer of 2.92×10^6 pfu/mL of phage. The 10 kb insert contains most of the yersiniose and part of the colitose biosynthetic genes.

(13) (a) Liu, D.; Verma, N. K.; Romana, L. K.; Reeves, P. R. *J. Bacteriol.* **1991**, *173*, 4814−4819. (b) Reeves, P. R. *Trends Genet.* **1993**, *9*, 17−22. (c) Kessler, A. C.; Hasse, A.; Reeves, P. R. *J. Bacteriol.* **1993**, *175*, 1412−1422. (d) Xiang, S.-h.; Hobbs, M.; Reeves, P. R. *J. Bacteriol.* **1994**, *176*, 4357−4365.

(14) While no significant homologues of YerF were found in the database, the YerG sequence displays moderate homology to many glycosyltransferases.

(15) (a) Godon, J. J.; Chopin, M. C.; Ehrlich, S. D. *J. Bacteriol.* **1992**, *174*, 6580−6589. (b) Chipman, D.; Barak, Z.; Schloss, J. V. *Biochim. Biophys. Acta* **1998**, *1385*, 401−419.

(16) Spectral data of **5**: ^{1}H NMR (D_2O) δ 0.81 (3H, d, J = 6.5 Hz, 5″-Me), 1.68 (1H, dd, J = 13.0, 4.5, 3″-H_{eq}), 1.95 (1H, t, J = 13.0, 3″-H_{ax}), 2.10 (3H, s, COMe), 3.80 (1H, m, 2″-H), 4.04−4.16 (3H, m, 4′-H, 5′-H), 4.19 (1H, dd, J = 5.0, 6.0, 3′-H), 4.22 (1H, dd, J = 5.0, 3.5, 2′-H), 4.30 (1H, q, J = 6.5, 5″-H), 5.38 (1H, dd, J = 6.8, 3.2, 1″-H), 5.76 (1H, d, J = 3.0, 1′-H). 5.94 (1H, d, J = 7.5, 5H), 7.84 (1H, d, J = 7.5, 6H). ^{13}C NMR (D_2O) δ 13.2, 22.6, 34.2, 63.4 (d, J = 8.4), 64.1 (d, J = 5.3), 67.3, 68.5, 74.0, 80.9, 82.0 (d, J = 9.2), 89.5, 94.8 (d, J = 6.6), 96.2, 141.1, 157.4, 165.9, 213.3.

(17) While the FAD coenzyme is not essential for the TPP-dependent chain-attachment step, whether it plays a role in the subsequent keto reduction catalyzed by an as yet unidentified reductase must await further investigation.

(18) Schellenberger, A. *Biochim. Biophys. Acta.* **1998**, *1385*, 177−186 and references therein.