Cloning and Characterization of the *Streptomyces peucetius dnmZUV* Genes Encoding Three Enzymes Required for Biosynthesis of the Daunorubicin Precursor Thymidine Diphospho-L-Daunosamine

SHAREE L. OTTEN,† MARK A. GALLO,‡ KRISHNAMURTHY MADDURI,∥ XIAOCHUN LIU,∥ AND C. RICHARD HUTCHINSON‡

School of Pharmacy† and Department of Bacteriology,∥ University of Wisconsin, Madison, Wisconsin 53706

Received 6 December 1996/Accepted 23 April 1997

Daunorubicin (DNR) and its C-14-hydroxylated derivative doxorubicin (DXR) (Fig. 1) are clinically important antitumor agents, and like many microbial secondary metabolites, they require a deoxyhexose component for their biological activity (7). These deoxyhexose constituents are commonly 6-deoxyhexoses including 2,6- and 4,6-dideoxy or trideoxy amino hexoses, as is the case for DNR and DXR, which contain the 2,3,6-trideoxy-3-aminoheptose daunosamine. The biosynthesis of these biologically important deoxy sugars is not well understood, but progress is being made, fueled in part by the availability of DNA sequence data. These data facilitate the construction of mutant strains that are disrupted in potential sugar genes and permit the assignment of reasonable functions to deduced gene products by sequence comparisons to proteins of known function, such as those for the biosynthesis of rhamnose (16) and the 3,6-dideoxyhexoses (14) that are found in the lipopolysaccharides of gram-negative bacteria. Thus a putative deoxyhexulose reductase, respectively, of the daunosamine biosynthesis (Fig. 1), has been identified in the DNR gene cluster of the wild-type *Streptomyces peucetius* (8), which presumably governs the first step of daunosamine biosynthesis (Fig. 1), has been identified in the DNR gene cluster of the wild-type *Streptomyces peucetius* ATCC 29050 strain, and a thymidine-diphospho (TDP)-glucose-4,6-dehydratase for the formation of the 4-keto-6-deoxyglucose-3(5)-epimerase and TDP-4-keto-6-deoxyglucose-3(5)-epimerase and TDP-4-ketodeoxyhexulose reductase, respectively, of the daunosamine biosynthetic pathway.

**Sequence analysis of the dnmZ, dnmU, and dnmV genes.** DNA sequencing of both strands of a 2,750-nucleotide (nt) AlwI-BglII fragment (Fig. 2), located between *dnr* and *dnm* in the DXR biosynthesis gene cluster (Fig. 3), was carried out with single-stranded DNA templates subcloned in M13 vectors as previously described (20). The CODONPREFERENCE program (6) was used to identify three complete open reading frames (ORFs), *dnmZ*, *dnmU*, and *dnmV*, and the C-terminal end of *dnmV* (25) on the basis of the characteristic third-position G+C bias and codon usage typical of *Streptomyces* genes. The *dnmZ* gene has a likely start codon (ATG) at nt 25 preceded by a probable ribosome binding site (RBS), GAGG, at nt 17. A stop codon (TGA) located at nt 1053 suggests that this ORF would encode a polypeptide having 341 amino acid residues (ignoring the formylmethionyl) and an Mr of 36,751. Similarly, the *dnmU* gene has a likely start codon (ATG) at nt 1100 that is preceded by a probable RBS, GAGG, at nt 1088. A stop codon (TGA) located at nt 1726 suggests that this ORF would encode a polypeptide having 208 amino acid residues and an Mr of 22,501. A likely start codon (ATG) for the *dnmV* gene is located at nt 1734 and is preceded by a probable RBS, GAGG, at nt 1725. A stop codon (TAG) located at nt 2657 suggests that *dnmV* encodes a polypeptide having 307 amino acid residues and an Mr of 32,492. The adjacent *dnmJ* gene (25) is transcribed convergently to *dnmV*; however, the stop codons of *dnmV* and *dnmJ* are not separated by a transcriptional terminator; rather, they are separated by a single base pair. This gene arrangement is consistent with the transcriptional mapping of the convergent *dnmZUV* and *dnr1dnmJ* transcripts which showed that the 3′ ends of the transcripts overlap extensively to produce complementary mRNA in the *dnr1dnmJ/dnmZUV* region (17) (Fig. 3).
Characterization of the deduced product of \textit{dnmZ}. Database comparisons of the deduced product of \textit{dnmZ} with the BLASTP and TBLASTN programs (1) showed that DnmZ shares significant similarity (22 to 26\% identity) with acyl coenzyme A (acyl-CoA) dehydrogenases (27), particularly short-chain fatty acid and isovaleryl-CoA dehydrogenases as found in rat (18) and human (19) cells. It also has a 26\% identity with AidB, an adaptive response protein of \textit{Escherichia coli} (13), and 74\% identity with the Orf3 protein from the \textit{lkmB} region of the DNR biosynthetic genes of \textit{Streptomyces griseus} JA3933 (12). However, these relationships might mean that DnmZ is only a flavoprotein and not an acyl-CoA dehydrogenase.

Gene disruption and replacement of \textit{dnmZ}. To test the role of \textit{dnmZ} in DNR biosynthesis, a 1.2-kb \textit{SmaI} fragment from \textit{pUC4-KIXX} (Pharmacia) containing the neomycin-kanamycin resistance gene was inserted into the \textit{MscI} site (Fig. 2) in \textit{dnmZ}. This disrupted copy of \textit{dnmZ} was moved into the temperature-sensitive \textit{pKC1139} vector (3) as a 4.6-kb \textit{BamHI} segment to give \textit{pWHM215}. \textit{S. peucetius} 29050 (\textit{pWHM215}) transformants were selected on solid R2YE medium (11) at 30\°C in the presence of kanamycin and apramycin. The cells were then grown in liquid R2YE medium containing no antibiotic at 39\°C for 2 days, and samples of the culture were plated on solid R2YE medium containing kanamycin and incubated at 30\°C to allow screening for apramycin-sensitive, kanamycin-resistant clones indicative of pWHM215 integration by homologous recombination at the \textit{dnmZ} locus and loss of the vector sequence. Southern hybridization of \textit{BamHI}-digested genomic DNA from a representative apramycin-sensitive, kanamycin-resistant clone, WMH1629, established that the 3.4-kb \textit{BamHI} fragment in the 29050 strain had shifted its mobility to 4.6 kb, as expected for replacement of the \textit{dnmZ} gene with its disrupted form (data not shown). The WMH1629 \textit{dnmZ::aphII} strain accumulated \epsilon-rhodomyacinone (RHO) when grown in a DNR production medium and analyzed for anthracycline metabolite production as previously described (20). Introduction of the \textit{dnmZ} (a \textit{MluI}-\textit{BamHI} segment [the \textit{MluI} site is about 180-bp upstream of the \textit{BspHI} site in \textit{dnmZ} shown in Fig. 2] cloned in pWHM3 as pWHM216) or \textit{dnmUV} (as pWHM218; see below) genes into this mutant by transformation did not restore DNR production, but this antibiotic was produced upon introduction of the \textit{dnmZUV} genes (a \textit{MluI}-\textit{BglII} segment cloned in pWHM3 as pWHM219). Thus, \textit{dnmZ} is essential for synthesis or attachment of daunosamine to RHO and the \textit{dnmZ::aphII} mutation appears to be polar on \textit{dnmUV}. (Since the actual transcription events for these genes have not been characterized beyond the information in reference 17, alternative explanations of the latter observation are possible.)

If the DnmZ protein acts as an acyl-CoA dehydrogenase, it is not clear on the basis of current information (14) how this activity would form part of daunosamine biosynthesis. To explore this idea further, we tested the ability of the 29050 and WMH1629 strains to utilize fatty acids as the sole carbon and
characteristics of the deduced product of dnmU. Database comparisons of the deduced product of dnmU with the BLASTP and TBLASTN programs (1) revealed similarities between DnmU and the NDP-4-keto-deoxyhexulose-3(5)-epimerase from sugar biosynthetic pathways that invert the configuration at C-3 and/or C-5 of the 4-keto-6-deoxy-hexulose derivatives which are common intermediates in the biosynthesis of many deoxysugars. The deduced product of dnmU is most similar to the product of ORF4 (84% identity by GAP analysis (6)) of S. griseus JAA3933 (12) and StrM (45% identity), the putative TDP-4-keto-6-deoxyglucose-3(5)-epimerase from the streptomycin gene cluster of S. griseus (22). In addition, DnmU is 34% identical to RfbD, the TDP-4-keto-6-deoxyglucose-3(5)-epimerase, for rhannose biosynthesis (16) and 31% identical to AscE, the CDP-4-keto-3,6-dideoxyhexulos-5-epi-
merase, for rhamnose biosynthesis (16) and 31% identical to AscE, the CDP-4-keto-3,6-dideoxyhexulose-5-epimerase from the ascarosyl biosynthetic pathway of Yersinia pseudotuberculosis (29, 31).

Characterization of the deduced product of dnmU. The deduced product of dnmU gene contained within a 4.4-kb BglII fragment was disrupted by replacing a 488-bp BspHI fragment containing most of dnmU with the kanamycin resistance gene from pUC4-KIXX. The disrupted dnmU gene, subcloned in the single-stranded DNA vector pPH5 (10), was introduced by transformation into the pseudotuberculosis (29, 31).

Characterization of the deduced product of dnmU. The deduced product of dnmU gene contained within a 4.4-kb BglII fragment was disrupted by replacing a 488-bp BspHI fragment containing most of dnmU with the kanamycin resistance gene from pUC4-KIXX. The disrupted dnmU gene, subcloned in the single-stranded DNA vector pPH5 (10), was introduced by transformation into the S. peucetius 29050 strain. Transformants containing the integrat-
ed plasmid were screened to identify colonies with the kana-
mycin-resistant, thiomycro-sensitive phenotype resulting from gene replacement by a double crossover as previously described (20). A putative dnmU::aphII mutant strain (WHM1673) was obtained, and the NolI-digested chromosomal DNA from this strain was subjected to Southern analysis using a 1.1-kb NolI fragment containing dnmU as the probe. The probe DNA hybridized to the predicted 1.8-kb NolI fragment (composed of the 0.6-kb inactivated dnmU gene and the 1.2-kb aphII insert), verifying the disruption of dnmU. The WHM1673 strain was analyzed for antibiotic production (21) and was found to pro-
duce no DNR but instead accumulated RHO (69.6 g/ml ± 17.42, n = 3), indicating that DnmU is required for daunos-
amine biosynthesis. The dnmU::aphII mutant could be com-
plicated by the dnmU gene cloned as pWHMS52 as a 1.7-kb BamHI-BglII fragment carrying the low-copy-number vector pWHM601 (9) or by dnmU cloned as pWHM53 as a 0.8-kb BamHI-NcoI fragment in pWHM601. The complementation by dnmU alone was surprising since the mutation in dnmU was expected to have a polar effect on dnmV. However, analysis of plasmid DNAs by restriction endonuclease diges-
tions and repetition of the complementation experiment con-

FIG. 2. The DNA sequence of the region containing dnmZUV. The proposed translational start and stop sites of the dnmZ, dnmU, and dnmV genes and the putative RBSSs are doubly and singly underlined, respectively. Amino acid trans-
lations are shown for the dnmZ, dnmU, and dnmV genes and the C-terminal end of the dnmV (17, 25) gene product; stop codons are indicated with an asterisk. The underlined amino acid sequence in the dnmV gene product is a dinucleotide-binding domain. Only the restriction sites discussed in the text are shown above their recognition sites.

FIG. 3. Restriction map of the DNR gene cluster of S. peucetius 29050 showing the relative locations of all the genes believed to be involved in daunosamine (dnm) biosynthesis. The dnm gene designations are represented by letters. Arrows indicate direction of transcription. Restriction enzyme abbreviations: B, BamHI; Bg, BglII.
firmed that the *dnmU*::*aphII* mutant strain was complemented by *dnmU* alone. Although this result raises the possibility that *dnmU* and *dnmV* are in separate transcription units, we did not investigate this matter.

The phenotype and complementation of the WHM1673 mutant strain indicate that DnmU is required for DNR biosynthesis. Together with the sequence similarities between the deduced product of *dnmU* and the NDP-4-keto-6-deoxyhexulose-3(5)-epimerases from other deoxysugar biosynthetic pathways, the results suggest that DnmU functions as the TDP-4-keto-6-deoxyglucose-3(5)-epimerase for the formation of TDP-daunosamine, although we do not know if TDP-4-keto-6-deoxyglucose is the actual substrate for this enzyme.

**Characterization of *dnmV***. Database comparisons of DnmV revealed strong similarity (64% identity) to the deduced product of *ORF5*, a gene of unknown function from the *lkmB* region of *S. griseus* (12) that has been suggested to be involved in the biosynthesis of daunosamine. Surprisingly, the deduced *ORF5* gene product lacks the C-terminal 122 amino acids of DnmV and contains the premature amber codon corresponding to nt 2291 of the *dnmUV* sequence (Fig. 2). The significance of this difference is not known; however, the length of the deduced DnmV peptide is similar to that of related proteins discussed below, and careful examination of the sequence data in this region did not reveal the presence of a sequencing error.

DnmV is also distantly related to a group of proteins that are involved in the metabolism of nucleotide-activated hexoses. These proteins include the UDP-glucose 4-epimerases such as ExoB from *Rhizobium meliloti* (5), the NDP-glucose 4,6-dehydratases that catalyze 6-deoxygenation of nucleotide-activated hexoses such as StrE from *S. griseus* (22), and RIBC (16), the TDP-4-ketohamnose ketoreductase that catalyzes the final step of rhamnose biosynthesis. The unifying properties of these proteins (which have different functions) are that they target the C-4 position of a nucleotide-activated hexose and possess a dinucleotide-binding domain (16, 32). This dinucleotide-binding domain is located near the N terminus of the protein and is characterized by a GXGXGXG motif which is present in StrE, DnmV, and RIBC and corresponds to nt 2291 of the *dnmUV* sequence (Fig. 2). The significance of this difference is not known; however, the length of the deduced DnmV peptide is similar to that of related proteins discussed below, and careful examination of the sequence data in this region did not reveal the presence of a sequencing error.

**Concluding remarks.** The important and diverse biological functions of the deoxysacharides (14), has led to a growing interest in the genetics of deoxysacharide biosynthesis, and much information has been gained from comparisons of the deduced amino acid sequences of genes required for the synthesis of structurally related sugars. The amino acid sequences of the NDP-hexose 4,6-dehydratases from various sources are found to be substantially conserved, reflecting the expected conserved biochemical mechanism of 6-deoxygenation. The product of this reaction, an NDP-4-keto-6-deoxyhexulose, is thought to be the precursor of many other deoxy and trideoxy sugars (14). In the synthesis of the 6-deoxyhexose rhamnose, the next step is a 3(5)-epimerization catalyzed by RbdB (16). The sequence similarity between RbdB and the products of *dnmU* and other putative 3(5)-epimerase genes suggests that they too may act on an NDP-4-keto-6-deoxyhexulose. Relative to the TDP-hexose-4,6-dehydratases and -3(5)-epimerases, DnmV and other putative 4-ketoreductases that have been identified so far share less overall homology. This observation may be a consequence of the different stereochemical functions of 4-ketoreductases or may indicate that 4-keto reduction occurs at a later step in the pathway when substrates have become structurally diverse.

Homologs of several of the *dnm* genes are present in the cluster of erythromycin biosynthesis genes in *Saccharopolyspora erythraea* (26). There is a close similarity between the steps in the daunosamine pathway and the formation of 2,6-dideoxy-1-mycarosyl, whereas the formation of 3,4,6-trideoxy-3-amino-D-desosamine, the second deoxysugar present in erythromycin, has fewer features in common with daunosamine biosynthesis. Two enigmatic differences stand out as follows: (i) The homologous *dnmQ* and *eryCII* genes both are associated with deoxyaminoglycoside biosynthesis but are believed to involve C-2 deoxygenation and C-4 deoxygenation, respectively, and (ii) A *dnmZ* homolog is not involved in erythromycin biosynthesis (or in oleandrose formation [15]). The mechanisms of C-2, C-3, and C-4 deoxygenation seem to be different in the two pathways even though all of the other steps of deoxysugar formation are similar.

**Nucleotide sequence accession number.** The DNA sequence data described in this paper have been deposited at EMBL and GenBank with accession number AF006633.

We thank Doug MacNeil and Leonard Katz for unpublished sequence data and helpful discussions about deoxysugar biosynthesis.

*This work was supported by grants from the American Cancer Society (CH-524) and Farmitalia Carlo Erba (Pharmacia S.p.a.)."