The \textit{dnrM} gene in \textit{Streptomyces peucetius} contains a naturally occurring frameshift mutation that is suppressed by another locus outside of the daunorubicin-production gene cluster

Mark A. Gallo,\textsuperscript{1} Joanne Ward\textsuperscript{1\dagger} and C. R. Hutchinson\textsuperscript{1,2}

Author for correspondence: C. R. Hutchinson. Tel.: +1 608 262 7582. Fax: +1 608 262 3134. e-mail: crhutchina@facstaff.wisc.edu

School of Pharmacy\textsuperscript{1} and Department of Bacteriology\textsuperscript{2}, University of Wisconsin, Madison, WI 53706, USA

A 2.7 \textit{kb} BamHI fragment of the daunorubicin biosynthetic cluster in \textit{Streptomyces peucetius} ATCC 29050 was shown to contain two ORFs, \textit{dnrL} and \textit{dnrM}, whose deduced products exhibit a high sequence similarity to a number of glucose-1-phosphate thymidylyl transferases and TDP-D-glucose dehydratases, respectively. Although these genes were believed to be necessary for the synthesis of the deoxyaminosugar, daunosamine, a constituent of daunorubicin, the \textit{dnrM} gene contains a frameshift in the DNA sequence that causes the premature termination of translation. A gene encoding another TDP-glucose 4,6-dehydratase, previously isolated from \textit{S. peucetius}, was identified by PCR amplification of genomic DNA. The presence of this gene explains why a \textit{dnrM::aphII} mutation did not block daunorubicin production.

\textbf{Keywords:} daunosamine, deoxyamino sugar, inactive gene, TDP-D-glucose 4,6-dehydratase, glucose-1-phosphate thymidylyl transferase

\section*{INTRODUCTION}

The \textit{Streptomyces} are filamentous, spore-forming soil bacteria that produce an array of products termed secondary metabolites. Many are useful in agriculture and medicine, for example as antiparasitic, antimicrobial, or anticancer agents. Relatively little is known in general about the biochemistry, physiology and genetics of production of secondary metabolites. The anthracyclines are a prevalent class of such compounds isolated from \textit{Streptomyces} sp. and other actinomycetes. Two anthracyclines, daunorubicin (DNR) and doxorubicin (DXR), have received considerable attention due to their potent antitumour activity (Arcamone, 1981). Our laboratory has been interested in determining the biochemical and genetic processes involved in the synthesis of DNR and DXR in \textit{Streptomyces peucetius} ATCC 29050.

Previous reports have identified a 45 \textit{kb} region that presumably contains all the genes required for the synthesis of and resistance to DNR and DXR (Otten et al., 1990; Stutzman-Engwall & Hutchinson, 1989; Madduri & Hutchinson, 1995a). DNR and DXR consist of two parts, a polyketide-derived aglycone and the trideoxyamino sugar daunosamine. All of the genes required for synthesis of the aglycone carbon skeleton have been characterized (Guilfoile & Hutchinson, 1991; Grimm et al., 1994; Madduri & Hutchinson, 1995b), and most of them share a high degree of homology with other type II polyketide synthases and cyclases at the level of DNA and deduced protein sequence (Hopwood & Sherman, 1990; Katz & Donadio, 1993; Hutchinson & Fujii, 1995).

By contrast, little is known about the genes involved in biosynthesis of daunosamine (Krugel et al., 1993; Thorson et al., 1993; Otten et al., 1995b). In this paper we report the sequence of a 2682 bp DNA segment containing two putative ORFs, \textit{dnrL} and \textit{dnrM} (Fig. 1), located within the known boundaries of the 45 \textit{kb} gene cluster. We show that \textit{dnrL} likely encodes a glucose-1-phosphate thymidylyl transferase and \textit{dnrM} encodes a thymidylyldiphospho(TDP)-D-glucose 4,6-dehydratase, on the basis of sequence comparisons to other known transferases and
dehydratases involved in deoxysugar metabolism. However, the \textit{dnrM} sequence was unexpectedly found to contain a frameshift that results in the formation of a truncated protein. Thus, inactivation of \textit{dnrM} in \textit{S. peucetius} 29050 did not prevent DNR or DXR production, suggesting that another dehydratase-encoding gene existed outside the 45 kb \textit{dnr} region. Southern analysis of chromosomal DNA using \textit{dnrM} as a probe showed, in addition to \textit{dnrM}, several weakly hybridizing bands; one of these appeared to be the gene for another dehydratase that, from the results of initial PCR amplification and DNA sequence analysis, encodes a dehydratase previously purified from \textit{S. peucetius} (Thompson et al., 1992).

**METHODS**

**Bacterial strains and plasmids.** \textit{S. peucetius} ATCC 29050 was obtained from the American Type Culture Collection. \textit{Escherichia coli} strains DH5aMCR (Life Technologies) and JM105 (Yanisch-Perron et al., 1985) were used for routine subcloning and preparation of ssDNA. The glutathione S-transferase (GST) vector pGEX-4 was obtained from Pharmacia. The high-copy number \textit{E. coli}-\textit{Streptomyces} shuttle vector pWHM3 was from Vara et al. (1989). M13 phage-derived mp18 and mp19 vectors (Yanisch-Perron et al., 1985) were used for sequencing by previously published methods (Summers et al., 1992). The pUC4-KIXX plasmid containing the neomycin/kanamycin resistance gene (\textit{aphII}) from \textit{Tn}5 was obtained from Pharmacia. The plasmid pDH5 (Hillemann et al., 1991) was used to generate ssDNA for transformation of \textit{Streptomyces} sp. and gene disruption; pWHM3 was used to clone the wild-type \textit{dnrLM} genes.

**Media and growth conditions.** \textit{S. peucetius} was grown at 30 °C on R2YE (Hopwood et al., 1985) for propagation of cells, DNA isolation, and protoplast preparation. ISP4 medium (Difco) was used to produce spores. Transformants of \textit{S. peucetius} were selected on 10 µg thiostrepton ml$^{-1}$, 25 µg apramycin ml$^{-1}$ or 50 µg neomycin ml$^{-1}$. \textit{E. coli} transformants were selected with 100 µg ampicillin ml$^{-1}$, 100 µg apramycin ml$^{-1}$ or 50 µg neomycin ml$^{-1}$. \textit{S. peucetius} strains were grown in 5 ml seed medium (Guilfoile & Hutchinson, 1991) for 2 d, then transferred to GPS production medium (Dekleva et al., 1985) and grown in 50 ml broth in a 250 ml baffled-bottom flask for an additional 3 d at 30 °C with shaking at 300 r.p.m. Cultures were acidified with oxalic acid, heated at 60 °C for 45 min, adjusted to pH 8.5, and extracted with chloroform (Otten et al., 1990). The extracted anthracycline metabolites were analysed by TLC or HPLC (Otten et al., 1990).

**DNA isolation and manipulation.** DNA isolation, restriction endonuclease digestions, and ligations were performed according to standard techniques (Sambrook et al., 1989). Hybridization analysis was carried out using the Genius non-radioactive kit (Boehringer Mannheim) according to the manufacturer's instructions. DNA was electrophoresed on 0.7% agarose gels, and alkali transfer to Hybond-N (Amersham) was performed for Southern analysis.

**DNA sequencing.** DNA fragments were subcloned into M13mp18 and M13mp19, and single-stranded templates were sequenced by the dideoxy chain-termination method using $[\alpha^{32}P]dCTP$ and Sequenase version 2.0 (USB) according to the manufacturer's instructions. 7-Deaza-dGTP was used instead of dGTP to avoid compressions.

**Disruption of \textit{dnrM}.** A 725 bp \textit{BglII} fragment internal to \textit{dnrM} was removed from the 2.7 kb \textit{BamHI} fragment containing \textit{dnrL} and \textit{dnrM}, and the DNA was blunt-ended by treatment with the Klenow fragment and ligated to a 1.1 kb \textit{SalI} fragment containing \textit{aphII}. The resulting 3.1 kb fragment was subcloned into pDH5 (Hillemann et al., 1991) to give pWHM209, and ssDNA was produced and used to transform \textit{S. peucetius} 29050 protoplasts. Transformants were selected on R2YE plates containing thiostrepton or neomycin to verify integration into the \textit{S. peucetius} genome by homologous recombination. Several of the transformants were subjected to two rounds of sporulation, and single colonies were picked and scored for the expected neomycin-resistant thiostrepton-sensitive phenotype of a double crossover recombination event. Four mutants were analysed by Southern analysis to verify gene disruption and replacement by digesting total chromosomal DNA with \textit{BamHI} and probing with either the 725 bp \textit{BglII} or the 2.7 kb \textit{BamHI} fragment. All neomycin-resistant, thiostrepton-sensitive mutants lacked the 725 bp \textit{BglII} fragment and contained a \textit{BamHI} fragment of 3.1 kb; properties one would expect of gene replacement mutants (data not shown). One \textit{dnrM::aphII} clone was chosen and named WMH1590.

**PCR amplification.** PCR was used to amplify an additional putative dehydratase from \textit{S. peucetius} 29050 chromosomal DNA. The PCR mixture contained 1 × PCR buffer [50 mM KCl, 10mM Tris/HCl, pH 9.0 (at 25 °C), 0.1% Triton X-100, and 2.5 mM MgCl$_2$] supplied by Promega. Each primer was present at 0.5 µM, and each dNTP was present at 0.1 mM. The codon bias of \textit{Streptomyces} genes (Wright & Bibb, 1992) was taken into account in designing the following two primers: GTCAACGTCACCGTACCGGCCGCGCGCGGAGATGCCGG(G/C)TAGCGC/G/CT, which corresponds to the N-terminal peptide sequence reported previously (Thompson et al., 1992); and TCA(G/C)AGG(G/C)GGCTCCACTACG/(G/-C)(A/T)(G/C)CG, which corresponds to the complementary sequence of a conserved peptide in the C-termini of TDP-D-glucose dehydratases. Taq polymerase (1 U, Promega) and

---
approximately 100 ng target DNA was added in a final reaction volume of 100 μl. Amplification was performed in a thermal cycler (model 480, Perkin-Elmer Cetus) by denaturing the samples at 100 °C for 5 min and then subjecting them to 25 cycles of denaturing (97 °C, 30 s), annealing (55 °C, 45 s) and elongation (70 °C, 90 s). Amplification products were blunt-ended with Klenow and 0.1 mM dNTPs, separated by agarose gel electrophoresis, and purified with Qiaex resin (Qiagen). The resulting DNA fragment was ligated into M13mp18 and 19 for sequencing.

Expression of dnrM in E. coli. The GST gene fusion system (Pharmacia) was used to express dnrM in E. coli. Plasmid pGEX-4T-1 was cut with XhoI, and linkers were added to introduce an SstI site. This modified plasmid was cut with EcoRI, blunt-ended by treatment with the Klenow fragment, then the DNA was cut with SstI and ligated to a 1 kb BspHI/blunt-SstI fragment containing dnrM to give pWHM204. Expression of GST:DnrM in E. coli(pWHM204) strains was induced with 1 mM IPTG. The formation of a novel, unique protein was demonstrated by SDS-PAGE analysis.

RESULTS

Sequence analysis of dnrL and dnrM

Sequence analysis by CODONPREFERENCE (Devereux et al., 1984) of a 2.7 kb BamHI fragment from within the dnr gene cluster (Fig. 1) revealed two ORFs, dnrL and dnrM. The region encompassing nt 1–655 appeared to contain no long ORFs with any recognizable similarity to known proteins by TFASTA analysis (Devereux et al., 1984). The most likely translational start site for dnrL is the ATG beginning at nt 655 and ending with TGA at nt 1708–1710, producing a polypeptide of 351 aa. The translational start site for dnrM is believed to be the ATG beginning at nt 1707. This start site overlaps the TGA stop of dnrL, implying a possible translational coupling of these two proteins. A TGA stop codon beginning at nt 1890 allowed for production of a 61 aa polypeptide (6572 Da). CODONPREFERENCE analysis revealed a switch in reading frames approximately 97 nt beyond the ATG, at nt 1707. If the DNA sequence data are interpreted in a manner that allows introduction of a frameshift, a TGA stop codon beginning at nt 2680 is found in this new frame, implying production of a polypeptide of 324 aa (35 177 Da) for the combined ORFs from nt 1707 to 2679. Codon usage analysis of this frame showed very few rare Streptomyces codons (Wright & Bibb, 1992).

Characteristics of the deduced gene products of dnrL and dnrM

The deduced product of dnrL has a high sequence similarity to several glucose-1-phosphate thymidylyl transferases (Table 1), enzymes that are involved in transferring a TDP moiety to glucose. This irreversible reaction is an important first step in the activation of sugars for eventual transfer to other molecules (Liu & Thorson, 1994). Consequently, dnrL most likely provides this enzyme for daunosamine biosynthesis.

The deduced product of the truncated dnrM gene shares a significant sequence similarity to the N-terminal portion of several TDP-glucose 4,6-dehydratases (Table 1). Translation of the second reading frame identified by CODONPREFERENCE analysis reveals a remarkable sequence similarity to the remaining amino acid residues of this family of dehydratases (Table 1). Therefore, we believe that these two frames should be linked (Fig. 2). A simple sequencing error was not responsible for the shift in translational sequence revealed in Fig. 2 because the region was sequenced several times in both directions from several

Table 1. Comparisons between the amino acid sequences of DnrL and DnrM with other deoxysugar biosynthesis enzymes having similar functions

<table>
<thead>
<tr>
<th>Protein</th>
<th>Number of amino acids</th>
<th>Similarity (%)</th>
<th>Identity (%)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DnrL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StrD</td>
<td>355</td>
<td>76-3</td>
<td>59-7</td>
<td>Pissowotzki et al. (1991)</td>
</tr>
<tr>
<td>TylA1</td>
<td>304</td>
<td>59-3</td>
<td>36-9</td>
<td>Merson-Davies &amp; Cundliffe (1994)</td>
</tr>
<tr>
<td>RfbA</td>
<td>293</td>
<td>55-6</td>
<td>31-5</td>
<td>Stevenson et al. (1994); Yao &amp; Valvano (1994)</td>
</tr>
<tr>
<td>DnrM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StrE</td>
<td>328</td>
<td>76-2</td>
<td>61-6</td>
<td>Pissowotzki et al. (1991)</td>
</tr>
<tr>
<td>RfbB</td>
<td>361</td>
<td>63-0</td>
<td>46-7</td>
<td>Stevenson et al. (1994)</td>
</tr>
<tr>
<td>Gdh</td>
<td>329</td>
<td>71-4</td>
<td>58-1</td>
<td>Linton et al. (1995)</td>
</tr>
</tbody>
</table>
M. A. GALLO, J. WARD and C. R. HUTCHINSON

Fig. 2. Nucleotide sequence of the 2682 bp BamHI fragment containing dnrL and dnrM. Putative RBSs are underlined. The translated amino acid sequences are shown below the DNA sequence. The presumed site of the frameshift in dnrM is indicated by a '*' and asterisks indicate stop codons.

M13 plaques derived from independent primary clones. Additionally, the region was cloned by PCR from S. peucetius 29050 chromosomal DNA and again shown to contain the same frameshift by DNA sequencing.

Expression of dnrM in E. coli

A GST::DnrM fusion protein (Methods) appeared as a 32.6 kDa band (Fig. 3), which is 6.6 kDa larger than the band corresponding to GST alone. This indicates that a polypeptide of 61 aa is present at the C-terminal end of GST, in agreement with the predicted size if the truncated DnrM protein is fused to GST. A fusion protein of 60 kDa [equal to GST (26 kDa) plus a full-length DnrM (34 kDa)] was not detected (Fig. 3). This result is consistent with the sequence analysis and corroborates the conclusion that dnrM contains a naturally occurring frameshift mutation.

Disruption of dnrM

The chromosomal copy of dnrM in S. peucetius 29050 was inactivated by insertion of aphH to produce the WMH1590 strain (Methods). Wild-type S. peucetius 29050
formation of TDP-\(\text{-}\)d-glucose is not a rate-limiting step in the conversion of RHO to DNR.

**Isolation of another TDP-\(\text{-}\)d-glucose 4,6-dehydratase gene**

Thompson et al. (1992) have partially purified a TDP-\(\text{-}\)d-glucose 4,6-dehydratase from *S. peucetius*. The N-terminal sequence of their protein does not agree with that predicted for DnrM, although it does show some similarity to this family of dehydratases (Liu & Thorson, 1994). We synthesized degenerate oligodeoxynucleotides based on highly conserved regions at the N-terminus of the reported amino acid sequence (Thompson et al., 1992) and the reverse complement of the C-terminus of such dehydratases (Liu & Thorson, 1994), then amplified by PCR the targeted region in the DNA from *S. peucetius* 29050. The DNA sequence of a small portion of the 1-1 kb fragment amplified was determined. Its translated sequence is: VNVTVTGAAGQIGYALVTVGAAALIG YALQRT; residues in bold are highly conserved in TDP-\(\text{-}\)d-glucose 4,6-dehydratases (Liu & Thorson, 1994). This sequence is not the same as that proposed for DnrM (Fig. 2), but the first 14 aa are identical to those reported by Thompson et al. (1992). It thus appears that we have located the gene for a functional TDP-\(\text{-}\)d-glucose 4,6-dehydratase of *S. peucetius*. The DNA sequence was not identical to any part of the 45 kb region of the *dnr* gene cluster sequenced, indicating that this other dehydratase is not encoded by a gene inside the defined *dnr* cluster. We were also unable to amplify this other dehydratase gene from pWHM341 (Otten et al., 1990), a cosm containing a 27 kb DNA segment to the left of the defined *dnr* cluster (Madduri & Hutchinson, 1995a), as in Fig. 1.

**DISCUSSION**

The genes for the biosynthesis, regulation, and resistance to secondary metabolites are normally clustered in *Streptomyces* species, perhaps because this arrangement confers a selective advantage in maintenance and regulation. Coordinate expression of such genes has often been found (Chater & Bibb, 1995). In this regard, we have shown that a transcript containing *dnrL* and *dnrM* is present in *S. peucetius* 29050, and is regulated by the *dnr*1 gene encoding a transcription factor (Madduri & Hutchinson, 1995a).

On the basis of the above, we were surprised to find an inactive *dnrM* gene in the *dnr* cluster. The lack of a functional gene encoding a TDP-\(\text{-}\)d-glucose 4,6-dehydratase in this gene cluster implies the presence of another functional gene elsewhere in the organism, as suggested by the results of our exploratory studies and the work of Thompson et al. (1992). Moreover, the existence of a *dnrL* homologue at another locus that provides a functional glucose-1-phosphate thymidylyl transferase is not excluded by our data. Since there are many other glycosylated anthacyclines produced by *Streptomyces* species, it would be interesting to investigate whether other DNR-producing species, or descendants of strain 29050, such as *S. peucetius* subsp. *caesius* ATCC 27952 (Arcamone et al., 1969), also contain a mutant *dnrM* homologue. (It

---

**Table 2. Effects on secondary metabolite production of the introduction of *dnrM* into *S. peucetius* strains 29050 and WMH1590**

Values are means of three samples for RHO (±4.17) and DNR (±1.02) and are in µg ml\(^{-1}\).

<table>
<thead>
<tr>
<th>Strain/plasmid</th>
<th>RHO</th>
<th>DNR</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>29050/pWHM3</td>
<td>27.7</td>
<td>2.9</td>
<td>9.6</td>
</tr>
<tr>
<td>29050/pWHM205</td>
<td>34.5</td>
<td>3.6</td>
<td>9.6</td>
</tr>
<tr>
<td>WMH1590/pWHM3</td>
<td>35.5</td>
<td>2.4</td>
<td>14.8</td>
</tr>
<tr>
<td>WMH1590/pWHM205</td>
<td>21.2</td>
<td>1.7</td>
<td>12.5</td>
</tr>
</tbody>
</table>
would be surprising to find that dnrM was mutated only in the isolate of the 29050 strain in our collection, but we have not tested this possibility.)

Although it is unusual to find nonfunctional genes in *Streptomyces* operons, nonfunctional ketoreductase domains in the products of polyketide synthase genes have been reported in the erythromycin-producing *Saccharopolyspora erythraea* (Donadio & Katz, 1992) and the avermectin-producing *Streptomyces avermitilis* (MacNeil et al., 1994). DNR biosynthesis is unique in this regard since a nonfunctional gene is present, but another gene product from outside the biosynthetic gene cluster complements the missing activity. In erythromycin biosynthesis, the cluster of production genes, as currently defined, lacks genes encoding glucose-1-phosphate thymidyl transerase and TDP-D-glucose 4,6-dehydratase, but genes for these enzymes have been found elsewhere in *S. erythraea* (Linton et al., 1995; S. Zotchev & C. R. Hutchinson, unpublished results).

The unglycosylated intermediate RHO is approximately 10-fold more prevalent than DNR when *S. peucetius* is grown in GPS medium (Dekleva et al., 1985). Therefore, it appears quite likely that one of the steps of daunosamine biosynthesis or attachment is rate-limiting. It is also possible that the TDP-D-glucose 4,6-dehydratase that actually catalyses the second step of daunosamine biosynthesis in *S. peucetius* is not well-adapted for this function and hence is forming a metabolic bottleneck. Consequently, it would be interesting to determine the effect on DNR production in *S. peucetius* 29050 of repairing the frameshift in dnrM or placing the other TDP-D-glucose 4,6-dehydrogenase gene under control of dnrI (Madduri & Hutchinson, 1995a), or introducing additional copies of a known TDP-D-glucose 4,6-dehydratase.

**ACKNOWLEDGEMENTS**

We are grateful to K. Linton and B. Jarvis for sharing sequence data prior to publication. The authors would also like to thank B. Jarvis and S. L. Otten for helpful discussions. This work was supported by grants from the National Institutes of Health (GM 25799 and 31925) and Pharmacia Spa, Milan, Italy.

**REFERENCES**


Received 27 June 1995; revised 14 September 1995; accepted 28 September 1995.