Streptomyces peucetius daunorubicin biosynthesis gene, dnrF: sequence and heterologous expression

Silvia Filippini, Maria Manuela Solinas, Umberto Breme, Manuela Birgit Schlüter, Davide Gabellini, Giuseppe Biamonti, Anna Luisa Colombo and Luisa Garofano

Author for correspondence: Silvia Filippini. Tel: +39 331 583763. Fax: +39 331 583755.

1 BioScience Center
PHARMACIA, Via Papa Giovanni XXIII 23, 20014 Nerviano (Mi), Italy
2 Istituto di Genetica Biochimica ed Evoluzionistica del CNR, Via Abbiati 207, 27100 Pavia, Italy

The dnrF gene, responsible for conversion of aklavinone to ε-rhodomycinone via C-11 hydroxylation, was mapped in the daunorubicin (Dnr) gene cluster of Streptomyces peucetius ATCC 29050, close to drrAB, one of the anthracycline-resistance genes. The dnrF gene was sequenced and should encode a protein of 489 amino acids with a molecular mass of 52 kDa. The deduced DnrF protein shows significant similarities with bacterial FAD- and NADPH-dependent hydroxylases either required to introduce hydroxyl groups into polycyclic aromatic polyketide antibiotics or involved in catabolism of aromatic compounds. Heterologous expression of dnrF in Streptomyces lividans TK23 and in Escherichia coli demonstrated that the gene encodes a NADPH-dependent hydroxylase catalysing the hydroxylation of aklavinone to yield ε-rhodomycinone. The enzyme is inactive on anthracyclines glycosylated at position C-7 and its activity decreases to a different extent with other substrate modifications, indicating that DnrF has a significant substrate specificity.

Keywords: aklavinone, hydroxylase, dnrF gene, anthracycline, Streptomyces peucetius

INTRODUCTION

Daunorubicin (Dnr) and especially its C-14 hydroxylated derivative doxorubicin (Dxr) are important antitumour agents used for the treatment of many different cancers (Arcamone, 1981). Unfortunately, besides their antineoplastic activity, they have long-term undesirable effects such as cardiotoxicity (Myers et al., 1988). For this reason, studies are being done to develop new Dnr derivatives or new anthracyclines showing an increased antitumour spectrum and, at the same time, reduced cardiotoxicity (Cassini et al., 1978; Crespi-Perellino et al., 1982; White & Stroshane, 1984).

New possibilities for structural changes, resulting from recent progress in the molecular genetics of Streptomyces, have opened a promising line in the development of hybrid antibiotics (Floss & Strohl, 1991; Hopwood et al., 1985b; Hutchinson, 1992; Strohl et al., 1989). This approach requires the understanding of the anthracycline biosynthetic pathway as well as the function of the single genes and enzymes involved.

Stutzman-Engwall & Hutchinson (1989) and Otten et al. (1990) have cloned, in a series of cosmids, the entire Dnr biosynthesis gene cluster of Streptomyces peucetius ATCC 29050. Single genes of the cluster have been isolated recently, sequenced and expressed in heterologous systems to enable partial characterization of Dnr biosynthesis enzymes (Madduri et al., 1993).

Previously we reported the isolation of a Dxr-Dnr resistance gene contained in a 5′ kb fragment (riel) from S. peucetius subsp. castens mutant 7600 derived from ATCC 27952 (Colombo et al., 1992). riel belongs to the Dnr cluster described by Stutzman-Engwall & Hutchinson (1989) and the identified resistance gene corresponds to the drrAB gene (Guilfoile & Hutchinson, 1991). We also presented preliminary evidence that in the same fragment, close to drrAB, is a gene responsible for conversion of aklavinone to ε-rhodomycinone via hydroxylation at C-11. This finding is consistent with the observation, made by Guilfoile & Hutchinson (1991), that, in S. peucetius ATCC 29050, a spacer of 233 nt separates the drrAB gene from an open reading frame (ORF) transcribed in the

Abbreviations: Dnr, daunorubicin; Dxr, doxorubicin; GST, glutathione S-transferase.

The GenBank accession number of the sequence reported in this paper is U18082.
opposite direction. The ORF was only partially sequenced but comparison of the deduced first 79 amino acids revealed a significant similarity with FAD- and NADPH-dependent bacterial hydroxylases (Blanco et al., 1993).

In this paper we report the sequence of the S. peucetius ATCC 29050 gene for the aklavinone C-11 hydroxylase (dnrF) and its heterologous expression in Escherichia coli and Streptomyces lividans.

METHODS

**Bacterial strains and plasmids.** The strains *Streptomyces lividans* TK23 (obtained from D. A. Hopwood, John Innes Institute, Norwich, United Kingdom) and *Streptomyces peucetius* 29050 (Grein et al., 1963), 7600, 7800 and 9700 (Colombo et al., 1992) and the plasmids pWHM3 (Vara et al., 1989), pWHM603 (Guilfoile & Hutchinson, 1991) and ric1 (Colombo et al., 1992) are described elsewhere. The *ermE* promoter (Gramajo et al., 1991) was isolated from pJ4070 kindly provided by M. J. Bibb (unpublished results). pUC19, M13mp18 and M13mp19 were prepared from New England Biolabs. pT7-7 was from USB and pGEX-2T was from Pharmacia Biotech. Plasmid pGPI-2 was provided by C. R. Hutchinson. The *Escherichia coli* strains used were JM109 (Yanisch-Perron et al., 1985), DH5α (BRL) and DH1 (Hanahan, 1983).

Plasmids pFI 15 and pLA-4 contain the *dnrF* gene, obtained by PCR amplification as detailed later, cloned into the EcoRI site of pT7-7 and pGEX-2T vectors, respectively.

Plasmid pFI 14 contains a region of 2.2 kb, obtained by partial digestion of pWHM603 with BamHI, cloned into the BamHI site of the pWHM3 vector (see Fig. 1). The 1.7 kb XbaI fragment containing the *dnrF* gene from pFI 15 was cloned into the XbaI site of pJ4070 downstream of the *ermE* promoter, obtaining plasmid pFI 16. pFI 16 was double digested with SaeI and HindIII and the 2 kb fragment, corresponding to the *ermE*-*-dnrF* hybrid gene, was inserted into pWHM3 digested with the same enzymes to obtain pFI 20.

**Chemicals.** Thiostrepton was obtained from Behring Diagnostic. Aklavinone, 1,0-decarbomethoxy-aklavinone, 11-deoxystreptactinone and 11-deoxythiostrepton were purchased from standard commercial sources. All other chemicals and biochemicals, unless specified, were obtained from Sigma.

**DNA sequence analysis.** The 1.6 kb and 0.6 kb BamHI fragments from pWHM603 (see Fig. 1) containing the *dnrF* gene were subcloned in pUC19, M13mp18 and M13mp19 vectors. Plasmids for sequencing reactions were prepared using Qiagen columns. Single-stranded DNA templates were isolated in 100 mM potassium phosphate buffer (pH 7.7) containing 1 M NaCl, 0.2 M EDTA, 0.1% SDS and 0.5 M LiCl. Restriction enzymes and other molecular biology materials were purchased from standard commercial sources. All other chemicals and biochemicals, unless specified, were obtained from Sigma.

**Enzymic assay for the DnrF protein with *E. coli* crude extracts.** Induced *E. coli* DH1(pGPI-2/pFI 15) cells were harvested by centrifugation at 3000 g for 20 min in 100 ml of 2x YT medium (Sambrook et al., 1989) (made with 80 µg ml⁻¹ of both kanamycin and ampicillin). At OD₅₇₀ = 34 the temperature was raised to 42 °C for 30 min, then the cells were grown for 120 min further at 30 °C. For the bioconversion assay antracyclines dissolved in DMSO at 50 µg ml⁻¹ and NADPH (5 mM final concentration) were added to a DH1(pGPI-2/pFI 15) induced culture and cells were grown for a further 60 min at 30 °C. The cultures were extracted and analysed by HPLC.

**Enzymic assay for the DnrF protein with *S. lividans* crude extracts.** *S. lividans* (pFI 20) was grown in YEME medium (Hopwood et al., 1985) for 76 h, harvested by centrifugation, washed in PBS, resuspended in 4 vols lysing buffer and disrupted by sonication. After centrifugation at 15000 g the supernatant was used as a crude enzyme preparation.

**HPLC analysis.** Acetonitrile/methanol (1:1, v/v) was used to extract the metabolites. The extracts were filtered by a Mille-
We previously reported the isolation and characterization of Streptomyces strains, prepared and analysed by HPLC. The metabolites were identified by HPLC with a reverse-phase Hypersil ODS-microbore column (Shandon). A linear gradient of 23-67% acetonitrile buffered at pH 2.5 with 125 mM sodium monobasic phosphate was applied for 52 min at a flow rate of 0.125 ml min⁻¹. The elution profile was monitored with a Hewlett-Packard diode array detector at 488, 424 or 254 nm.

**Streptomyces transformation and bioconversion experiments.** Protoplasts were produced from S. lividans TK23 grown in YEME medium supplemented with 0.5% glycine at 28°C as described by Hopwood et al. (1985a) and regenerated on R2YE agar. S. lividans transformants were selected and maintained on thiostrepton at 50 µg ml⁻¹. S. lividans TK23(pFI 14) or S. lividans TK23(pFI 20) was grown for 48 h in YEME medium at 30°C. One millilitre of the culture was used to inoculate 25 ml fermentation medium described by McGuire et al. (1979). Anthracyclines were added to 68 h cultures and, after 24 h, 650 mg oxalic acid was added. Cultures were incubated at 50°C for 45 min, then the crude extracts were prepared and analysed by HPLC.

**Preparation of antiserum to the GST-DnrF fusion protein.** Antibodies against the purified recombinant protein were produced in 6 week old Balb/c female mice according to the following protocol: the first intraperitoneal shot contained 50 µg purified recombinant protein in 50 µl sterile PBS plus 50 µl complete Freund adjuvant. Four more injections, at 2 week intervals, were done with the same amount of antigen but without adjuvant. In Western blot analysis the serum was used at a 1:1000 dilution.

**Protein analysis.** SDS-PAGE and Western blot analysis were performed as previously described (Buvoli et al., 1990). Mouse anti-GST-DnrF antibodies were used at a 1:1000 dilution and revealed by alkaline-phosphatase-conjugated anti-mouse IgG (Promega).

**RESULTS**

**Cloning and sequencing of the aklavinone 11-hydroxylase gene**

We previously reported the isolation and characterization of ricl, a 5.7 kb DNA fragment from the S. roseosporus 7600 mutant, which contains the Dxr-Dnr resistance gene (drrAB). Bioconversion experiments with S. lividans TK23 (ricl) suggested the presence, close to drrAB, of the dnrF gene, which converts aklavinone (a yellow compound) to e-rhodomycinone (red) via 11-hydroxylation (Colombo et al., 1992).

We have now mapped, by means of the same bioconversion assay, the dnrF gene within pWHM603 (Guilfoile & Hutchinson, 1991), which contains part of the Dnr cluster from S. peucetius ATCC29050 (Fig. 1). The region conferring the ability to S. lividans TK23 to convert the added aklavinone to e-rhodomycinone (80% conversion) is located in a 2.2 kb fragment (plasmid pFI 14) (Fig. 1).

The pFI 14 insert was sequenced and, as shown in Fig. 2, the first 118 nt correspond to the 5’ end of the drrAB gene (non-coding strand) (Guilfoile & Hutchinson, 1991). The ORF for the dnrF gene was identified on the basis of previously reported considerations (Bibb et al., 1984) about the distribution of G+C content across codons in Streptomyces genes. Thus a region spanning from position 352 to 1822 contains the dnrF gene (Fig. 2), beginning with a putative GTG translation start codon and ending with a TGA stop codon. A plausible ribosome-binding site (GGAGG) (Strohl, 1992) is located 6 nt upstream of the GTG codon. The first 237 nt of the ORF were previously sequenced by Guilfoile & Hutchinson (1991) and proposed to be part of the dnrF gene.

The drrAB and dnrF genes are separated by a spacer of 233 nt that contains the drrAB promoter (Guilfoile & Hutchinson, 1991) and most likely the divergent dnrF promoter. This region is significantly AT-richer than the remaining part of the clone (A+T 46% vs 30%).

**Analysis of the deduced dnrF gene product**

The dnrF gene encodes a protein of 489 amino acids with a predicted molecular mass of 52 kDa. Sequence comparison with the GenBank database performed by the FASTA and TFASTA programs showed significant
resemblance to two groups of hydroxylase enzymes (Fig. 3): those catalysing the introduction of hydroxyl groups into polycyclic aromatic polyketide antibiotics (Blanco et al., 1993; Decker et al., 1993) and those involved in catabolism of aromatic compounds in different organisms (Blanco et al., 1993; Kalin et al., 1992; Nurk et al., 1991; Perkins et al., 1990; Weijer et al., 1982). No similarity with P-450 cytochromes (Andersen et al., 1993; Stassi et al., 1993; Weber et al., 1991) involved in the hydroxylation of erythromycin, or with actA gene products (Caballero et al., 1991), responsible for hydroxylation of the ring structures in actinorhodin biosynthesis, was found. As shown in Fig. 3, the DnrF enzyme contains two motifs that are common to a number of FAD- and NADPH-dependent enzymes. The first motif, located at the N-terminus of the protein (amino acids 3-47), is the so called

Fig. 2. Nucleotide sequence of the 2.2 kb insert from plasmid pFl14 with the deduced amino acid sequence of the dnrF ORF in the single-letter code. The putative RBS of the dnrF gene is boxed. The first 118 nt are complementary to the 5' end of the drrAB ORF (codon arrangement is represented).
Fig. 3. Comparison of the deduced amino acid sequence of the \textit{dnrF} gene with sequences of amino acids encoded by other hydroxylase genes. The alignment shows the similarity among the \textit{dnrF} product and different hydroxylases in the 'ADP-binding site' and in the 'FAD-binding site' motifs, present in many FAD- and NADPH-dependent enzymes. The amino acids in the consensus sequence (on top of each motif) are the ones present in at least half of the compared protein sequences. Asterisks indicate the non-conserved positions. Dashes in the protein sequences represent amino acids identical to the consensus. Grey boxes highlight the more conserved regions. The hydroxylase proteins compared are as follows: SchC (Blanco et al., 1993); WhiE (Blanco et al., 1993); TfdB (Perkins et al., 1990); PheA (Nurk et al., 1991); TcmG (Decker et al., 1993); PclO (Blanco et al., 1993); PhyA (Kalin et al., 1992); Hbh (Weijer et al., 1982); OtcC (Blanco et al., 1993).

\[ \beta\alpha\beta \text{ fold and is involved in binding of the ADP moiety of FAD (Wierenga et al., 1986). The second motif seems to be important in binding the ribityl chain of the flavin moiety of FAD (Eggink et al., 1990; Russel & Model, 1988). These similarities further support the conclusion that the \textit{dnrF} gene encodes for the aklavinone 11-hydroxylase.} \]

Expression of the \textit{dnrF} gene in \textit{Streptomyces} and in \textit{E. coli}

We stated above that, upon transformation with pFI 14, \textit{S. lividans} acquired the ability to bioconvert aklavinone to \textit{e}-rhodomycinone. Incubations for 1 h gave 11-hydroxylation of 10\% of the aklavinone (50 \mu g ml\(^{-1}\)) added to the medium. Complete conversion was seen after 6 h. Since this could be due to low levels of gene expression, we cloned the \textit{dnrF} gene (1605 nt) into the pWHM3 vector under the control of the strong \textit{ermE} \* promoter (pFI 20 plasmid; see Methods). \textit{S. lividans}(pFI 20) converted 100\% of aklavinone (50 \mu g ml\(^{-1}\)) to \textit{e}-rhodomycinone in 1 h incubations. Moreover, this transformant also catalysed the hydroxylation of compounds biochemically derived from aklavinone (see Table 1 and Fig. 4), but it had no activity with molecules other than anthracyclines (Table 1).

For an initial biochemical characterization we expressed the \textit{dnrF} gene in \textit{E. coli} as a fusion with GST. To this aim the \textit{dnrF} gene was PCR-amplified and cloned into the \textit{EcoRI} site of the pGEX-2T vector (see Methods). The
Table 1. Percentage in vitro and in vivo conversion of different chemical species

<table>
<thead>
<tr>
<th>Chemical species</th>
<th>Concn (µg ml⁻¹)</th>
<th>In vitro conversion</th>
<th>In vivo cell bioconversion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>E. coli pFI 15</td>
<td>S. lividans pFI 20</td>
</tr>
<tr>
<td>Aklavinone</td>
<td>50</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>10-Decarbomethoxy-aklavinone</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11-Deoxykarminomycinone</td>
<td>50</td>
<td>70%</td>
<td>100%</td>
</tr>
<tr>
<td>11-Deoxydauromycinone</td>
<td>50</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>11-Deoxy-4-demethyl-adriamycinone</td>
<td>30</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>11-Deoxykarminomycin</td>
<td>30</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>11-Deoxydauromycin</td>
<td>30</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>11-Deoxy-4-demethyl-daunomycin</td>
<td>30</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>1,8-Dihydroxyanthraquinone</td>
<td>50</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

GST–DnrF fusion, purified on a glutathione agarose column, was assayed for its ability to convert aklavinone to ε-rhodomycinone and used as antigen to raise polyclonal antibodies in a mouse. Since the fusion protein did not show any hydroxylase activity (data not shown), we cloned the dnrF gene into the pT7-7 vector that allowed the expression of an unfused DnrF protein. Interestingly, upon transformation with this construct (pFI 15), E. coli was able to bioconvert aklavinone to ε-rhodomycinone (Table 1). Furthermore, the cell extract prepared from E. coli(pFI 15) (see Methods) catalysed the aklavinone 11-hydroxylation in a NADPH-dependent reaction, as determined by HPLC analysis (Fig. 5). Both E. coli(pFI 15) cells and crude extracts were active on substrates other than aklavinone, and showed the same specificity observed with S. lividans(pFI 20) cells and extracts (Table 1).

To confirm the expression of the DnrF protein in heterologous systems we carried out Western blot experiments using the antibodies raised to the GST–DnrF fusion. As shown in Fig. 6, a protein with an apparent molecular mass of 52 kDa was recognized in extracts from S. peucetius ATCC 29050, E. coli DH1(pFI 15) and S. lividans TK23(pFI 20) while no signal was detected in extracts from E. coli DH1(pT7-7) and S. lividans TK23. The same protein was present in the S. peucetius 7600 strain but absent in the non-producing S. peucetius 9700 mutant. Moreover, no protein band was recognized in extracts of the S. lividans TK23(pFI 14) strain, which indicates a low level of dnrF expression. This result is compatible with the reduced rate of bioconversion observed.

DISCUSSION

In this paper we report the sequencing and expression of the dnrF gene from S. peucetius ATCC 29050. The gene belongs to the Dxr-Dnr biosynthetic gene cluster and, as previously suggested (Colombo et al., 1992), it is closely associated with the Dxr-Dnr resistance gene drrAB. The dnrF gene encodes an enzyme of 489 amino acids that catalyses the conversion of aklavinone to ε-rhodomycinone via 11-hydroxylation. Accordingly, the analysis of the deduced DnrF protein sequence revealed the presence of two motifs common to a number of FAD- and NADPH-dependent hydroxylases involved in catabolism of aromatic compounds. As shown in Fig. 3, both regions are formed by alternating divergent and conserved sequences. In particular, the more C-terminal motif seems to be formed by three conserved regions separated by amino acid stretches containing Gly and Pro, two residues very often associated with protein bending. The first conserved box is the only one where Cys or His can be found. It is intriguing to observe that DnrF is the only protein among those shown that lacks these residues. We think that these differences could reflect enzymic specificity and, at the same time, can be used as a measure of the relatedness of different members of this protein family.

For preliminary biochemical characterization we expressed the dnrF gene in both S. lividans and E. coli cells. In vitro conversion assays carried out with crude extracts demonstrated that the 11-hydroxylation reaction is NADPH-dependent. DnrF has a significant substrate specificity, since it is completely inactive on molecules other than anthracyclines (see Table 1).

Aklavinone, the natural substrate, is a key intermediate in the biosynthesis of Dxr and Dnr and requires several successive chemical modifications such as hydroxylation, glycosylation, oxidation and methylation. DnrF catalyses the first among these modifications, i.e. 11-hydroxylation of aklavinone (yellow) to ε-rhodomycinone (red). The fact that DnrF acts early in the metabolic pathway of Dnr biosynthesis does not mean that 11-hydroxylation is required for all the successive modifications to take place. For instance, the dnrF 7800 mutant is able to produce 11-deoxydaunorubicin (yellow) and 11-deoxydoxorubicin (yellow). These results support the alternative hypothesis that the temporal order is simply linked to the substrate specificity of the enzyme because DnrF utilizes the 11-
Fig. 4. Schematic Dnr-Dxr biosynthetic pathway (red compounds) and the structural formula of each compound. The enzymic reaction carried out by the dnrF gene product is indicated. Broken arrows indicate yellow compounds derived from the biosynthetic intermediates of the yellow mutant S. peucetius 7800. The box contains the aglycons compounds.
S. FILIPPINI and OTHERS

Fig. 5. HPLC analysis of aklavinone bioconversion by E. coli(pFl 15) (a) and S. lividans(pFl20) (b) at 254 nm. (c) Aklavinone and e-rhodomycinone relative spectra of the samples from (a). A, Aklavinone; B, e-rhodomycinone.

Fig. 6. Western blot analysis of E. coli and Streptomyces protein extracts with anti-GST-DnrF mouse polyclonal antibodies. Lanes (a): GST, purified glutathione S-transferase; pLA-4, GST-DnrF fusion protein purified from E. coli DH5α(pLA-4); pT7-7, E. coli DH1(pT7-7) total extract; pFl 15, E. coli DH1(pFl 15) total extract. Lanes (b): TK23, total extract from S. lividans TK23; pFl 14, total extract from S. lividans TK23(pFl 14); pFl 20, total extract from S. lividans TK23(pFl 20); 29050, total extract from S. peucetius ATCC 29050; 7600 and 9700, total extracts from 7600 and 9700 mutants derived from S. peucetius ATCC 27952. The positions of the molecular mass markers are indicated on the left of each panel.
where the promoters of both genes are located. It is conceivable that this arrangement can be relevant for a possible coordinate expression of the two genes, as already demonstrated for many prokaryotic and eukaryotic divergent promoters (Beck & Warren, 1988). One indication supporting this idea derives from the dnrF 7800 mutant that produces 11-deoxydaunorubicin and 11-deoxydoxorubicin. This mutant fails to express the drrAB gene and is sensitive to Dxr and Dnr. It is reasonable to think that the drrAB gene is primarily involved in conferring resistance to 11-hydroxylated anthracyclines and therefore is absolutely required after the activation of the dnrF gene.

Altogether these considerations point to the drrAB–dnrF gene system as a possible target for a major control point in the anthracycline biosynthetic pathway, a hypothesis that deserves further analysis.

ACKNOWLEDGEMENTS

This work was supported by a Biotechnology P.N.R. contract between M.U.R.S.T. and Farmitalia Carlo Erba. We thank G. Rivola for the 11-deoxanthracyclines and for helpful discussions. We also thank C. R. Hutchinson for providing pWHM603 and for helpful suggestions. D.G. and M.S. are supported by a fellowship of the National Research Council Progetto Finalizzato Ingegneria Genetica and ‘A. Buzzi Traverso’ foundation, respectively.

REFERENCES


(pheA) encoding phenol monooxygenase from Pseudomonas sp. EST1001: expression in Escherichia coli and Pseudomonas putida. Gene 102, 13–18.


Received 18 November 1994; accepted 23 December 1994.