Organization of a Cluster of Erythromycin Genes in Saccharopolyspora erythraea

J. MARK WEBER,* JUDITH O. LEUNG, GREGORY T. MAINE, RICA H. B. POTENZ, THOMAS J. PAULUS, AND JANET P. DEWITT


Received 1 September 1989/Accepted 16 January 1990

Erythromycin A is a medically important macrolide anti-biotic produced by the gram-positive, sporforming bacterium Saccharopolyspora erythraea (formerly Streptomyces erythraeus NRRL 2338) (29). As is the case for other antibiotic-producing organisms, the genes for the biosynthesis of erythromycin (ery) are thought to be clustered about the gene for resistance to erythromycin, ermE (22, 27, 29); however, much remains to be learned about the organization and location of the many individual genes predicted to be involved in the pathway (6, 18).

The ermE gene was originally cloned by Thompson et al. (24) and later sequenced by Uchiyama and Weisblum (25); its promoter and the promoter of an adjacent upstream open reading frame were characterized by Bibb et al. (3, 4). The adjacent open reading frame has since been sequenced by Dhillon et al. (7), inactivated by gene disruption, and found to be involved in erythromycin biosynthesis as an eryC-type gene. Independently, the open reading frame was shown by Vara et al. (25a) to be an eryC-type gene through complementation of the eryCl60 allele, and we obtained the same results as Dhillon et al. (7) by using linker insertion mutagenesis (see below). Other landmarks in the ery gene cluster are (i) eryG, the O-methyltransferase located 6.5 kilobase pairs (kb) downstream of the ermE promoter region (15a, 28); (ii) eryAI, the macrolactone synthetase, a large locus beginning 3 kb upstream of eryG (J. Tuan, J. M. Weber, M. Staver, J. O. Leung, and L. Katz, Gene, in press); and (iii) the eryB25, eryB26, and eryD24 genes, mapped within an 18-kb region upstream of ermE (25a).

Despite this recent progress, large areas of the ermE region were functionally uncharacterized prior to this study, and the C-6 hydroxylase (see Fig. 3, step 2) was genetically uncharacterized owing to a lack of mutations in this step. In the work reported here we investigated the function of these uncharacterized regions between the known genes by using S. erythraea and two methods we developed for this purpose which make further use of the plT702-based plasmid integration and excision system (27, 28). The first approach, scanning gene disruption, was modeled after mutational cloning described by Chater and Bruton (5) for the analysis of the methylenomycin gene cluster in Streptomyces coelicolor and it is analogous to a similar system described for Saccharomyces cerevisiae (20). The second approach, linker scanning, was modeled after linker insertion (8, 15, 23), which has been used in both procaryotic and eucaryotic systems for mutational analysis of complex genetic loci.

We report the creation of a set of mutations covering a region of approximately 30 kb and the identification of five new loci involved in synthesis or attachment of the erythromycin sugars, eryB1, eryBII, eryCl, eryCI, and eryH. Mutations in eryH block both the synthesis or attachment of mycarose (as in an eryB mutant) and the previous step, the C-6 hydroxylation reaction. We also present additional information on the previously identified genes in the cluster, including the first mutational characterization of the ermE gene in S. erythraea.

MATERIALS AND METHODS

Bacterial strains, cloning vectors, and growth media. Wild-type derivatives of S. erythraea were used in this study. The host strains for integrative transformation were S. erythraea UW110 (met-4 leu-18 rif-63) and UW267 (arg-6 eryB25). Cross-feeding and precursor feeding test strains were as follows: eryA, UW352 (arg-14 eryA34); eryB, UW267 (arg-6 eryB25); eryC, UW254 (eryC160); and eryD, UW280 (arg-8 eryD24), described previously (29). The host strain for autonomously replicating plT702 derivatives was Streptomyces lividans TK21 (kindly provided by D. A. Hopwood, John Innes Institute, Norwich, United Kingdom). The host strain for Escherichia coli-derived plasmid vectors was E. coli DH5α (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) or JM109 (31). The erythromycin assay strain Bacillus

* Corresponding author.
### TABLE 1. Plasmids used for gene disruption or gene replacement analyses

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Insert or linker no.</th>
<th>Parent vector</th>
<th>Insert (+) or deletion (−) size (kb)</th>
<th>Comments or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGM7</td>
<td>E-7</td>
<td>pGM420</td>
<td>−0.8 + BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM8</td>
<td>Cl-8</td>
<td>pGM420</td>
<td>−0.4 + BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM14</td>
<td>E-14</td>
<td>pGM420</td>
<td>+BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM16</td>
<td>CI-16</td>
<td>pGM420</td>
<td>+BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM18</td>
<td>CI-18</td>
<td>pGM420</td>
<td>+BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM406</td>
<td>+R-000</td>
<td>pIJ702</td>
<td>+0.45</td>
<td>w.t. insertion</td>
</tr>
<tr>
<td>pGM434</td>
<td>H-434</td>
<td>pGM420</td>
<td>+BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM436</td>
<td>+R-036</td>
<td>pGM420</td>
<td>+BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM449</td>
<td>+/G-449</td>
<td>pIJ487</td>
<td>+1.8</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pGM450</td>
<td>+R-050</td>
<td>pGM420</td>
<td>−0.05 + BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM454</td>
<td>+R-054</td>
<td>pGM420</td>
<td>−0.79 + BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM457</td>
<td>+R-057</td>
<td>pGM420</td>
<td>−2.26 + BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM458</td>
<td>+R-058</td>
<td>pGM420</td>
<td>+BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM459</td>
<td>+R-059</td>
<td>pGM420</td>
<td>+BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM460</td>
<td>+R-060</td>
<td>pGM420</td>
<td>+BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pGM461</td>
<td>+R-061</td>
<td>pGM420</td>
<td>+BglII</td>
<td>Gene replacement</td>
</tr>
<tr>
<td>pMW2</td>
<td>−1</td>
<td>pIJ702</td>
<td>+5</td>
<td>w.t. insertion (26)</td>
</tr>
<tr>
<td>pMW6</td>
<td>+6</td>
<td>pIJ702</td>
<td>+1.6</td>
<td>w.t. insertion</td>
</tr>
<tr>
<td>pMW11-A</td>
<td>A-11</td>
<td>pIJ702</td>
<td>+0.4</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW37-3</td>
<td>A-3</td>
<td>pIJ702</td>
<td>+2.5</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW37-17</td>
<td>A-17</td>
<td>pIJ450</td>
<td>+1.7</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW37-18</td>
<td>A-18</td>
<td>pMW27</td>
<td>+1.9</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW40-15</td>
<td>A-15</td>
<td>pMW27</td>
<td>+1.75</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW52-9</td>
<td>BII-9</td>
<td>pMW27</td>
<td>+0.5</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW42-14</td>
<td>CII-14</td>
<td>pMW27</td>
<td>+0.9</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW42-15</td>
<td>CII-15</td>
<td>pMW27</td>
<td>+0.9</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW42-32</td>
<td>+32</td>
<td>pMW27</td>
<td>+0.98</td>
<td>w.t. insertion</td>
</tr>
<tr>
<td>pMW42-36</td>
<td>CII-36</td>
<td>pMW27</td>
<td>+0.9</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW44</td>
<td>+G-44</td>
<td>pGM420</td>
<td>+5.3</td>
<td>w.t. insertion</td>
</tr>
<tr>
<td>pMW55-26</td>
<td>H-26</td>
<td>pMW27</td>
<td>+0.37</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW55-27</td>
<td>BII-27</td>
<td>pMW27</td>
<td>+0.7</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW56-3</td>
<td>+3</td>
<td>pMW27</td>
<td>+0.53</td>
<td>w.t. insertion</td>
</tr>
<tr>
<td>pMW56-19</td>
<td>BII-19</td>
<td>pMW27</td>
<td>+0.55</td>
<td>Gene disruption</td>
</tr>
<tr>
<td>pMW56-33</td>
<td>H-33</td>
<td>pMW27</td>
<td>+0.43</td>
<td>Gene disruption</td>
</tr>
</tbody>
</table>

* For more details see Materials and Methods.
* See Fig. 4.
* +BglII linker inserted in place of the deleted DNA; deletion indicated by −.
* All plasmids were transformed into strain UW110 (met4 leu18 rif63), except pGM7 and pGM14, which were transformed into UW267 (arg-l eryB25).
* w.t., Wild type.

* subtilis PY79 was donated by P. Youngman, University of Pennsylvania, Philadelphia.

Plasmids pIJ702 (11), pIJ450, pIJ486, and pIJ487 (26) were obtained from D. Hopwood and M. Bibb, John Innes Institute. Cosmids p7A2 and pJ1 and subclones containing the erythromycin resistance gene and fragments of the *eryA1* locus were generously provided by L. Katz, J. Tuan, and M. Staver, Abbott Laboratories; a description of the isolation of *ermE*-containing cosmids is described elsewhere (Tuan et al., in press). A list of plasmids mentioned in this report is presented in Table 1.

*S. erythraea* strains were grown in liquid culture in 50-ml tryptic soy broth (Difco Laboratories, Detroit, Mich.) or in 500-ml shake flasks at 32°C. For solid cultures, plates of tryptic soy broth–2% Bacto-Agar (Difco) were used for cross-feeding and precursor feeding experiments, or R72 agar plates (28) were used for regeneration of mycelia from protoplasts or as a general sporulation medium. Thiostrepton was added at 5 μg/ml (liquid culture) or 10 μg/ml (solid culture) for selection of *S. erythraea* transformants. *Streptomyces lividans* was handled under conditions specified by Hopwood et al. (10). *E. coli* was grown on LB plates or broth; ampicillin was added at 100 μg/ml when needed.

**Antibiotics, enzymes, and chemicals.** Thiostrepton was a gift of S. J. Lucania, E. R. Squibb & Sons, Princeton, N.J.; ampicillin was purchased from Sigma Chemical Co., St. Louis, Mo.

Commercially available restriction endonucleases and T4 DNA ligase were purchased from Bethesda Research Laboratories or New England Biolabs, Inc., Beverly, Mass. Calf intestinal alkaline phosphatase and DNase I were purchased from Boehringer Mannheim Biochemicals, Indianapolis, Ind. Reagents and chemicals for DNA labeling were purchased from Bethesda Research Laboratories (Nick Translation System). [2P]dCTP was purchased from Amersham Corp., Arlington Heights, Ill.

2,3,5-Triphenyltetrazolium chloride, p-aniisaldehyde, dextran sulfate, Ficoll (type 400), polyvinylpyrrolidone, salmon sperm DNA, and bovine serum albumin were purchased from Sigma.

**New multifunctional parent vectors for integrative transformation in *S. erythraea*.** Two new E. coli-Streptomyces-Saccharopolyspora shuttle vectors, pMW27 and pGM420, were created for this study (Fig. 1). pMW27 was constructed from pIJ702 (10, 11) and pUC 18 (31) joined at their unique *KpnI* sites in the orientation with the *EcoRI* site of pUC 18 closer to the *BglII* site of pIJ702. pGM420 was constructed from pIJ702 and pUC 18 and then digested with *BglII* sites in the orientation with the *HindIII* site of pUC 18 closer to the *BglII* site of pIJ702.

**Transformation of *S. erythraea*.** The method of Weber and Losick (27) was used for transformation of *S. erythraea*. A similar technique is described by Yamamoto et al. (30).

**Scanning gene disruption.** Scanning gene disruption involved the following general approach. A cloned DNA fragment from the *S. erythraea* chromosome covering the region of interest was isolated by gel electrophoresis, purified from the agarose gel with Geneclean (Bio101, La Jolla, Calif.) and partially digested with *SalI*. DNA fragments in the size range of approximately 0.2 to 1.0 kb were isolated by gel electrophoresis, purified from the gel with Geneclean, and ligated to pMW27 which had then been digested with *BglII* and treated with calf intestinal alkaline phosphatase.

*E. coli* DH5α cells were transformed with the products of the ligation, and transformants were selected with 100 μg of ampicillin per ml on LB agar. Transformants carrying plasmids with inserts were identified by colony blot hybridization (14) with the original DNA fragment or subfragment as a radioactive probe. Plasmids with inserts (and intact pIJ702 sequences) were integratively transformed (separately) into protoplasts of *S. erythraea* UW110. Primary transformants of *S. erythraea* were selected with 10 μg of thiostrepton per ml on R2T2 plates. The spores of primary transformants (a mixture of thiostrepton-resistant and thiostrepton-sensitive spores) were harvested and combined in 1 ml of 20% glycerol. Integrated transformants were isolated from the mixture by either (i) plating 0.2 ml of the combined primary transformant spores onto R2T2, harvesting the lawn of spores, and replating for single colonies on R2T2 plus thiostrepton (20 μg/ml) or (ii) passing the combined spores of primary transformants through two to three single colony purifications on R2T2 plus thiostrepton (20 μg/ml).

To determine the location of the plasmid insertion, chromosomal DNA was prepared from the integratively transformed strains and the parent strain. The DNAs were compared by Southern analysis, probing with the DNA fragment covering the region into which insertions were
being made. Integration by homologous recombination was implied by the absence of a band comigrating with the parental DNA band and by the presence of two new junction bands in the transformant DNA (see Fig. 2A and B). A more precise mapping of the insertion site was obtained through analysis of the size of the junction fragments and through restriction and Southern analyses of the plasmid.

Construction of other integrative plasmids for gene disruption. In the course of this analysis, several other integrative plasmids were constructed.

pMW11-A was used for insertional inactivation of the eryG gene, based on the experiment described previously (28). It was constructed from the 0.4-kb PstI fragment isolated from pGM403 (described below) cloned into the unique PstI site of pIJ702.

pMW37-18, pMW37-17, and pMW37-3 were constructed for the initial gene disruption experiments in eryAI; their construction and characterization are described elsewhere (Tuan et al., in press) and described in this report with permission from the authors.

pMW44 is an integrative plasmid that contains the +/−44 insert (see Fig. 4). It was constructed in E. coli DH5α cells (Bethesda Research Laboratories) in a two-step process. First, pMW43 was prepared from the 5.3-kb XhoI fragment between sites 18 and 30 (see Fig. 4) obtained from cosmid pJJ (Tuan et al., in press) cloned into the unique XhoI site of pBI30 (International Biotechnologies, Inc., New Haven, Conn.). Next, pMW43 was digested with BamHI and HindIII, releasing the clone with new ends for subcloning into BgII-HindIII-digested pGM420 (Fig. 1B) to create pMW44.

pGM403 was an intermediate in the construction of pGM406 and pGM449. It was constructed from cosmid p7A2 (Tuan et al., in press) digested with HindIII and EcoRI, releasing a 9.1-kb HindIII-EcoRI fragment (see Fig. 4, sites 1 to 21) containing the eryG gene, which was purified by gel electrophoresis. pBR322 was digested with EcoRI and BamHI, and the resulting DNA was ligated to the 4.5-kb EcoRI-BamHI fragment containing eryG (see Fig. 4, sites 11 to 21), obtained by digesting the purified 9.1-kb HindIII-EcoRI fragment with BamHI. The resulting plasmid, pGM403, contains the eryG gene cloned into pBR322.

pGM406 contains the +406 insert (see Fig. 4); it was constructed from KpnI-digested pGM403 that had been treated with calf intestinal alkaline phosphatase and ligated to KpnI-digested pIJ702.

pGM449 was constructed from pGM403 digested with XhoI and treated with DNA polymerase I (Klenow fragment) to create blunt ends. The resulting DNA was then digested with EcoRI, and the 1.8-kb fragment that was produced was purified by gel electrophoresis. pJ487 (26) was digested with BamHI, treated with Klenow fragment to form blunt ends, and then digested with EcoRI. The 6.2-kb fragment that was produced was ligated to the 1.8-kb fragment from pGM403 to create pGM449.

Linker scanning. Linker scanning in S. erythraea combines the in vitro mutagenic technique of linker mutagenesis (8) with a method for gene replacement. In S. erythraea, gene replacement can be performed by exploiting the pIJ702 based integration-excision system (28).

Plasmids were constructed of 6- to 14-kb DNA fragments from the ermE region of the S. erythraea chromosome cloned into pUC18 (Bethesda Research Laboratories). The plasmids were treated with DNAseI to generate single-cut, full-length linear plasmid DNA whose ends were made blunt through treatment with DNA polymerase I (Klenow fragment). Phosphorylated BgII linkers (New England BioLabs) were ligated to the plasmid DNA in 10-fold molar excess during a 16-h incubation at 15°C. The ligation products were heated at 70°C for 10 min and then digested with an excess of BgII for 4 h at 37°C. Linearized DNA was gel purified to remove plasmids not containing BgII linkers, religated at

FIG. 1. Restriction maps for plasmids pMW27 and pGM420. These plasmids were constructed from pIJ702 and pUC18 as described in Materials and Methods and were used as autonomously replicating high-copy vectors in E. coli or Streptomyces lividans and as parent vectors for the formation of integrative plasmids in S. erythraea. Sites available for cloning and preserving the trinuc functioning of pMW27 are XbaI, HindIII, EcoRI, SstI, and BglII; in pGM420 the sites are KpnI, PstI, SphI, BgII, XbaI, HindIII, and EcoRI.
15°C overnight, and transformed into *E. coli* JM109 cells selecting for ampicillin resistance. Plasmids with new *BglII* sites located within the *S. erythraea* DNA were identified, subcloned into pGM420 (Fig. 1B), and integratively transformed into *S. erythraea* UW110. Purified integrated transformants were physically characterized by chromosomal Southern analyses to confirm the presence of new DNA sequences in the chromosome (see Fig. 2C and D) and biochemically analyzed to determine the erythromycin phenotype.

Integrated strains were passed nonselectively through a cycle of growth and sporulation, allowing for spontaneous excision of the integrated vector. Depending on the site of the second crossover with respect to the mutation, excision of the plasmid led to the stable incorporation of the mutation into the host genome or reconstruction of the original wild-type sequence. Chromosomal Southern analysis of the excised transformant derivatives was used to physically identify the strains in which gene replacement had occurred (see Fig. 2C and D).

**Techniques for the characterization of the Ery− phenotype.** Bioassays for erythromycin and the erythromycin precursor feeding technique are described by Weber et al. (29). Identification of erythromycin pathway intermediates by thin-layer chromatography and the cross-feeding technique for characterization of *S. erythraea* mutants are described by Weber et al. (29) (also see Fig. 5 and 6).

**RESULTS**

**Mutational analysis of the ermE region.** Scanning gene disruption in *S. erythraea* is based on the ability of pIJ702-derived plasmids bearing segments of *S. erythraea* chromosomal DNA to integrate into the chromosome by single-reciprocal (Campbell) recombination (27, 28). In principle, if the cloned chromosomal DNA is internal to a transcription unit, integration of the plasmid should be mutagenic (an example is presented in Fig. 2A and B). If it overlaps a transcriptional boundary, insertion is predicted to be non-mutagenic.

Linker scanning in *S. erythraea* relies on the pIJ702-based gene replacement technique (described below) and the linker method for in vitro mutagenesis of cloned DNA (8). It was used for the introduction of in vitro-mutagenized DNA sequences into the chromosome in place of the wild-type allele at its natural location. Variations on the method of gene replacement have been described for other antibiotic-producing species by Piret and Chater (16), Anzai et al. (1), and Yagi et al. (Y. Yagi, M. W. Smith, J. L. Lampi, H. A. Whaley, and C. L. Haber, Abstr. Annu. Meet. Am. Soc. Microbiol. 1989, O-15, p. 307), although the process was originally described for *Saccharomyces cerevisiae* by Scherer and Davis (17).

To characterize gene replacement in *S. erythraea*, we analyzed total DNA during three stages of the process: before and after integration of the plasmid, and after excision of the plasmid. An example involving the *eryCII6* mutation is presented (Fig. 2D). The predicted transformant structures fell into four classes (Fig. 2C). Classes I and II resulted from single reciprocal crossovers on opposite sides of the linker insertion; classes III and IV indicated the operation of a homogenetization (gene conversion) process functioning in both directions, as well as integration. Analysis of the phenotypes of integrative intermediates in gene replacement indicated whether the transcript was entirely contained within the cloned fragment or overlapped one or both ends of the insert.

We found that the percentage and class of homogenotes varied depending on the mutation. In the experiment described, three of four transformants were homogenotes, although only one representative of the two different classes is shown (Fig. 2D). Another type of transformant was sometimes observed that contained tandem duplications of the plasmid in the chromosome; however, this phenomenon has not yet been further characterized.

In the final stage of gene replacement, the vector sequences were allowed to spontaneously excise from the chromosome. This process occurred when the cells were grown through a cycle of sporulation in the absence of drug selection. These strains were usually found by screening 100 to 200 colonies for the loss of the plasmid-associated thioestrepton resistance and then by Southern analysis to show loss of the plasmid and introduction of a new restriction site into the chromosome.

To simplify the handling of larger numbers of mutagenic plasmids needed for scanning mutagenesis, we developed plasmids pMW27 and pGM420, which could be used as high-copy plasmids in *E. coli* and *Streptomyces* spp. and could integrate at high frequency in *S. erythraea* (Fig. 1A and B). The efficiency of integrative transformation was such that integration could be achieved with plasmids bearing as little as 300 base pairs (bp) of homology to the *S. erythraea* chromosome. This improved the odds of identifying small transcriptional units in the *ermE* region by scanning gene disruption.

Although both gene disruption and gene replacement in *S. erythraea* utilize pIJ702-based vectors, little is known about the cause of the instability of pIJ702 in *S. erythraea*. To determine whether the pIJ702 replicon was required for integrative transformation, we transformed *S. erythraea* with pIJ702-based bifunctional plasmids isolated from *E. coli* which contained *S. erythraea* DNA inserted into the pIJ702 replicon at the *BamHI* site (12, 13). In parallel, transformations were performed with pIJ702-based plasmids carrying *S. erythraea* DNA at *BglII*, the usual cloning site. Transformants reproducibly appeared with plasmids carrying *BglII* inserts at frequencies between 10³ and 10⁴ primary transformants per μg of DNA, but never appeared with plasmids carrying *BamHI* inserts. We speculate that the pIJ702 replicon must function at least partially to obtain primary transformants and that perhaps it is able to be maintained relatively stably in the chromosome of this species because its replication functions are defective in *S. erythraea*.

The *eryAI* locus. *eryA* mutants are blocked early in the biosynthetic pathway (Fig. 3), either in the synthesis of the erythronolide ring (6-deoxyerythronolide B [DEB]) or in the synthesis of its precursors, methylmalonyl coenzyme A or propionyl coenzyme A (6, 29). The enzymes involved in the later steps of the pathway, however, are functional in *eryA* strains, as inferred from the results of cross-feeding and precursor-feeding experiments (29) in which the intermediates erythronolide B (EB) and 3-a-nycarosyl erythronolide B (MEB) (Fig. 3) are converted into erythromycin.

The *eryAI* locus was first defined through complementation of the *eryAI941* mutation and through gene disruption experiments in the area around *eryAI941*, showing the minimum size of the locus to be 5 kb (Tuan et al., in press).

We have continued the gene disruption analyses and in so doing have found the left end (closest to *ermE*) of *eryAI* but not the right end of the locus after analysis of an additional 4 kb.

From the results of scanning gene disruption, the *eryAI* mutation was identified as the insertion farthest to the left in
FIG. 2. Analysis of DNA from gene disruption and gene replacement experiments. (A) Integration of the 8.8-kb recombinant plasmid pMW55-26 into the S. erythraea chromosome. The recombinant plasmid (circle) derived from pMW27 is shown with two BamHI sites (B) and two NcoI sites (N) and is carrying a 0.37-kb Sau3A fragment of S. erythraea chromosomal DNA (©). The homologous sequence on the parental S. erythraea chromosome (©) is also shown, as well as the BamHI and NcoI sites in this region of the chromosome. The postulated eryH transcript is depicted (—). The chromosomal structure predicted for the integratively transformed strain is shown below the bold arrow. This structure shows disruption of the postulated transcription unit for eryH (—©©) and (i) formation of two new 5.3- and 5.5-kb BamHI fragments composed of vector and chromosomal sequences and a third BamHI fragment (0.5 kb) made up entirely of vector sequences, and (ii) formation of two new NcoI fragments (3.0 and 13 kb) composed of vector and chromosomal DNA and a 1.1-kb third fragment from vector DNA alone. This model is consistent with the Southern data shown in panel B. (B) Southern analysis of chromosomal DNA from the parental...
FIG. 3. Erythromycin A and its biosynthetic pathway. (A) The erythromycin biosynthetic pathway showing the different type of ery mutations which block the pathway at the indicated step(s). Step 1, All reactions involved with formation of the first complete erythronolide, DEB; step 2, hydroxylation of DEB to form EB; step 3, synthesis and attachment of mycarose to EB to form MEB; step 4, synthesis and attachment of desosamine to MEB to form erythromycin D (ErD), the first bioactive intermediate in the pathway; step 5, hydroxylation of erythromycin D to form erythromycin C (ErC); step 6, methylation of erythromycin C to form the end product erythromycin A (ErA). (B) Chemical structure of DEB. (C) Structure of erythromycins A, B, C, and D.
The **eryAI** locus (Fig. 4). The *Sau3A* insert AI-1 carried by pMW40-15 extended from a point 0.1 kb to the left of site 27 to approximately 1.75 kb to the right of site 27. This insert overlapped with inserts AI-17 and AI-18; however, it did not overlap with insert CII-36, which was the first insertion mutation to the left of *eryAI* that did not confer the EryA phenotype.

We expected to be able to find the right boundary of the *eryAI* locus by using the same strategy as used to find the left boundary. We therefore removed a 5-kb region from site 27 (Fig. 4), overlapping the right end of the previous *eryAI* boundary (J. Tuan et al., in press) by 1 kb. Forty integrated transformants were obtained and analyzed in this region, and all were *eryA* mutants. Our results indicate that *eryA* genes extend throughout this 9-kb region, and we are in the process of analyzing the adjacent sequences to find the end of the cluster.

**Genes for formation or attachment of the erythromycin sugars: *eryB*, *eryBII*, *eryC*, and *eryCII**. To analyze the function of DNA sequences in the 2.3-kb EcoRI fragment overlapping the left end of *eryAI* (Fig. 4, sites 21 to 27), we made random plasmid insertions into this region of strain UW110. The integrated strains fell into three phenotypic classes: EryP, EryC, and wild-type. The EryP and EryC mutants have been described previously (29); they are postulated to carry defects in the biosynthesis or attachment of the two erythromycin sugar groups, m ycarnose and desosamine, respectively. This is inferred from the results of precursor accumulation and feeding experiments (29). Similar types of experiments were performed on the insertion mutants produced in this study (Fig. 5 and 6). We noticed however, that the EryC phenotype varies slightly from one *eryC* mutant to the next in the proportions of the two precursors, EB and MEB, that are accumulated.

To map the inserts of six representative plasmids in this region, we digested them to release the insert, blotted them by the method of Southern (22), and probed them with the right and left NcoI subfragments of the 2.3-kb EcoRI fragment. The right probe hybridized to the three *eryC* inserts (CII-14, CII-15, and CII-36; all were 60 kb), and the two wild-type inserts (+32, 0.98 kb; +5, 0.81 kb). The left probe hybridized to the *eryB* insert (BII-9, 0.5 kb) and the wild-type inserts. The mutants were further analyzed through Southern analysis of chromosomal DNA to show that they actually carried the integrated plasmid through recombination at the site of homology (data not shown).

These results indicated that an *eryC* locus, herein designated *eryCII*, is completely contained within this 2.3-kb EcoRI fragment; it falls to the left of the Al-1 fragment and to the right of or partially overlapping fragment +32. An *eryB* locus designated *eryBII* is also possibly located here, to the left of the CII-36 fragment but probably overlapping the left end of fragment +32. We infer from the location of fragment +32 that *eryBII* and *eryCII* are likely to be on separate transcripts.

Scanning further to the left of site 17 by gene disruption, we found another *eryB* insertion mutant, *eryBII19*. The 0.55-kb insert from this plasmid mapped to the region overlapping site 20, which is about 0.5 kb to the right of the transcription start point of *eryG* (28). We do not know yet whether the *eryBII* and *eryBII19* mutations are in the same or adjacent transcripts.

A second pair of *eryB* and *eryC* loci were found in the intermediate vicinity of *ermE*. The second *eryC* locus, designated *eryCI*, was identified by linker scanning in the region immediately upstream (to the left) of the *ermE* gene (Fig. 4). This locus had been partially sequenced in an earlier study; however, no function had been determined for the gene (4). In the present study three *eryCII* mutants were generated by gene replacement and characterized. The first mutation, *eryCl8*, was a deletion extending from 50 bp upstream of site 0 to 400 bp farther upstream (Fig. 4). The second and third mutations, *eryCII6* and *eryCII8*, were BglII linker insertions at sites 500 and 850 bp upstream of site 0, respectively. At standard biochemical and morphological analyses, the *eryCII* phenotype was indistinguishable from *eryC1*. Our results are consistent with results of experiments performed independently by (i) Dhillon et al. (7), who identified the *eryC1* gene through gene disruption, and (ii) by Vara et al. (25a), who identified the *eryC1* gene through complementation of the *eryC160* allele.

The second *eryB* locus, *eryBII*, was identified by scanning gene disruption in the region immediately downstream of *ermE*. The mutagenic plasmid pMW55-27 carried an insert (Fig. 4, BI-27) of approximately 0.8 kb that overlapped a pair of adjacent *KpnI* sites (Fig. 4, sites 3 and 4) separated by approximately 0.68 kb. Although this insertion fell very close to the 3' end of *ermE*, it did not affect the resistance of the strain to erythromycin. The results therefore indicate that there are at least two *eryB* and *eryC* loci and that they are significantly separated from one another in the *ery* gene cluster.

**Insertions in the *eryG* region.** We repeated the construction of the insertion mutation which was originally used to describe the *eryG* gene (Fig. 4, *eryGII*) (28) and confirmed that a strain bearing it accumulates erythromycin C (Fig. 5, Er-C). In addition, four new insertions in the *eryG* region were produced and analyzed to further characterize the *eryG* transcript.

One insert, +406 (Fig. 4), was 4.5 kb, covered the entire *eryG* region between sites 12 and 21, and resulted in the wild-type, erythromycin-accumulating phenotype. Another insert, +3, was 0.5 kb and mapped to the 2-kb interval on the left side (Fig. 4, *ermE*-proximal side) between sites 12 and 16; it was also wild-type. The two other sites partially overlapped *eryG*: +/G-449 was between sites 18 and 21, and +/G-44 was between sites 18 and 30. Strains bearing these mutations had a leaky EryG phenotype indicated as +/G; the *eryG*+/G449 strain accumulated slightly more erythromycin C than erythromycin A when grown in TSB; the *eryG*+/G44 strain accumulated mostly erythromycin A but showed higher levels of accumulation of erythromycin C than did the wild-type strain.

Since the +406 insertion was wild type, at least one end of the *eryG* transcript must be contained in the 4.5-kb region between sites 12 and 21 (Fig. 4). The wild-type +3 insertion narrowed the left end down to the 2-kb interval between sites 12 and 16. This was consistent with previously reported RNA studies showing the 3' end of *eryG* to fall in this region (28). The +/G-449 and +/G-44 insertions were more complex. Because these strains did produce significant levels of erythromycin A, the inserts did overlap the 5' end of at least one *eryG* transcript; nevertheless, a significant disruption in the efficiency of conversion of erythromycin C to erythromycin A was created. This was especially true for the *eryG*+/G449 mutation. It is possible that another transcript covering *eryG* begins to the right of site 17. We are in the process of investigating this phenomenon further.

**Mutants bearing defects in a new locus, *eryf*, accumulate DEB.** To test for the presence of *ery* genes between *eryG* and *ermE*, the 5-kb fragment between sites 1 and 11 (Fig. 4) was analyzed by gene disruption. Two Ery− strains were identi-
S. ERYTHRAEA ERYTHROMYCIN GENE CLUSTER
the accumulation of DEB, as determined by thin-layer chromatography, proton magnetic resonance spectroscopy, and mass spectrometry (J. McAlpine, personal communication), and did not accumulate detectable quantities of any of the other erythromycin intermediates (Fig. 5). In precursor-feeding experiments, both mutant strains were unable to convert exogenously supplied DEB, as expected, but they were also unable to convert exogenously supplied EB, indicating that they were also defective in some aspect of the synthesis or attachment of mycarose to DEB (also true for eryB mutants). Cross-feeding experiments did not distinguish between these mutants and eryB mutants; both types of strains cross-secreted to eryA mutants; cross-secreted with eryC mutants, and did not cross-feed with other eryB or eryD mutants (Fig. 6).

The biochemical data indicated that strains bearing mutations in this locus are defective in two consecutive steps of the biosynthetic pathway; first in the C-6 hydroxylation at step 2, and second in the addition or synthesis of mycarose to the macrolide ring at step 3 of the pathway (Fig. 3). This new type of mutant is being given the designation EryH. The two alleles described previously are eryH26 and eryH33, bearing the 0.37- and 0.43-kb inserts, respectively.

In related linker-scanning experiments with the BamHI fragment between sites 6 and 11 (Fig. 4), a third eryH mutation was found, eryH434. Like the previously described eryH mutants, the eryH434 mutant was also blocked in steps 2 and 3 of the pathway (Fig. 3).

For these experiments plasmid pGM434 was constructed, which contained an in vitro-mutagenized derivative of the BamHI fragment (sites 6 to 11) cloned into the BglII site of pGM420, and transformed into strain UW110. Three classes of transformants, I, II, and V, were obtained and analyzed (see convention established in previous example [Fig. 2C]). The results showed class I integrants to be Ery⁺ and class II integrants to be Ery⁻. Recall that in the previous example, both integrated classes I and II were Ery⁺, presumably because one good copy of the eryCl gene was present in the chromosome of each class. The results obtained in this example with the eryH434 mutation, however, indicate that the cloned fragment contained only the left end of the transcript. When this type of plasmid integrates, one complete transcript (to the right of the vector sequences) and one incomplete transcript (to the left of the vector sequences) are generated. Only when the crossover causes the mutation to fall into the right side with the complete gene (class II) will integration be mutagenic. In the case of eryH434, the mutant phenotype occurred only when the linker mutation fell to the right (class II). The complete transcription unit, therefore, is predicted to overlap the right side of the cloned DNA (site 11) and extend to the left into the cloned fragment at least as far as the BglII linker (H-434) but not as far as site 6.

In summary, the combined results from two mutagenic plasmid insertions and a single gene replacement suggest that the eryH locus begins on the left side in the interval between site 6 and the H-434 insert and ends to the right of site 11 before reaching the right end of the +3 insertion (Fig. 4). Further support for this interpretation was obtained through another linker insertion, H-436, which is 1 kb to the left of H-434 and close to the right of site 6. It was found to have no effect on erythromycin biosynthesis. This mutation therefore lies beyond the left end of the eryH locus.

ErmE mutants are highly sensitive to erythromycin A. In previous studies the ermE gene was cloned from S. erythraea and shown to confer resistance to erythromycin in Streptomyces lividans (24). The ermE gene has been tran-
scriptionally characterized as well (4, 25). In this study we conducted experiments to determine whether \textit{ermE} is the only determinant for erythromycin resistance in \textit{S. erythraea} and whether expression of \textit{ermE} is required for expression of the biosynthetic pathway genes. To do this, \textit{ermE} mutants were created in an \textit{eryB25} (EB-accumulating) background (strain UW267; see Materials and Methods) by gene replacement of the wild-type allele with either the \textit{ermE7} mutation, a 0.8-kb deletion from site 1 to the right into \textit{ermE}, or the \textit{ermE14} mutation, a BglII linker insertion at a site approximately 100 bp to the right of site 0. Strains carrying the \textit{eryB25} mutation alone were resistant up to 10 mg of erythromycin A per ml; however, double mutants carrying both \textit{eryB25} and either \textit{ermE7} and \textit{ermE14} were resistant up to only 10 \textmu g of erythromycin A per ml. The \textit{ermE} mutants, therefore, were 1,000-fold more sensitive to erythromycin A than the parent strain was. Although these \textit{ermE} \textit{eryB} mutants were now highly sensitive to erythromycin, they (i) continued to produce and accumulate the precursor EB and (ii) were now highly sensitive to the precursor MEB at levels as low as 10 \textmu g/ml, presumably because they were able to convert this intermediate to erythromycin. On the basis of the results described above, we suggest that \textit{ermE} is the only constitutively functioning erythromycin resistance gene in \textit{S. erythraea} and that it is not essential for expression of most, if not all, of the erythromycin biosynthetic genes (only step 3, EB to MEB, remains to be tested).

Phenotypically silent mutations \textit{5' of \textit{ermE}}. To further investigate the function of sequences \textit{5' of \textit{ermE}} (to the left in Fig. 4), we analyzed the 10-kb BglII fragment between sites \textit{9} and \textit{14} by linker scanning. Although seven linker insertion-deletion mutants were constructed in this interval, none of them were found to be defective in erythromycin biosynthesis. The 3-kb interval between \textit{kb 7.5} and \textit{10.7} was particularly well characterized in that it contained one 2-kb deletion (+-457) and four linker insertions separated from each other by less than 500 bp (+-458, +-459, +-460, and +-461). This interval therefore represents the left end or a significant gap in the \textit{ery} gene cluster.

Vara et al. (25a) suggest from the results of complementation experiments that at least one other \textit{ery} gene defined by the \textit{eryB25} and \textit{eryB26} alleles maps farther to the left of this region. Our results based on two additional deletions to the left (+-454 and +-450) suggest that this gene is not between \textit{kb 12} and \textit{15.2} (+-454) or at \textit{kb 15.2} (+-450). The \textit{eryB} gene of Vara et al. therefore must be in the spaces between these mutations or farther to the left.

No other phenotypes were associated with insertions in this region, indicating that this region does not contain genes involved in primary metabolism, morphological differentiation, or pigment production.

**DISCUSSION**

The principal contribution of this study was to identify and map new genes in the region close to the erythromycin resistance gene (\textit{ermE}) of \textit{S. erythraea}. To do this we used pIJ702-based gene disruption (28) and developed a linker insertion technique, also by using pIJ702-based vectors, to selectively mutagenize an area of approximately 30 kb. We found five new loci, \textit{eryB1}, \textit{eryBII}, \textit{eryCII}, \textit{eryCII}, and \textit{eryH}, all involved in some aspect of deoxysugar biosynthesis or attachment to the macroactone ring. Despite their common functions, they were not clustered, except for \textit{eryBII} and \textit{eryCII}, which were adjacent but appear to be on separate transcription units. This is in contrast to the clustered arrangement for the genes required for macrolactone-ring synthesis (\textit{eryAI} genes), which appear to be in a single transcription unit of more than 9 kb.

We compared our new mutants from gene disruption and gene replacement with the previously described mutants from chemical treatment and/or UV radiation (29; J. M. Weber, Ph.D. dissertation, University of Wisconsin, Madison, 1983) and found them to be phenotypically similar. We have since learned that the \textit{eryCII} and \textit{eryCII} loci described in this report are actually the sites of two of the five previously described \textit{eryC} mutations, \textit{eryC160} (also referred to as \textit{ery-60} or \textit{eryCII-60}) and \textit{eryC210} (also referred to as \textit{ery-210}), respectively (7, 25a, 29; Weber, Ph.D. dissertation). The \textit{eryC160} allele has been recently shown by Vara et al. (25a) to be complemented by DNA from the \textit{eryC} locus; and we have recently observed complementation of the \textit{eryC210} allele by the 5.3-kb \textit{XhoI} fragment (Fig. 4) containing the \textit{eryCII} locus (data not shown). The other reported stable \textit{eryC} allele, \textit{eryC55} (Weber, Ph.D. dissertation), has not yet been tested in complementation experiments.

The mapping of the previously described \textit{eryB} alleles is addressed in the recent work of Vara et al. (25a). Although two \textit{eryB} alleles, \textit{eryB25} and \textit{eryB26}, were mapped to the region within 18 kb upstream of \textit{ermE}, two others, \textit{eryB40} and \textit{eryB46}, were not mapped there, but also did not map to the \textit{eryB} or \textit{eryBII} loci described in this report. Our data show that there are at least two separate \textit{eryB} loci; however, we have yet to determine whether they are the sites of any previously described \textit{eryB} alleles.

In addition to the known \textit{eryB} and \textit{eryC} strains, we found mutants with mutations in a new locus, designated \textit{eryH}. The \textit{eryH} mutants were actually blocked in two apparently unrelated steps of the pathway: (i) in the biosynthesis or attachment of mycarose to the macrolactone ring (like \textit{eryB} mutants), and (ii) in the C-6 hydroxylation of \textit{DEB} to form \textit{EB}. This unusual phenotype could result from our insertion mutations if they disrupted a regulatory gene affecting both steps 2 and 3 (Fig. 3) of the pathway, or if the insertion mutations had a polar effect through disruption of an operon containing genes for both steps 2 and 3. Since the C-6 hydroxylase has been recently purified (19), it should soon be possible from the amino acid sequence to generate synthetic oligonucleotide probes and screen the region directly for the C-6 hydroxylase gene.

Erythromycin resistance is critical to the survival of \textit{S. erythraea}. An important finding of ours was that the \textit{ermE} gene is the only constitutively expressed determinant of this phenotype in \textit{S. erythraea}. We also show that the \textit{ermE} gene product does not play a structural or regulatory role in the production of erythromycin from the intermediate MEB, and it is not involved in the operation of the early pathway either, up to synthesis of the intermediate EB. Apparently not all antibiotic producers are the same in this regard. It is known, for example, that the producer of tylosin, \textit{Streptomyces fradiae}, carries at least two genes which can confer resistance to tylosin (2, 32) and that other antibiotic resistance genes may be involved in regulatory or enzymatic functions required for antibiotic biosynthesis (9).

In summary, we have more clearly defined the organization and function of a 30-kb region of the \textit{S. erythraea} chromosome around the \textit{ermE} gene. We found eight loci contained within a 22-kb portion of this region, extending from 1 kb upstream of \textit{ermE} to 21 kb downstream. Starting to the far right, \textit{eryAI} was the largest transcription unit identified so far, at 9 kb. It is known to be transcribed from right to left, convergent with \textit{ermE} (Fig. 4) (Tuan et al., in...
press). The next three loci, \( ery\text{CII} \), \( ery\text{BIN} \), and \( ery\text{G} \), were smaller and appear to be on separate transcripts; \( ery\text{G} \) transcription is also convergent with \( \text{ermE} \). The 5-kb interval between \( ery\text{G} \) and \( \text{ermE} \), only the \( ery\text{H} \) locus has been identified so far, and it appears to contain a regulatory gene or an operon of at least two structural biosynthetic genes involved in steps 2 and 3 of the pathway (Fig. 3). Approximately 3 kb of adjacent DNA in this region is still mutagenically uncharacterized. Further to the left are the \( ery\text{BI} \), \( \text{ermE} \), and \( ery\text{CII} \) genes. The \( \text{ermE} \) gene and \( ery\text{CII} \) have been shown in previous work to be adjacent to one another and transcribed divergently (4). To the left of \( ery\text{CII} \) is a 6-kb region that has not been mutagenically characterized, although it is said to contain at least the \( ery\text{D} \) allele (25a). This area is followed by a 3- to 8-kb region which does not contain genes involved in erythromycin biosynthesis and represents either a gap in the \( ery \) cluster or the left end of the cluster. We expect that even more \( ery \) genes may be found as the region is more completely mutagenized, for example, the regions between \( \text{ermE} \) and \( ery\text{G} \), and the sequences at both ends of the cluster. We expect that this can be reliably accomplished in future studies by using the gene disruption and linker scanning techniques that we have developed and described in this report.

**ACKNOWLEDGMENTS**

We thank our colleagues at Abbott Laboratories, especially J. Tuan, M. Staver, S. Donadio, D. Brown, and L. Katz, for providing cosmids and plasmid subclones of \( S. \) *erythreus* DNA and for many helpful discussions; J. McAlpine for purification and analysis of erythromycin precursors by proton magnetic resonance spectroscopy and nuclear magnetic resonance spectroscopy; S. Donadio for critical reading of the manuscript; and Al Reetz for artwork. Also, we thank D. Hopwood, M. Bibb, K. Chater, and co-workers at the John Innes Institute, Norwich, United Kingdom, for cloning vectors pJ702, pJ486, pJ487, and pJ450 and for helpful discussions.

**LITERATURE CITED**


