Overproduction and Characterization of the Erythromycin C-12 Hydroxylase, EryK[†,‡]

Ralph H. Lambalot and David E. Cane*

Department of Chemistry, Box H, Brown University, Providence, Rhode Island 02912

Jose J. Aparicio and Leonard Katz

Pharmaceutical Discovery Research, Abbott Laboratories, Abbott Park, Illinois 60064

Received October 14, 1994[⊗]

ABSTRACT: Hydroxylation of C-12 is one of the final steps in the biosynthesis of erythromycin A (ErA). A point of uncertainty in the erythromycin pathway has been whether the C-12 hydroxylase operates on each of two possible substrates, erythromycin B (ErB) and erythromycin D (ErD). Stassi *et al.* have cloned the gene, designated *eryK*, which encodes the P-450 monooxygenase responsible for erythromycin C-12 hydroxylation in *Saccharopolyspora erythraea* [Stassi, D., Donadio, S., Staver, M. J., & Katz, L. (1993) *J. Bacteriol. 175*, 182−189]. We report the overproduction of EryK in *Escherichia coli* as insoluble inclusion bodies; the solubilization, refolding, and reconstitution of active holo-EryK; and kinetic confirmation of a 1200−1900-fold preference of the enzyme for ErD over the alternative C-12 hydroxylase substrate ErB. Our results indicate that ErB is a shunt metabolite in the erythromycin biosynthetic pathway.

Erythromycin A (ErA)[1] is a macrolide antibiotic produced by *Saccharopolyspora erythraea*. Its structural features include a 14-membered macrocyclic lactone ring which bears the sugar residues cladinose (C-3) and desosamine (C-5) (Figure 1). By selectively inhibiting polypeptide synthesis in gram-positive bacteria without altering mammalian ribosomal function (Corcoran, 1984), ErA has been used extensively as an effective clinical antibiotic for more than 30 years. The biosynthesis of ErA requires two hydroxylations (Figure 2). The initial hydroxylation occurs at C-6 of 6-deoxyerythronolide B (6-dEB), the first identifiable intermediate in the ErA pathway (Corcoran & Vygantas, 1982). The P-450 monooxygenase responsible for this modification (EryF) has been cloned and overproduced in *Escherichia coli,* and recognition elements critical to the EryF−6-dEB complex have been defined (Andersen & Hutchinson, 1992; Andersen *et al.*, 1993). While exploring new sequence information from the *S. erythraea* genome, Stassi *et al.* (1993) identified an open reading frame (ORF) designated *eryK*, located in the erythromycin biosynthesis cluster, about 50 kb downstream from the erythromycin resistance gene, *ermE*. On the basis of sequence homologies with members of the P-450 superfamily, Stassi *et al.* were able to identify the *eryK* gene product as a P-450 monooxygenase. Furthermore, the introduction of a mutation in *eryK*

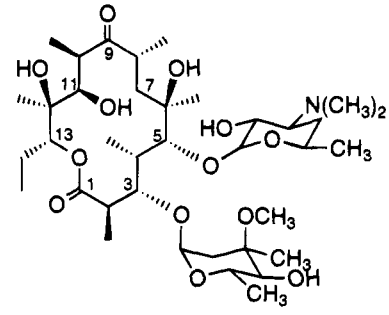

FIGURE 1: Structure of the macrolide antibiotic erythromycin A (ErA).

caused the accumulation of erythromycins B (ErB) and D (ErD), establishing EryK as the C-12 hydroxylase active in the final stages of ErA biosynthesis. Stassi *et al.* were thus able to clone and sequence the second P-450 monooxygenase of the erythromycin pathway.

A point of uncertainty in the biosynthesis of ErA has been whether ErD is converted to ErA through initial hydroxylation at C-12 followed by *O*-methylation of ErC; whether *O*-methylation of ErD first produces ErB, which is then hydroxylated at C-12; or whether both pathways are in operation. In their initial studies of the C-12 hydroxylase, Corcoran and Vygantas (1977) reported that a crude *S. erythraea* lysate was capable of hydroxylating a reduced [9-^{3}H]-(9*S*)-dihydro-ErD derivative but incapable of hydroxylating [9-^{3}H]-(9*R*)-dihydro-ErD or *either* of the analogous ErB derivatives. This bias for the (9*S*)-dihydro-ErD epimer by the C-12 hydroxylase agrees with the modeling studies of Andersen *et al.* (1993) which demonstrated that the C-9 hydroxyl of (9*S*)-dihydro-6-dEB has the same orientation as the C-9 carbonyl in the native macrocyclic lactone, rendering the (9*S*) epimer the preferred substrate. However, the significance of the early C-12 hydroxylase studies is uncertain since the authors concluded on the basis of limited data that the hydroxylase was not a P-450

[†] This work was supported by NIH Grant GM22172 and an ACS Organic Division Fellowship to R.H.L.

[‡] The revised sequence for *eryK* has been deposited at GenBank, Accession Number L05776.

* To whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* January 15, 1995.

[1] Abbreviations: Er, erythromycin; 6-dEB, 6-deoxyerythronolide B; ORF, open reading frame; SAM, *S*-adenosylmethionine; ECPCR, expression cassette polymerase chain reaction; IPTG, isopropyl β-D-thiogalactopyranoside; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; Tris, tris(hydroxymethyl)aminomethane; DTT, dithiothreitol; dNTP, deoxynucleotide triphosphate; LB, Luria−Bertani medium; EDTA, ethylenediaminetetraacetic acid.

0006-2960/95/0434-1858$09.00/0

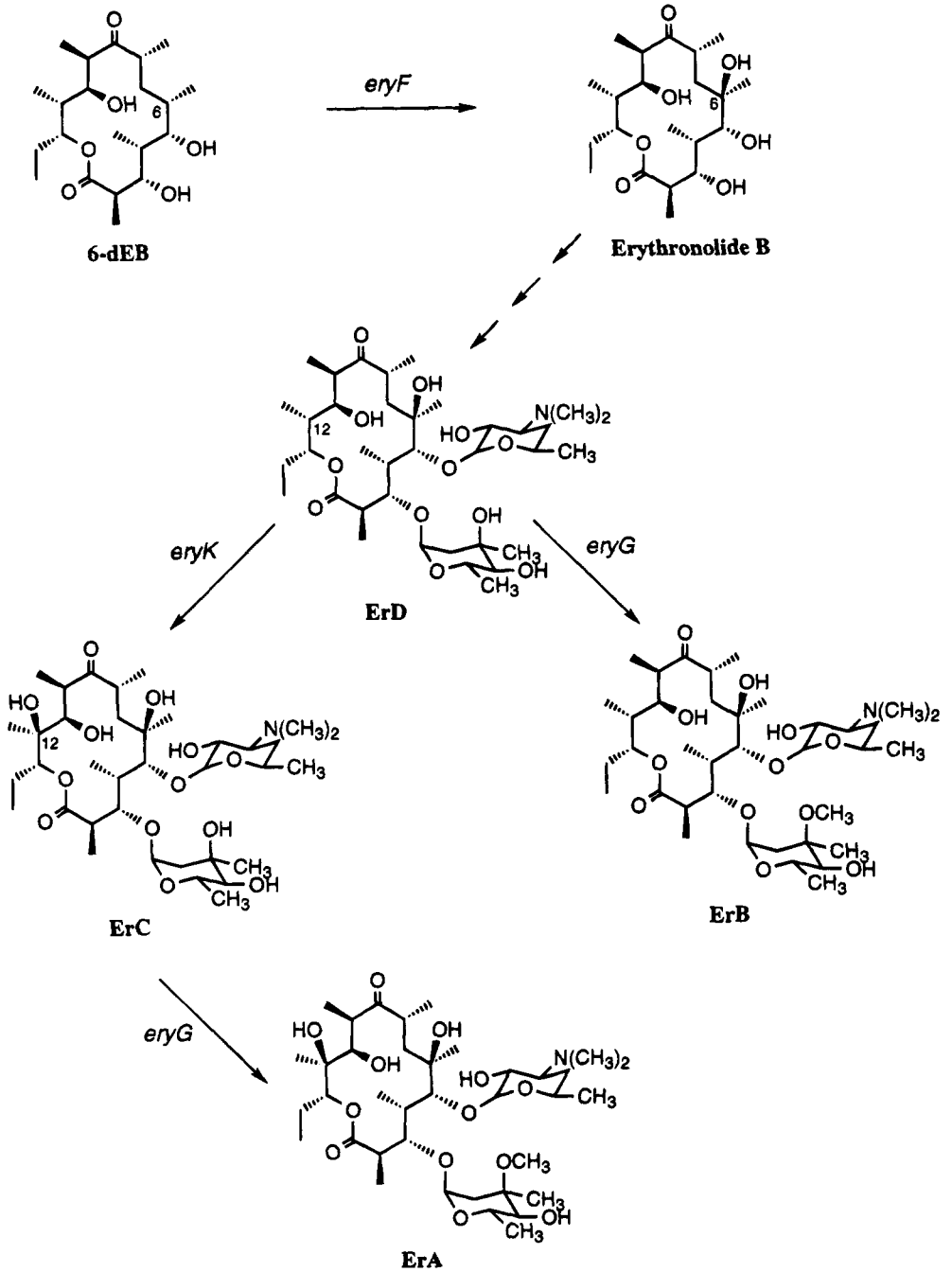

FIGURE 2: Final stages of the biosynthetic pathway leading to the erythromycins. The *eryF* gene product hydroxylates 6-dEB to yield erythonolide B. Following the attachment of mycarose and desosamine, only ErD is hydroxylated at C-12 by the *eryK* gene product to yield ErC. Both ErD and ErC may be methylated on the mycarosyl moiety by the *eryG* gene product to yield ErB and ErA, respectively.

monooxygenase. We report here the heterologous overproduction of EryK in *E. coli* as insoluble inclusion bodies; the solubilization, refolding, and reconstitution of active holo-EryK; and kinetic confirmation of a 1200−1900-fold preference for ErD over the alternative C-12 hydroxylase substrate ErB. Our conclusions support Corcoran's assertion that ErB is a shunt metabolite in the erythromycin biosynthetic pathway and establish definitively that EryK is a P-450 monooxygenase.

EXPERIMENTAL PROCEDURES

Purity of the erythromycins was confirmed by HPLC and 400-MHz ^{1}H NMR. Oligonucleotide primers were prepared on an ABI394 synthesizer (Applied Biosystems, Foster City, CA) or purchased from Integrated DNA Technologies, Inc. (Coralville, IA). The pLM1 expression vector was a gift from Prof. Gregory L. Verdine of Harvard University (Cambridge, MA) (Sodeoka *et al.*, 1993). *Escherichia coli* strain XL1-Blue was obtained from Stratagene (San Diego, CA), JM109(DE3) was from Promega (Madison, WI), DH5α was from Gibco BRL (Gaithersburg, MD), and BL21(DE3) was from Novagen (Madison, WI). Plasmid pUC19 (Norrander *et al.*, 1983) was purchased from Gibco BRL. All growth was perfomed at 37 °C. Nucleoside triphosphates, *Eco*RI, *Hin*dIII, and T4 DNA ligase were purchased from Promega (Madison, WI). λ DNA/*Hin*dIII, ΦX174/*Hae*III, and *Pfu* DNA polymerase were purchased from Stratagene. Sequenase 2 was obtained from U.S. Biochemicals (Cleveland, OH). The 1-kb and 100-bp DNA ladders and isopropyl β-D-thiogalactopyranoside (IPTG) were purchased from

Gibco BRL. Ultra-Free Pro-Bind filters and Ultrafree-MC filter units were purchased from Millipore (Milford, MA). All other materials used for recombinant DNA manipulations, enzymatic assays, and protein purification were of the highest quality available. All buffers and media were prepared with double-deionized nanopure water.

PCR and ligation reactions were run in a Coy Laboratory Products Model 50/60 Tempcycler (Ann Arbor, MI). Preparative volume centrifugations were performed using a DuPont Sorvall RC5 centrifuge at 4 °C (Wilmington, DE). Trace volumes of ethanol were thoroughly evaporated from DNA residues using a Savant Speed-Vac. Isocratic HPLC purification of substrates and assay mixtures was performed as described using a Waters Radial-Pak μBondapak C-18 reversed-phase cartridge loaded into a Z-Module radial compression unit. Discontinuous 10% SDS−PAGE was performed using either 0.75 × 5 × 7 cm or 0.75 mm × 14 cm × 14 cm gels by the method initially described by Laemmli (1970). Protein concentrations were determined either spectrophotometrically or by using the Bradford assay reagent concentrate and bovine serum albumin standard available from Bio-Rad (Melville, NY). UV−vis spectra were obtained using a Hewlett-Packard HP 8452A diode array spectrophotometer interfaced to a Gateway 2000 4DX-33 personal computer. Enzymatic assays were performed at constant temperature (30 °C) using an HP thermal cell holder and a VWR 1130 constant temperature water circulator. Kinetic constants were obtained by fitting data to the appropriate functions using Grafit version 3 by Erithacus Software (London). Mass spectra were obtained using an HP 5890 Series II gas chromatograph interfaced to an HP 5988A quadrupole mass spectrometer.

PCR Amplification of eryK. Plasmid pEVEH8 was constructed by subcloning an 8-kb *Eco*RI−*Hin*dIII fragment containing *ery*K (Stassi *et al.*, 1993) in pUC19. The optimal reaction mixture for PCR amplification of *eryK* from the *Eco*RI-linearized pEVEH8 template was as follows: 200 nM DNA primers, 1 ng of pEVEH8 DNA template, 250 μM dNTPs, 1 mM $MgCl_2$, and a 10× dilution of Stratagene *Pfu* DNA polymerase reaction buffer 1 (20 mM Tris·HCl, 10 mM KCl, 6 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, and 0.1% Triton X-100, pH 8.2) in 3 × 100 μL reaction mixtures. The 52-mer forward primer incorporated an *Eco*RI restriction site (underlined), the T7 gene 10 ribosome binding site and translational leader sequence (italics), and a point mutation which alters the start codon from a leucine (TTG) to a methionine (boldface): 5′-ATG ATT ACG AAT TCA *GGA GAT ATA CAT* **ATG** ACC ACC ATC GAC GAA GTT CCC G-3′. The 37-mer reverse primer incorporated a *Hin*dIII restriction site (underlined) 5′ of the CTA complement of the normal TAG stop codon (boldface): 5′-C AGT GCC AAG CTT **CTA** CGC CGA CTG CCT CGG CGA GGA-3′. Following an initial 5-min incubation at 96 °C, the reaction mixtures were transferred to an ice bath; 2.5 units of *Pfu* polymerase was added to each tube, and the reaction mixtures were overlayed with 20 μL of mineral oil, centrifuged for 15 s at 12000*g*, and returned to the thermal cycler for 3 min at 65 °C. Thirty-five cycles, each consisting of a 3-min extension at 76 °C, a 2-min denaturation at 95 °C, and a 2-min annealing at 65 °C, were performed. The reaction program concluded with a final 7-min extension at 76 °C before holding at 4 °C until the reaction vials were removed from the cycler. Passage through a Pro-Bind filter by centrifugation at 12000*g* for 30 s, ethanol precipitation with 2.5 vol of ice-cold ethanol, and centrifugation at 12000*g* for 30 min yielded a total of 6 μg of a 1200-bp PCR-amplified product as determined by 1.5% agarose gel electrophoresis using 150 ng of ΦX174/*Hae*III and 150 ng of 100-bp ladder as standards.

Preparation of the Overexpression Plasmid pRL05. The PCR product and the pLM01 vector (Sodeoka *et al.*, 1993) were each digested with both *Eco*RI and *Hin*dIII by incubating 3 μg of DNA with 3 μL of endonuclease (30 units of *Eco*RI or 36 units of *Hin*dIII) in 40 μL of Promega endonuclease reaction buffer (final concentrations for *Eco*RI, 90 mM Tris·HCl, 6 mM $MgCl_2$, and 50 mM NaCl, pH 7.5; final concentrations for *Hin*dIII, 10 mM Tris·HCl, 6 mM $MgCl_2$, and 50 mM NaCl, pH 7.5). Reaction mixtures were incubated for 4 h at 37 °C, and complete digestion of the plasmid was confirmed by agarose gel electrophoresis. Following digestion of the PCR product, reactions were diluted to 240 μL with H_2O and passed through a Pro-Bind filter by centrifugation at 12000*g* for 30 s. Since the Pro-Bind filters are better adapted to removal of proteins from low-MW DNA solutions, pLM1 digestions were extracted with phenol/chloroform followed by ethanol precipitation.

Prior to ligation, the DNA fragments were purified by filtration using Ultra-Free MC filter units. The cleanup procedure consisted of diluting the DNA to 400 μL with H_2O in a 1.5-mL Eppendorf tube, heating to 65 °C for 5 min in a heat block, chilling the tube on ice, and filtering by centrifugation at 5000 rpm for 10 min. The filter-collected residues were again taken up in 400 μL of H_2O and transferred to clean 1.5-mL Eppendorf tubes. This cleanup procedure was repeated twice for a total of three washings. After the final filtration, the residues were dissolved in 50 μL of H_2O. Following determination of the DNA concentrations by 0.7% agarose gel electrophoresis, the *eryK* gene insert and the pLM1 expression vector were ligated to each other. The ligation mixture contained 100 ng of *eryK*/*Hin*dIII/*Eco*RI (2 μL of 50 ng/μL stock, 140 fmol), 200 ng of pLM1/*Hin*dIII/*Eco*RI (4 μL of 50 ng/μL stock, 93 fmol), and 2 μL of H_2O. This mixture was heated to 65 °C for 5 min and then chilled on ice before the addition of 1 μL of 10× Promega T4 DNA ligase reaction buffer (300 mM Tris·HCl, 100 mM $MgCl_2$, 100 mM DTT, and 10 mM ATP, pH 7.8) and 1 μL of T4 DNA ligase (1 unit/μL). The reaction mixture was incubated for 16 h at 14 °C. This crude reaction mixture was used directly for the transformation of *E. coli* XL1-Blue.

Transformation of E. coli XL1-Blue and BL21(DE3) Strains. Transformation of competent *E. coli* cells employing selection for ampicillin resistance was performed following published procedures (Sambrook *et al.*, 1989). Plasmids isolated from XL1-Blue/pRL05 and BL21(DE3)/pRL05 cells were confirmed for correct size after miniprep isolation of plasmid DNA followed by restriction analysis on agarose gels. XL1-Blue/pRL05 was first constructed in order to isolate sufficient plasmid DNA for the effective transformation of the BL21(DE3) overexpression host.

Overexpression, Solubilization, and Purification of Apo-EryK. Two 500-mL batches of LBA medium (100 μg/mL ampicillin) were inoculated with 5 mL each of an overnight culture of *E. coli* BL21(DE3)/pRL05. The cultures were grown for 5 h in 2.5-L baffled flasks at 250 rpm until the $OD_{600\ nm}$ reached 2.0. Expression of the T7 RNA polymerase

gene in the lysogenic host strain was derepressed by the addition of IPTG to 1 mM, and growth was allowed to continue for an additional 5 h. Cells were harvested by a 15-min centrifugation at 10000*g* and 4 °C and washed once with 50 mM Tris·HCl, 150 mM NaCl, 10 mM DTT, and 1 mM EDTA, pH 7.5. Centrifugation of the washed cells yielded 5 g of wet cells. Lysis was effected by resuspending the pellet in 50 mM Tris·HCl, 150 mM NaCl, 10 mM DTT, and 1 mM EDTA, pH 7.5 (3 mL/g of wet cell mass); adding lysozyme to 2 mg/mL; and incubating at room temperature for 30 min with occasional stirring. Triton X-100 was then added to 0.1% (250 μL of a 10% stock), and the now viscous mixture was incubated for 10 min at room temperature with stirring. The viscosity was reduced by adding DNase to 50 μg/mL and incubating for 10 min at room temperature. The inclusion-bound material was collected with the cellular debris by centrifugation for 15 min at 15000*g* and 4 °C. The pellet was resuspended in 50 mM Tris·HCl, 10 mM DTT, 1 mM EDTA, 1 mg/mL deoxycholate, and 2 mg/mL lysozyme, which was followed by centrifugation for 15 min at 15000*g*. The insoluble pellet (4 g) was then solubilized in 30 mL of 8 M urea, 50 mM Tris·HCl, and 10 mM DTT, pH 7.5 (solubilization buffer). Centrifugation for 15 min at 15000*g* removed 1 g of urea-insoluble material.

The solubilized material was loaded onto a 30-mL Q-Sepharose column which had been equilibrated with solubilization buffer. The column was washed with 200 mL of solubilization buffer (0.35 mL/min) followed by elution with a 300-mL linear gradient of 0−250 mM NaCl in solubilization buffer. Apo-EryK began to elute at 50 mM NaCl and continued for 10 7-mL fractions. These fractions were pooled and diluted 4-fold with 50 mM Tris·HCl, pH 7.5, to yield a 2.5 mg/mL protein solution (Bradford assay). Refolding of apo-EryK was achieved by dividing this 350-mL solution equally among five 4 × 45 cm dialysis membranes (Spectropor; MW cutoff, 6000−8000; 5 mL/cm capacity) and dialyzing overnight at 4 °C against 3.5 L of 50 mM Tris·HCl, pH 7.5, sparged continuously with argon. The buffer was changed once and dialyzed for another 4 h under the same conditions. The renatured protein was then dialyzed against 3.5 L of 0.1 M histidine and 20% (w/v) glycerol, pH 8.0 (reconstitution buffer), for 6 h at 4 °C with continuous argon sparging. In this manner approximately 200 mg of renatured apo-EryK was obtained.

Reconstitution of Holo-EryK. The dialyzed fractions were pooled in a 250-mL Wheaton bottle to yield 175 mL of a 1.1 mg/mL apo-EryK solution as determined spectrophotometrically, assuming A_{280}(1 mg/mL) = 1. Solid DTT was added to 20 mM along with 3 mg of hemin chloride (4.2 μmol, 1.1 equiv) which had been dissolved in a minimal volume of NH_4OH (50 μL) and diluted to 1 mL with H_2O. The bottle was equipped with a magnetic stir bar, capped with a 24/40 latex septum, wrapped in aluminum foil, and incubated at room temperature with stirring under a positive argon pressure. Reconstitution reached a maximum within 20 h as evidenced by CO binding difference spectroscopy. Addition of more hemin chloride (1.5 mg) and further incubation (16 h) at room temperature did not increase the concentration of the P-450 holo-EryK (P-450$_{EryK}$) species above its maximum of 10.3 μM as determined spectrophotometrically, assuming $\epsilon_{mM}^{452} = 88$ (Yu & Gunsalus, 1974). In a glove bag under an argon flow, the rust-colored holo-EryK solution was dispensed in 17 × 10 mL aliquots into 15-mL screw cap Falcon tubes, which were then sealed with parafilm, wrapped in foil, and quick frozen in an acetone/dry ice bath. Storage at −80 °C provided a source of enzyme which remained stable for months.

CO Binding Difference Spectra. Using a diode-array spectrophotometer in overlay mode, a spectrum of the holo-EryK solution in reconstitution buffer (500 μL in a 1-cm quartz cuvette) was recorded from 190 to 820 nm. The ferric heme was then reduced by adding 5 μL of a freshly prepared saturated solution of dithionite ($Na_2S_2O_4$), and the sample was mixed by inversion. A spectrum of the reduced enzyme was then recorded. The spectrally shifted CO-bound P-450 species was generated by bubbling CO (Aldrich, 10% in helium) through a narrow-bore gel-loading pipet tip into the solution for 30 s (**CAUTION**: Carbon monoxide is a toxic gas and should only be used in a well-ventilated area). Spectra were recorded every 15 s until a maximum absorbance at 450 nm was achieved. The P-450 species eventually decayed to the enzymatically inactive P-420 species due to the strong reducing agent. The CO difference spectrum was obtained by subtracting the CO-bound P-450 spectrum from the dithionite-reduced spectrum. The concentration of active P-450 species was determined assuming $\epsilon_{mM}^{452} = 88$ relative to baseline absorbance at 490 nm (Yu & Gunsalus, 1974). Alternatively, concentrations of both P-450$_{EryK}$ and P-420$_{EryK}$ were determined using the molar extinction coefficients listed in Table 1.

Spectrophotometric Substrate Binding Assay. The spectrum of a 500-μL solution of holo-EryK was recorded from 190 to 820 nm with a diode array spectrophotometer in overlay mode. Spectra were obtained following each successive addition of 10 10-μL aliquots of substrate dissolved in 50 mM Tris·HCl, pH 7.5,, covering a suitable range of concentrations (0.1−15 K_D if solubility permitted). It was imperative that the substrate solutions be free of methanol or ethanol, as these contaminants produced erratic results. A series of blank spectra were generated by repeating the above procedure with buffer solution alone. Difference spectra were then obtained by subtracting the appropriate blank from each of the substrate-bound spectra. This eliminated any effects caused by dilution and the addition of Tris·HCl to the reconstitution buffer. Characteristic type I binding spectra were observed in the 350−500-nm region. Delta values (ΔAU) were determined by subtracting the trough absorbance at 426 nm from the peak absorbance at 392 nm. The data were fit by nonlinear least squares regression to the equation $\Delta AU_{obsd} = \Delta AU_{max}[S]/([S] + K_D)$ to obtain K_D, representing the concentration at which binding is half-maximal.

HPLC Assay for C-12 Hydroxylase Activity. A typical assay consisted of 1 μM holo-EryK, 9 μM spinach ferredoxin, 2.5 μM spinach ferredoxin $NADP^+$ reductase, 50 mM Tris·HCl, pH 7.5, and substrate (0.9−90 μM) in a total volume of 500 μL. The reactions were run in a 30 °C constant temperature bath and were halted by adding 2 drops of 5% NaOH and extracting three times with 1 vol of EtOAc. The organic extracts were pooled and back extracted once with water. Evaporation to dryness under a stream of nitrogen yielded a residue suitable for HPLC analysis using a Waters Radial-Pak μBondapak C-18 reversed-phase cartridge loaded into a Z-Module radial compression unit and

MeOH/H_2O/NH_4OH (80:20:0.1) as an isocratic mobile phase at a flow rate of 2 mL/min (Pellegatta *et al.*, 1983). Components of the assay mixtures were identified by comparison to retention times of authentic standards: ErD, 10.7 min; ErC, 8.1 min; ErB, 13.5 min; and ErA, 9.9 min.

Continuous Spectrophotometric Assay for Initial Rate Kinetics. In a total volume of 500 μL, 1 μM holo-EryK, 9 μM spinach ferredoxin, 2.5 μM spinach ferredoxin $NADP^+$ reductase, 50 mM Tris·HCl, pH 7.5 and substrate (0.9−90 μM) were mixed together in a 1-cm quartz cuvette mounted in a constant temperature cell holder equilibrated to 30 °C. Spectral acquisition at 340 nm using an HP 8452A diode array spectrophotometer in kinetics mode was initiated using a 1-s cycle time, a 0.5-s integration time, and a 200-s run time. After 2 cycles, NADPH was added in 5 μL of 50 mM Tris·HCl, pH 7.5, to 170 μM, and data were acquired for the full run time. Initial rates were determined by applying a quadratic fit to the linear region of the time trace, determined to be the first 20 s.

DNA Sequencing. Plasmid pRL05 was transformed into *E. coli* DH5α competent cells. A single colony was picked from an LB−ampicillin (100 μg/mL) plate to inoculate a 100-mL LB medium and grown overnight. Plasmid DNA was isolated by Qiagen Midi Prep column chromatography (Qiagen Inc., Chatsworth, CA) following the manufacturer's protocol. The yield and final concentration of the plasmid were determined by agarose gel electrophoresis. Sequencing of double-stranded templates (Zhang *et al.*, 1988) with Sequenase 2 employed the dideoxy chain termination method (Sanger *et al.*, 1977) and the terminal deoxynucleotidyl transferase modification protocol (Fawcett & Bartlett, 1990) to fill in prematurely terminated chains and reduce nonspecific stops. Oligonucleotide primers were made to extend the sequence across both strands of the *eryK* gene. Sequence analysis employed the GCG (Genetics Computer Group, Madison, WI) software package (Devereux *et al.*, 1984).

RESULTS

Overproduction of EryK in E. coli. The polymerase chain reaction was used to amplify the *eryK* gene and to provide it with a strong, T7 polymerase-regulated promoter. The amplified gene was digested with *Eco*RI and *Hin*dIII and ligated into the complementary restriction-digested pLM1 expression vector (Figure 3). The ligated product was first cloned into *E. coli* XL1-Blue, and a supercoiled recombinant vector of the correct size and with the expected restriction map (pRL05) was isolated from a single transformed colony. The overexpression host strain, *E. coli* BL21(DE3), was next transformed with pRL05, and a single colony carrying the correct vector was selected. IPTG induction (1 mM) of the BL21(DE3)/pRL05 transformant yielded high-level overexpression as determined by discontinuous 10% SDS−PAGE of the crude whole-cell lysate. Unfortunately, the overproduced protein was determined to reside in the 15000*g* insoluble lysate fraction. Varying the temperature of induction (25, 30, and 37 °C) and the concentration of IPTG used (0, 0.5, and 1.0 mM) as well as introduction of pRL05 into the host strain JM109(DE3) all failed to produce soluble enzyme. Various detergents and chaotropic reagents were explored for their ability to solubilize overproduced EryK. Triton X-100 (≤5.0%), CHAPS (≤50 mM), octyl glucopyranoside (≤50 mM), deoxycholate (≤50 mM), and cholate (≤10 mM) all proved incapable of solubilizing EryK as evidenced by SDS−PAGE analysis. Only 50 mM cholate, ≥3 M guanidinium hydrochloride, and urea could solubilize the inclusion-bound EryK.

Table 1: Extinction Coefficients, ε (cm^{-1} mM^{-1}), for $P\text{-}450_{cam}$ and $P\text{-}420_{cam}$[a]

wavelength (nm)	$P\text{-}450_{cam}$	$P\text{-}420_{cam}$
421	31.2	179.0
446	120.0	21.6

[a] Values apply to the absolute absorbance spectra of the reduced CO bound complexes (Gunsalus & Wagner, 1978).

Purification, Resolubilization, and Reconstitution of EryK. Purification of the recombinant apo-EryK to >90% homogeneity was achieved in a single step. Following lysis, the 15000*g* insoluble pellet was washed with 1 mg/mL deoxycholate and 2 mg/mL lysozyme in 50 mM Tris·HCl, 10 mM DTT, and 1 mM EDTA, pH 7.5. Solubilization of the washed inclusion bodies in 8 M urea, 50 mM Tris·HCl, and 20 mM DTT, pH 7.5, followed by Q-Sepharose purification (0−250 mM NaCl, linear gradient), yielded 200 mg of apo-EryK from a 1-L culture. The denatured EryK was refolded by dialyzing a 2.5 mg/mL solution against 50 mM Tris·HCl, pH 7.5. Reconstitution of the heme-containing holo-EryK was achieved by dialyzing the refolded apo-EryK against 0.1 M histidine and 20% glycerol, pH 8.0, followed by addition of a stoichiometric amount of hemin chloride. Reconstitution was performed under an argon atmosphere for 20−36 h. Aliquots were removed at various time intervals and assayed spectrophotometrically for their ability to produce the reduced CO-bound chromophore at 450 nm (Figure 4) (Omura & Sato, 1964a,b). In accord with Wagner's observations (1981), heme recombination with apo-EryK initially produced an inactive P-420 form of the holoenzyme ($P\text{-}420_{EryK}$) as determined by a ferrous heme−CO complex absorbance maximum at 420 nm. This P-420 species presumably arises from both the improper binding of heme to the active site (180° rotation about the α-γ-*meso*-carbon axis) and oxidation of the axial cysteine ligand. Conversion to the active P-450 form ($P\text{-}450_{EryK}$) was slow and reached a maximum within 20 h. Complete conversion of the P-420 to the P-450 species was never achieved as evidenced by the presence of both 420- and 450-nm maxima in the absolute CO-bound spectrum (Figure 4C).

By applying the extinction coefficients reported for the $P\text{-}450_{cam}$ and $P\text{-}420_{cam}$ reduced CO-bound species (Table 1), it was possible to estimate the concentrations of the three mixture components: $[P\text{-}450_{EryK}]$ = 10.3 μM; $[P\text{-}420_{EryK}]$ = 6.9 μM; and [apo-EryK] = 7.8 μM. Thus, a 40% conversion of apo-EryK to active $P\text{-}450_{EryK}$ had been achieved. This reconstituted preparation was used in characterizing the kinetics of erythromycin C-12 hydroxylation and remained stable for months when stored at −80 °C in histidine/glycerol/DTT. Attempts to separate $P\text{-}450_{EryK}$ from the other two species by Mono-Q and Superose-12 FPLC were unsuccessful. Comparison of the Superose-12 EryK chromatogram to gel filtration MW standards indicated that the reconstituted holoenzyme exists as a monomer with an apparent molecular mass of 54 kDa.

Confirmation of the Recombinant Nucleotide Sequence. The recombinant *eryK* gene was sequenced to determine whether any mutations had been introduced during PCR

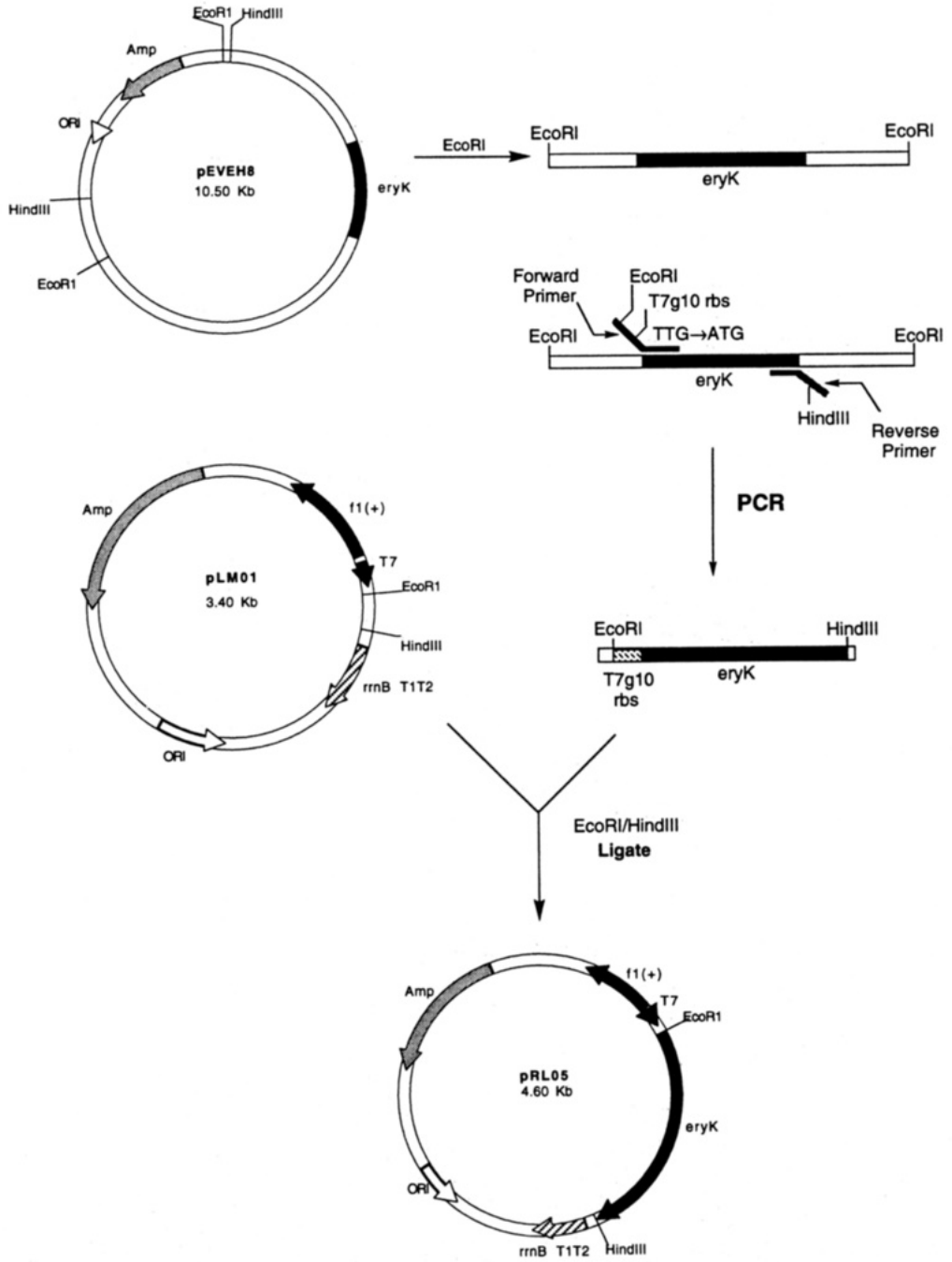

FIGURE 3: Cloning strategy for *eryK*. The *eryK* gene was amplified from the *Eco*RI linearized pEVEH8 template by using ECPCR. The forward primer incorporated an *Eco*RI restriction site, a T7 gene 10 ribosome binding site and translational leader sequence, and a point mutation which altered the start codon from a leucine to a methionine. The reverse primer incorporated a *Hin*dIII restriction site. The PCR product and the pLM01 vector were digested with both *Eco*RI and *Hin*dIII. The doubly digested insert and vector were then ligated using T4 DNA ligase.

amplification. Comparison of the recombinant sequence to the genomic sequence indicated that discrepancies were present. Upon resequencing of the original genomic clone, it was determined that errors were present in the initial report of the *eryK* sequence (Stassi *et al.*, 1993). Comparison of the recombinant sequence to the revised *eryK* sequence confirmed that no mutations had been introduced by PCR apart from the desired start codon mutation. N-Terminal amino acid sequencing confirmed that the first five amino acids of the recombinant protein are Met-Thr-Thr-Ile-Asp as expected. The analysis also indicated that a minor component of sequence Thr-(Thr)-Ile-Asp-(His/Glu/Leu/Ala/Asp/Arg) was also present in which residues in parentheses were assigned with low confidence. This N-terminal processed secondary sequence was present in approximately 15% abundance.

Assay for Erythromycin C-12 Hydroxylase Activity. Activity of the reconstituted P-450$_{EryK}$ was initially confirmed by incubating 70 μg of ErD with 1.6 μg of P-450$_{EryK}$ in the presence of 20 μM spinach ferredoxin, 0.2 units of spinach NADP$^+$ reductase, 1 mM NADP$^+$, 7.5 mM glucose 6-phosphate, and 0.8 unit of glucose-6-phosphate reductase in a final volume of 620 μL of 50 mM Tris·HCl, pH 7.5, for 60 min at 30 °C. C-18 reversed-phase HPLC analysis (Pellegatta *et al.*, 1983) of the assay products indicated that 40% of the ErD (t_R = 11.2 min) had been converted to a product with the same retention time as authentic ErC (t_R = 8.3 min). Assays of P-450$_{EryK}$ solutions prepared under the optimized

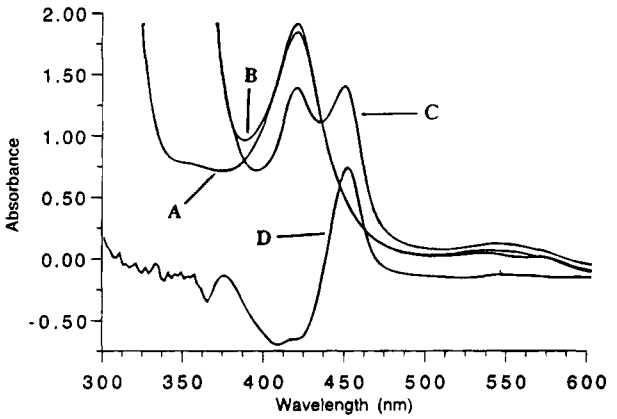

FIGURE 4: CO difference spectra of reconstituted P-450$_{EryK}$. (A) Reconstituted holo-EryK. (B) Sample A reduced with sodium dithionite. (C) Sample B sparged gently with CO. (D) Reduced CO-bound difference spectrum (spectrum B subtracted from spectrum C).

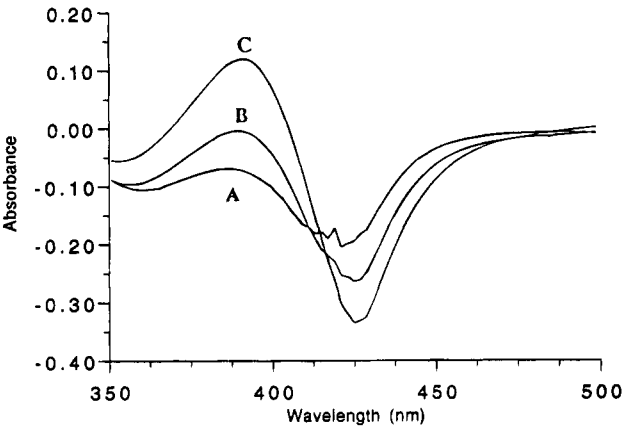

FIGURE 5: Type I binding spectra of ErD bound to P-450$_{EryK}$ at increasing concentrations. (A) 1.7 μM ErD. (B) 10.3 μM ErD. (C) 98.4 μM ErD.

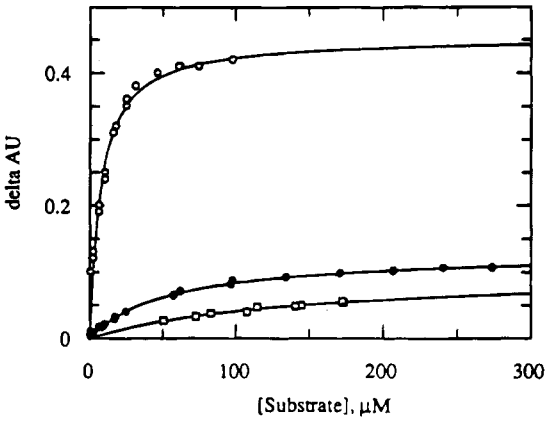

FIGURE 6: Saturation curves showing binding of substrates to P-450$_{EryK}$: ErB (□), ErC (●), and ErD (○). Dissociation constants, K_D, were determined by fitting the ΔAU values derived from the type I binding spectra to the equation $\Delta AU_{obsd} = \Delta AU_{max}[S]/(K_D + [S])$.

reconstitution conditions described in Experimental Procedures yielded 100% conversion of ErD to ErC in less than 10 min. By contrast, incubation of 95 μM ErB with 1 μM optimally reconstituted P-450$_{EryK}$ along with the necessary cofactors and enzymes yielded 40% conversion of ErB to a peak with HPLC retention comparable to authentic ErA only after an extensive incubation time of 11 h. Electron impact mass spectrometric analysis of the product-containing fraction confirmed the presence of an erythromycin derivative. MS analysis of the erythromycins is hampered by the facile dissociation of a desosamine ion yielding a base peak of *m/z* 158 (Jaret *et al.*, 1973). Virtually all of the high-mass/high-information peaks between *m/z* 158 and the molecular ion are below 1% relative abundance. However, it was possible to create a MS library using authentic ErD and ErC and the HP GC/MS software package. A spectral comparison algorithm confirmed the identity of the enzymatic product as ErC.

Measurement of Substrate Dissociation Constants. The dissociation constants for P-450 substrates and inhibitors can be determined by spectrophotometric assay. Binding of an efficient substrate to the P-450 active site induces a spectral shift from a low-spin, ferric heme species (420 nm) to a high-spin species (390 nm). This effect is saturable and can be measured by the UV difference spectrum obtained in the presence and absence of substrate (Figure 5) (Jefcoate, 1978). A plot of the measured delta value, where $\Delta AU = A_{390nm} - A_{420nm}$, vs substrate concentration yields a typical saturation curve (Figure 6). The data can be fit by nonlinear least squares methods to eq 1 to obtain the dissociation constant,

$$\Delta AU_{obsd} = \Delta AU_{max}[S]/(K_D + [S]) \quad (1)$$

K_D, for the substrate tested. Performing this assay with ErB, ErC, and ErD provided measures of their ability to bind to the P-450$_{EryK}$ active site (Table 2).

Steady-State Kinetics of Erythromycin C-12 Hydroxylation. Assuming a 1:1 stoichiometric relationship between NADPH consumption and substrate hydroxylation, initial velocities were measured spectrophotometrically at various substrate concentrations (1−100 μM) by monitoring the rate of absorbance decay at 340 nm over time. In this manner it was possible to accurately determine the steady-state kinetic parameters for ErD (Table 2). Initial rates for ErD hydroxylation were obtained by applying a quadratic fit directly to the initial 20-s linear region of the NADPH consumption curve. Due to the inherent curvature of the rate curve, quadratic fits yield a more accurate extrapolation to the rate at $t = 0$ than a simple linear function. A double-reciprocal plot of the data obtained for ErD indicated an asymptotic rise in $1/V$ at low values of 1/[S] (increasing substrate concentration). This type of behavior is suggestive of substrate inhibition. For this reason we fit the initial velocity data to eq 2, which takes substrate inhibition into account

$$v = V_{max}/[1 + (K_M/[S]) + ([S]/K_i)] \quad (2)$$

(Figure 7). The calculated steady-state parameters were k_{cat} = 375 ± 70 min^{-1}, K_M = 43.9 ± 10.3 μM, and K_i = 16.6 ± 4.9 μM (Table 2). Alternatively, the data at low concentrations, where substrate inhibition was negligible, could also be fit directly to the standard Michaelis−Menten equation to obtain apparent values for k_{cat} and K_M of 108 ± 11 min^{-1} and 8.0 ± 2.3 μM, respectively. In the latter case, the K_M for ErD agreed closely with the K_D of 7.27 ± 0.67 μM obtained from the spectrophotometric binding assay. Also note that the specificity constants obtained from both fits, 8.5 ± 2.5 and 13.5 ± 4.1, are in close agreement with each other.

Incubation of ErB at concentrations of 40−200 μM with EryK did not result in any detectable consumption of

Table 2: Hydroxylation and Binding of Erythromycins by EryK

S	k_{cat}, min^{-1} [a]	K_M, μM	k_{cat}/K_M	K_i, μM	K_D, μM	ΔAU_{max}
ErD	375 ± 70	43.9 ± 10.3	8.5 ± 2.5	16.6 ± 4.9	7.27 ± 0.67	0.45
	108 ± 11[b]	8.0 ± 2.3[b]	13.5 ± 4.1			
ErC					55.3 ± 2.7	0.13
ErB	<1		0.007 ± 0.001[c]		141 ± 25	0.10

[a] k_{cat} is corrected for the calculated concentration of active P-450$_{EryK}$. [b] Apparent values obtained from fitting the Michaelis–Menten equation to the low-concentration data in the linear region of the saturation curve. [c] Assumes $k_{cat}/K_M = k_{cat}/K_D$.

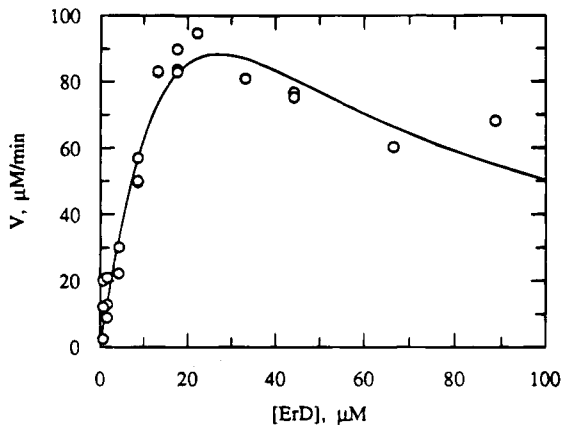

FIGURE 7: Least squares fit of initial rate data obtained with ErD. Initial rates were measured using the continuous spectrophotometric assay described in the text. Initial rates were determined by applying a quadratic fit to the linear region of the time trace, determined to be the first 20 s. The data were then fit to the equation $v = V_{max}/[1 + (K_M/[S]) + ([S]/K_i)]$.

NADPH over the background rate. An upper limit of $k_{cat} = 1$ min^{-1} could be assigned, however, by using larger quantities of enzyme in combination with an NADPH-regenerating system and monitoring ErA production by HPLC. Assuming that the K_D values estimated for ErB by the binding assay are accurate reflections of K_M in the absence of substrate inhibition (Jefcoate, 1978), ErB should have a K_M of 141 ± 25 μM. On the basis of these assumptions, the preference of EryK for ErD over ErB as substrate can be estimated from the ratio of specificity constants, $[k_{cat}/K_M(ErD)]/[k_{cat}/K_M(ErB)] = 1200-1900$.

DISCUSSION

In their report of the *eryK* sequence Stassi *et al.* (1993) had found an open reading frame ending at TAG-1303 but whose precise beginning could not be established by *Streptomyces* codon bias alone. Potential starts were GTG-67, GTG-79, TTG-110, and ATG-137; of these, only the latter two were preceded by potential ribosome binding sites. Inspection of several P-450 sequences indicated that 9 out of the 13 compared began with either Met-Thr or Thr. Both of the other reported *S. erythraea* P-450s, EryF and ORF 405, start with Met-Thr-Thr. These observations led Stassi *et al.* to conclude that TTG-110, which is followed by two Thr codons, is the most likely start for *eryK*. This would render *eryK* 399 codons long, encoding a protein of deduced molecular mass of 43.9 kDa and p*I* of 4.69. Applying these deductions seemed the most reasonable place to begin our subcloning of *eryK*.

We employed the expression cassette polymerase chain reaction (ECPCR) to amplify the *eryK* gene from the cloned genomic fragment (Figure 3) (MacFerrin *et al.*, 1990; Schreiber & Verdine, 1991). A forward primer was designed which altered the *Streptomyces* TTG start codon to an *E. coli* ATG start codon and introduced an *Eco*RI restriction site as well as the T7 gene 10 promoter and translational leader sequence. The reverse primer was designed to introduce the *Hin*dIII restriction site. Following ligation of the *eryK* gene into the pLM01 vector, heterologous overexpression of EryK followed in a straightforward fashion.

Reconstitution of the inclusion-bound apo-EryK to active holo-EryK was greatly facilitated by previous work with the camphor-induced hydroxylase of *Pseudomonas putida*. In the study of P-450$_{cam}$, researchers became interested in elucidating the mechanistic details of camphor hydroxylation by removing the heme prosthetic group from the active holoenzyme and replacing it with novel hemes containing modified substituents or enriched with Fe isotopes. In the course of their studies, Gunsalus and co-workers developed a process for reconstituting apo-P-450$_{cam}$, prepared by an acid/butanone extraction of holo-P-450$_{cam}$ (Wagner *et al.*, 1981). The procedure as applied to EryK involved (1) refolding the urea-solubilized apo-EryK by dialysis against 50 mM Tris·HCl, pH 7.5; (2) dialyzing the refolded apoenzyme against 0.1 M histidine and 20% glycerol, pH 8.0, under anaerobic conditions; and (3) maintaining the integrity of the active site cysteine ligand by addition of solid DTT to a final concentration of 20 mM. This histidine/glycerol/DTT mixture offers an advantage over other reconstitution methods by producing soluble and monomeric forms of both apo-EryK and heme which do not tend to aggregate over the time period of reconstitution.

The spectrophotometric binding assay indicates that ErD is bound 20-fold more tightly than ErB (Table 2). In fact, ErB is bound even more loosely than the enzyme-released product ErC. That the dissociation constant for ErB is 2.5-fold higher than the K_D for ErC is in itself compelling evidence that ErD is the natural substrate of the C-12 hydroxylase. It is evident that methylation of a distal mycarosyl hydroxyl group to form a cladinose sugar (i.e., ErB) has a pronounced effect upon the binding of the macrolide to the EryK active site. This profound effect on the binding of ErB manifests itself as a 1200–1900-fold kinetic preference for the unmethylated ErD.

CONCLUSIONS

In summary, the erythromycin C-12 hydroxylase (EryK) has been overproduced in *E. coli*. The initially isolated protein was obtained as insoluble inclusion bodies, which were solubilized in urea, purified to 90% homogeneity by anion-exchange chromatography, and refolded to give soluble apo-EryK. Reconstitution of holo-EryK with the heme cofactor yielded 100-mg quantities of active P-450 monooxygenase. By use of recombinant overproduced P-450$_{EryK}$, the preferred pathway for ErA biosynthesis has been confirmed to involve oxidation of ErD to ErC, which

undergoes methylation to give ErA. ErB is therefore conclusively shown to be a shunt metabolite of the normal biosynthetic pathway. In addition, the mycarosyl hydroxyl of ErD has been identified as a critical feature of productive substrate binding and perhaps a primary recognition element acting early in the C-12 hydroxylase reaction. Considered together with the C-6 hydroxylase EryF, in which a distal ethyl substituent was shown to be necessary for effective substrate binding (Andersen *et al.*, 1993), both *S. erythraea* monooxygenases involved in the biosynthesis of erythromycin have been shown to have a strong substrate specificity.

REFERENCES

Andersen, J. F., & Hutchinson, C. R. (1992) *J. Bacteriol. 174*, 725−735.

Andersen, J. F., Tatsuta, K., Gunji, H., Ishiyama, T., & Hutchinson, C. R. (1993) *Biochemistry 32*, 1905−1913.

Corcoran, J. W. (1981) in *Antibiotics IV: Biosynthesis* (Corcoran, J. W., Ed.) pp 132−174, Springer-Verlag, New York.

Corcoran, J. W. (1984) in *Macrolide Antibiotics: Chemistry, Biology, and Practice* (Omura, S., Ed.) pp 231−259, Academic Press, Orlando, FL.

Corcoran, J. W., & Vygantas, A. M. (1977) *Fed. Proc. 36*, 663.

Corcoran, J. W., & Vygantas, A. M. (1982) *Biochemistry 21*, 263−269.

Devereux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387−395.

Fawcett, T. W., & Bartlett, S. G. (1990) *BioTechniques 9*, 46−48.

Gunsalus, I. C., & Wagner, G. C. (1978) *Methods Enzymol. 52*, 166−188.

Jaret, R. S., Mallams, A. K., & Vernay, H. F. (1973) *J. Chem. Soc., Perkin Trans. 1*, 1389−1400.

Jefcoate, C. R. (1978) *Methods Enzymol. 52*, 258−279.

Laemmli, U. K. (1970) *Nature 227*, 680−685.

MacFerrin, K. D., Terranova, M. P., Schreiber, S. L., & Verdine, G. L. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 1937−1941.

Norrander, J., Kempe, T., & Messing, J. (1983) *Gene 26*, 101−106.

Omura, T., & Sato, R. (1964a) *J. Biol. Chem. 239*, 2370−2378.

Omura, T., & Sato, R. (1964b) *J. Biol. Chem. 239*, 2379−2385.

Pellegatta, G., Carugati, G. P., & Coppi, G. (1983) *J. Chromatogr. 269*, 33−39.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Nat. Acad. Sci. U.S.A. 74*, 5463−5467.

Schreiber, S. L., & Verdine, G. L. (1991) *Tetrahedron 47*, 2543−2562.

Sodeoka, M., Larson, C. J., Chen, L., LeClair, K. P., & Verdine, G. L. (1993) *Bioorg. Med. Chem. Lett. 3*, 1095−1100.

Stassi, D., Donadio, S., Staver, M. J., & Katz, L. (1993) *J. Bacteriol. 175*, 182−189.

Wagner, G. C., Perez, M., Toscano, W. A., & Gunsalus, I. C. (1981) *J. Biol. Chem. 256*, 6262−6265.

Yu, C. A., & Gunsalus, I. C. (1974) *J. Biol. Chem. 249*, 107−110.

Zhang, H., Scholl, R., Browse, J., & Somerville, C. (1988) *Nucleic Acids Res. 16*, 1220.

BI942394A