JOURNAL OF BACTERIOLOGY, Feb. 1989, p. 754–760
0021-9193/89/020754-07$02.00/0

Vol. 171, No. 2

Cloning, Characterization, and Expression in *Escherichia coli* of the *Streptomyces clavuligerus* Gene Encoding Deacetoxycephalosporin C Synthetase

STEVEN KOVACEVIC, BARBARA J. WEIGEL, MATTHEW B. TOBIN, THOMAS D. INGOLIA, AND JAMES R. MILLER*

Department of Molecular Genetics, Lilly Research Laboratories, Lilly Corporate Center, Indianapolis, Indiana 46285

Received 15 August 1988/Accepted 14 November 1988

Biosynthesis of cephalosporin antibiotics involves an expansion of the five-membered thiazolidine ring of penicillin N to the six-membered dihydrothiazine ring of deacetoxycephalosporin C by a deacetoxycephalosporin C synthetase (DAOCS) enzyme activity. Hydroxylation of deacetoxycephalosporin C to form deacetylcephalosporin C by a deacetylcephalosporin C synthetase (DACS) activity is the next step in biosynthesis of cephalosporins. In *Cephalosporium acremonium*, both of these catalytic activities are exhibited by a bifunctional enzyme, DAOCS-DACS, encoded by a single gene, *cefEF*. In *Streptomyces clavuligerus*, separable enzymes, DAOCS (expandase) and DACS (hydroxylase), catalyze these respective reactions. We have cloned, sequenced, and expressed in *E. coli* an *S. clavuligerus* gene, designated *cefE*, which encodes DAOCS but not DACS. The deduced amino acid sequence of DAOCS from *S. clavuligerus* (calculated M_r of 34,519) shows marked similarity (~57%) to the deduced sequence of DAOCS-DACS from *C. acremonium*; however, the latter sequence is longer by 21 amino acid residues.

Cephalosporin antibiotics are of great importance clinically in the treatment of infections caused by penicillinase-producing bacterial pathogens such as *Staphylococcus aureus*. Cephalosporin C, a cephalosporin produced by *Cephalosporium acremonium*, and cephamycin C, a 7-alpha-methoxy cephalosporin produced by *Streptomyces lactamdurans* and *Streptomyces clavuligerus*, are intermediates in the manufacture of clinically important parenteral cephalosporin antibiotics. Recent recombinant DNA studies have significantly increased our knowledge of two enzyme activities involved in cephalosporin biosynthesis, isopenicillin N synthetase (IPNS) and deacetoxycephalosporin C synthetase (DAOCS).

IPNS is the enzyme responsible for formation of the beta-lactam ring in both penicillin and cephalosporin antibiotics. The gene encoding IPNS, *pcbC*, was cloned first from *C. acremonium* (19) and then from *Penicillium chrysogenum* (3), the producer of penicillins G and V. The deduced amino acid sequences (3, 19) provided important new information because direct biochemical determinations were not available from limited amounts of homogeneous IPNS (10, 16). The high yield of IPNS from expression of these cloned genes in *Escherichia coli* (3, 19) provided ready access to material for physical and chemical characterizations. Availability of the *pcbC* gene for in vitro mutagenesis allowed comparison of natural and altered IPNS molecules and resultant independent assessment of the role in catalysis for each of two cysteine residues in *C. acremonium* IPNS (20). Subsequently, purified IPNS and synthetic oligonucleotides were used to clone the *pcbC* gene from *S. clavuligerus* (13) and *Streptomyces lipmanii* (25).

Cloning, sequencing, and expression in *E. coli* of *pcbC* genes from *Aspergillus nidulans* (17, 25) and *S. lipmanii* (25) have recently been reported. DNA sequence information from the previously cloned fungal genes allowed the isolation of another fungal *pcbC* gene by cross-hybridization. Thus, sequence comparisons can now be made among related beta-lactam antibiotic-producing strains of both fungi and *Streptomyces* spp.

In the biosynthesis of cephalosporin C by *C. acremonium*, two sequential steps in the pathway, ring expansion of penicillin N to deacetoxycephalosporin C and hydroxylation of deacetoxycephalosporin C to form deacetylcephalosporin C, are catalyzed by a bifunctional enzyme, DAOCS-deacetoxycephalosporin C synthetase (DAOCS-DACS). DAOCS-DACS, also referred to as expandase-hydroxylase, was purified to homogeneity (6). By using information from amino acid sequences determined for tryptic peptides of this pure protein, the *cefEF* gene which encodes DAOCS-DACS in *C. acremonium* was cloned from a cosmid library and sequenced (21). An M_r of 36,462 was calculated from the deduced amino acid sequence. Expression of the cloned gene in *E. coli* demonstrated both DAOCS and DACS activities in the bacterium, thereby unequivocally establishing the bifunctional nature of the enzyme. Previously, bifunctionality had been suggested by the inability to separate the expandase and hydroxylase activities during purification of the enzyme from *C. acremonium* extracts (6, 22). DAOCS activity may be the rate-limiting step in the production of cephalosporin C in certain strains of *C. acremonium* (22). Indeed, a *C. acremonium* strain that contains an extra cloned copy of the *cefEF* gene produces less penicillin N and more cephalosporin C than does its nonrecombinant parent; large-scale experiments with the improved strain are pending (8). In *S. clavuligerus*, DAOCS and DACS are separable enzymes (reported M_r values of 29,500 and 26,200, respectively) (12). DAOCS and DACS are also separable enzymes in *S. lactamdurans* (4).

The *S. clavuligerus* gene encoding DAOCS was of interest from several standpoints. As discussed elsewhere (T. D. Ingolia and S. W. Queener, Med. Chem. Rev., in press), a gene encoding DAOCS and not DACS is potentially useful for improving procedures for preparing clinically important oral antibiotics. Because of the functional difference be-

* Corresponding author.

tween DAOCS-DACS in *C. acremonium* and DAOCS in *S. clavuligerus*, comparison between the *C. acremonium* DAOCS-DACS gene and an *S. clavuligerus* DAOCS gene and comparison of the corresponding encoded proteins could help locate the active site(s) responsible for ring expansion of penicillin N and hydroxylation of deacetoxycephalosporin C. Finally, comparison of the DNA sequence of the *S. clavuligerus* DAOCS gene with the *C. acremonium* DAOCS-DACS gene offered a means of further testing a suggested evolutionary relationship of the cephalosporin pathways in fungi and bacteria (3, 17, 25; Ingolia and Queener, in press).

In this paper, we report the cloning, sequencing, and expression in *E. coli* of an *S. clavuligerus* gene, designated *cefE*, that encodes DAOCS, but not DACS, activity when expressed in the bacterium. These results unambiguously demonstrate the separate nature of expandase and hydroxylase in *S. clavuligerus*.

MATERIALS AND METHODS

Bacterial strains and growth conditions. *E. coli* strains were grown in TY broth (19) supplemented with tetracycline (5 μg/ml) or ampicillin (80 μg/ml). DNA techniques followed standard procedures (15). *S. clavuligerus* ATCC 27064 was grown in tryptic soy broth (Difco) for isolation of chromosomal DNA and purification of the DAOCS protein.

Biochemical techniques. Purification of the *S. clavuligerus* DAOCS protein will be reported elsewhere (J. E. Dotzlaf and W.-K. Yeh, submitted for publication). The purified DAOCS protein was sequenced by using a gas-phase sequencer (model 470A; Applied Biosystems, Inc.) by established procedures (9). The 66-base "guessmer" and oligonucleotide for site-directed mutagenesis were synthesized by using an Applied Biosystems DNA synthesizer (model 380A) according to the protocols recommended by the manufacturer.

Cloning of the *S. clavuligerus* DAOCS gene. A 0.5-kilobase pair (kbp) *Pst*I-*Kpn*I fragment from the *S. lipmanii* IPNS gene (25) was nick translated and used to probe an *S. clavuligerus* DNA library constructed in the cosmid pKC462a, which has been described previously (18). The 3 of 1,200 apramycin-resistant transformants that hybridized were further characterized. Concurrently, a 66-base "guessmer" was designed and synthesized on the assumption of typical *Streptomyces* codon bias (7; E. T. Seno and R. H. Baltz, *in* S. Shapiro, ed., *Regulation of Secondary Metabolism in Actinomycetes*, in press) which would encode amino acids 1 to 22 of the experimentally derived amino-terminal amino acid sequence. The guessmer was radioactively labeled with T4 polynucleotide kinase and added to restriction digests of the three cosmids that hybridized to the IPNS gene. Under stringent hybridization conditions, one of the cosmids was observed to hybridize strongly to the guessmer. After further restriction enzyme fragmentation and hybridization with guessmer (Southern analysis), an approximately 3-kbp *Bam*HI fragment was subcloned into pUC8 (pOW380; Fig. 1).

DNA sequencing. The DNA sequence of the *S. clavuligerus* DAOCS gene was determined by using an Applied Biosystems model 370A automatic DNA sequencer. Fragments were subcloned into M13 vectors for sequencing.

Construction and induction of the *E. coli* expression vector pOW382. The potential translation initiation codon was selected on the basis of several criteria (see Results). An *Nde*I site was incorporated into the putative *S. clavuligerus*

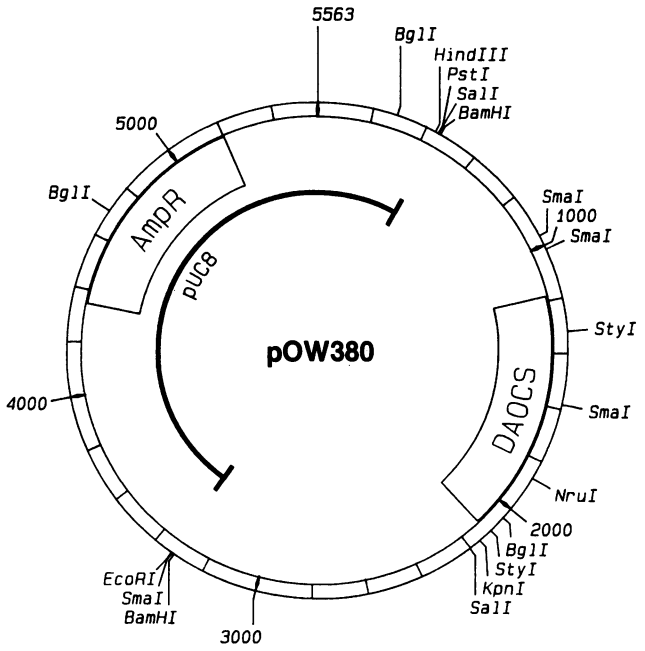

FIG. 1. Restriction map of pOW380. A 3-kbp *Bam*HI fragment was subcloned from a cosmid vector into the unique site in pUC8. Location of the DAOCS gene is as indicated and is inserted in a clockwise direction. Only sites confirmed by sequencing of the insert are shown.

DAOCS gene by site-directed mutagenesis (1) of an M13 vector carrying the 3-kbp *Bam*HI fragment from pOW380. After confirmation of the alteration, the approximately 2.6-kbp *Nde*I-*Bam*HI fragment was cloned into pCZR336 (21), an expression vector containing the λ p_L promoter and temperature-sensitive repressor *c*I857 (pOW382; see Results). *E. coli* JM109(pOW382) cells were grown at 30°C overnight in tetracycline-supplemented TY broth, diluted 100-fold in fresh tetracycline-supplemented TY broth, and induced at 42°C for 6 h for protein production. After 6 h, granules were clearly visible by phase-contrast microscopy. Crude extracts with granules were solubilized as described previously (23) and subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis to determine molecular weight and amount of the plasmid-expressed protein.

DAOCS assays. After induction at 42°C for 6 h, 250 ml of the *E. coli* culture was harvested by centrifugation, and the cell pellet was washed in 30 ml of solution A (15 mM Tris [pH 7.5], 10% ethanol, 10% glycerol, 10 mM dithiothreitol) plus 1 M KCl. The cells were subsequently washed once in 30 ml of solution A without KCl and suspended in 7 ml of solution A plus 10 mM ascorbic acid. Cells were broken by sonication on ice with three 20-s pulses, using a cell disruptor (model 200; Branson Sonic Power Co., Danbury Conn.) with microtip. Phenylmethylsulfonyl fluoride was added to a final concentration of 2.5 mM, and the sonic extract was clarified by centrifugation for 30 min at 10,000 rpm (Sorvall RC-5B, SS34 rotor; Dupont Instruments). Both cell debris and supernatant were stored at −70°C. These fractions were assayed for DAOCS and DACS activity as previously described (21).

RESULTS

Cloning of the *S. clavuligerus* DAOCS gene. Genes for the biosynthesis of an antibiotic by a procaryote may be clus-

tered (physically linked) (7, 14, 24; Seno and Baltz, in press). The isolation of one gene of the pathway may lead directly to the isolation of others. Weigel et al. (25) have recently reported the cloning and expression of IPNS from *S. lipmanii*. Moreover, comparison of the IPNS genes and proteins from *Streptomyces* spp., *C. acremonium*, *P. chrysogenum*, and *A. nidulans* showed that the IPNS genes are closely related (3, 13, 17, 19, 25). Therefore, the IPNS gene of one species might hybridize to the IPNS gene of another. On the assumption that beta-lactam biosynthetic genes would be linked, a 0.5-kbp *Pst*I-*Kpn*I fragment from the coding region of the *S. lipmanii* IPNS gene was nick translated and used as an initial screen of the *S. clavuligerus* cosmid library. Three cosmids were obtained and used for further characterization.

A 66-base guessmer that would encode the first 22 amino acids of the expandase gene was designed and synthesized. The DNA probe was designed to use preferred codons from *Streptomyces* genes which have a high G+C content (Seno and Baltz, in press). One of the three clones that hybridized to the *S. lipmanii* IPNS gene probe was found to hybridize strongly to the guessmer. Further Southern analysis indicated that the expandase gene was encoded within an approximately 3-kbp *Bam*HI fragment. After being subcloned into pUC8, the insert DNA was sequenced. The DNA sequence of the expandase gene matched the guessmer in 55 of 66 bases (83.3%).

Sequence analysis of the *S. clavuligerus* DAOCS gene is shown in Fig. 2. Comparison of the *S. clavuligerus* DAOCS gene and the *C. acremonium* DAOCS-DACS (21) gene showed 67% identity (data not shown). Alignment of the amino acids from the NH_2 terminus resulted in 56.7% identity between the proteins (Fig. 3). Interestingly, the amino and carboxyl ends of the proteins diverged the most between the two. The bacterial protein was about 20 amino acids shorter than the fungal protein (determination of exact differences must await protein sequencing). Several regions, especially those from amino acids 50 to 130, 150 to 210, and 225 to 270 (67 to 73% identity), were highly similar. Reducing the amino acids to functionally similar families increased the overall identity to 75% (data not shown).

The G+C content of the DAOCS-coding sequence was found to be 67.7%. The DNA sequence was examined by using an algorithm described by Bibb et al. (2) that predicts the reading frame of the protein by the G+C content at each position within a codon. The plot of G+C content versus base position (Fig. 4) indicated that a typical *Streptomyces* coding sequence was present. Review of the DNA sequence upstream of the putative initiation codon revealed the presence of a potential ribosome-binding site (G_AGAGG) located 17 bases from the translation start (the spacing is defined according to Seno and Baltz [in press]). Because the expandase gene is expressed during secondary metabolism, typical primary metabolic promoter regions (−35 and −10) were understandably not evident, since promoters used during secondary metabolism are distinct from these promoters (Seno and Baltz, in press).

Expression of the DAOCS gene in *E. coli*. To prove unambiguously that the cloned open reading frame coded for DAOCS activity, the sequence was cloned into an expression vector and placed into an *E. coli* host. An *Nde*I site was placed at the putative initiation codon by site-directed mutagenesis. The entire open reading frame was introduced into the expression vector pCZR336 on an *Nde*I-*Bam*HI fragment (pOW382; Fig. 5).

After induction of the λ p_L promoter, cell extracts were analyzed for protein content, apparent molecular weight, and expandase activity. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Fig. 6) revealed that more than 10% of the total protein was present as a band with an apparent molecular weight of 36,000. This protein band comigrated with expandase purified from *S. clavuligerus* (data not shown). The calculated molecular weight for the expandase encoded by the open reading frame was 34,519, in close agreement with the size judged by gel electrophoresis.

Crude *E. coli* extracts were assayed for expandase as well as hydroxylase activity. These extracts readily converted penicillin N to deacetoxycephalosporin C but did not convert deacetoxycephalosporin C to deacetylcephalosporin C (Table 1). Because the DAOCS-DACS protein from *C. acremonium* has been previously shown to carry out both conversions in *E. coli* under identical assay conditions (6), we conclude that DAOCS from *S. clavuligerus* does not contain DACS.

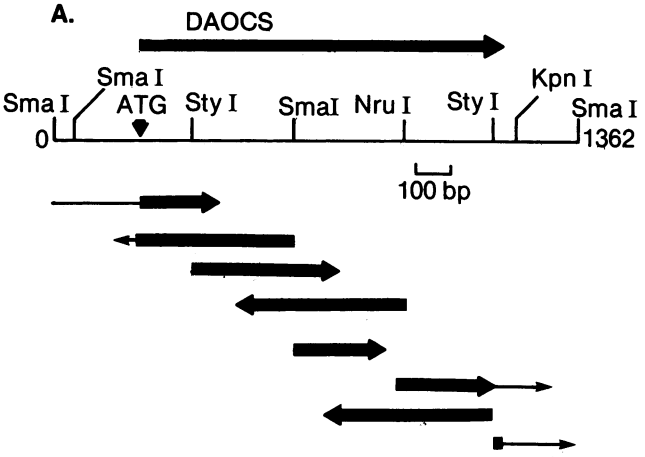

FIG. 2. (A) Sequencing strategy of the DAOCS gene. Southern hybridization with the guessmer and restriction site mapping located a *Sty*I fragment and a *Sma*I site within the putative coding region. Fragments were isolated and in some instances blunt ended with Klenow enzyme for cloning into M13mp18 or M13mp19 and subsequent sequence analysis. The large arrows indicate sequence within the reading frame; nearly all of the coding region has been sequenced on both strands. The DAOCS gene and the direction of transcription are also shown. (B) Complete nucleotide and predicted amino acid sequence of the DAOCS gene in pOW380. The translation start site is at position +1; asterisks above the sequence show the bases that were also present in the guessmer. Cycles 2 and 18 of the sequenator run were ambiguous but suggested the possibility of histidine residues at each position.

DISCUSSION

We have reported the isolation, sequencing, and characterization of the DAOCS gene that encodes the protein responsible for the ring expansion of penicillins to cephalosporins in *S. clavuligerus*. In the preliminary screen of the cosmid library, an internal restriction fragment of the IPNS gene from *S. lipmanii* (25) as well as a guessmer to the DAOCS gene were used. One of the three cosmid clones selected with the IPNS gene fragment contained the DAOCS gene and was subcloned.

Identification of the DAOCS gene was judged by several lines of evidence. Analysis of the derived DNA sequence distinguished the NH_2-terminal region. Although this region had some similarity to that of the DAOCS-DACS protein

B.

```
        -30                         -10                     10                          30
                                                    *** **  **  **  **  **  *** ***  *  **  **
CAC CCG GTT CAC GGA TTA CGA GAG GAT CAG TGA GAG TTG ATG GAC ACG ACG GTG CCC ACC TTC AGC CTG GCC
                                                    Met Asp Thr Thr Val Pro Thr Phe Ser Leu Ala
                                                                                        10

                        50                          70                      90
**  *** *** *** *** **  *** *** *** *** ***
GAA CTC CAG CAG GGC CTG CAC CAG GAC GAG TTC CGC AGG TGT CTG AGG GAC AAG GGC CTC TTC TAT CTG ACG
Glu Leu Gln Gln Gly Leu His Gln Asp Glu Phe Arg Arg Cys Leu Arg Asp Lys Gly Leu Phe Tyr Leu Thr
                                                                        30

  110                         130                         150                         170
GAC TGC GGT CTG ACC GAC ACC GAG CTG AAG TCG GCC AAG GAC ATC GTC ATC GAC TTC TTC GAG CAC GGC AGC
Asp Cys Gly Leu Thr Asp Thr Glu Leu Lys Ser Ala Lys Asp Ile Val Ile Asp Phe Phe Glu His Gly Ser
                                                        50

        190                     210                         230
GAG GCG GAG AAG CGC GCC GTC ACC TCG CCC GTC CCC ACC ATG CGC CGC GGC TTC ACC GGG CTG GAG TCG GAG
Glu Ala Glu Lys Arg Ala Val Thr Ser Pro Val Pro Thr Met Arg Arg Gly Phe Thr Gly Leu Glu Ser Glu
                                        70

250                     270                         290                         310
AGC ACC GCC CAG ATC ACC AAT ACC GGC AGC TAC TCC GAC TAC TCG ATG TGC TAC TCG ATG GGC ACC GCG GAC
Ser Thr Ala Gln Ile Thr Asn Thr Gly Ser Tyr Ser Asp Tyr Ser Met Cys Tyr Ser Met Gly Thr Ala Asp
                        90

    330                         350                         370                         390
AAC CTC TTC CCG TCC GGT GAC TTC GAG CGG ATC TGG ACC CAG TAC TTC GAC CGC CAG TAC ACC GCC TCC CGC
Asn Leu Phe Pro Ser Gly Asp Phe Glu Arg Ile Trp Thr Gln Tyr Phe Asp Arg Gln Tyr Thr Ala Ser Arg
                                                                                        130

                    410                         430                     450
GCG GTC GCC CGG GAG GTC CTG CGG GCG ACC GGG ACC GAG CCC GAC GGC GGG GTC GAG GCC TTC CTC GAC TGC
Ala Val Ala Arg Glu Val Leu Arg Ala Thr Gly Thr Glu Pro Asp Gly Gly Val Glu Ala Phe Leu Asp Cys
                                                                        150

  470                         490                         510                         530
GAG CCG CTG CTG CGG TTC CGC TAC TTC CCG CAG GTC CCC GAG CAC CGC AGC GCC GAG GAG CAG CCC CTG CGG
Glu Pro Leu Leu Arg Phe Arg Tyr Phe Pro Gln Val Pro Glu His Arg Ser Ala Glu Glu Gln Pro Leu Arg
                                                        170

              550                     570                         590
ATG GCG CCG CAC TAC GAC CTG TCG ATG GTC ACC CTC ATC CAG CAG ACA CCC TGC GCC AAC GGC TTC GTC AGC
Met Ala Pro His Tyr Asp Leu Ser Met Val Thr Leu Ile Gln Gln Thr Pro Cys Ala Asn Gly Phe Val Ser
                                        190

600                     630                         650                         670
CTC CAG GCC GAG GTC GGC GGC GCG TTC ACG GAC CTG CCC TAC CGT CCG GAC GCC GTC CTC GTC TTC TGC GGC
Leu Gln Ala Glu Val Gly Gly Ala Phe Thr Asp Leu Pro Tyr Arg Pro Asp Ala Val Leu Val Phe Cys Gly
                        210

    690                         710                         730                         750
GCC ATC GCG ACC CTG GTG ACC GGC GGC CAG GTC AAG GCC CCC CGG CAC CAT GTC GCG GCC CCC CGC AGG GAC
Ala Ile Ala Thr Leu Val Thr Gly Gly Gln Val Lys Ala Pro Arg His His Val Ala Ala Pro Arg Arg Asp
    230                                                                                 250

                    770                         790                     810
CAG ATA GCG GGC AGC AGC CGC ACC TCC AGT GTG TTC TTC CTC CGT CCC AAC GCG GAC TTC ACC TTC TCC GTC
Gln Ile Ala Gly Ser Ser Arg Thr Ser Ser Val Phe Phe Leu Arg Pro Asn Ala Asp Phe Thr Phe Ser Val
                                                                        270

  830                         850                         870                         890
CCG CTG GCG CGC GAG TGC GGC TTC GAT GTC AGC CTG GAC GGC GAG ACC GCC ACG TTC CAG GAT TGG ATC GGG
Pro Leu Ala Arg Glu Cys Gly Phe Asp Val Ser Leu Asp Gly Glu Thr Ala Thr Phe Gln Asp Trp Ile Gly
                                                        290

        910                     930
GGC AAC TAC GTG AAC ATC CGC CGC ACA TCC AAG GCA TAG
Gly Asn Tyr Val Asn Ile Arg Arg Thr Ser Lys Ala End
                                        310
```

```
S. clavuligerus   1  MDTTVPTFSLAELQQGLHQDEFRRCLRDKGLFYLTDCGLTDTELKSAKDI  50
C. acremonium     1  MTSKVPVFRLDDLKSGKVLTELAEAVTTKGIFYLTESGLVDDDHTSARET  50

                 51  VIDFFEHGSEAEKRAVTSPVPTMRRGFTGLESESTAQITNTGSYSDYSMC  100
                 51  CVDFFKNGSEEEKRAVTLADRNARRGFSALEWESTAVVTETGKYSDYSTC  100

                101  YSMGTADNLFPSGDFERIWTQYFDRQYTASRAVAREVLRATGTE.PDGGV  149
                101  YSMGIGGNLFPNRGFEDVWQDYFDRMYGAAKDVARAVLNSVGAPLAGEDI  150

                150  EAFLDCEPLLRFRYFPQVPEHRSAEEQPLRMAPHYDLSMVTLIQQTPCAN  199
                151  DDFVECDPLLRLRYFPEVPEDRVAEEEPLRMGPHYDLSTITLVHQTACAN  200

                200  GFVSLQAEVGGAFTDLPYRPDAVLVFCGAIATLVTGGQVKAPRHHVAAPR  249
                201  GFVSLQCEVDGEFVDLPTLPGAMVVFCGAVGTLATGGKVKAPKHRVKSPG  250

                250  RDQIAGSSRTSSVFFLRPNADFTFSVPLARECGFDVSLDGETATFQDWIG  299
                251  RDQRVGSSRTSSVFFLRPKPDFSFNVQQSREWGFNVRIPSERTTFREWLG  300

                300  GNYVNIRRTSKA                                        311
                301  GNYVNMRRDKPAAAEAAVPAAAPVSTAAPIAT                    332
```

FIG. 3. Similarities between the *S. clavuligerus* DAOCS and *C. acremonium* DAOCS-DACS proteins. Both proteins are aligned starting with the initial methionine residue. Identities are boxed. The top sequence (*S. clavuligerus*) has a gap of 1 amino acid at position 145 and is shorter than the *C. acremonium* DAOCS-DACS protein by 21 amino acids (20 residues occurring at the C terminus).

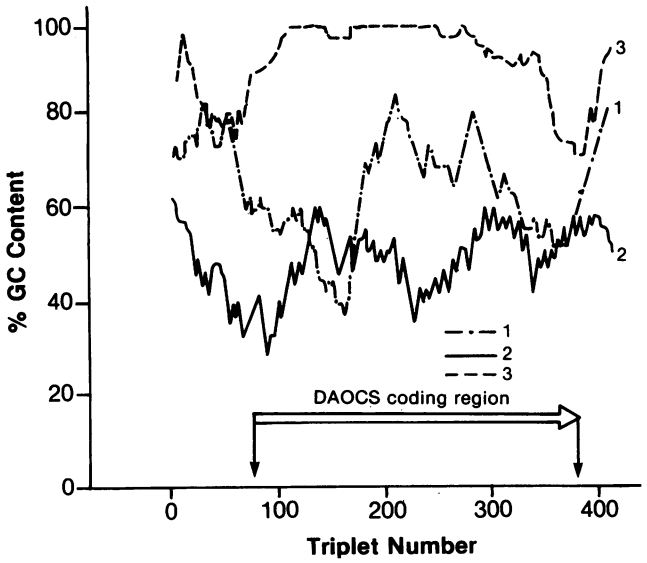

FIG. 4. Analysis of G+C content as a function of position in the codon. Each curve represents the G+C content of one of the three codon positions. The percent G+C at each base is calculated as a forward average of 45 codons according to the algorithm of Bibb et al. (2). The arrow along the axis indicates the limits of the DAOCS protein-coding region.

from *C. acremonium*, regions distal to the amino terminus showed a higher similarity (~70%, amino acids 61 to 105; 80%, amino acids 161 to 190). These similarities indicated the proper reading frame of the gene until the sequence was confirmed.

The proper reading frame was also indicated by the preferential base frequency normally found in *Streptomyces* structural genes. Applying the algorithm reported by Bibb et al. (2) led to the assignment of the initiation codon and the potential stop codon. The calculated molecular weight is 34,519, in close agreement with the M_r of 36,000 for the expandase purified from *S. clavuligerus* (W.-K. Yeh, personal communication). A previous value for purified expandase of 29,500 was reported by Jensen et al. (12); the same-sized protein was reportedly made by *S. lividans* when this naive host was transformed with a *Streptomyces* vector containing fragmented *S. clavuligerus* DNA (M. K. R. Burnham, J. E. Hodgson, and I. D. Normansell, European Patent Publication 233715, August 1987). Whether the difference in size is trivial or represents an altered or different expandase species is not known. Nevertheless, unambiguous proof of expandase cloning was demonstrated by the expression of expandase in *E. coli*. This expandase protein from *E. coli* has the same mobility as that of expandase purified from *S. clavuligerus* and, unlike DAOCS-DACS from *C. acremonium*, has only DAOCS activity, which confirms the data of Jensen et al. (12).

The similarity of the *S. clavuligerus* DAOCS and *C. acremonium* DAOCS-DACS genes suggests a close evolu-

TABLE 1. Expandase and hydroxylase activities from various sources[a]

Enzyme source	Sp act (U/mg of protein)	
	Expandase	Hydroxylase
Cephalosporium expandase-hydroxylase		
C. acremonium crude extract	0.016	0.002
Recombinant *E. coli*		
Crude extract	0.009	0.0009
Granular preparation	0.29	+[b]
Streptomyces expandase		
S. clavuligerus crude extract	0.004	0.00023
Recombinant *E. coli*		
Crude extract	0.009	ND[c]
Granular preparation	0.052[d]	ND

[a] Contributed by W.-K. Yeh.
[b] Activity was present but not quantitated.
[c] ND, None detected.
[d] Recovery of the expandase activity was not optimized.

tionary relationship. On the basis of the similarities among IPNS genes (17, 21, 25; Ingolia and Queener, in press), we and others previously proposed a horizontal transfer of beta-lactam biosynthetic genes, probably from a procaryote to a eucaryote, about 370 million years ago. This hypothesis predicts that comparison of other fungal and bacterial beta-lactam biosynthetic genes will yield about the same degree of similarity if all of the genes were transferred concomitantly. The data presented here support this notion. The *C. acremonium* IPNS DNA sequence is 64.9% identical to the *S. lipmanii* IPNS DNA sequence, and the amino acid sequences are 56.9% identical (25). We have shown that the *C.*

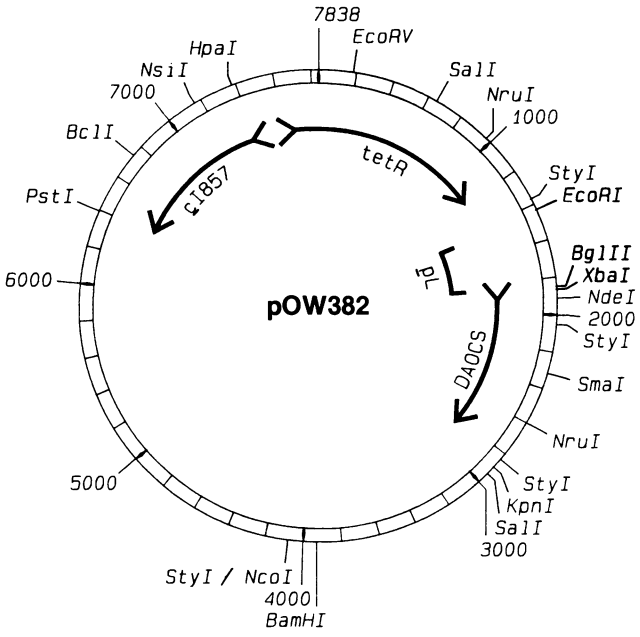

FIG. 5. Restriction map of pOW382. After site-directed mutagenesis of the 3-kbp *Bam*HI fragment of pOW380 (Fig. 1) to place an *Nde*I site at the translation initiation codon, the *Nde*I-to-*Bam*HI fragment was ligated to the *Nde*I-*Bam*HI fragment of pCZR336 (21) for expression in *E. coli* (see Materials and Methods). *p*L, λ p_L promoter containing transcription and translation initiation signals; tetR, tetracycline resistance coding region; *c*I857, λ repressor protein-coding region. Only selected unique sites of pCZR336 are shown.

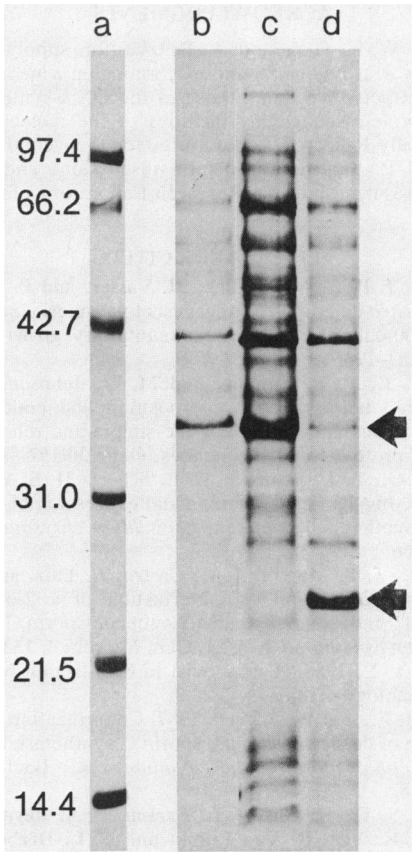

FIG. 6. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of *E. coli* JM109(pOW382) induced for expandase production. Growth and induction conditions are described in Materials and Methods. Lanes: a, molecular weight standards (β-galactosidase, 97,400; bovine serum albumin, 66,200; ovalbumin, 42,700; carbonic anhydrase, 31,000; soybean trypsin inhibitor, 21,500; lysozyme, 14,400); b and c, 10 and 20 μl, respectively, of JM109(pOW382) extracts; d, crude extract of induced JM109(pCZR336). pCZR336 and pOW382 differ only by the *Nde*I-*Bam*HI insert. pCZR336 contains the human growth hormone gene; pOW382 contains the DAOCS gene. Their gene products are shown by the lower and upper arrows, respectively.

acremonium DAOCS-DACS DNA sequence is 67% identical to the *S. clavuligerus* DAOCS DNA sequence and that the amino acid sequences are 56.7% identical. The remarkable closeness of the percent identities for these pairs of biosynthetic genes suggests that they diverged at the same time, probably via a horizontal transfer of the entire pathway. Since the fungal beta-lactam biosynthetic genes are believed to be genetically dispersed (it is known that the IPNS and DAOCS-DACS genes are on different chromosomes in *C. acremonium*; P. L. Skatrud and S. W. Queener, personal communication) and *Streptomyces* biosynthetic pathways are typically genetically linked, it is most likely that the genes were transferred from *Streptomyces* spp. to the fungi.

Finally, the availability of the *S. clavuligerus cefE* gene for site-directed mutagenesis has practical significance. Genetic modification of the *cefE* gene to encode an expandase which would ring expand penicillin G or V, particularly if expressed in *P. chrysogenum*, could provide the basis for improved manufacturing processes for clinical cephalosporin antibiotics (5; Ingolia and Queener, in press).

ACKNOWLEDGMENTS

We thank W.-K. Yeh and J. Dotzlaf for supplying purified DAOCS from *S. clavuligerus* and performing enzyme assays on *E. coli* crude extracts. We thank those of the DNA-synthesizing and DNA- and protein-sequencing facilities of the molecular biology division of Lilly Research Laboratories for their contributions, S. Queener and P. Skatrud for helpful discussions and sharing of unpublished results, and Lilly Research Laboratories for continued support.

LITERATURE CITED

1. **Adelman, J. P., J. S. Hayflick, M. Vasser, and P. H. Seeburg.** 1983. *In vitro* deletional mutagenesis for bacterial production of the 20,000-dalton form of human pituitary growth hormone. DNA **2**:183–193.
2. **Bibb, M. J., P. R. Findlay, and M. W. Johnson.** 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene **30**:157–166.
3. **Carr, L. G., P. L. Skatrud, M. E. Scheetz II, S. W. Queener, and T. D. Ingolia.** 1986. Cloning and expression of the isopenicillin N synthetase gene from *Penicillium chrysogenum*. Gene **48**:257–266.
4. **Cortés, J., J. F. Martín, J. M. Castro, L. Laiz, and P. Liras.** 1987. Purification and characterization of a 2-oxoglutarate-linked ATP-independent deacetoxycephalosporin C synthase of *Streptomyces lactamdurans*. J. Gen. Microbiol. **133**:3165–3174.
5. **Demain, A. L.** 1986. A new twist in beta-lactam biosynthesis. Bio/Technology **4**:18.
6. **Dotzlaf, J. E., and W.-K. Yeh.** 1987. Copurification and characterization of deacetoxycephalosporin C synthetase/hydroxylase from *Cephalosporium acremonium*. J. Bacteriol. **169**: 1611–1618.
7. **Fishman, S. E., K. Cox, J. L. Larson, P. A. Reynolds, W. T. Seno, W.-K. Yeh, R. Van Frank, and C. L. Herberger.** 1987. Cloning genes for the biosynthesis of a macrolide antibiotic. Proc. Natl. Acad. Sci. USA **84**:8248–8252.
8. **Fogelsong, M.** 1987. Large-scale production involving *Cephalosporium acremonium* strain LU4-79-6. Fed. Regist. **52**:48660.
9. **Hewick, R. M., M. W. Hunkapiller, L. E. Hood, and W. J. Dreyer.** 1981. A gas-liquid solid phase peptide and protein sequenator. J. Biol. Chem. **256**:7990–7997.
10. **Hollander, I. J., Y.-Q. Shen, J. Heim, and A. L. Demain.** 1984. A pure enzyme catalyzing penicillin biosynthesis. Science **224**: 610–612.
11. **Jensen, S. E.** 1985. Biosynthesis of cephalosporins. CRC Crit. Rev. Biotechnol. **3**:277–301.
12. **Jensen, S. E., D. W. S. Westlake, and S. Wolfe.** 1985. Deacetoxycephalosporin C synthetase and deacetoxycephalosporin C hydroxylase are two separate enzymes in *Streptomyces clavuligerus*. J. Antibiot. **38**:263–265.
13. **Leskiw, B. K., Y. Aharonowitz, M. Mevarech, S. Wolfe, L. C. Vining, D. W. S. Westlake, and S. E. Jensen.** 1988. Cloning and nucleotide sequence of the IPNS gene from *Streptomyces clavuligerus*. Gene **62**:187–196.
14. **Malpartida, F., and D. A. Hopwood.** 1984. Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. Nature (London) **309**:462–464.
15. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Pang, C.-P., B. Chakravarti, R. M. Adlington, H. H. Ting, R. L. White, G. S. Jayatilake, J. Baldwin, and E. P. Abraham.** 1984. Purification of isopenicillin N synthetase. Biochem. J. **222**: 789–795.
17. **Ramón, D., L. Carramolino, C. Patiño, F. Sánchez, and M. A. Peñalva.** 1987. Cloning and characterization of the isopenicillin N synthetase gene mediating the formation of the β-lactam ring in *Aspergillus nidulans*. Gene **57**:171–181.
18. **Rao, R. N., M. A. Richardson, and S. Kuhstoss.** 1987. Cosmid shuttle vectors for cloning and analysis of *Streptomyces* DNA. Methods Enzymol. **153**:166–198.
19. **Samson, S. M., R. Belagaje, D. T. Blankenship, J. L. Chapman, D. Perry, P. L. Skatrud, R. M. Van Frank, E. P. Abraham, J. E. Baldwin, S. W. Queener, and T. D. Ingolia.** 1985. Isolation, sequence determination, and expression in *Escherichia coli* of the isopenicillin N synthetase gene from *Cephalosporium acremonium*. Nature (London) **318**:191–194.
20. **Samson, S. M., J. L. Chapman, R. Belagaje, S. W. Queener, and T. D. Ingolia.** 1987. Analysis of the role of cysteine residues in isopenicillin N synthetase activity by site-directed mutagenesis. Proc. Natl. Acad. Sci. USA **84**:5705–5709.
21. **Samson, S. M., J. E. Dotzlaf, M. L. Slisz, G. W. Becker, R. M. Van Frank, L. E. Veal, W.-K. Yeh, J. R. Miller, S. W. Queener, and T. D. Ingolia.** 1987. Cloning and expression of the fungal expandase/hydroxylase gene involved in cephalosporin biosynthesis. Bio/Technology **5**:1207–1214.
22. **Scheidegger, A., M. T. Küenzi, and J. Nüesch.** 1984. Partial purification and catalytic properties of a bifunctional enzyme in the biosynthetic pathway of β-lactams in *Cephalosporium acremonium*. J. Antibiot. **37**:522–531.
23. **Schoner, B. E., R. M. Belagaje, and R. G. Schoner.** 1986. Translation of a synthetic two-cistron mRNA in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **83**:8506–8510.
24. **Stanzak, R., P. Matsushima, R. H. Baltz, and R. N. Rao.** 1986. Cloning and expression in *Streptomyces lividans* of clustered erythromycin biosynthesis genes from *Streptomyces erythreus*. Bio/Technology **4**:229–232.
25. **Weigel, B. J., S. G. Burgett, V. J. Chen, P. L. Skatrud, C. A. Frolik, S. W. Queener, and T. D. Ingolia.** 1988. Cloning and expression in *Escherichia coli* of isopenicillin N synthetase genes from *Streptomyces lipmanii* and *Aspergillus nidulans*. J. Bacteriol. **170**:3817–3826.