JOURNAL OF BACTERIOLOGY, Apr. 1991, p. 2354–2365
0021-9193/91/072354-12$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 173, No. 7

Characterization of the *Cephalosporium acremonium pcbAB* Gene Encoding α-Aminoadipyl-Cysteinyl-Valine Synthetase, a Large Multidomain Peptide Synthetase: Linkage to the *pcbC* Gene as a Cluster of Early Cephalosporin Biosynthetic Genes and Evidence of Multiple Functional Domains

SANTIAGO GUTIÉRREZ, BRUNO DÍEZ, EDUARDO MONTENEGRO, AND JUAN F. MARTÍN*

Section of Microbiology, Department of Ecology, Genetics and Microbiology, University of León, 24071 León, Spain

Received 28 September 1990/Accepted 14 January 1991

A 24-kb region of *Cephalosporium acremonium* C10 DNA was cloned by hybridization with the *pcbAB* and *pcbC* genes of *Penicillium chrysogenum*. A 3.2-kb *Bam*HI fragment of this region complemented the mutation in the structural *pcbC* gene of the *C. acremonium* N2 mutant, resulting in cephalosporin production. A functional α-aminoadipyl-cysteinyl-valine (ACV) synthetase was encoded by a 15.6-kb *Eco*RI-*Bam*HI DNA fragment, as shown by complementation of an ACV synthetase-deficient mutant of *P. chrysogenum*. Two transcripts of 1.15 and 11.4 kb were found by Northern (RNA blot) hybridization with probes internal to the *pcbC* and *pcbAB* genes, respectively. An open reading frame of 11,136 bp was located upstream of the *pcbC* gene that matched the 11.4-kb transcript initiation and termination regions. It encoded a protein of 3,712 amino acids with a deduced M_r of 414,791. The nucleotide sequence of the gene showed 62.9% similarity to the *pcbAB* gene encoding the ACV synthetase of *P. chrysogenum*; 54.9% of the amino acids were identical in both ACV synthetases. Three highly repetitive regions occur in the deduced amino acid sequence of *C. acremonium* ACV synthetase. Each is similar to the three repetitive domains in the deduced sequence of *P. chrysogenum* ACV synthetase and also to the amino acid sequence of gramicidin synthetase I and tyrocidine synthetase I of *Bacillus brevis*. These regions probably correspond to amino acid activating domains in the ACV synthetase protein. In addition, a thioesterase domain was present in the ACV synthetases of both fungi. A similarity has been found between the domains existing in multienzyme nonribosomal peptide synthetases and polyketide and fatty acid synthetases. The *pcbAB* gene is linked to the *pcbC* gene, forming a cluster of early cephalosporin-biosynthetic genes.

Cephalosporin C is a β-lactam antibiotic formed in *Cephalosporium acremonium* (*Acremonium chrysogenum*; the name *C. acremonium* is used here to avoid confusion between *Penicillium chrysogenum* and *A. chrysogenum*) and other filamentous fungi by condensation of L-α-aminoadipic acid, L-cysteine, and L-valine (11, 12). The first step in the cephalosporin-biosynthetic pathway involves a complex set of biochemical reactions (Fig. 1), including the activation and condensation of the three precursor amino acids and racemization of the L-valine to the D configuration to form the tripeptide δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV). All these reactions appear to be carried out by the multifunctional enzyme ACV synthetase (38). The ACV formed is then cyclized by the enzyme isopenicillin N synthase, encoded by the gene *pcbC*, to form isopenicillin N, the first intermediate with antibiotic activity. The pathway in *C. acremonium* proceeds with the isomerization of isopenicillin N to penicillin N and expansion of the thiazolidine ring of penicillin N to the dihydrothiazine ring of deacetoxycephalosporin C, which two steps later gives rise to cephalosporin C (23).

Most of the enzymes of the cephalosporin-biosynthetic pathway have been studied, but until now it has been extremely difficult to characterize the ACV synthetase owing to its high molecular weight. Cell-free systems of *C. acremonium* (3, 4) and *Streptomyces clavuligerus* (15) catalyzing ACV formation have been described. This enzyme has been purified from extracts of *Aspergillus nidulans* and shown to be a multifunctional ATP-dependent peptide synthetase (38). The purified ACV synthetase of *C. acremonium* is about 300 kDa (1), although this estimated figure is probably low (8a). Attempts to obtain the N-terminal amino acid sequence proved unsuccessful, but several internal amino acid sequences were obtained by tryptic digestions (see below) (1a).

The gene (*pcbC*) encoding isopenicillin N synthase of *C. acremonium* was cloned a few years ago (32). More recently, the gene (*cefEF*) for the ring-expanding enzyme (deacetoxycephalosporin C synthetase) was cloned; this gene encodes a polypeptide with deacetoxycephalosporin C hydroxylase activity in addition to the deacetoxycephalosporin C synthetase (33). However, the gene encoding the *C. acremonium* ACV synthetase has not been cloned. The availability of this gene would be extremely useful in understanding the molecular mechanisms of ACV synthesis, the first step of the biosynthetic pathway.

Recently, the gene *pcbAB* encoding the ACV synthetase of *P. chrysogenum* was cloned in our laboratory by complementation of blocked mutants deficient in ACV synthetase (10). It shows an open reading frame (ORF) of 11,376 nucleotides that encodes a large protein with a deduced M_r

* Corresponding author.

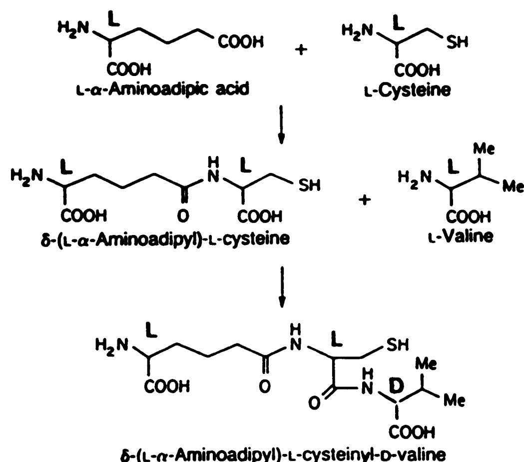


FIG. 1. First step in the cephalosporin-biosynthetic pathway: activation of amino acids, condensation, and isomerization reactions carried out by the ACV synthetase. Note the change in configuration from L-valine (large L) to D-valine (large D). Me, Methyl group.

of 425,971. The availability of the *pcbAB* gene of *P. chrysogenum* for heterologous hybridization prompted us to clone the equivalent gene of the cephalosporin producer *C. acremonium*. In this article, we report the characterization of the *pcbAB* gene of *C. acremonium* C10 and the homology of the multifunctional ACV synthetase encoded by this gene with the ACV synthetase of *P. chrysogenum* and the gramicidin synthetase I (GS1) and the tyrocidine synthetase I (TY1) of *Bacillus brevis*. The *pcbAB* gene of *C. acremonium* is linked to the *pcbC* gene, forming a cluster of early cephalosporin-biosynthetic genes.

MATERIALS AND METHODS

Microorganisms and vectors used. *C. acremonium* C10, a high-cephalosporin-producing strain released by Panlabs Inc. (3, 4, 29), was used as the source of DNA. *C. acremonium* N2, a cephalosporin nonproducer mutant (36) which is deficient in isopenicillin N synthase (29) due to a mutation in the structural *pcbC* gene, was used in complementation studies. The *pcbAB* and *pcbC* genes of *P. chrysogenum* were cloned previously (6, 10) from the DNA of *P. chrysogenum* AS-P-78, a strain provided by Antibióticos, S.A. (León, Spain).

Escherichia coli DH5α was used as the recipient strain for high-frequency plasmid transformation, and *E. coli* Q-359 served as the host for lambda EMBL3 phage derivatives. *E. coli* WK6 was used as the host for obtaining single-stranded DNA from pBluescript plasmids.

Construction of a gene library of *C. acremonium* C10 DNA. A gene library of *C. acremonium* C10 was constructed in the *ble*-EMBL3 vector, an EMBL3 phage derivative that carries the bleomycin-phleomycin resistance gene (*ble*) of *Streptoalloteichus hindustanus* (13a). *ble*-EMBL3 DNA was digested with *Bam*HI to separate the arms of the phage, which were purified by sucrose gradient centrifugation. About 0.4 μg of DNA was mixed with 0.5 μg of partially *Sau*3AI-digested total *C. acremonium* DNA (after selecting fragments of 13 to 17 kb by sucrose gradient ultracentrifugation). The ligation mixture was packaged in vitro by using the lambda phage packaging system (Amersham, Buckinghamshire, U.K.) and used to infect *E. coli* Q-359. A total of approximately 70,000 phage plaques were studied.

Southern hybridization with the *pcbC* probe. About 14,000 phage plaques per petri dish were blotted to nitrocellulose paper and hybridized as described before (6) with a 1.0-kb *Nco*I fragment that carries the promoterless *pcbC* gene of *P. chrysogenum*. The probe was labeled by nick translation by standard methods (6, 20).

mRNA isolation and Northern (RNA) blotting. Total RNA from *C. acremonium* was obtained by the phenol-sodium dodecyl sulfate (SDS) method as described before for *P. chrysogenum* (6, 9, 10). RNA (5 μg) was run in a 0.7% agarose–formaldehyde gel with RNA molecular weight markers (*E. coli* 16S and 23S rRNAs and sets I and III from Boehringer Mannheim). The gel was blotted onto a nitrocellulose filter by standard methods, baked in a vacuum oven at 80°C for 2 h, prehybridized at 42°C overnight in 50% formamide–5× Denhardt's solution–5× SSPE (20)–0.1% SDS–500 μg of denatured salmon sperm DNA per ml and hybridized in the same buffer containing 100 μg of denatured salmon sperm DNA per ml at 42°C for 24 h (1× Denhardt's solution contains 0.02% each bovine serum albumin, Ficoll, and polyvinylpyrrolidone). Afterwards, the filter was washed twice in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS at room temperature for 15 min, twice again in 0.1× SSC–0.1% SDS at room temperature for 15 min, and once more in 0.1× SSC–0.1% SDS at 55°C and autoradiographed with Amersham X-ray film.

DNA sequencing. DNA was sequenced by two different strategies. The head and tail fragments of the gene (see Results) were sequenced after subcloning into small fragments by the dideoxynucleotide termination method (34), with either Sequenase (U.S. Biochemicals, Cleveland, Ohio) or *Taq* polymerase (Promega, Madison, Wis.) as described previously (6, 26).

The internal fragments of the gene subcloned into pBluescript KS(+) were sequenced by generating ordered sets of deletions with the Erase-a-Base system (Promega) by digestion with exonuclease III from appropriate ends (see Results), followed by S1 exonuclease removal of single-stranded DNA and treatment with Klenow DNA polymerase to fill the gaps introduced by the former enzymes (5). Sequencing of the ordered sets of fragments was carried out by the dideoxynucleotide method.

Computer analysis of nucleotide and amino acid sequences was made with the DNASTAR Program (DNA STAR, London, U.K.).

Complementation of an ACV synthetase-deficient mutant of *P. chrysogenum*. Protoplasts of *P. chrysogenum npe5*, a mutant obtained from *P. chrysogenum* Wis 54-1255, which is deficient in ACV synthetase (13b), were transformed with 5 μg of pULJL43, a vector carrying the phleomycin resistance marker, or pULSC1, a derivative of pULJL43 that carries the ACV synthase and isopenicillin N synthase genes in addition to the phleomycin resistance marker. Protoplasts were obtained and transformed protoplasts were regenerated in Czapek-sorbitol medium as described before (6a). Transformants were selected by resistance to phleomycin (30 μg/ml).

RESULTS

Cloning of the region upstream of the *pcbC* gene. Since evidence from our work on the cloning of the ACV synthetase gene of *P. chrysogenum* indicated that this gene was located upstream of the isopenicillin N synthase gene, the

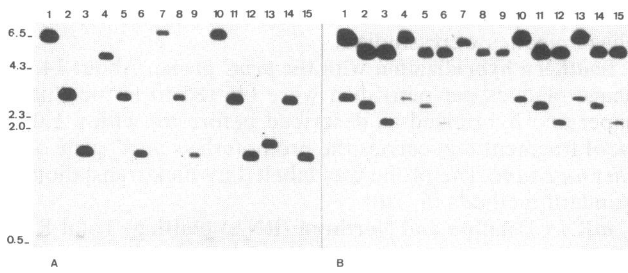


FIG. 2. Hybridization of DNA from five recombinant phages with probes internal to the *pcbC* gene (A) or the *pcbAB* gene (B). DNAs were digested with *Sal*I (lanes 1, 4, 7, 10, and 13), *Bam*HI (lanes 2, 5, 8, 11, and 14), or *Sal*I plus *Bam*HI (lanes 3, 6, 9, 12, and 15). Size markers are indicated on the left (in kilobases).

DNA regions flanking the *pcbC* gene of *C. acremonium* were cloned by hybridization with probes containing internal fragments of the *pcbC* and *pcbAB* genes of *P. chrysogenum*.

A gene library of the high-cephalosporin-producing strain *C. acremonium* C10 was prepared in the *ble*-EMBL3 substitution vector and screened with a 1.0-kb *Nco*I probe carrying the *pcbC* gene of *P. chrysogenum* without its promoter region (6, 7).

Eight phage plaques gave positive hybridization with the *pcbC* probe. They corresponded to phages with five different inserts of *C. acremonium* DNA that were named F1, F3, F4, F5, and F6. The DNA of these five phages was purified, digested with *Sal*I, *Bam*HI, and *Sal*I plus *Bam*HI, and hybridized with the same *pcbC* probe as above. Although the phages showed *Sal*I bands of different sizes that hybridized with the probe, all of them contained a 3.2-kb *Bam*HI fragment which hybridized with the *pcbC* gene (Fig. 2A). The 3.2-kb *Bam*HI fragment was mapped with several restriction enzymes subcloned in pBluescript KS(+) and partially sequenced (1.62 kb). The nucleotide sequence obtained was very similar to that of the upstream region of the *pcbC* gene of *P. chrysogenum* and included 0.4 kb of the previously reported *pcbC* gene of *C. acremonium* (32).

Complementation of the *C. acremonium* N2 mutant blocked in isopenicillin N synthase formation. The 3.2-kb *Bam*HI fragment was subcloned in the fungal vector pULJL43 (5a, 10) and used to transform *C. acremonium* N2, a strain with a point mutation in the structural *pcbC* gene (29, 30). Similarly, transformants were obtained with a 7.2-kb *Hin*dIII fragment of the insert in phage F3 subcloned in pULJL43, which also carries the *pcbC* gene region.

Transformants were isolated by resistance to phleomycin and assayed for cephalosporin production by the agar plug method. Sixteen transformants out of 18 tested gave positive complementation resulting in cephalosporin production, confirming that the 3.2-kb *Bam*HI fragment (internal to the 7.2-kb *Hin*dIII fragment) common to all five recombinant phages carried the functional *C. acremonium pcbC* gene.

***pcbAB* gene of *C. acremonium* is linked to the *pcbC* gene.** A 6.0-kb *Sal*I internal fragment of the *pcbAB* gene of *P. chrysogenum* was used as a probe to hybridize with the same blots that had been hybridized with the *pcbC* gene. All five recombinant phages showed strong hybridization (Fig. 2B) with the *pcbAB* probe, suggesting that the two genes were linked in the phage inserts. Mapping with several restriction enzymes followed by hybridization with the *pcbAB* probe showed that the hybridizing region was located upstream of the *pcbC* gene, as occurs in *P. chrysogenum* (10).

The largest insert in phage F3 (Fig. 3) extends 11.5 kb upstream of the *pcbC* gene. Since this amount of genetic sequence might not be enough to encode a gene of the expected size of *pcbAB* (11 to 12 kb), additional phages were isolated from the library of *C. acremonium* C10 with a homologous 0.935-kb *Xho*I probe internal to the region of the *pcbAB* gene (probe P_6 in Fig. 4). Seventeen new recombinant phages were isolated that carried regions that extended upstream of the *pcbC* gene. The maps of phages F12, F17, F18, F20, F21, F23, F24, and F25 together with those of F1, F3, F4, F5, and F6 are shown in Fig. 3.

Transcriptional map of the region upstream of the *pcbC* gene. Total RNA from *C. acremonium* C10 was hybridized as indicated in Materials and Methods with eight different consecutive DNA fragments of the region that contained the *pcbAB* and *pcbC* genes (probes P_1 to P_8, Fig. 4) to establish the sizes of the transcripts in that region.

The results (Fig. 5) of the different hybridizations showed that two transcripts of 1.15 and 11.4 kb were found in the region. Probes P_1, P_2, P_5, P_6, and P_7 gave hybridization with the 11.4-kb transcript, whereas P_3, P_4, and P_8 did not. Probes P_3 and P_4 hybridized with the 1.15-kb transcript that corresponded to the *pcbC* gene. These results indicated that a large gene that might encode a polypeptide of the size of the ACV synthetase was located upstream of the *pcbC* gene.

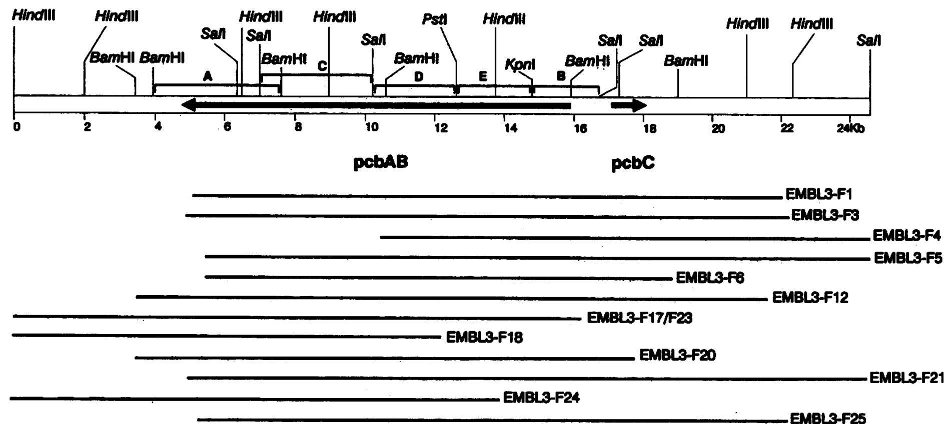
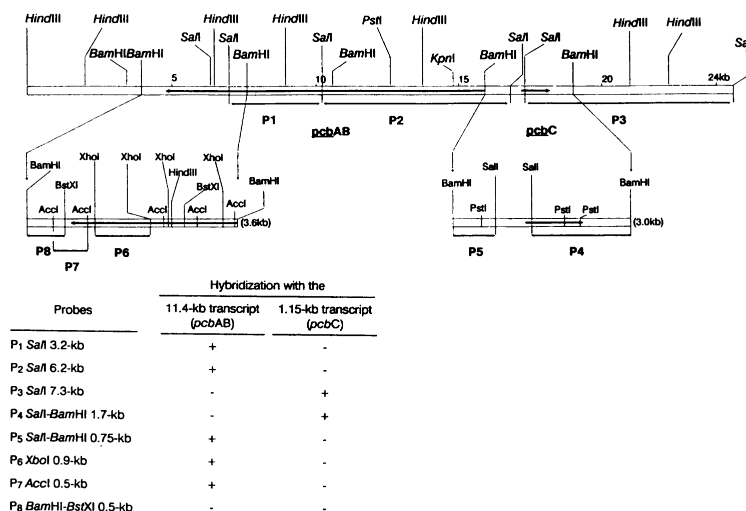


FIG. 3. Restriction maps of the inserts in 12 different phages. The relative positions and orientations of the genes *pcbC* and *pcbAB* are indicated by thick arrows. The DNA fragments A, B, C, D, and E used for sequencing are indicated by horizontal brackets.



Probes	Hybridization with the 11.4-kb transcript (pcbAB)	Hybridization with the 1.15-kb transcript (pcbC)
P_1 *SalI* 3.2-kb	+	-
P_2 *SalI* 6.2-kb	+	-
P_3 *SalI* 7.3-kb	-	+
P_4 *SalI-BamHI* 1.7-kb	-	+
P_5 *SalI-BamHI* 0.75-kb	+	-
P_6 *XhoI* 0.9-kb	+	-
P_7 *AccI* 0.5-kb	+	-
P_8 *BamHI-BstXI* 0.5-kb	-	-

FIG. 4. Transcriptional map of the region enclosing the *pcbAB* and *pcbC* genes, showing the DNA fragments used in Northern hybridizations to map the transcripts (probes P_1 to P_8). The relative positions of the *pcbAB* and *pcbC* genes are indicated by arrows.

Transcript initiation and termination regions. To establish the end of the 11.4-kb transcript, hybridizations were carried out with probes corresponding to consecutive small DNA fragments of the regions of the expected ends of the transcript. The distal end of the transcript, with respect to the *pcbC* gene, was mapped by using three probes, P_6, P_7, and P_8. Probes P_6 and P_7 gave positive hybridization, whereas the adjacent probe P_8 did not. This result indicated that the distal end of the 11.4-kb transcript was located in the 0.5-kb *AccI* fragment.

Similarly, the proximal end of the 11.4-kb transcript was mapped with probes P_4 and P_5. P_4 did not gave hybridization, whereas P_5 showed faint hybridization, indicating that the 0.75-kb *SalI-BamHI* fragment corresponded to one end of the transcript. This result was confirmed by nucleotide sequence determination, which identified a GTG translation initiation triplet at the beginning of an ORF in this region.

The orientation of the gene was established by homology with the *P. chrysogenum* gene and also by correlation with the amino acid sequences of internal peptides (see below); it is transcribed in the opposite direction from the *pcbC* gene, as occurs in *P. chrysogenum*.

***pcbAB* gene has an uninterrupted reading frame of 11,136 bp.** The entire region enclosing the *pcbAB* gene was initially subcloned in dephosphorylated pBluescript KS(+) in five fragments (Fig. 3): A, 3.6-kb *BamHI*; B, 1.7-kb *SalI-KpnI* with filled-in ends; C, 3.2-kb *SalI*; D, 2.4-kb *SalI-PstI*; and E, 2.0-kb *PstI-KpnI*.

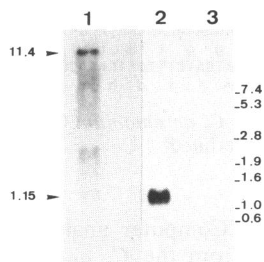


FIG. 5. Northern hybridization of total RNA from *C. acremonium* C10 (48-h culture) obtained as indicated in the text with probes P_1 (lane 1), P_4 (lane 2), and P_8 (lane 3) (see Fig. 4). The hybridizing bands corresponding to the *pcbAB* (11.4 kb) and *pcbC* (1.15 kb) transcripts are indicated. RNA size markers (set I from Boehringer) are indicated on the right (in kilobases).

Two different strategies were used to sequence these five fragments. The head and tail fragments B and A were subcloned in both orientations in pBluescript KS(+) digested with *EcoRV* or *BamHI*, respectively. Fragment A was subdivided into 23 small fragments, and fragment B was subcloned into 14 small fragments. These overlapping inserts were sequenced by the dideoxynucleotide termination method (34).

The three internal large fragments of the *pcbAB* gene (C, D, and E) were sequenced by the Erase-a-Base method (Promega). Fragment C, subcloned in *SalI*-digested pBluescript KS(+) in both orientations, was digested with *BstXI* and *XbaI* to obtain appropriate ends to generate the deletions. Fragment D was subcloned (filled-in ends) into *EcoRV*-digested pBluescript KS(+) and digested with *PstI* and *EcoRI*. Finally, fragment E (filled-in ends) was also subcloned in the *EcoRV* site of pBluescript KS(+) and digested with *PstI* and *EcoRI*. Clones overlapping the junctions of the five fragments were sequenced; the entire region (13.6 kb) was sequenced in both strands.

The nucleotide sequence of an 11.6-kb region corresponding to the *pcbAB* gene is shown in Fig. 6. A very long ORF of 11,136 bp was found that matched the transcript initiation and termination regions identified by Northern hybridization. The G+C content of the ORF was 58.16%, and the codon usage was similar to that of the cloned genes from *P. chrysogenum* (10) and other filamentous fungi (2). The protein encoded consisted of 3,712 amino acids with a deduced M_r of 414,791. Internal peptides 72, 37.5, 42.8, and 63.4 described by Baldwin et al. (1a) were found in amino acid positions 556 to 575, 192 to 200, 1754 to 1763, and 2694 to 2703, respectively, with homologies to the amino acid sequence of the *C. acremonium* ACV synthetase ranging from 80 to 100%.

The nucleotide and deduced amino acid sequences were very similar to those of the ACV synthetase of *P. chrysogenum* (10) (Fig. 7), which showed 62.9% similarity at the nucleotide level and 54.9% homology at the amino acid level.

```
GGATGCATGCATTGGCCTGTATCAAAGGTCCGGGATTCCCCAGGAGTATAAGACGTTCGTGCTGGGAGATCTAGCGACGTGTTGGGAAATATCGGCCGTA   100
GAGTGCGAAAAAGAACTGGCGGAAATATTTCTCCTTGGACTCGGTCACACTCAGTCAGTAGTGGACTGCCAGTCTATCATACACCTTTGATATCAACATG   200
ACTATCCTTACAGGTGCCGACGACGCCTCGTCATACCACAGGTATGTCTTCACAGCCTCTGGAAAGCGCAGTTGGGAGCTATCTCTAACATTACCACATC   300
AGGCGCAATGGAAGCTCTGATATCCCAAAAGGTGCCATCCACCGCAACGGCTTCGCAGCCGCAGCCCCTGACTGCTGGATCCGGTCC
                                                                                     GTGGCCCTGGAAC   400
                                                                                      m  a  l  e  q
AGTGGAAGACTACGGTCCAGTCCGTCTCGGAGCGGTGCGATCTGAGCGGGCTGAGCCAGCATCCCACCGACTACCAGCTGGCCTCTACGGGCGTGAAGGG   500
 w  k  t  t  v  q  s  v  s  e  r  c  d  l  s  g  l  s  q  h  p  t  d  y  q  l  a  s  t  g  v  k  g
CGCAGGCGGTAGCAGCATCGAGGAGCGCAGTGCCATCGTCTCAGACGAGTTGTTCTCGAGTCTGCGAGACGTGTGCTCACAGAGACAGCTGGACCCTCGG   600
 a  g  g  s  s  i  e  e  r  s  a  i  v  s  d  e  l  f  s  s  l  r  d  v  c  s  q  r  q  l  d  p  r
TCACTCATGCTGTTTTCCGTGCACCAGATGCTCAAGAGGTTCGGAAACGGATCTCACACCGTCGTGGCGTCACTCGTAACTTCATCAGAGGGATGCCCTT   700
 s  l  m  l  f  s  v  h  q  m  l  k  r  f  g  n  g  s  h  t  v  v  a  s  l  v  t  s  s  e  g  c  p  s
CAACTTCGGCCTGGAGGGCCATCCCCTCCGTCATCCATCATATAGAGGGCGGAGACAACAACAACACAGTCGCCTCTGCCGTGGAACAGGCGGCGAATCT   800
 t  s  a  w  r  a  i  p  s  v  i  h  h  i  e  g  g  d  n  n  n  t  v  a  s  a  v  e  q  a  a  n  l
CCTGAACTCAGAAGGATCGGGACAGGACCTTCTGATTCCCATCGGACTCACTGAGCTCGTCAAGTCGGAGCTGATTGACCTCCTGGTCATCTTCGACGAC   900
 l  n  s  e  g  s  g  q  d  l  l  i  p  i  g  l  t  e  l  v  k  s  e  l  i  d  l  l  v  i  f  d  d
GAGACAAATAACATACGACTGCCGCAGGACTTCCCACTTATCCTGCGGATACATCAGCGGCAAGACCACTGGCAGCTGTCAGTCCGGTATCCCTCGCCCC   1000
 e  t  n  n  i  r  l  p  q  d  f  p  l  i  l  r  i  h  q  r  q  d  h  w  q  l  s  v  r  y  p  s  p  l
TTTTCGACACCATGGTCATCGACAGCTTTCTGAGCGCACTTCACAACCTGTTGTCCGCGGTGACAAAACCCTCCCAGCTCGTGCGCGACATCGAGCTGCT   1100
 f  d  t  m  v  i  d  s  f  l  s  a  l  h  n  l  l  s  a  v  t  k  p  s  q  l  v  r  d  i  e  l  l
CCCAGAATACCAGGTCGCTCAGCTGGAGAAGTGGAACAACACAGACGGCGACTACCCCACCGAGAAGCGGCTACATCATCTGTTCGAGGAGGCAGCAGTG   1200
 p  e  y  q  v  a  q  l  e  k  w  n  n  t  d  g  d  y  p  t  e  k  r  l  h  h  l  f  e  e  a  a  v
CGTCGTCCCCAACACGTTGCCCTCATCTGCGGCGACAAGCGCATCACCTATGAGGAGTTGAATGCTATGGCGAATCGCCTGGCCCACCATCTGGTATCCT   1300
 r  r  p  q  h  v  a  l  i  c  g  d  k  r  i  t  y  e  e  l  n  a  m  a  n  r  l  a  h  h  l  v  s  s
CGGGTATCCAGACTGAGCAGCTCGTCGGTCTCTTCCTCGACAAGACCGAGCTCATGATCGCTACTATTCTGGGCATCTGGAAATCTGGTGCCGCGCATGT   1400
 g  i  q  t  e  q  l  v  g  l  f  l  d  k  t  e  l  m  i  a  t  i  l  g  i  w  k  s  g  a  a  h  v
ACCTATCGACCCTGGGTACCCGGACGAGCGTGTCAAGTTCGTCCTGAATGATACGAAGGCGCAAGTGGTCATTGCTAGTCAGAGGCACGTCGATCGACTG   1500
 p  i  d  p  g  y  p  d  e  r  v  k  f  v  l  n  d  t  k  a  q  v  v  i  a  s  q  r  h  v  d  r  l
CGGGCTGAGGCTGTTGGCGGCCAGCATCTTCGCATCATCGGTCTCGAATCTCTGTTCGACAACCTTGCTCAACAGACACAACACTCACCAGAGACGTCGG   1600
 r  a  e  a  v  g  g  q  h  l  r  i  i  g  l  e  s  l  f  d  n  l  a  q  q  t  q  h  s  p  e  t  s  g
GCAATTTGACCCATCTGCCCCTGAACAGCAAACAGCTTGCGTACGTGACATACACCTCGGGCACCACGGGCTTCCCGAAAGGCATCTACAAGGAGCACAC   1700
 n  l  t  h  l  p  l  n  s  k  q  l  a  y  v  t  y  t  s  g  t  t  g  f  p  k  g  i  y  k  e  h  t
AAGCGTCGTTAACAGCATCACCGATCTGTCTGCTCGGTACGGTGTGGCCGGGGAGGACGACGAGGTGATACTCGTCTTCTCCGCCTACGTCTTCGAGCCA   1800
 s  v  v  n  s  i  t  d  l  s  a  r  y  g  v  a  g  e  d  d  e  v  i  l  v  f  s  a  y  v  f  e  p
TTCGTGCGCCAGATGCTCATGGCCCTGACCACGGGCAACTCTCTCGCCATCATCAGCGACGAGGACAAGTTCGACCCTGACACCCTTATTCCCTTCATCC   1900
 f  v  r  q  m  l  m  a  l  t  t  g  n  s  l  a  i  i  s  d  e  d  k  f  d  p  d  t  l  i  p  f  i  q
AAAAACACAAAGTCACTTACATCCACGCCACCTCGTCAGTGTTGCAGGAGTACGACTTCGGGTCCTGCCCCTCGTTGAAACGCATGATTCTGGTGGGAGA   2000
 k  h  k  v  t  y  i  h  a  t  s  s  v  l  q  e  y  d  f  g  s  c  p  s  l  k  r  m  i  l  v  g  e
GAACTTGACAGAGCCGCGCTACGAGGCCCTGAGGCAGCGCTTCAAGTCGCGCATCCTGAATGAATATGGCTTCACCGAGTCTGCGTTTGTGACGGCGCTC   2100
 n  l  t  e  p  r  y  e  a  l  r  q  r  f  k  s  r  i  l  n  e  y  g  f  t  e  s  a  f  v  t  a  l
AACATATTCGAGCCTACCTCACAGAGGAAGGACATGAGTCTGGGAAGGCCGGTGCGCAACGTCAAGTGCTATATCTTGGATGCCAACCTCAAGAGAGTCC   2200
 n  i  f  e  p  t  s  q  r  k  d  m  s  l  g  r  p  v  r  n  v  k  c  y  i  l  d  a  n  l  k  r  v  p
CCATCGGTGTTACAGGGGAGCTGCACATCGGTGGCTTGGGTATATCCCGGGGGTACATGAATAGGGAGGAGCTCACAAGGCAGAAGTTCCTCCCGAACCC   2300
 i  g  v  t  g  e  l  h  i  g  g  l  g  i  s  r  g  y  m  n  r  e  e  l  t  r  q  k  f  l  p  n  p
CTACCAGACCGATAAGGAGCGCCAACGGGGTGTCAACTCAACCATGTACAAGACAGGAGATCTGGCCCGCTGGCTACCCAGTGGCGAAGTCGAGTATCTC   2400
 y  q  t  d  k  e  r  q  r  g  v  n  s  t  m  y  k  t  g  d  l  a  r  w  l  p  s  g  e  v  e  y  l
GGCCGTGCCGACTTCCAGATCAAGCTGCGCGGCATTCGAATTGAGCCCGGCGAGATCGAGTCCACTCTCGCCATGTATCCCGGAATCAGGGCCAGCATCG   2500
 g  r  a  d  f  q  i  k  l  r  g  i  r  i  e  p  g  e  i  e  s  t  l  a  m  y  p  g  i  r  a  s  i  v
TCGTGTCAAAGAAGCTTCTCAGTCAGGGGCAGGAGACGATCCAAGACCACCTTGTGGGGTACTATGTTTGCGATGAGGGCCACATCCCCGAGGGTGACCT   2600
 v  s  k  k  l  l  s  q  g  q  e  t  i  q  d  h  l  v  g  y  y  v  c  d  e  g  h  i  p  e  g  d  l
GCTGAGCTTCCTGGAGAAGAAGCTACCTCGGTACATGGTCCCGACGCGCCTTGTCCAACTGGCTCAGATTCCAACCAATATCAACGGCAAGGCGGATCTG   2700
 l  s  f  l  e  k  k  l  p  r  y  m  v  p  t  r  l  v  q  l  a  q  i  p  t  n  i  n  g  k  a  d  l
CGTGCTCTTCCTGCCGTCGAAGTCGCCGTAGCTCCCACCCACAAGCAGGATGGCGAGCGAGGAAACCAGCTGGAGAGCGACCTGGCTGCCATATGGGGCA   2800
 r  a  l  p  a  v  e  v  a  v  a  p  t  h  k  q  d  g  e  r  g  n  q  l  e  s  d  l  a  a  i  w  g  n
ACATTTTGAGTGTTCCCGCTCAAGACATTGGGTCTGAATCCAACTTCTTCCGCCTGGGTGGCCACAGTATTGCATGCATCCAGCTCATTGCTCGTGTGCG   2900
 i  l  s  v  p  a  q  d  i  g  s  e  s  n  f  f  r  l  g  g  h  s  i  a  c  i  q  l  i  a  r  v  r
ACAGCAGCTAGGCCAGGGGATTACCCTCGAGGAGGTCTTCCAGACCAAGACGTTGCGAGCTATGGCTGCCCTCTTGTCGGAAAAGTACACGAAGGCGTCG   3000
 q  q  l  g  q  g  i  t  l  e  e  v  f  q  t  k  t  l  r  a  m  a  a  l  l  s  e  k  y  t  k  a  s
AATGGGACGAACGGAGTGACCAACGGCACTGCTCACGTCAACGGCCACGCAGCGAACGGCCATGTCAGCGACAGCTACGTGGCCAGCAGTTTGCAGCAAG   3100
 n  g  t  n  g  v  t  n  g  t  a  h  v  n  g  h  a  a  n  g  h  v  s  d  s  y  v  a  s  s  l  q  q  g
GCTTTGTTTACCATTCACTCAAGAACGAACTGTCCGAGGCGTACACCATGCAATCCATGATCCACTATGGTGTGCCCCTGAAACGGGATATTTACCAAGC   3200
 f  v  y  h  s  l  k  n  e  l  s  e  a  y  t  m  q  s  m  i  h  y  g  v  p  l  k  r  d  i  y  q  a
GGCATGGCAGAGGGTACAGGGGGAGCACCCTGCACTGCGGCTTCGGTTCACATGGGAGGCCGAAGTGATGCAGATCGTGGACCCGAAATCTGAACTCGAC   3300
 a  w  q  r  v  q  g  e  h  p  a  l  r  l  r  f  t  w  e  a  e  v  m  q  i  v  d  p  k  s  e  l  d
TGGCGTGTTGTTGACTGGACCGATGTTTCGAGCCGGGAGAAGCAGCTGGTTGCGCTGGAGCAACTCCAAACGGAGGACCTTGCTAAGGTCTACCATCTCG   3400
 w  r  v  v  d  w  t  d  v  s  s  r  e  k  q  l  v  a  l  e  q  l  q  t  e  d  l  a  k  v  y  h  l  d
ATAAGGGGCCCCTTATGCGACTATACCTCATCCTGCTTCCGGACTCAAAGTACTCCTGTCTGTTCAGCTGCCACCATGCCATTCTCGATGGGTGGAGTCT   3500
 k  g  p  l  m  r  l  y  l  i  l  l  p  d  s  k  y  s  c  l  f  s  c  h  h  a  i  l  d  g  w  s  l
GCCCCTGCTCTTCAACAATGTCCACCAGGCCTACCTCGATCTCGTCGAAGGCACTGCTTCGCCCGTCGAGCAGGACGCTACCTACCTACTCGGCCAGCAG   3600
 p  l  l  f  n  n  v  h  q  a  y  l  d  l  v  e  g  t  a  s  p  v  e  q  d  a  t  y  l  l  g  q  q
TACCTGCAGAGCCACAGGGACGACCATCTCGACTTCTGGGCCGAGCAGATCGGCAGGATCGAAGAGCGCTGCGACATGAATGCGCTGCTGAATGAGGCCA   3700
 y  l  q  s  h  r  d  d  h  l  d  f  w  a  e  q  i  g  r  i  e  e  r  c  d  m  n  a  l  l  n  e  a  s
GCCGATACAAGGTGCCCCTGGCCGACTATGACCAAGTCCGCGAGCAGAGGCAGCAGACCATCAGTCTGCCCTGGAACAACTCCATGGACGCTGGTGTGCG   3800
 r  y  k  v  p  l  a  d  y  d  q  v  r  e  q  r  q  q  t  i  s  l  p  w  n  n  s  m  d  a  g  v  r
GGAAGAACTCTCCAGTCGTGGCATCACCCTTCATTCCATTCTACAGACGGTCTGGCACCTGGTCCTCCACTCTTATGGAGGAGGCACCCACACGATCACC   3900
 e  e  l  s  s  r  g  i  t  l  h  s  i  l  q  t  v  w  h  l  v  l  h  s  y  g  g  g  t  h  t  i  t
GGCACCACCATCTCCGGCCGTCACCTGCCCGTCCCCGGAATTGAGCGCTCTGTTGGTCTCTTCATCAACACACTCCCTATGATCTTTGATCACACCGTCT   4000
 g  t  t  i  s  g  r  h  l  p  v  p  g  i  e  r  s  v  g  l  f  i  n  t  l  p  m  i  f  d  h  t  v  c
```

FIG. 6. Nucleotide and deduced amino acid sequences of an 11.6-kb fragment enclosing the *C. acremonium* C10 ACV synthetase gene. The translation initiation triplet is boxed, and a TATA sequence in the upstream region is underlined.

The two amino acid sequences were perfectly aligned with the DNASTAR Program with the introduction of infrequent gaps. A single in-frame GTG translation initiation codon was found in the 5′ region of the gene by homology with the amino acid sequence in the amino-terminal region of the *pcbAB* gene of *P. chrysogenum*.

Amino acid sequence of the ACV synthetase contains three regions similar to each other and to regions found in other peptide synthetases. Computer analysis of the amino acid sequence deduced from the *C. acremonium pcbAB* gene detected three regions that were similar to each other. The analysis suggests that these regions correspond to repeated domains in the ACV synthetase protein (Fig. 8) similar to those observed in the ACV synthetase of *P. chrysogenum*. The first domain extends from amino acids 234 to 1062, the second from amino acids 1335 to 2162, and the third from

```
GCCAGGATATGACAGCGCTCGAGGCCATTGAGCATGTCCAAGGCCAAGTCAACGCCATGAACTCCCGGGGCAACGTCGAGCTCGGACGCATGAGCAAGAA   4100
 q  d  m  t  a  l  e  a  i  e  h  v  q  g  q  v  n  a  m  n  s  r  g  n  v  e  l  g  r  m  s  k  n
CGACCTCAAGCACGGGCTCTTCGACACCCTCTTCGTCCTCGAGAACTACCCAAACCTCGACACGGAGCAGCGGGAGAAGCACGAGGAGAAGCTCAAGTTC   4200
 d  l  k  h  g  l  f  d  t  l  f  v  l  e  n  y  p  n  l  d  t  e  q  r  e  k  h  e  e  k  l  k  f
ACCATCAAGGGTGGCACGGAGAAGCTCAGTTACCCGCTGGCCGTGATTGCCCAAGAGGACGGCGACAGCGGATGCTCGTTTACGCTCTGCTATGCGGGCG   4300
t  i  k  g  g  t  e  k  l  s  y  p  l  a  v  i  a  q  e  d  g  d  s  g  c  s  f  t  l  c  y  a  g  e
AGCTCTTCACGGATGAGTCCATCCAGGCGCTCCTGGACACTGTCCGGGACACCCTGAGTGATATTCTCGGGAACATCCATGCCCCTATCCGCAACATGGA   4400
 l  f  t  d  e  s  i  q  a  l  l  d  t  v  r  d  t  l  s  d  i  l  g  n  i  h  a  p  i  r  n  m  e
GTACCTCTCCTCGAACCAGACGGCGCAGCTCGACAAGTGGAATGCCACCGCCTTCGAGTACCCCAACACCACACTGCACGCCATGTTCGAGTCCGAGGCG   4500
 y  l  s  s  n  q  t  a  q  l  d  k  w  n  a  t  a  f  e  y  p  n  t  t  l  h  a  m  f  e  s  e  a
CAGCAGAAGCCGGACAAGGTGGCCGTGGTGTACGAGGATATCAGGCTGACCTACCGCGAGCTCAACAGCCGTGCCAATGCCCTGGCGTTCTACCTCCTCT   4600
q  q  k  p  d  k  v  a  v  v  y  e  d  i  r  l  t  y  r  e  l  n  s  r  a  n  a  l  a  f  y  l  l  s
CCCAGGCGGCTATCCAACCGAACAAGCTGGTCGGGCTGATCATGGACAAGAGCGAGCACATGATCACGAGCATCCTCGCGGTCTGGAAAACGGGTGGAGC   4700
 q  a  a  i  q  p  n  k  l  v  g  l  i  m  d  k  s  e  h  m  i  t  s  i  l  a  v  w  k  t  g  g  a
CTACGTCCCGATCGACCCTCGATACCCTGACCAGCGTATCCAGTATATCCTGGAGGATACGGCGGCTCTCGCAGTCATCACGGACAGTCCTCATATTGAC   4800
 y  v  p  i  d  p  r  y  p  d  q  r  i  q  y  i  l  e  d  t  a  a  l  a  v  i  t  d  s  p  h  i  d
CGTCTGCGCAGCATCACCAACAACCGCCTTCCTGTTATCCAGTCGGACTTTGCTCTCCAACTCCCGCCCAGCCCAGTTCATCCCGTCTCAAACTGCAAGC   4900
r  l  r  s  i  t  n  n  r  l  p  v  i  q  s  d  f  a  l  q  l  p  p  s  p  v  h  p  v  s  n  c  k  p
CAAGCGACCTCGCCTACATCATGTACACATCCGGCACCACTGGCAACCCCAAGGGTGTCATGGTGGAGCACCACGGTGTAGTGAATCTGTGCGTTTCACT   5000
 s  d  l  a  y  i  m  y  t  s  g  t  t  g  n  p  k  g  v  m  v  e  h  h  g  v  v  n  l  c  v  s  l
CTGCCGGCTCTTCGGCCTTCGGAACACAGATGACGAGGTCATCCTCTCGTTCTCGAACTACGTCTTCGACCACTTTGTCGAGCAGATGACGGATGCCCTT   5100
 c  r  l  f  g  l  r  n  t  d  d  e  v  i  l  s  f  s  n  y  v  f  d  h  f  v  e  q  m  t  d  a  l
CTCAACGGTCAGACTCTTGTGGTCCTCAACGACGAGATGCGTGGCGACAAGGAGAGGCTTTACAGATACATCGAGACCAACCGCGTCACGTACCTCTCGG   5200
 l  n  g  q  t  l  v  v  l  n  d  e  m  r  g  d  k  e  r  l  y  r  y  i  e  t  n  r  v  t  y  l  s  g
GGACACCTTCCGTCATCTCCATGTACGAGTTCGACCGGTTCCGCGACCACCTGCGGCGCGTGGATTGCGTCGGCGAGGCCTTCAGCGAGCCGGTATTCGA   5300
 t  p  s  v  i  s  m  y  e  f  d  r  f  r  d  h  l  r  r  v  d  c  v  g  e  a  f  s  e  p  v  f  d
CAAGATCCGCGAGACGTTCCCGGGTCTCATCATCAACGGTTATGGCCCGACTGAGGTGTCTATCACTACCCACAAGCGGCCCTACCCGTTCCCGGAGCGC   5400
k  i  r  e  t  f  p  g  l  i  i  n  g  y  g  p  t  e  v  s  i  t  t  h  k  r  p  y  p  f  p  e  r  r
GCACAGACAAGAGCATCGGTTGCCAGCTGGACAACAGCACGAGCTACGTCCTCAACGATGACATGAAGCGCGTGCCCATCGGGGCCGTGGGAGAGCTGTA   5500
 t  d  k  s  i  g  c  q  l  d  n  s  t  s  y  v  l  n  d  d  m  k  r  v  p  i  g  a  v  g  e  l  y
CCTTGGTGGCGATGGCGTCGCTCGCGGATACCACAACCGGCCAGACCTGACGGCTGACCGGTTCCCTGCCAACCCCTTCCAGACGGAGCAGGAGAGACTT   5600
 l  g  g  d  g  v  a  r  g  y  h  n  r  p  d  l  t  a  d  r  f  p  a  n  p  f  q  t  e  q  e  r  l
GAGGGCCGAAATGCGCGTCTGTATAAGACTGGTGACTTGGTTCGCTGGATCCACAATGCAAACGGCGATGGTGAGATCGAGTACCTCGGCCGCAACGACT   5700
e  g  r  n  a  r  l  y  k  t  g  d  l  v  r  w  i  h  n  a  n  g  d  g  e  i  e  y  l  g  r  n  d  f
TCCAGGTCAAGATTCGAGGCCAGAGAATCGAGCTGGGAGAGATCGAGGCCGTGCTTTCATCCTATCCGGGCATCAAACAATCCGTCGTCCTGGCCAAGGA   5800
 q  v  k  i  r  g  q  r  i  e  l  g  e  i  e  a  v  l  s  s  y  p  g  i  k  q  s  v  v  l  a  k  d
CCGCAAGAATGACGGGCAGAAGTACCTCGTCGGCTACTTCGTCTCCTCAGCAGGGTCCCTGTCCGCCCAGGCCATCCGCCGCTTCATGCTCACGAGCCTG   5900
 r  k  n  d  g  q  k  y  l  v  g  y  f  v  s  s  a  g  s  l  s  a  q  a  i  r  r  f  m  l  t  s  l
CCCGATTACATGGTTCCTGCGCAGCTGGTGCCCATCGCCAAGTTCCCCGTCACCGTGAGCGGGAAGCTCGATGCCAAGGCCTTGCCCGTGCCAGACGATA   6000
p  d  y  m  v  p  a  q  l  v  p  i  a  k  f  p  v  t  v  s  g  k  l  d  a  k  a  l  p  v  p  d  d  t
CAGTCGAGGATGACATTGTGCCACCGCGTACCGAGGTTGAGCGCATCCTAGCTGGGATCTGGTCTGAGCTGTTGGAGATACCGGTCGACAGGATCAGCAT   6100
 v  e  d  d  i  v  p  p  r  t  e  v  e  r  i  l  a  g  i  w  s  e  l  l  e  i  p  v  d  r  i  s  i
CTACAGTGACTTCTTCAGTCTGGGCGGCGACAGTCTCAAGAGTACCAAGCTGTCCTTTGCTGCCACGCGGGCTCTCGGTGTGGCCGTCAGTGTCCGCAAC   6200
 y  s  d  f  f  s  l  g  g  d  s  l  k  s  t  k  l  s  f  a  a  t  r  a  l  g  v  a  v  s  v  r  n
TTGTTCAGCCATCCGACTATCGAAGCCTTGTCTCAGTGGATTATCAGGGGTTCGAACGAGGTCAAGGATGTGGCTGTGGTGAAGGGCGGTGCCAGTCTTG   6300
l  f  s  h  p  t  i  e  a  l  s  q  w  i  i  r  g  s  n  e  v  k  d  v  a  v  v  k  g  g  a  s  l  d
ATATCCCCCTATCCCCTGCCCAGGAAAGACTCATGTTCATCCACGAGTTCGGCCATAGCGGCGAGGATACTGGTGCTTACAATGTGCCTTTGCAGCTGCA   6400
 i  p  l  s  p  a  q  e  r  l  m  f  i  h  e  f  g  h  s  g  e  d  t  g  a  y  n  v  p  l  q  l  q
GCTTCACCATGATGTCTGTCTCGAGTCGCTTGAGAAGGCTCTGCGGGATGTCGTCTCGAGACACGAGGCTCTCCGGACCTTGATCACCAGGACCCAGAAG   6500
 l  h  h  d  v  c  l  e  s  l  e  k  a  l  r  d  v  v  s  r  h  e  a  l  r  t  l  i  t  r  t  q  k
TCCTCCGTGCACTGCCAGAAGATCCTCGACGCCGAAGAAGCGCAAAAGCTCTTCTCTGTTGATGTTCTGCGCCTGACCTCGGAGACGGAGATGCAGGGCA   6600
s  s  v  h  c  q  k  i  l  d  a  e  e  a  q  k  l  f  s  v  d  v  l  r  l  t  s  e  t  e  m  q  g  r
GGATGGCCGAGAGTACCGCCCACGCCTTCAAGCTCGACGAGGAACTCCCGATTCATGTACGCCTGTACCAGGTTGTACGTGATGGCCGCACGCTCAGCTT   6700
 m  a  e  s  t  a  h  a  f  k  l  d  e  e  l  p  i  h  v  r  l  y  q  v  v  r  d  g  r  t  l  s  f
TGCCAGCATCGTCTGCCACCATCTGGCGTTTGACGCGTGGTCATGGGATGTGTTCCAGAGGGACTTGGACGCCTTCTATGCCGTCCATACGAAGCACAAG   6800
 a  s  i  v  c  h  h  l  a  f  d  a  w  s  w  d  v  f  q  r  d  l  d  a  f  y  a  v  h  t  k  h  k
GCTGCCGCCAACCTGCCAACCCTCCGCGTGCAATATAAGGAGTATGCGATAGAGCACCGCCGGGCTCTCCGCGCTGAGCAACACCGTGTTCTCGCGGACT   6900
a  a  a  n  l  p  t  l  r  v  q  y  k  e  y  a  i  e  h  r  r  a  l  r  a  e  q  h  r  v  l  a  d  y
ACTGGCTGCGCAAGCTCAGTGACATGGAGGCGTCTTATCTGGTCCCCGATCGCCCTCGACCGGCGCAGTTTGACTATACCGGGAACGATCTCCAGTTCTC   7000
 w  l  r  k  l  s  d  m  e  a  s  y  l  v  p  d  r  p  r  p  a  q  f  d  y  t  g  n  d  l  q  f  s
AACTACTCCCGAGACCACCGCGCAGTTGAAGGAGCTGGCCAAGCGCGAGGGTTCAAGCCTCTACACCGTTGTGGCGGCGGCGTACTTTCTGCTTCTCTAC   7100
 t  t  p  e  t  t  a  q  l  k  e  l  a  k  r  e  g  s  s  l  y  t  v  v  a  a  a  y  f  l  l  l  y
GTGTACACCAACCAGCGGGATATCACGATTGGTATTCCCGTTGCGCACCGTAACCATCCGGACTTTGAGTCGGTTGTCGGCTTCTTTGTCAACTTGCTCC   7200
v  y  t  n  q  r  d  i  t  i  g  i  p  v  a  h  r  n  h  p  d  f  e  s  v  v  g  f  f  v  n  l  l  p
CTCTGCGGGTCAACGTGTCTCAGTCGGACATTCATGGACTTATCCAGGCAGTGCAGAAAGAGCTTGTCGATGCCCAGATCCATCAGGACTTGCCATTCCA   7300
 l  r  v  n  v  s  q  s  d  i  h  g  l  i  q  a  v  q  k  e  l  v  d  a  q  i  h  q  d  l  p  f  q
GGAGATCACCAAGCTTCTTCATGTGCAGCACGATCCAAGCCGCCATCCCCTTCTCCAGGCCGTGTTCAACTGGGAAAACGTACCCGCCAATGTCCACGAG   7400
 e  i  t  k  l  l  h  v  q  h  d  p  s  r  h  p  l  l  q  a  v  f  n  w  e  n  v  p  a  n  v  h  e
GAGCAGCTGCTTCAGGAGTACAAGCCGCCCTCGCCTCTGCCTTCGGCGGCCAAGTTTGATCTCAACGTCACGGTGAAAGAGAGCGTCAATTCGCTCAACG   7500
e  q  l  l  q  e  y  k  p  p  s  p  l  p  s  a  a  k  f  d  l  n  v  t  v  k  e  s  v  n  s  l  n  v
TCAACTTCAACTATCCTACCAGCCTCTTCGAGGAGGAGACCGTTCAGGGGTTCATGGAAACCTTCCATCTCCTTCTTCGACAACTGGCCCACAACAAGGC   7600
 n  f  n  y  p  t  s  l  f  e  e  e  t  v  q  g  f  m  e  t  f  h  l  l  l  r  q  l  a  h  n  k  a
TAGCACAAGCCTCTCGAAGCTGTCGGTTGAAGATGGAGTGTTGAATCCAGAGCCGACTAACCTTCAGCCCTCAAGCCGGGACAGCGGAAATTCACTCCAT   7700
 s  t  s  l  s  k  l  s  v  e  d  g  v  l  n  p  e  p  t  n  l  q  p  s  s  r  d  s  g  n  s  l  h
GGGCTCTTCGAGGACATCGTGGCCTCGACCCCGGACCGCATCGCAATTGCTGACGGCACCAGGAGTCTCTCGTACTCCGAACTCAACGAGCGGGCAAACC   7800
g  l  f  e  d  i  v  a  s  t  p  d  r  i  a  i  a  d  g  t  r  s  l  s  y  s  e  l  n  e  r  a  n  q
AGCTGGTACATTTGATCATCTCTTCTGCCAGTATTGTAGCAGACGACCGCATCGCTCTTCTTTTGGACAAGAGCATCGATATGGTGATTGCTCTCCTGGC   7900
 l  v  h  l  i  i  s  s  a  s  i  v  a  d  d  r  i  a  l  l  l  d  k  s  i  d  m  v  i  a  l  l  a
```

FIG. 6—*Continued.*

residues 2409 to 3387. Computer alignment of these three regions with each other and with the three domains found in the ACV synthetase of *P. chrysogenum* was very good (see Discussion). However, there was no significant homology in the three interdomain regions of the protein.

The three domains of the ACV synthetase of *C. acremonium* showed extensive amino acid homology with TY1 and GS1 of *B. brevis*, encoded by the genes *tycA* and *grsA*, respectively (17, 25), but they did not show significant homology with the τ-glutamylcysteine synthetase or glutathione synthetase of *E. coli* (13, 39).

Thioesterase domains in the ACV synthetases of *C. acremonium* and *P. chrysogenum*. An interesting result was the finding of a thioesterase domain in the amino acid sequence of the carboxyl-terminal regions of the ACV synthetases of *P. chrysogenum* (previously unreported) and *C. acremonium* by comparison with the thioesterase domain of the fatty acid synthetases of rat, chicken, rabbit, and goose. The

```
AGTTTGGAAGGCCGGTGCCGCATATGTGCCCCTTGACCCGACATATCCGTCGCAGAGGACTGAGCTCATCTTGGAGGAATCTAGTGCCAGGACGCTCATC   8000
 v  w  k  a  g  a  a  y  v  p  l  d  p  t  y  p  s  q  r  t  e  l  i  l  e  e  s  s  a  r  t  l  i
ACCACTAGAAAGCACACGCCGAGGGGAGGAACAGTCGCAAATGTTCCAGCCGTGGTCCTTGACAGCCCCGAGACCCTAGCCTGCCTCAACCAGCAGTCAA   8100
 t  t  r  k  h  t  p  r  g  g  t  v  a  n  v  p  s  v  v  l  d  s  p  e  t  l  a  c  l  n  q  q  s  k
AGGAAAACCCGACAACGTCAACGCAGAAACCGTCCGACCTCGCATATGTCATCTTCACCTCGGGAACCACAGGCAAGCCCAAGGGGGTTCTGGTGGAGCA   8200
  e  n  p  t  t  s  t  q  k  p  s  d  l  a  y  v  i  f  t  s  g  t  t  g  k  p  k  g  v  l  v  e  h
CCAGAGCGTAGTCCAGCTGCGCAATTCCCTCATCGAGCGATACTTCGGCGAGACCAACGGGTCTCACGCCGTGCTCTTCCTGTCCAACTACGTCTTCGAC   8300
 q  s  v  v  q  l  r  n  s  l  i  e  r  y  f  g  e  t  n  g  s  h  a  v  l  f  l  s  n  y  v  f  d
TTCTCTCTTGAACAGCTCTGTCTCTCAGTCTTGGGTGGAAACAAGCTCATCATTCCACCAGAGGAGGGTCTCACGCACGAGGCATTCTACGACATCGGCC   8400
 f  s  l  e  q  l  c  l  s  v  l  g  g  n  k  l  i  i  p  p  e  e  g  l  t  h  e  a  f  y  d  i  g  r
GCAGGGAGAAGCTATCCTATCTCAGCGGGACGCCCTCGGTGCTGCAGCAGATTGAGCTCTCCCGTCTGCCGCATCTTCACATGGTCACCGCTGCGGGCGA   8500
  r  e  k  l  s  y  l  s  g  t  p  s  v  l  q  q  i  e  l  s  r  l  p  h  l  h  m  v  t  a  a  g  e
GGAGTTCCACGCTAGTCAGTTTGAGAAGATGCGCTCCCAGTTCGCGGGCCAGATCAACAACGCCTATGGTATCACTGAGACGACCGTGTACAACATCATC   8600
 e  f  h  a  s  q  f  e  k  m  r  s  q  f  a  g  q  i  n  n  a  y  g  i  t  e  t  t  v  y  n  i  i
ACCACGTTCAAGGGCGATGCCCCCTTTACCAAGGCACTCTGCCACGGGATCCCCGGAAGTCACGTCTACGTCCTGAACGACCGACTTCAGCGTGTTCCTT   8700
 t  t  f  k  g  d  a  p  f  t  k  a  l  c  h  g  i  p  g  s  h  v  y  v  l  n  d  r  l  q  r  v  p  f
TCAACGCTGTTGGCGAGCTCTACTTGGGCGGTGACTGCCTTGCTCGCGGGTACCTCAACCAGGATGCCCTGACCAACGAGCGATTCATCCCCAACCCTTT   8800
  n  a  v  g  e  l  y  l  g  g  d  c  l  a  r  g  y  l  n  q  d  a  l  t  n  e  r  f  i  p  n  p  f
CTACGAGCCGAAACAGGCAAGTGACAGTCGTCCCCAGAGACTCTACAAGACTGGAGATCTGGTGCGCTTCCGTGGACCCCACCATCTCGAGTATCTCGGC   8900
 y  e  p  k  q  a  s  d  s  r  p  q  r  l  y  k  t  g  d  l  v  r  f  r  g  p  h  h  l  e  y  l  g
CGCAAGGACCAGCAGGTCAAGCTGAGGGGCTTCCGCATCGAGCTCTCCGAGGTGCGGGATGCCGTCCTAGCCATCTCTGCTGTTAAGGAGGCTGCCGTCA   9000
 r  k  d  q  q  v  k  l  r  g  f  r  i  e  l  s  e  v  r  d  a  v  l  a  i  s  a  v  k  e  a  a  v  i
TCCCCAAGTATGACGAGGATGGCTCCGATTCACGAAGGGTCAGCGCCATCGTCTGCTACTACACGCTCAACGCCGGAACTGTGTGCGAAGCATCGAGTAT   9100
  p  k  y  d  e  d  g  s  d  s  r  r  v  s  a  i  v  c  y  y  t  l  n  a  g  t  v  c  e  a  s  s  i
CCGTGACCACCTGCACGCCAACCTTCCCCCGTACATGGTCCCAAGTCAGATCCACCAGTTGGAGGGATCTCTCCCCGTGACCGTCAATGGGAAGCTCGAC   9200
 r  d  h  l  h  a  n  l  p  p  y  m  v  p  s  q  i  h  q  l  e  g  s  l  p  v  t  v  n  g  k  l  d
CTGAACAGGCTCTCCACAACTCAAGTCTCGCAGCCAGAGCTTTACACCGCTCCACGAAATTCGACAGAGGAAACCTTGTGCCAGCTTTGGGCATCTCTCC   9300
 l  n  r  l  s  t  t  q  v  s  q  p  e  l  y  t  a  p  r  n  s  t  e  e  t  l  c  q  l  w  a  s  l  l
TAGGCGTCGACCACTGCGGCATTGACGACGACCTGTTTGCCCGAGGCGGCGACAGCATCTCCTCTCTCCGACTAGTGGGTGACATCTACCGCGCGCTAGG   9400
  g  v  d  h  c  g  i  d  d  d  l  f  a  r  g  g  d  s  i  s  s  l  r  l  v  g  d  i  y  r  a  l  g
ACGCAAGGTCACCGTCAAGGACATCTACCTCCACCGCAGCGTCCGAGCCCTAAGCGAAAATGTCCTGACCGACCAGAAGGATAAGGGTACTCTGCCAGCG   9500
 r  k  v  t  v  k  d  i  y  l  h  r  s  v  r  a  l  s  e  n  v  l  t  d  q  k  d  k  g  t  l  p  a
TCTCCTCCCCTCCAGCGAGCGGAGCAGGGCCAGGTTGAGGGCGACGCACCGCTTCTCCCCATCCAGGACTGGTTCCTTTCCAAGCCCCTGGATAACCCCG   9600
 s  p  p  l  q  r  a  e  q  g  q  v  e  g  d  a  p  l  l  p  i  q  d  w  f  l  s  k  p  l  d  n  p  a
CTTACTGGAACCACTGCTTCACCATTCGAACCGGGGCACTCTCCGTCGAAGGGCTCCGGGGTGCTCTGAAGCTGCTGCAGGAGCGCCACGACGTGCTGCG   9700
  y  w  n  h  c  f  t  i  r  t  g  a  l  s  v  e  g  l  r  g  a  l  k  l  l  q  e  r  h  d  v  l  r
TCTGAGACTGCAACGCCGGGACGAAGGTCGCCATGTTCAGACCTTTGCGCGTGACTGCGCGCAACCTCGCTTGACTGTGCTAGACCGACGAAGCTTCGAG   9800
 l  r  l  q  r  r  d  e  g  r  h  v  q  t  f  a  r  d  c  a  q  p  r  l  t  v  l  d  r  r  s  f  e
GACGCAGAGGATGTACAGGAGGCTCTCTGCGAGATCCAATCTCATTTCGACCTCGAGAATGGACCCCTCTACACAGTGGCGTACATCCACGGTTACGAGG   9900
 d  a  e  d  v  q  e  a  l  c  e  i  q  s  h  f  d  l  e  n  g  p  l  y  t  v  a  y  i  h  g  y  e  d
ACGGCTCCGCCCGAGTGTGGTTTGCCTGCCATCACGTCATGGTCGACACTGTGAGCTGGAACATTATACTGCAAGACCTGCAGGCTCTCTATCATGGAGA  10000
  g  s  a  r  v  w  f  a  c  h  h  v  m  v  d  t  v  s  w  n  i  i  l  q  d  l  q  a  l  y  h  g  d
CAGCCTTGGTCCCAAGAGCAGCAGCGTGCAGCAGTGGTCGCTAGCTGTCAGCGACTACAAAATGCCACTGTCGGAGAGGGCGCATTGGAATGTGCTCAGG  10100
  s  l  g  p  k  s  s  s  v  q  q  w  s  l  a  v  s  d  y  k  m  p  l  s  e  r  a  h  w  n  v  l  r
AAGACAGTCGCCCAGAGCTTCGAGACCCTGCCTATCTGCATGGGCGGCGTGCTCCAGTGCCAGGAGAAGTTCTCGAGGGAAACGACAACAGCTCTGCTCT  10200
 k  t  v  a  q  s  f  e  t  l  p  i  c  m  g  g  v  l  q  c  q  e  k  f  s  r  e  t  t  t  a  l  l  s
CCAAGGCCTGCCCTGCCTTGGACTCCGGTATGCATGAGATCCTTCTCATGGCCGTGGGCTCCGCGCTGCAGAAGGCGGCAGGGGATGTCCCTCAGGTCGT  10300
  k  a  c  p  a  l  d  s  g  m  h  e  i  l  l  m  a  v  g  s  a  l  q  k  a  a  g  d  v  p  q  v  v
CACGATAGAGGGTCACGGGCGCGAAGATACTATCGACGCAACTCTGGACGTCAGCCGGACAGTCGGCTGGTTCACGAGCATGTACCCCTTCGAGATCCCC  10400
 t  i  e  g  h  g  r  e  d  t  i  d  a  t  l  d  v  s  r  t  v  g  w  f  t  s  m  y  p  f  e  i  p
AAAGTGACCGACCCCGCTCAGGGCGTCGTCGATGTCAAGGAGGCGATGCGTCGCGTGCCGAATAGGGGTGTCGGTTACGGTCCAGCCTACGGATACGGCG  10500
 k  v  t  d  p  a  q  g  v  v  d  v  k  e  a  m  r  r  v  p  n  r  g  v  g  y  g  p  a  y  g  y  g  g
GATCGTGCCTGCCCGCGGTGAGCTTCAACTACCTTGGTCGCCTGGACCAGGCTTCCTCGGGGGCTCAAAGGGACTGGACGCTGGTCATGGATGAAGACGA  10600
  s  c  l  p  a  v  s  f  n  y  l  g  r  l  d  q  a  s  s  g  a  q  r  d  w  t  l  v  m  d  e  d  e
GTATCCGGTCGGACTGTGCACCAGCGCTGAGGACTCGGGACGAAGCTCCTCCATGGTGGATTTCACCTTCTCTATCTCTGGCGGCCAGCTTGTCATGGAT  10700
 y  p  v  g  l  c  t  s  a  e  d  s  g  r  s  s  s  m  v  d  f  t  f  s  i  s  g  g  q  l  v  m  d
ATGAGTAGCAGCTGGGGCCACGGCGCACGAAATGAATTCGTTCGCACAGTTCGTAACACACTAGATGACTTGATCAAAACAACGAGCAGCAGGGACTTCA  10800
 m  s  s  s  w  g  h  g  a  r  n  e  f  v  r  t  v  r  n  t  l  d  d  l  i  k  t  t  s  s  r  d  f  s
GCGCACCTCTGCCTCCGTCGGATCAGGAGTCCAGCTTCACCCCTTATTTTGTCTTCGAAGAGGGCGAGCGACACGGCGCTCCGCTCTTCCTGCTCCCACC  10900
  a  p  l  p  p  s  d  q  e  s  s  f  t  p  y  f  v  f  e  e  g  e  r  h  g  a  p  l  f  l  l  p  p
TGGCGAAGGCGGAGCGGAGAGCTACTTCCACAACATTGTCAAGGGTCTCCCGAACCGCAATCTTGTCGTGTTCAACAATCATTACCGCGAGGAGAAGACG  11000
  g  e  g  g  a  e  s  y  f  h  n  i  v  k  g  l  p  n  r  n  l  v  v  f  n  n  h  y  r  e  e  k  t
CTCCGGACCATCGAGGCGCTGGCCGAGTACTACCTGTCGCACATCCGATCCATCCAGCCGGAGGGGCCATACCACATCCTCGGCTGGAGTTTCGGAGGCA  11100
 l  r  t  i  e  a  l  a  e  y  y  l  s  h  i  r  s  i  q  p  e  g  p  y  h  i  l  g  w  s  f  g  g  i
TCCTCGGTCTCGAGGCGGCAAAGCGATTGACTGGCGAGGGTCACAAGATTGCCACGCTGGCACTTATCGATCCGTACTTTGACATCCCGTCCGCGTCCAA  11200
  l  g  l  e  a  a  k  r  l  t  g  e  g  h  k  i  a  t  l  a  l  i  d  p  y  f  d  i  p  s  a  s  k
GGCCATCGGCCAACCTGACGATGCCTGCGTCTTGGACCCCATATACCACGTCTACCACCCGTCGCCGGAGAGCTTCAGGACGGTGTCATCTCTCACTAAT  11300
  a  i  g  q  p  d  d  a  c  v  l  d  p  i  y  h  v  y  h  p  s  p  e  s  f  r  t  v  s  s  l  t  n
CACATAGCCCTGTTCAAGGCTACCGAGACGAATGACCAGCATGGCAATGCCACGCAGCAGGCCCTGTATGAGTGGTTTGCCACGTGCCCTTTGAACAACC  11400
 h  i  a  l  f  k  a  t  e  t  n  d  q  h  g  n  a  t  q  q  a  l  y  e  w  f  a  t  c  p  l  n  n  l
TGGACAAGTTTTTGGCGGCCGACACGATCAAGGTGGTTCCTCTGGAGGGTACACATTTTACCTGGGTGCACCACCCGGAGCAGGTGCGCTCAATGTGCAC  11500
  d  k  f  l  a  a  d  t  i  k  v  v  p  l  e  g  t  h  f  t  w  v  h  h  p  e  q  v  r  s  m  c  t
TATGCTGGATGAATGGCTTGGG
  m  l  d  e  w  l  g
                         TGAACGAGGCAGTTGCTGTGAGAGAATGAGAATGAGACACAAAACGCGGGCGGAAGAGAGACTTCCTCGGACGGCGGG  11600
```

FIG. 6—*Continued.*

amino acid sequence GWSFGGIL is fully conserved in *P. chrysogenum* and *C. acremonium* ACV synthetases (Fig. 9) and is in agreement with the GXSXG consensus motif, (where X is any amino acid) in the thioesterase domains of fatty acid synthetases (see Discussion).

Phosphopantetheine-binding amino acid sequences. A search of the consensus phosphopantetheine-binding sequence described for acyl-carrier proteins (DLGXDS*LXX VEV/I) revealed the presence of three partially conserved sequences at amino acid positions 820 to 828, 1909 to 1917, and 2983 to 2991 (boxed in Fig. 7), one in each of the three repeated domains of the ACV synthetase (see Discussion). All of them contain a serine residue; the hydroxyl group of the serine residue labeled with an asterisk is known to be covalently linked to the 4′-phosphopantetheine residue.

Linkage of *pcbAB* and *pcbC* genes. To confirm the evidence obtained by hybridization, the region extending from the *pcbAB* to the *pcbC* gene was sequenced. The intergenic region consisted of 1,233 bp, and there were no long ORFs in it. We did not find any transcript originating from the intergenic region. A total of 400 bp of the *pcbC* gene from strain C10 was also sequenced, and the nucleotide sequence

A.chrysogenum ACV Synthetase (1,3712)
P.chrysogenum ACV Synthetase (62,3781)

FIG. 7. Comparison of the amino acid sequences of the ACV synthetases of *C. acremonium* and *P. chrysogenum*. Beginning and ending amino acid numbers are shown separated by a comma in parentheses. Identical amino acids are shaded, and functionally equivalent amino acids are indicated by dots. The positions of the three phosphopantetheine-binding sequences are boxed.

(not shown) was identical to that described for a different *C. acremonium* strain by Samson and coworkers (32). The homology in the intergenic region between the nucleotide sequences of *C. acremonium* and *P. chrysogenum* (9) was very low compared with the homology in the regions corresponding to the ORFs of the *pcbAB* and *pcbC* genes. Sequencing studies proved unequivocally that *pcbAB* and *pcbC* were linked, forming part of the cluster of early cephalosporin-biosynthetic genes.

***pcbAB* gene of *C. acremonium* complements an ACV synthetase-deficient mutant of *P. chrysogenum*.** Protoplasts of *P. chrysogenum* mutant *npe5* were transformed with pULSC1,

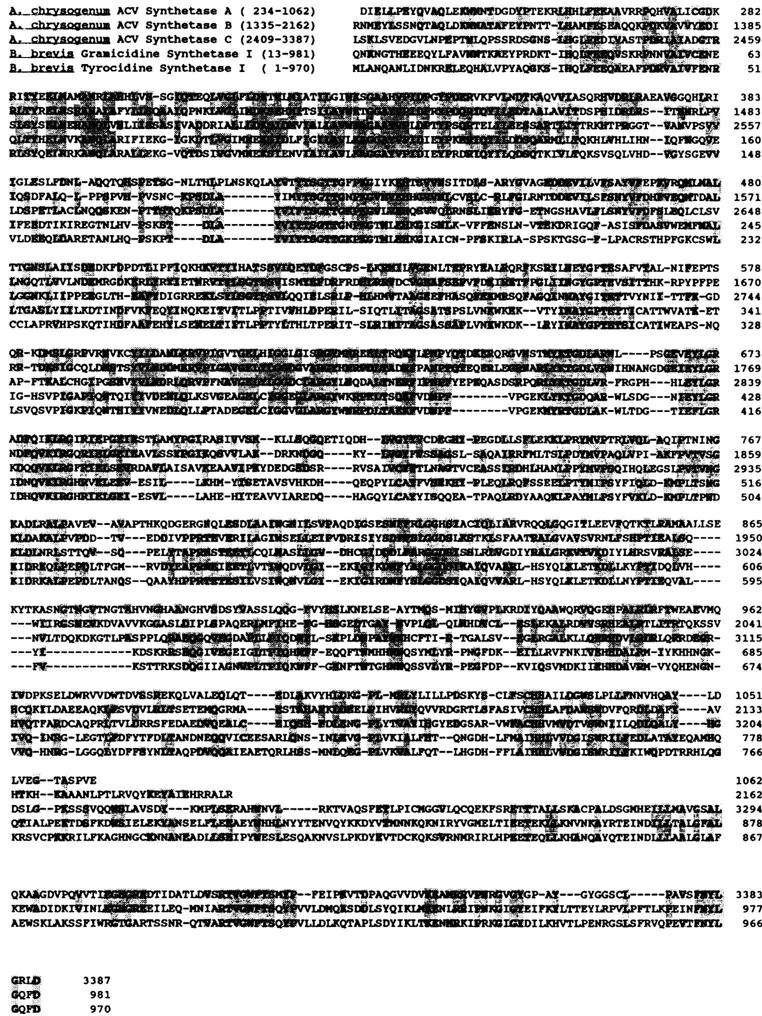


FIG. 8. Comparison of the three repeated domains of the *C. acremonium* ACV synthetase with each other and with the amino acid sequences of TY1 and GS1 of *B. brevis*. Gaps (dashes) have been introduced for maximal alignment. Conserved amino acid sequences are shaded.

which carries the *pcbAB* and *pcbC* genes, as indicated in Materials and Methods. Three clones out of 38 phleomycin-resistant transformants tested gave levels of penicillin production comparable to those of *P. chrysogenum* Wis 54-1255, the parental strain from which mutant *npe5* was obtained. No complementation of the mutation was observed when the protoplasts were transformed with the control vector pULJL43 without inserts. This result indicates that a func-

C.acremonium ACV synthetase (3518,3590)
P.chrysogenum ACV synthetase (3572,3653)
RAT FAS Thioesterase (81,151)
Chick FAS Thioesterase (95,165)

Goose FAS Thioesterase
Chick FASThioesterase
RAT FAS Thioesterase
RAT FAS Thioesterase II
P.chrysogenum ACV synthetase
C.acremonium ACV synthetase

FIG. 9. Thioesterase domains of the ACV synthetases of *C. acremonium* and *P. chrysogenum* and several fatty acid synthetases (FAS). The conserved region (light shading) around the consensus motif G-X-S-X-G (heavy shading) is enlarged. Beginning and ending amino acid numbers are shown separated by a comma in parentheses. The reactive serine residue in the active centers is indicated by an asterisk. Dashes indicate gaps introduced to maximize alignment.

tional ACV synthetase gene was located in the 15,603-bp *Eco*RI-*Bam*HI fragment of *C. acremonium* DNA subcloned in plasmid pULSC1.

The antibiotic produced by the pULSC1-transformed clones was identified as penicillin G by its antimicrobial spectrum and high-pressure liquid chromatography retention time, which agrees with the fact that only the ACV synthetase and isopenicillin N synthase genes of *C. acremonium* (but not other late genes of the cephalosporin-biosynthetic pathway) were located in the fragment used for transformation.

DISCUSSION

The formation of ACV is similar in many aspects to the synthesis of other microbial peptides which are synthesized by the nonribosomal enzyme-thiotemplate mechanism (16). Peptide synthetases require a complex spatial organization to direct the activation reactions and sequential polymerization of the component amino acids. This requires either a very high molecular weight multienzyme polypeptide or a set of small polypeptides each carrying a specific function assembled together in an enzyme complex.

The purified ACV synthetases of *Aspergillus nidulans* (38), *C. acremonium* (1, 8a), and *P. chrysogenum* (12a) have very high molecular masses, in the range of 300 to 400 kDa. The molecular mass of the ACV synthetase of *A. nidulans* has been reported as 220 kDa (38), but it is unclear whether the 220-kDa protein arises by posttranslational processing of a larger polypeptide. The purified ACV synthetases of *A. nidulans* and *C. acremonium* carry out the condensation of the three precursor amino acids, including activation of L-α-aminoadipic acid, L-cysteine, and L-valine, racemization of L-valine to D-valine, and the polymerization steps to form the tripeptide (4, 38). Although there is still no firm evidence indicating that a single polypeptide carries all the reactions involved, the finding of (i) a long ORF (11.1 kb) in the upstream region of the *pcbC* gene which has high homology with the *pcbAB* gene of *P. chrysogenum* (which is known to encode a functional ACV synthetase), (ii) an 11.4-kb transcript, and (iii) the presence of four internal peptides obtained by tryptic digestion in different positions of the deduced amino acid sequence strongly argue in favor of the synthesis of a large polypeptide with all the functions required for the synthesis of the LLD-ACV tripeptide. The deduced M_r of the *C. acremonium* ACV synthetase protein, 414,791, is similar to the M_r of the corresponding polypeptide of *P. chrysogenum* AS-P-78, 425,971 (10). The homology between these proteins is high; 54.9% of the amino acids are identical, and the figure is higher when functionally equivalent amino acids are considered.

The *C. acremonium* ACV synthetase contains three repeated regions with conserved amino acid sequences. These regions show high similarity in the amino acid sequence with the three equivalent domains of the ACV synthetase of *P. chrysogenum* and with *B. brevis* TY1 and GS1. The conserved amino acid sequences probably represent centers involved in ATP-mediated activation of amino acids, since both TY1 and GS1 are involved in activation and racemization of the amino acid phenylalanine (17, 25). The presence of three amino-acid-activating domains in the ACV synthetase is consistent with a similar organization that exists in the heavy tyrocidine synthetases II (230 kDa) and III (460 kDa), which activate three and six amino acids, respectively, during tyrocidine biosynthesis.

We have found a thioesterase domain in the ACV synthetases of *C. acremonium* and *P. chrysogenum* which has a conserved sequence of amino acids similar to that of the thioesterase domains of the fatty acid synthetases of rat (27), chicken (40), rabbit (14), and goose (28). The subunits of fatty acid synthetases from *E. coli* and most other procaryotes can be dissociated, whereas the thioesterase domain in the ACV synthetases of *C. acremonium* and *P. chrysogenum* appears to be part of a single long polypeptide. This arrangement resembles the organization of the fatty acid synthetases in vertebrates, in which all the components are combined within a single octafunctional polypeptide chain. In *Saccharomyces cerevisiae* and other lower fungi, the fatty acid synthetase consists of tri- and pentafunctional, respectively, subunits α and β (35).

The similarity between the thioesterase domains of ACV synthetases and fatty acid synthetases is greater in the region around the consensus motif G-X-S*-X-G found in several serine active-site (marked by the asterisk) esterases (8). The serine has been identified by labeling experiments as the reactive residue of thioesterase (31).

The homology between polypeptide synthesis with the thiotemplate mechanism and polyketide and fatty acid synthesis is very interesting from the evolutionary point of view. A certain homology between these processes was observed by Lipmann (18). Polypeptide synthesis by the nonribosomal thiotemplate mechanism is probably a very primitive mechanism of protein synthesis that preceded the sophisticated mechanism of ribosomal protein formation. Activation and polymerization of the amino acids as thioesters in the thiotemplate mechanism strongly resemble the mechanisms of polyketide biosynthesis and fatty acid synthesis (21, 22). The finding of a thioesterase domain and three sequences similar to the phosphopantetheine-binding consensus sequence of the acyl-carrier protein provides support for a common ancestral gene from which polypeptide synthetases and polyketide synthetases (including fatty acid synthetases) derived. The relevance of the three phosphopantetheine-binding sequences in amino acid polymerization is unclear, since biochemical data indicate that about 1 mol of pantethenic acid could be liberated per mol of ACV synthetase (1).

The *pcbAB* gene is clustered with the *pcbC* gene; both are early genes of the cephalosporin-biosynthetic pathway. *cefEF* (and probably other late genes of the pathway) was located in chromosome II, whereas the *pcbC* gene was found in chromosome VI (37). It seems that most genes of the cephalosporin-biosynthetic pathway are clustered into two groups. *pcbAB* and *pcbC* belong to the cluster of early biosynthetic genes, and *cefEF* is located in the second (or late) cluster of genes. We have found that the three genes encoding the entire penicillin-biosynthetic pathway are linked in a single cluster in *P. chrysogenum* (9, 10) and *A. nidulans* (19, 26). The linkage of antibiotic-biosynthetic genes is a well-known phenomenon in many antibiotic-producing organisms (24) and suggests that the linkage has occurred during evolution owing to an ecological selective advantage conferred by coordinated expression of the clustered genes.

ACKNOWLEDGMENTS

This work was supported in part by a grant from Gist-Brocades, Delft, Holland. B. Díez received a fellowship from the Diputación León, Spain, and E. Montenegro was supported by a fellowship from the Ministry of Education and Science, Madrid.

We acknowledge the excellent technical assistance of M. I.

Corrales, S. Llamas, and R. Barrientos. We thank C. Esmahan for critical reading of the manuscript.

REFERENCES

1. **Baldwin, J. E., J. W. Bird, R. A. Field, N. M. O'Callaghan, and C. J. Schofield.** 1990. Isolation and partial characterisation of ACV synthetase from *Cephalosporium acremonium* and *Streptomyces clavuligerus*. J. Antibiot. **43**:1055–1057.

1a. **Baldwin, J. E., J. W. Bird, R. A. Field, N. M. O'Callaghan, C. J. Schofield, and A. C. Willis.** 1990. Symposium, "Fifty Years of Penicillin Application." Berlin, Federal Republic of Germany.

2. **Ballance, D. J.** 1986. Sequences important for gene expression in filamentous fungi. Yeast **2**:229–236.

3. **Banko, G., and A. L. Demain.** 1987. δ-(L-α-Aminoadipyl)-L-cysteinyl-D-valine synthetase (ACV synthetase): a multifunctional enzyme with broad substrate specificity for the synthesis of penicillin and cephalosporin precursors. J. Am. Chem. Soc. **109**:2858–2860.

4. **Banko, G., S. Wolfe, and A. L. Demain.** 1986. Cell-free synthesis of δ-(L-α-aminoadipyl)-L-cysteine, the first intermediate of penicillin and cephalosporin biosynthesis. Biochem. Biophys. Res. Commun. **137**:528–535.

5. **Barnes, W. M., M. Bevan, and P. H. Son.** 1983. Kilosequencing: creation of an ordered nest of asymmetric deletions across a large target sequence carried on phage M13. Methods Enzymol. **101**:98–122.

5a. **Barredo, J. L., et al.** Unpublished data.

6. **Barredo, J. L., J. M. Cantoral, E. Alvarez, B. Díez, and J. F. Martin.** 1989. Cloning, sequence analysis and transcriptional study of the isopenicillin N synthase of *Penicillium chrysogenum* AS-P-78. Mol. Gen. Genet. **216**:91–98.

6a. **Cantoral, J. M., B. Díez, J. L. Barredo, A. Alvarez, and J. F. Martin.** 1987. High-frequency transformation of *Penicillium chrysogenum*. Bio/Technology **5**:494–497.

7. **Carr, L. G., P. L. Skatrud, M. E. Scheetz III, S. W. Queener, and T. D. Ingolia.** 1986. Cloning and expression of the isopenicillin N synthetase gene from *Penicillium chrysogenum*. Gene **48**:257–266.

8. **Dayhoff, M. O., W. C. Barker, and J. K. Hardman.** 1976. Atlas of protein sequencing and structure, vol. 5, p. 53–66. National Biomedical Research Foundation, Washington, D.C.

8a. **Demain, A. L.** Personal communication.

9. **Díez, B., J. L. Barredo, E. Alvarez, J. M. Cantoral, P. van Solingen, M. A. M. Groenen, A. E. Veenstra, and J. F. Martín.** 1989. Two genes involved in penicillin biosynthesis are linked in a 5.1 kb *Sal*I fragment in the genome of *Penicillium chrysogenum*. Mol. Gen. Genet. **218**:572–576.

10. **Díez, B., S. Gutiérrez, J. L. Barredo, P. van Solingen, L. H. M. van der Voort, and J. F. Martin.** 1990. The cluster of penicillin biosynthetic genes. Identification and characterization of the *pcb*AB gene encoding the α-aminoadipyl-cysteinyl-valine synthetase and linkage to the *pcb*C and *pen*DE genes. J. Biol. Chem., **265**:16358–16365.

11. **Eriquez, L. A., and M. A. Pisano.** 1979. Isolation and nature of intracellular alpha-aminoadipic acid-containing peptides from *Paecilomyces persicinus* P-10. Antimicrob. Agents Chemother. **16**:392–397.

12. **Fawcett, P. A., J. J. Usher, J. A. Huddleston, R. C. Bleany, J. J. Nisbet, and E. P. Abraham.** 1976. Synthesis of δ-(α-aminoadipyl)cysteinyl valine and its role in penicillin biosynthesis. Biochem. J. **157**:651–660.

12a. **Gil, S., and J. F. Martín.** Unpublished data.

13. **Gushima, H., S. Yasuda, E. Soeda, M. Yokota, M. Kondo, and A. Kimura.** 1984. Complete nucleotide sequence of the *E. coli* glutathione synthetase *gsh*-II. Nucleic Acids Res. **12**:9299–9307.

13a. **Gutiérrez, S., et al.** Unpublished data.

13b. **Gutiérrez, S., and J. M. Cantoral.** Unpublished data.

14. **Hardie, D. G., K. B. Dewart, A. Aitken, and A. D. McCarthy.** 1985. Amino acid sequence around the reactive serine residue of the thioesterase domain of rabbit fatty acid synthase. Biochim. Biophys. Acta **828**:380–382.

15. **Jensen, S., D. W. S. Westlake, and S. Wolfe.** 1988. Production of the penicillin precursor δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine (ACV) by cell-free extracts from *Streptomyces clavuligerus*. FEMS Microbiol. Lett. **49**:213–218.

16. **Kleinkauf, H., and H. von Döhren.** 1987. Biosynthesis of peptide antibiotics. Annu. Rev. Microbiol. **41**:259–289.

17. **Krause, M., and M. A. Marahiel.** 1988. Organization of the biosynthesis genes for the peptide antibiotic gramicidin S. J. Bacteriol. **170**:4669–4674.

18. **Lipmann, F.** 1971. Attempts to map a process evolution of peptide biosynthesis. Science **173**:875–884.

19. **MacCabe, A. P., M. B. R. Riach, S. E. Unkles, and J. R. Kinghorn.** 1990. The *Aspergillus nidulans npe*A locus consists of three contiguous genes required for penicillin biosynthesis. EMBO J. **9**:279–287.

20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

21. **Martín, J. F.** 1976. Fatty acids and polyene macrolide biosynthesis by *Streptomycetes*. Dev. Ind. Microbiol. **17**:223–231.

22. **Martín, J. F.** 1977. Biosynthesis of polyene macrolide antibiotics. Annu. Rev. Microbiol. **31**:13–38.

23. **Martín, J. F., and P. Liras.** 1989. Enzymes involved in penicillin, cephalosporin and cephamycin biosynthesis, p. 153–187. *In* A. Fiechter (ed.), Advances in biochemical engineering/biotechnology, vol. 39. Springer-Verlag, Berlin.

24. **Martín, J. F., and P. Liras.** 1989. Organization and expression of genes involved in the biosynthesis of antibiotics and other secondary metabolites. Annu. Rev. Microbiol. **43**:173–206.

25. **Mittenhuber, G., R. Weckermann, and M. A. Marahiel.** 1989. Gene cluster containing the genes for tyrocidine synthetases 1 and 2 from *Bacillus brevis*: evidence for an operon. J. Bacteriol. **171**:4881–4887.

26. **Montenegro, E., J. L. Barredo, S. Gutiérrez, B. Díez, E. Alvarez, and J. F. Martín.** 1990. Cloning, characterization of the acyl-CoA:6-amino penicillanic acid acyltransferase gene of *Aspergillus nidulans* and linkage to the isopenicillin N synthase gene. Mol. Gen. Genet. **221**:322–330.

27. **Naggert, J., A. Witkowski, J. Mikkelsen, and S. Smith.** 1988. Molecular cloning and sequencing of a cDNA encoding the thioesterase domain of the rat fatty acid synthetase. J. Biol. Chem. **263**:1146–1150.

28. **Poulose, A. J., L. Rogers, and P. E. Kolattukudy.** 1981. Primary structure of a chymotryptic peptide containing the "active serine" of the thioesterase domain of fatty acid synthetase. Biochem. Biophys. Res. Commun. **103**:377–382.

29. **Ramos, F. R., M. J. López-Nieto, and J. F. Martín.** 1986. Coordinate increase of isopenicillin N synthetase, isopenicillin N epimerase and deacetoxycephalosporin C synthetase in a high cephalosporin-producing mutant of *Acremonium chrysogenum* and simultaneous loss of the three enzymes in a non-producing mutant. FEMS Microbiol. Lett. **35**:123–127.

30. **Ramsdem, M., B. A. McQuade, K. Saunders, M. K. Turner, and S. Harford.** 1989. Characterization of a loss-of-function mutation in the isopenicillin N synthetase gene of *Acremonium chrysogenum*. Gene **85**:267–273.

31. **Randhawa, Z. I., J. Naggert, R. W. Blacher, and S. Smith.** 1987. Amino acid sequence of the serine active-site region of the medium-chain S-acyl fatty acid synthetase thioester hydrolase from rat mammary gland. Eur. J. Biochem. **162**:577–581.

32. **Samson, S. M., R. Belagaje, D. T. Blankenship, J. L. Chapman, D. Perry, P. L. Skatrud, R. M. Vanfrank, 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.

33. **Samson, S. M., J. F. 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.

34. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

35. **Schweizer, E., K. Werkmeister, and M. K. Jain.** 1978. Fatty acid biosynthesis in yeasts. Mol. Cell. Biochem. **21**:95–106.
36. **Shirafuji, H., Y. Fujisawa, M. Kida, T. Kanzaki, and M. Yoneda.** 1979. Accumulation of tripeptide derivatives by mutants of *Cephalosporium acremonium*. Agric. Biol. Chem. **43**: 155–160.
37. **Skatrud, P. L., and S. W. Queener.** 1989. An electrophoretic molecular karyotype for an industrial strain of *Cephalosporium acremonium*. Gene **79**:331–338.
38. **Van Liempt, H., H. von Döhren, and H. Kleinkauf.** 1989. δ-(L-α-Aminoadipyl)-L-cysteinyl-D-valine synthetase from *Aspergillus nidulans*. J. Biol. Chem. **264**:3680–3684.
39. **Watanabe, K., Y. Yumano, K. Murata, and A. Kimura.** 1986. The nucleotide sequence of the gene for τ-glutamylcysteine synthetase of *Escherichia coli*. Nucleic Acids Res. **14**:4393–4400.
40. **Yuan, Z., W. Liu, and G. G. Hammes.** 1988. Molecular cloning and sequencing of DNA complementary to chicken liver fatty acid synthase mRNA. Proc. Natl. Acad. Sci. USA **85**:6328–6331.