Two Sets of Paralogous Genes Encode the Enzymes Involved in the Early Stages of Clavulanic Acid and Clavam Metabolite Biosynthesis in Streptomyces clavuligerus

Kapil Tahlan,* Hyeon Ung Park,† Annie Wong, Perrin H. Beatty, and Susan E. Jensen

Department of Biological Sciences, University of Alberta, Edmonton T6G 2E9, Canada

Received 2 June 2003/Returned for modification 30 September 2003/Accepted 21 October 2003

Recently, a second copy of a gene encoding proclavaminate amidinohydrolase (pah1), an enzyme involved in the early stages of clavulanic acid and clavam metabolite biosynthesis in Streptomyces clavuligerus, was identified and isolated. Using Southern analysis, we have now isolated second copies of the genes encoding the carboxyethylarginine synthase (cea5) and β-lactam synthetase (bls) enzymes. These new paralogues are given the gene designations ceaS1 and bls1 and are located immediately upstream of pah1 on the chromosome. Furthermore, sequence analysis of the region downstream of pah1 revealed a second copy of a gene encoding ornithine acetyltransferase (oat1), thus indicating the presence of a cluster of paralogue genes. ceaS1, bls1, and oat1 display 73, 60, and 63% identities, respectively, at the nucleotide level to the original ceaS2, bls2, and oat2 genes from the clavulanic acid gene cluster. Single mutants defective in ceaS1, bls1, or oat1 were prepared and characterized and were found to be affected to variable degrees in their ability to produce clavulanic acid and clavam metabolites. Double mutants defective in both copies of the genes were also prepared and tested. The ceaS1/ceaS2 and the bls1/bls2 mutant strains were completely blocked in clavulanic acid and clavam metabolite biosynthesis. On the other hand, oat1/oat2 double mutants still produced some clavulanic acid and clavam metabolites. This may be attributed to the presence of the argI gene in S. clavuligerus, which encodes yet another ornithine acetyltransferase enzyme that may be able to compensate for the lack of OAT1 and -2 in the double mutants.

The genus Streptomyces comprises gram-positive filamentous soil-dwelling organisms, which are renowned for their ability to produce a wide variety of chemically distinct antibiotics and secondary metabolites (25). When grown in soy medium, Streptomyces clavuligerus produces a variety of β-lactam compounds, including cephemycin C, penicillin N, clavulanic acid, and at least four other clavams (9, 17). S. clavuligerus is used commercially for the production of the clinically important β-lactamase inhibitor clavulanic acid on an industrial scale (21). Although clavulanic acid and the other clavams are structurally related to each other, only clavulanic acid is inhibitory to β-lactamases (3). The β-lactamase inhibitory activity of clavulanic acid has been attributed to its 3R, 5S stereochemistry, which differs from the 5S stereochemistry of all of the other known clavams (hereafter referred to as 5S clavams) produced by S. clavuligerus (21).

Although separate pathways produce cephemycin C and clavulanic acid, the two gene clusters involved in their biosynthesis are found grouped on the S. clavuligerus chromosome, forming a supercluster (1, 19, 45). Clavulanic acid and the 5S clavams arise by the condensation of L-arginine and glyceroldehyde-3-phosphate (24). This reaction is catalyzed by the enzyme carboxyethylarginine synthase and leads to the formation of N2-(2-carboxyethyl)arginine, the first dedicated inter-
clavamic acid biosynthesis, a gene (orf6) apparently encoding an ornithine acetyltransferase (OAT) is also present within the clavulamic acid gene cluster. Although its function in clavulamic acid and clavam metabolite biosynthesis is unclear, mutation of orf6 has been shown to decrease metabolite production (20). Recently, orf6 was shown to encode a protein with OAT activity (23). OATs are normally involved in arginine biosynthesis (12), and since arginine is a precursor of clavulamic acid and the clavams on soy medium (34). The continued production of some clavamins on soy medium even when cas2 is disrupted by insertional inactivation indicated that cas2 is located elsewhere on the chromosome, whereas cas1 is located adjacent to the clavulamic acid gene cluster (1, 19, 45), and since arginine is a precursor of clavulamic acid and the clavams, perhaps orf6, hereafter called oat, functions to increase the flux of arginine into the pathway.

*S. clavuligerus* possesses two CAS isozymes, encoded by two separate paralogous genes, cas1 and cas2 (4, 27). cas2 is located within the clavulamic acid gene cluster (1, 19, 45), whereas cas1 is located elsewhere on the chromosome, surrounded by genes that are involved in 5S clavam but not clavulamic acid biosynthesis (27, 32). The two cas paralogues are regulated differently, and transcriptional studies have shown that cas2 is transcribed in both complex soy and defined starch-asparagine (SA) medium, whereas cas1 is transcribed exclusively in soy medium (34). The continued production of some clavamic acid and 5S clavams on soy medium even when cas2 was disrupted by insertional inactivation indicated that cas1 could partially complement the cas2 mutation (34). Very similar phenotypes were observed when individual mutants defective in each of ceaS, bls, pah, and oat from the clavulamic acid gene cluster were prepared and analyzed, suggesting that paralogues may also exist for these genes (20). This was recently shown to be true for pah, since a parologue was isolated and characterized (22a). The pah paralogues were designated as pah1 and pah2, with pah2 located adjacent to cas2 in the clavulamic acid gene cluster. Preliminary studies indicated that pah1, on the other hand, is not located in the vicinity of cas1 (22a).

In the present study we report the isolation and characterization of paralogues for the ceaS, bls, and oat genes, all of which are located in the region of the *S. clavuligerus* chromosome flanking pah1. Mutants defective in each of the paralogous genes, and double mutants defective in both genes, were prepared by targeted gene disruption and tested for their abilities to produce clavulamic acid and 5S clavam metabolites.

**MATERIALS AND METHODS**

**Bacterial strains, plasmids, media, and culture conditions.** All bacterial strains and plasmids used in this study are described in Table 1. Cultures of *Escherichia coli* were grown in liquid culture in Luria broth (LB) and maintained on LB agar medium at 37°C (42). Plasmid-containing cultures were supplemented with ampicillin (100 μg/ml), apramycin (50 μg/ml), chloramphenicol (25 μg/ml), or kanamycin (50 μg/ml). *S. clavuligerus* was maintained either on MYM (43) or ISP4 medium agar plates (Difco, Detroit, Mich.) at 28°C. Plasmid-bearing Streptomyces cultures were supplemented with apramycin (25 μg/ml), kanamycin (50 μg/ml), or thioestrepton (5 μg/ml for *S. clavuligerus* and 50 μg/ml for *S. lividans*). Cultures for the isolation of chromosomal DNA were grown in Trypticase soy broth supplemented with 1% starch. Cultures for analysis of clavulamic acid and 5S clavam metabolite production were grown on both SA medium and on soy medium as described previously (34). All Streptomyces liquid cultures were grown at 28°C on a rotary shaker at 250 rpm.

**DNA isolation, manipulation, and Southern analysis.** Plasmid DNA isolation from *Escherichia coli* cultures, restriction endonuclease digestion, ligation, generation of blunt-ended fragments, and *E. coli* transformation were carried out by standard procedures (42). In all subsequent procedures, when DNA fragments with incompatible ends were to be ligated, they were first made blunt by treatment with the Klenow fragment of DNA polymerase I. The QIAquick gel extraction kit (Qiagen, Inc.) was used for the isolation of DNA fragments separated by agarose gel electrophoresis. Plasmid and genomic DNA isolation from *Streptomyces* spp. and preparation and transformation of *Streptomyces lividans* protoplasts were conducted as described earlier (25). Preparation and transfor-
mation of *S. clavuligerus* protoplasts was as described by Paradkar and Jensen (34). Studies using the PCR were performed with the Expand High-Fidelity PCR system according to the manufacturer’s instructions (Roche). Conjugative transduction of DNA from *E. coli* into *S. clavuligerus* was carried out as described for *S. coelicolor* (25), except that AS-1 medium supplemented with 10 mM MgCl₂ was used for the isolation of exconjugants (7). Southern analysis of appropriate subclones isolated from the 4.3-kb *lacI* fragment deleted to remove *pah1* DNA sequence analysis. The nucleotide sequence data obtained were compiled and analyzed by using GeneWorks (IntelliGenetics, Dublin, Ohio) by the Molecular Biology Service Unit, University of Alberta.

**DNA sequence analysis.** The nucleotide sequence data obtained were compiled and analyzed by using GeneTools 1.0 (BioTools, Inc.). Prediction of open reading frames (ORFs) based on codon preference was done with the online program FramePlot 2.3.2 (http://watson.nih.gov/jcu/cgi-bin/framesplot.pl). Similarity and homology searches were performed by using the online basic local alignment search tool (BLAST) program at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/). The PROSITE online

---

**TABLE 1. Bacterial strains and plasmids used in this study**

| Strain, plasmid, or cosmid | Description* | Source reference |
|---------------------------|--------------|-----------------|
| 6G9                       | pWE15-derived cosmid carrying DNA fragment encoding *ceaS1*, *bls1*, *pah1*, and *oat1* and flanking sequences | 22a |
| 14E10                     | pWE15-derived cosmid carrying DNA fragment encoding *ceaS1*, *bls1*, *pah1*, and *oat1* and flanking sequences | 22a |
| p2.8-18                   | pUC18 containing 2.8-kb *EcoRI* fragment carrying part of *ceaS1* | This study |
| p5K-6                     | pUC18 containing 4.3-kb NcoI fragment carrying *pah1*, *oat1*, and upstream sequences but with an internal *KpnI* fragment deleted to remove *pah1* | This study |
| p5.7                      | pUC18 containing 5.7-kb *EcoRI* fragment carrying the 3’ end of *ceaS1*, all of *bls1*, *pah1*, and *oat1*, and flanking sequences | This study |
| p5.7-T                    | p5.7 containing *tsr* from pTSR#8 inserted at *FseI* site within *bls1* | This study |
| p5.7-TH                   | *E. coli*-Streptomyces shuttle vector formed by fusing p5.7T to pJOE829 | This study |
| pApOrf6                   | Disruption construct carrying *oat2* disrupted with *apr* | 20 |
| pCAD2-3                   | pUC120 carrying *ceaS2* and *bls2* | 20 |
| pCAD2-3(L1-5)             | pCAD2-3 but with a frameshift mutation in *ceaS2* | This study |
| pCAD2-3(L1-5)486          | *E. coli*-Streptomyces shuttle vector; pCAD2-3(L1-5) fused to pJO486 | This study |
| pFDNeo-S                  | pUC18 carrying *neo* from Tn5 | 13 |
| pJ486                     | Streptomyces plasmid vector; *Tsr*’ | D. A. Hopwood |
| pJJ773                    | Template plasmid for PCR targeting; *acc(3)IV* plus *oriT* | 16 |
| pJO829                    | Streptomyces plasmid vector; *HgyI* | J. Altenbucher, University of Stuttgart, Stuttgart, Germany |
| pNEO5K-6A                 | p5K-6 containing *neo* from pFDNeo-S inserted at *RsrII* site in *oat1* | This study |
| pTSR#8                    | pUC118 containing *tsr* gene from pJJ702 | 1 |
| pUC18                     | *E. coli* cloning vector; *Amp’* | Stratagene |
| pUC118                    | *E. coli* phagemid; *Amp’* | 44 |
| pUC120Apr                 | pUC120 containing the *apr* gene flanked by *Neo* sites | 34 |

*HgyI*, hygromycin resistance; *Amp’*, ampicillin resistance.
program at the ExPaSy home page was used to search for specific peptide motifs (http://ca.expasy.org/protsite/).

Creation of targeted gene replacement mutants. The plasmid pCAD2-3 was used to prepare cas2 frameshift mutants (cas2-2-Fs). pCAD2-3 was linearized by digestion at the unique NotI site located 674 bp from the proposed start codon of cas2. The linearized plasmid was made blunt by treatment with the Klenow fragment of DNA polymerase I and then recircularized to give pCAD2-3(L1-5) with a 4-bp insertion resulting in a +1 frameshift mutation in cas2. The Streptomyces vector plJ1846 was fused to pCAD2-3(L1-5) at the HindIII site to give the E. coli-Streptomyces shuttle vector pCAD2-3(L1-5)J1846, which was passed through S. lividans TK24 and then transformed within the S. clavuligerus cas2-apr mutant. 4B. cas2::apr-J1846 has a disruption in the cas2 gene resulting from insertion of an apramycin resistance gene cassette (apr) (20). Apramycin- and thiostrepton-resistant (Apr' and Tsr', respectively) transformants were allowed to sporulate twice on nonselective medium to isolate Apr' and Tsr' mutants. Southern analysis was then used to confirm the replacement of the original cas2::apr by cas2-2-Fs. An 855-bp EcoRI-NotI fragment which contained the 5' region of cas2 and some upstream sequence was used as the cas2-specific probe.

cas1 mutants were prepared by using the Redirect PCR targeting system described by Gust et al. (16). The Redirect PCR targeting materials were supplied by Plant Biosciences, Ltd., Norwich, United Kingdom. The primers KTA14 (5'-GGATCCGTCGACCTAACAGATGAAGAGGCCGCGGCTCAATGGTGAAA) and KTA15 (5'-CGGGGCCGGCGATGCTGAACTCGTCTCCAGGTTGCTAGTATGGCTGGAGCTGCTG) were used to amplify the disruption cassette from the template plasmid plJ773. The disruption cassette comprised the acc(3W) gene, conferring Apr', and an RK2 origin of transfer (oriT) flanked by DNA sequence homologous to regions immediately upstream and downstream of cas1. The cosmid 14E10 was used to prepare the cas1 disruption construct in E. coli BW25113/pApOrf6. BM4100 independence was confirmed by Southern analysis. The disruption construct was transformed into E. coli S17-1, and the entire gene and some upstream coding sequence, was used as the disruption cassette (Delta51::apr) to produce the mutant cosmid 14E10-AP. 14E10-AP was then introduced into wild-type S. clavuligerus by conjugation, and exconjugants were selected based on Apr' and kanamycin sensitivity (Kan') on AS-1 medium supplemented with 10 mM MgCl2. These isolates were allowed to sporulate under nonselective conditions to isolate unigenomic Apr' Kan' spores.

cas1/cas2 double mutants were prepared as described for the cas1 mutants, except that the mutant cosmid 14E10-AP was conjugated into the cas2-2-Fs mutant strain, O2FS, instead of the wild type. Disruption of the wild-type copy of cas1, in the cas1 mutants and in the cas1/cas2 double mutants, was confirmed by Southern hybridization. A 777-bp EcoRI-NotI fragment internal to cas1 was used as the cas1-specific probe.

bla1 mutants were prepared by isolating the thiostrepton resistance gene cassette (tra) from pTSR#8 as an EcoRI-Streptomyces shuttle vector pCAD2-3(L1-5)J1846, which was passed through S. clavuligerus ceaS2 to generate gene replacement mutants, essentially as described by Paradkar and Jensen (34). Southern analysis was used to confirm the replacement of the genomic wild-type copies of cas1, in the cas1 mutants and in the cas1/cas2 double mutants, was confirmed by Southern hybridization. A 777-bp EcoRI-NotI fragment internal to cas1 was used as the cas1-specific probe.

bla2 mutants were prepared by transformation of the bla2 mutant strain (originally called Bc::apr [20]) with the bla2 disruption construct. p5.7-TH. Apr' and Tsr' transformants were selected, and the bla2::gbl2 double mutation was confirmed by Southern analysis.

out1 was disrupted by linearization of the bla2 mutant strain (originally called Bc::apr [20]) with the bla2 disruption construct. p5.7-TH. Apr' and Tsr' transformants were selected, and the bla2::bla2 double mutation was confirmed by Southern analysis.

out1 was disrupted by linearization of the bla2 mutant strain (originally called Bc::apr [20]) with the bla2 disruption construct. p5.7-TH. Apr' and Tsr' transformants were selected, and the bla2::bla2 double mutation was confirmed by Southern analysis.

out1 was disrupted by linearization of the bla2 mutant strain (originally called Bc::apr [20]) with the bla2 disruption construct. p5.7-TH. Apr' and Tsr' transformants were selected, and the bla2::bla2 double mutation was confirmed by Southern analysis.

out1 was disrupted by linearization of the bla2 mutant strain (originally called Bc::apr [20]) with the bla2 disruption construct. p5.7-TH. Apr' and Tsr' transformants were selected, and the bla2::bla2 double mutation was confirmed by Southern analysis.

out1 was disrupted by linearization of the bla2 mutant strain (originally called Bc::apr [20]) with the bla2 disruption construct. p5.7-TH. Apr' and Tsr' transformants were selected, and the bla2::bla2 double mutation was confirmed by Southern analysis.
that the ceaS gene in the clavulanic acid gene cluster is carried on a 12-kb EcoRI fragment (20). The 2.85-kb EcoRI fragment was therefore postulated to encode a second copy of ceaS, the putative ceaS paralogue.

Previous DNA sequence analyses had not located any putative paralogues in the regions flanking cas1 (32), and pah1 is apparently not linked to cas1. However, it seemed possible that other paralogues, if they exist, might be clustered with pah1. Cosmids 14E10 and 6G9, previously shown to carry pah1 and flanking regions of the chromosome, were digested with EcoRI and subjected to Southern analysis with the ceaS-specific probe. A 2.85-kb hybridizing fragment was observed in both of the cosmids (Fig. 3), and the fragment was subcloned for further study. Sequence analysis indicated the presence of an incomplete ORF showing similarity to the 5' end of ceaS, located at one end of this fragment. The same EcoRI digests of the cosmids 14E10 and 6G9 were also screened with a probe specific for the gene encoding β-lactam synthetase (bls), to search for a bls paralogue in the same region (data not shown).

The 582-bp BglII-NcoI fragment from pCAD2-3 was used as the bls-specific probe. A 5.7-kb fragment seen in digests of the cosmids 6G9 but not in 14E10, hybridized to the probe, and was therefore subcloned and partially sequenced. At one end of this fragment the 3' end of the ceaS-like ORF was found. The complete ORF was 1,668 bp in size, and the predicted amino acid sequence of this ORF showed 66% identity (77% similarity) to CeaS. Downstream of the ceaS-like ORF, separated by an intergenic region of 23 bp, a 1,584-bp ORF was found that displayed 49% identity (59% similarity) to β-LS at the amino acid level (Table 2). The previously described pah1 gene was found to be located immediately downstream and in the same orientation as the bls-like ORF, separated by a gap of 314 bp (Fig. 4). These newly found ORFs were designated ceaS1 and bls1, respectively, due to their linkage to pah1.

The pah1 gene was initially found on a 4.3-kb NcoI fragment that also contained flanking sequences from the S. clavuligerus chromosome (22a). When the region downstream of pah1 was sequenced, an 1,176-bp ORF was found separated by an intergenic gap of 124 bp. The predicted amino acid sequence of this ORF showed 47% identity (58% similarity) to OAT from the clavulanic acid gene cluster. This ORF was therefore called oat1 to distinguish it from oat2, which lies adjacent to cas2 in the clavulanic acid gene cluster. Like OAT2, the predicted OAT1 protein also contains an ArgJ family domain, and oat1 was predicted to be transcribed in the orientation opposite to pah1 transcription (Fig. 4). S. clavuligerus also possesses another known OAT encoded by argJ, which is part of the arginine biosynthetic gene cluster (38) and the predicted OAT1 showed 29% identity to ArgJ.

The gene cluster comprising ceaS1, bls1, pah1, and oat1 is hereafter referred to as the paralogue gene cluster (Fig. 4) to distinguish it from the clavulanic acid gene cluster which com-

![FIG. 3. Southern analysis of cosmids 6G9 and 14E10 with a ceaS2-specific probe. Lane 1, ceaS2 probe (control); lanes 2 and 3, EcoRI-digested 6G9 and 14E10, respectively.](http://aac.asm.org/)

![FIG. 4. Diagram of the genes encoded by the paralogue gene cluster thought to be involved in the early stages of clavulanic acid and 5S clavam metabolite biosynthesis. Only restriction sites referred to in the text are shown. EcoRI* denotes an EcoRI site arising from the multiple cloning site of the cosmId vector, pWE15. (The diagram is not to scale.)](http://aac.asm.org/)
prises ceaS2, bsl2, pah2, cas2, and oat2 in addition to other genes involved in clavulanic acid and 5S clavam metabolite synthesis.

**Generation of a ceaS2 frameshift mutant.** In order to study the involvement of ceaS1 in clavulanic acid and 5S clavam metabolite biosynthesis, ceaS1 single and ceaS1/ceaS2 double mutants were prepared. However, as a first step, a new ceaS2 single-mutant strain was constructed in which the apr resistance cassette of the original ceaS2::apr mutant (20) was replaced by a simple frameshift mutation. This enabled the Apr^r gene cassette to be used in the preparation of the ceaS1 disruption mutant. The new ceaS2 mutant was generated by the introduction of a plasmid construct, pCAD2-3(L1-5)486 carrying a frameshifted mutant copy of ceaS2 (ceaS2-Fs), into the previously prepared ceaS2::apr mutant strain (20), and by screening for double-crossover events with a loss of Apr^r.

To confirm the replacement of the apr disrupted copy of ceaS2 by ceaS2::apr, genomic DNA from the new ceaS2 mutant and wild-type strains was analyzed by Southern hybridization after Southern hybridization after digestion with EcoRI and NruI. When the EcoRI-NruI fragment from pCAD2-3 (includes both the ceaS2 and bsl2 genes) was used as a ceaS2-specific probe, a 2.0-kb fragment hybridized to the probe in the wild-type and the new ceaS2-Fs mutant samples (Fig. 5). In contrast, the same probe hybridized to a 3.5-kb fragment in samples from the parental ceaS2::apr mutants (Fig. 5). When the same blot was stripped and reprobed with apr from pUC120Apr as a probe, no hybridizing bands were seen in lanes containing DNA from the wild type and the ceaS2-Fs mutants, whereas a 3.5-kb fragment hybridized to the probe in lanes containing DNA from the ceaS2::apr mutant (Fig. 5). These results were consistent with the replacement of ceaS2::apr by ceaS2-Fs on the *S. clavuligerus* chromosome.

Three ceaS2-Fs mutants were fermented in SA and soy media in single shake flask cultures, along with wild-type and parental ceaS2::apr strains. Supernatants from 72- and 96-h cultures were analyzed by HPLC and bioassays for clavulanic acid and 5S clavam metabolite production. HPLC analysis showed that the ceaS2-Fs mutants produced clavulanic acid at levels of up to 13% compared to the wild-type strain in soy medium. Various amounts of clavam metabolites were also detected in these culture supernatants, but production of all of the detectable clavam metabolites was depressed. As predicted earlier, both the ceaS2::apr and ceaS2-Fs mutants were completely blocked in clavulanic acid and clavam metabolite biosynthesis in SA medium (20).

**Generation of a ceaS1 mutant.** The ceaS1 mutant was created by using the recently described Redirect PCR targeting system (16). The PCR primers used were designed as such that ceaS1 was completely deleted and replaced by the acc(3)IV + oriT cassette (ΔceaS1::apr).

Five mutants were isolated and confirmed by Southern analysis (Table 3) and then characterized by fermentation in SA and soy media. On HPLC analysis of SA and soy culture supernatants, a reduction in clavulanic acid production compared to the wild-type strain was observed (Table 4). 5S clavam metabolite biosynthesis also varied in soy medium, but no specific pattern was identified. The ceaS1 disruption did not have any effect on alanylclavam or cephamycin production, as indicated by bioassays.

**Generation of an *S. clavuligerus* ceaS1 and ceaS2 double mutant.** Since both the ceaS1 and the ceaS2 single mutants produced some clavulanic acid and 5S clavams, a ceaS1/ceaS2 double-mutant strain was prepared to verify that the two genes are indeed true paralogues. The ceaS1/ceaS2 double mutant was prepared by conjugating the mutant cosmids 14E10-AP containing ΔceaS1::apr into the ceaS2-Fs mutant strain. Three parental ceaS2-Fs mutants were used to isolate six ceaS1/ceaS2

![FIG. 5. Southern analysis of the ceaS2-Fs mutant. (A) Diagram of the ceaS2 region of the *S. clavuligerus* chromosome in the wild type, the ceaS2::apr mutant, and the ceaS2-Fs mutant. The gray arrow represents the apr disruption cassette, and the open arrow represents ceaS2 with the direction of transcription represented by the direction of the arrowheads. The solid bar represents the frame shift mutation, and the fine lines represent the rest of the *S. clavuligerus* chromosome. (B) Southern analysis of EcoRI- and NruI-digested genomic DNA from *S. clavuligerus* wild-type and ceaS2 mutant strains. DNA from the wild type, from ceaS2::apr mutants 4B and 4B-C, and from ceaS2-Fs mutants FS3, FS7, FS8, and FS10 was probed with a ceaS2-specific probe and an apr-specific probe.](attachment:fig5.png)
TABLE 3. Hybridization profiles from Southern analyses of wild-type and mutant strains

| Mutant | Restriction endonuclease(s) | Hybridizing fragment (kb) |
|--------|-----------------------------|---------------------------|
|        | Parent^a                      | Mutant                    |
| ceaS2-Fs | EcoRI/NruI | ceaS2 | 3.5 | 2.0 |
|        |.neoI | apr | 3.5 | None |
| ΔceaS1::apr | NeoI | ceaS1 | 1.2^c | None^c |
|        | apr | None | 3.4 |
| ΔceaS1::apr/ceaS2-Fs | NeoI | ceaS1 | 1.2^c | None^c |
|        | apr | None | 3.4 |
| bsl1::tsr | NeoI | bsl1 | 1.9^c | 3.0^c |
|        | tsr | None | 3.0 |
| bsl1::tsr/bls2::apr | NeoI | bsl1 | 1.9^c | 3.0^c |
|        | tsr | None | 3.0 |
| oat1::neo | BglII | oat1 | 4.3 | 2.4 and 2.9 |
|        | neo | None | 2.4 and 2.9 |
| oat1::neo/oat2::apr | BglII | oat2 | 6.5 | 7.95 |
|        | neo | None | 7.95 |

^a Gene-specific probes are described in Materials and Methods. An approximately 1.5-kb fragment from pUC120Apr was used as the apramycin-specific probe. An approximately 1-kb EcoRI-HindIII fragment from pTSR#8 was used as the thiostrepton-specific probe. An approximately 1-kb BamHI-HindIII fragment from pFDNeo-S was used as the neomycin-specific probe.

^c Faint cross-hybridizing bands were observed which can be attributed to the presence of the respective paralogues.

The six isolated ceaS1::ceaS2 double mutants were analyzed for clavulanic acid, 5S clavam, and cephamycin production after 72 and 96 h of growth in SA and soy medium. No clavulanic acid or 5S clavam production was observed, though at levels only 40% of that produced by the wild-type strain (20). To establish unequivocally the involvement of both bsl1 and bsl2 in clavulanic acid biosynthesis, a bsl1/bls2 double-mutant strain was prepared and tested for its ability to produce clavulanic acid and clavam metabolites. On analysis of culture supernatants from bsl1/bls2 double mutants grown in soy and SA medium, no clavulanic acid or 5S clavam production was detected by HPLC or bioassays. Although clavulanic acid and clavam metabolite biosynthesis was completely abolished in these mutants, they still produced wild-type levels of cephapmycin.

The involvement of oat1, which lies immediately downstream of pah1 (Fig. 4), in clavulanic acid and clavam metabolite biosynthesis was investigated by preparing a mutant with neo disrupting the oat1 gene. The oat1 mutation was verified by Southern analysis (Table 3), and its effect was surveyed by fermenting different oat1 mutants in soy and SA medium with wild-type cured strains. After 72 h of growth, clavulanic acid production was between 28 to 86% and 14 to 87% in SA and soy medium, respectively, compared to cured wild-type controls. Cephamycin production was again unaffected in the bsl1 mutant strains.

Generation of a bsl1 mutant. The bsl1 mutant was prepared by insertion of a Tsr^g gene cassette into the bsl1 gene in the opposite orientation to bsl1. Nine mutants were isolated, verified by Southern analysis (Table 3), and analyzed for their ability to produce clavulanic acid and clavam metabolites. On HPLC analysis of SA and soy culture supernatants, the mutants appeared to be little affected in clavulanic acid or 5S clavam metabolite biosynthesis. In SA medium the bsl1 mutants produced between 102 to 312% and 103 to 162% of the wild-type levels of clavulanic acid after 72 and 96 h of growth, respectively. Similar results were obtained when soy culture supernatants were analyzed. After 72 and 96 h of growth in soy medium, the mutants produced between 50 to 247% and 107 to 173% of the wild-type levels clavulanic acid. Once again, the level of 5S clavams varied from mutant to mutant, with some bsl1 mutants accumulating elevated levels of 5S clavam metabolites in soy culture, whereas others produced much reduced levels. At 96 h clavam-2-carboxylate and 2-hydroxymethyl clavam production varied from 10 to 301% and 5 to 355%, respectively, compared to cured wild-type controls. Cephapmycin production was again unaffected in the bsl1 mutant strains.

Generation of an oat1 mutant. Eight single oat1 mutants were isolated, verified by Southern analysis (Table 3), and analyzed for their ability to produce clavulanic acid and clavam metabolites. On analysis of culture supernatants from the oat1 mutants, they still produced wild-type levels of cephapmycin.

Generation of an oat1 mutant. The involvement of oat1, which lies immediately downstream of pah1 (Fig. 4), in clavulanic acid and clavam metabolite biosynthesis was investigated by preparing a mutant with neo disrupting the oat1 gene. The oat1 mutation was verified by Southern analysis (Table 3), and its effect was surveyed by fermenting different oat1 mutants in soy and SA medium along with wild-type cured strains. After 72 h of growth, clavulamic acid production was between 28 to 86% and 14 to 87% in SA and soy medium, respectively, compared to cured wild-type controls. After 96 h of growth, even more variation in clavulamic acid production was observed in both media compared to the cured wild-type controls. In SA medium clavulamic acid production varied from 30 to 154% and in soy medium from 57 to 145% compared to the wild-type strain. No specific trend in 5S clavam production was observed, and cephapmycin production was unaffected in the oat1 mutants.

Generation of an oat1 and oat2 double mutant. Previously, when supernatants from oat2 mutant cultures were analyzed, the mutants still produced clavulamic acid in soy medium, although at levels only 40% of that produced by the wild-type strain (20). To determine whether the ability of the oat2 mutant to produce clavulamic acid could be attributed to the presence of oat1, we prepared oat1/oat2 double-mutant strains of S. clavuligerus and studied their effect on clavulamic acid and 5S clavam biosynthesis. This was done by introducing the previously prepared oat2::apr disruption (20) into the S. clavulg-
**DISCUSSION**

This study extends the work of Jensen et al. (20), who found that *S. clavuligerus* *ceaS*, *bls*, *pah1*, *cas2*, and *oat* mutants, when prepared individually, still retained some ability to produce clavulanic acid and 5S clavam metabolites. A similar decrease in clavulanic acid production was also observed in the culture supernatants analyzed after 96 h of growth (Table 4). Again, there was a high degree of variation in the levels of 5S clavam metabolites produced, and no specific trend was observed. Cephapemycin production was unaffected in the *oat1/oat2* mutants.

In the present study we report the isolation and characterization of three additional paralogues of genes encoding proteins involved in the early stages of clavulanic acid and 5S clavam biosynthesis. The genes *ceaS1*, *bls1*, and *oat1* were found grouped together, along with the previously reported *pah1* (22a), in a cluster designated the paralogue gene cluster. This paralogue cluster is distinct from the clavulanic acid gene cluster that encodes *ceaS2*, *bls2*, and *oat2* and other genes involved in clavulanic acid and clavam metabolite biosynthesis (20, 26, 29). *ceaS1*, *bls1*, and *oat1* are 73, 60, and 63% identical to *ceaS2*, *bls2*, and *oat2*, respectively, while the predicted proteins encoded by *ceaS1*, *bls1*, and *oat1* are 66, 49, and 47% identical to those encoded by *ceaS2*, *bls2*, and *oat2*, respectively. From the paralogue gene cluster, *pah1* shows the highest level of identity to its paralogous counterpart *pah2*, which was 71% at the amino acid level (22a).

OAT catalyzes the formation of ornithine by transferring an acetyl group from N-acetylornithine to glutamate, a key step in the biosynthesis of arginine (12). Recently, *orf6* from the clavulanic acid gene cluster, referred to as *oat2* in the present study, was shown to display OAT activity (23). In addition, *S. clavuligerus* contains another OAT, ArgJ, that is encoded by the arginine biosynthetic gene cluster (39), making the predicted OAT1 the third OAT so far known to be present in *S. clavuligerus*. Although ArgJ is presumably required for arginine biosynthesis, *oat2* and *oat1* are associated with genes involved in the biosynthesis of clavulanic acid and the 5S clavams, which utilize arginine as a precursor. To test whether *oat1* was involved in the biosynthesis of clavulanic acid or the clavams, an *oat1* mutant was prepared, but no marked decrease in clavulanic acid or 5S clavam production by the *oat1* mutant strain was observed. Since the *oat1* mutant still has wild-type copies of *oat2* and ArgJ, it is possible that OAT2 and ArgJ can compensate for the *oat1* mutation. Alternatively, OAT1 activity may not be essential for clavulanic acid or 5S clavam metabolism. To investigate the former hypothesis, we prepared an *oat1/oat2* double-mutant strain and saw that the levels of clavulanic acid and 5S clavams produced by the double mutant dropped in both SA and soy media compared to the wild-type strain, although some production still remained (Table 4). If *oat1* and *oat2* indeed encode OATs involved in providing arginine for clavulanic acid and clavam biosynthesis, then the residual production of these metabolites by the *oat1/oat2* double mutant may be attributed to the presence of wild-type ArgJ in this mutant. Although a role for *oat1* and *oat2* in providing increased precursor availability for metabolite synthesis seems most plausible, their role in some unidentified step in clavulanic acid or 5S clavam biosynthesis cannot be ruled out.

All of the known or putative OATs sequenced to date contain the autoprotoelytic cleavage motif KGXGMXXPX-(M/L) AT(M/L) with cleavage taking place between the alanine and threonine residues (11). OAT2 is expressed as a 42-kDa peptide that undergoes posttranslational autoprotoelytic cleavage to form a small 19-kDa subunit and a large 25-kDa subunit, which oligomerize to form an 84-kDa heterotetramer (23). The cleavage occurs between alanine 180 and threonine 181 residues of the motif 169KGVGMLEPDATML183 (23). A similar motif, 169KGAGMLAPGLATTLL182, is also found in *S. clavuligerus* ArgJ, but posttranslational cleavage has yet to be demonstrated. The predicted amino acid sequence of OAT1 in this region is 167KGPGTGPAEQDDRSTL182, which deviates from the consensus sequence, and it is also missing the adjacent alanine and threonine residues where cleavage is thought to take place. Therefore, more work is required to confirm whether *oat1* actually encodes an active OAT or if it could have some other, yet-unknown function in clavulanic acid or clavam biosynthesis.

The *ceaS1* and the *bls1* genes described here were isolated by screening *S. clavuligerus* chromosomal DNA, using sequences from *ceaS2* and *bls2* as probes, respectively. When *ceaS2* was first sequenced, it showed striking similarity to genes encoding acetohydroxyacid synthases (AHAS) based on observed homologies and conservation of five of the eight amino acids forming the active center of AHAS (35). Subsequently, *CeaS2* was shown to catalyze the thiamine pyrophosphate (TPP)-dependent condensation of glyceraldehyde-3-phosphate and l-arginine to form carboxyethylarginine, the first reaction in the clavulanic acid-clavam biosynthetic pathways (24). These five amino acids, associated with the active centers of AHAS enzymes, which are found in CeaS2 (49E,132Q,472G,499N, and 503G), are also conserved in CeaS1 (45E,112Q,446IGAQMARPDQPT, and 476G). Since CeaS2 utilizes TPP, it also contains the TPP binding motif (146IGAOMRPDPQTFLIAGDGG447), and a similar TPP binding motif (428MAAQIARPGEVFLIAGDGG) is also present in the predicted CeaS1.

β-Lactam synthetase, as the name suggests, is responsible for the formation of the β-lactam ring of clavulanic acid and the clavams (2, 28). It catalyzes the second reaction in the clavulanic acid and 5S clavam biosynthetic pathway and requires ATP and Mg2+ (2, 28). Crystallographic studies on β-LS2 have identified certain amino acids that are involved in substrate binding and catalysis (30, 31), and these amino acids were also found to be conserved in β-LS1 (Fig. 6). Both β-LS2
and β-LS2 show similarities to asparagine synthases (AS-B) from different organisms. Asparagine synthases belong to a family of enzymes called Ntn amidotransferases that have a conserved cysteine residue at their N terminus (8). This conserved cysteine residue is missing at the N terminus of both β-LS2 and β-LS1, respectively, similar functions were envisioned for these pairs of proteins. To test whether ceaS1 and bsl1 were involved in clavulamic acid or clavam biosynthesis, ceaS1 and bsl1 mutants were prepared and tested individually. Since the ceaS1 and the bsl1 mutants still contain functional wild-type copies of ceaS2 and bsl1, respectively, any clavulamic acid or clavams produced by these mutants can be attributed to the presence of these genes. Large variations in the levels of 5S clavam production were observed in the bsl1 mutant. It is possible that the production of these metabolites may be extremely sensitive to minor variations in culture conditions. Despite our best efforts to produce replicate cultures, these variations persist and are observed in all of our fermentation studies. To confirm the hypothesis that ceaS1 and ceaS2 and bsl1 and bsl2 were indeed true paralogues, ceaS1/ceaS2 and bsl1/bls2 double mutants were prepared and characterized. Both the ceaS1/ceaS2 and the bsl1/bls2 double mutants were found to be completely blocked in clavulamic acid and clavam biosynthesis, lending further evidence to the paralogue hypothesis.

Results from the present study and from other work (22a) show that the genes encoding enzymes involved in the early stages of clavulamic acid and clavam metabolite biosynthesis, at least up to the level of clavaminic acid, are duplicated in *S. clavuligerus*. At present, the paralogue gene cluster contains paralogues of four genes from the clavulamic acid gene cluster in a similar but not an identical arrangement to their counterparts in the clavulamic acid gene cluster (20). Surprisingly, cas1, the paralogous counterpart of cas2, is found elsewhere on the *S. clavuligerus* chromosome and is surrounded by genes that are involved exclusively in 5S clavam metabolite biosynthesis and not in the biosynthesis of clavulamic acid (32). In addition to the striking absence of the *cas1* paralogue, the relative orientation of *oat1* is opposite to that of *oat2* with respect to its neighboring genes. Therefore, the paralogue gene cluster does not appear to have arisen from a direct duplication of a portion of the clavulamic acid gene cluster; rather, a somewhat more complex evolution is indicated.

The reasons why *S. clavuligerus* has two sets of genes encoding enzymes involved in the early part pathway that is shared between clavulamic acid and the 5S clavams is not clear. One explanation could be to provide a gene dosage effect and thereby increase the level of clavaminic acid production, which is a precursor of both clavulamic acid and the 5S clavams. A second possibility is that the two sets of paralogues have arisen to serve separate, parallel biosynthetic pathways, which happen to share intermediates up to the level of clavaminic acid. To date, there is no evidence for paralogues of other genes in the clavulamic acid gene cluster, based on a study of other work (20), suggesting that these other genes are dedicated to the biosynthesis of either clavulamic acid or the 5S clavams with no need for duplication and increased gene dosage (20). Therefore, in conclusion, the genes involved in clavulamic acid and clavam biosynthesis are now shown to be grouped into three distinct gene clusters in *S. clavuligerus*, the clavulamic acid gene cluster, the *cas1*-associated clavam gene cluster and the paralogue gene cluster, with no evidence of linkage between the clusters. Duplication of all of the genes encoding enzymes involved in the early shared steps of the pathway is evident, with paralogues to the original clavulamic acid gene cluster genes located in both the *cas1* clavam gene cluster and in the newly described paralogue gene cluster.

Acknowledgments

This study was supported by grants from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada. K.T. was supported by a studentship from the Alberta Heritage Foundation for Medical Research.

References

1. Aidoo, K. A., A. Wong, D. C. Alexander, R. A. Rittamper, and S. E. Jensen. 1994. Cloning, sequencing, and disruption of a gene from Streptomyces clavuligerus involved in clavulamic acid biosynthesis. Gene 147:41–46.
2. Bachmann, B. O., R. Li, and C. A. Townsend. 1996. β-Lactam synthetase: a new biosynthetic enzyme. Proc. Natl. Acad. Sci. USA 93:4089–4093.
3. Baggaley, K. H., A. J. Brown, and C. J. Schofield. 1997. Chemistry and biosynthesis of clavulamic acid and other clavams. Nat. Prod. Rep. 14:309–333.
4. Baggaley, K. H., S. W. Elson, N. H. Nicholson, and J. T. Sime. 1990. Studies on the biosynthesis of clavulamic acid. Part 4. Synthetic routes to the monocyclic β-lactam precursor, proclavamic acid. J. Chem. Soc. Perkin Trans. 1 1990:1513–1520.
5. Baldwin, J. E., R. M. Adlington, J. S. Bryans, A. O. Bringhen, J. B. Coates, N. P. Crouch, M. D. Lloyd, C. J. Schofield, S. W. Elson, K. H. Baggaley, R. Cassels, and N. H. Nicholson. 1991. Isolation of dihydroclavamic acid, an intermediate in the biosynthesis of clavulamic acid. Tetrahedron 47:4089–4100.
6. Baldwin, J. E., M. D. Lloyd, B. Who-Son, C. J. Schofield, S. W. Elson, K. H. Baggaley, R. Cassels, and N. H. Nicholson.
