ORF17 from the Clavulanic Acid Biosynthesis Gene Cluster Catalyzes the ATP-dependent Formation of \(N\)-Glycyl-clavaminic Acid*  

Received for publication, July 15, 2005, and in revised form, October 25, 2005. Published, JBC Papers in Press, October 26, 2005, DOI 10.1074/jbc.M507711200

Haren Arulanantham†, Nadia J. Kershaw‡, Kirsty S. Hewitson§, Claire E. Hughes§, Jan E. Thirkettle§, and Christopher J. Schofield††

From the †Department of Chemistry and The Oxford Centre for Molecular Sciences, Chemistry Research Laboratory, University of Oxford, Mansfield Road, Oxford OX1 3TA and ‡Glxosmithkline Pharmaceuticals, Clarendon Road, Worthing, West Sussex BN14 8QH, United Kingdom

(3R,5R)-Clavulanic acid, a clinically used inhibitor of serine \(\beta\)-lactamases, is produced by fermentation of Streptomyces clavuligerus. The early steps in clavulanic acid biosynthesis leading to the bicyclic \(\beta\)-lactam intermediate \((3S,5S)\)-clavaminic acid have been defined. However, the mechanism by which \((3S,5S)\)-clavaminic acid is converted to the penultimate intermediate \((3R,5R)\)-clavaldehyde is unclear. Disruption of orf15 or orf16, of the clavulanic acid biosynthesis gene cluster, blocks clavulanic acid production and leads to the accumulation of \(N\)-acetyl-glycyl-clavaminic acid and \(N\)-glycyl-clavaminic acid, suggesting that these compounds are intermediates in the pathway. Two alternative start codons have been proposed for orf17 to encode for two possible polypeptides, one of which has 92 N-terminal residues less then the other. The shorter version of orf17 was successfully expressed in Escherichia coli and purified as a monomeric protein. Sequence analyses predicting the ORF17 protein to be a member of the ATP-grasp fold superfamily were supported by N-xylation mass spectrometric analyses that demonstrated binding of ATP to the ORF17 protein. Semisynthetic clavaminic acid, prepared by in vitro reconstitution of the biosynthetic pathway from the synthetically accessible intermediate proclavaminic acid, was shown by mass spectrometric analyses to be converted to \(N\)-glycyl-clavaminic acid in the presence of ORF17, ATP, and glycine. Under the same conditions \(N\)-acetyl-glycine and clavaminic acid were not converted to \(N\)-acytetyl-clavaminic acid. The specificity of ORF17 as an \(N\)-glycyl-clavaminic acid synthetase, together with the reported accumulation of \(N\)-glycyl-clavaminic acid in orf15 and orf16 disruption mutants, suggested that \(N\)-glycyl-clavaminic acid is an intermediate in clavulamic acid biosynthesis.

The bicyclic \(\beta\)-lactam clavamic acid is a potent inhibitor of Class A serine \(\beta\)-lactamases (penicillinases) and is used clinically in combination with semisynthetic penicillins such as amoxicillin (1, 2). Clavamic acid is one of a family of clavams produced as secondary metabolites by Streptomyces clavuligerus, but it is unusual among naturally occurring clavams in that it possesses the 3R,5R stereochemistry required for reaction with penicillin-binding proteins and \(\beta\)-lactamases. Although clavamic acid has only eight carbons and two chiral centers, its lability and density of functionalization renders its preparation via chemical synthesis difficult (2, 3). Instead, it is commercially isolated from fermentations of S. clavuligerus. There has been interest in clavam biosynthesis with a view to optimizing the fermentation of clavulamic acid as well as for engineering the pathway to produce new derivatives with a broad spectrum of antibacterial and \(\beta\)-lactamase inhibitory activities.

The clavulamic acid biosynthesis gene cluster in S. clavuligerus is currently thought to comprise ~18 genes that are directly involved in its biosynthesis, transport, and regulation (Fig. 1a, orfs 2–19 (4–7). Another differently regulated clavam gene cluster in S. clavuligerus has also been identified that contains a minimum of four genes homologous to orfs 2, 3, 4, and 6 of the better characterized cluster (8, 9). Furthermore, cas1 (clavaminic acid synthase 1), a homologue of cas2 (orf5), is located elsewhere in the genome (10, 11) and is flanked by genes associated with the production of other \((3S,5S)\)-clavam metabolites (12).

Steps in the pathway leading to \((3S,5S)\)-clavamic acid, the proposed branch point between the biosynthesis of clavulamic acid and the \((3S,5S)\)-clavams (13), have been characterized (Fig. 1b, and the enzymes that catalyze them have been identified. In the first step, l-arginine and d-glyceraldehyde-3-phosphate react to give \(N^2\)-carboxyethylarginine in a thiamin diphosphate-dependent reaction catalyzed by \(N^2\)-carboxyethylarginine synthase (CEAS) (14). \(\beta\)-Lactam formation is catalyzed by \(\beta\)-lactam synthetase (BLS) (15, 16), to give deoxydglycaminopropionolactam acid (17), which is then converted in four steps to \((3S,5S)\)-clavaminic acid. Three of these steps are catalyzed by a single 2-oxoglutarate (2-OG)-dependent oxygenase, clavaminic acid synthase (CAS2) (10, 18–20), and one is catalyzed by proclavaminate amidohydrolase (PAH) (21), to give the bicyclic clavam ring system (22–24). Crystal structures have been reported for the first four enzymes in the pathway (25–28) as well as for the gene product of orf6 (OAT2), an ornithine acetyltransferase (29, 30) proposed to be involved in the biosynthesis of the l-arginine feedstock for the pathway. The final step in the pathway has also been identified and shown to involve the NADPH-dependent reduction of the labile \((3R,5R)\)-clavaldehyde to give \((3R,5R)\)-clavulamic acid (31).

Despite knowledge of the sequences of the likely enzymes involved (4–6, 23, 32), there is little mechanistic information for the apparent double epimerization and oxidative deamination that must occur in the conversion of \((3S,5S)\)-clavamic acid to \((3R,5R)\)-clavaldehyde. Acylated (3S,5S)-derivatives of clavamic acid (Fig. 1c), \(N\)-acytetyl-cla-

* This work was supported by grants from the Biotechnology and Biological Sciences Research Council, the European Union, the GlaxoSmithKline PLC. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Tel.: 44-1865-275625; Fax: 44-1865-275654; E-mail: christopher.schofield@chem.ox.ac.uk.

§ The abbreviations used are: CEAS, \(N^2\)-carboxyethylarginine synthase; GCAS, \(N\)-glycyl-clavaminic acid synthase (ORF17); BLS, \(\beta\)-lactam synthetase; CAS, clavaminic acid synthase; PAH, proclavaminate amidohydrolase; 2-OG, 2-oxoglutarate; LC-MS, liquid chromatography-mass spectrometry; ESI-MS, electrospray ionization MS.
vaminic acid, N-acetyl-glycyl-clavaminic acid, and N-glycyl-clavaminic acid have been isolated from a mutant strain of *S. clavuligerus* deficient in clavulanic acid production (designated *dcl*8) (33). Recently, the latter two of these compounds have been reported in disruption mutants of *orf*15 and *orf*16 (32). These compounds have not been observed in disruption mutants of any of the other *orfs* of the clavulanic acid biosyn-
thetic gene cluster, and the reactions by which they are produced have not been defined.

One problem with attempts to define the reactions involved in the biosynthesis of (3R,5β)-clavaldehyde from (3S,5S)-clavaminic acid is the availability of potential substrates with which to challenge candidate enzymes. This problem is common to studies on other biosynthetic pathways leading to many secondary (and some primary) metabolites. Here we have described studies on the latter stages of clavulonic acid biosynthesis in which we approached the problem of a lack of availability of intermediates by utilizing the activity of characterized recombinant enzymes coupled to accessible synthetic intermediates. The results revealed that ORF17 can catalyze the biosynthesis of N-glycyl-clavaminic acid from clavaminic acid in an ATP-dependent manner, supporting the proposal that N-glycyl-clavaminic acid is an intermediate in the clavulonic acid biosynthesis pathway.

**EXPERIMENTAL PROCEDURES**

**Materials**—Unless otherwise stated, all chemicals were obtained from Sigma. DNA manipulations were carried out by standard protocols (34). Restriction enzymes were purchased from New England Biolabs Inc. Oligonucleotides were synthesized by SigmaGenosys Ltd. Expression vectors were purchased from Novagen. Protein concentrations were determined by the Bradford method.

**LC-MS Analyses**—The LC-MS system comprised a Waters 600 controller pump with a Waters 2700 sample manager in combination with a Micromass ZMD mass spectrometer using electrospray ionization mass spectrometry (ESI-MS). MassLynx version 3.5 was used for data analysis and processing. The following high performance liquid chromatography columns were used: C18 (150 × 4.60 mm, 5 μ) reverse-phase column (LUNA 5 μ, C18 (2) 100A, Phenomenex). C18 octadecylsilyl guard cartridges (SecurityGuard, Phenomenex) were used throughout and replaced as required. The method was modified from Jensen et al. (32). Samples were eluted with a gradient from 5 to 30% acetonitrile in 100 mM ammonium formate, pH 4.0, at a flow rate of 1 ml/min over 15 min followed by a wash of 30% acetonitrile in 100 mM ammonium formate, pH 4.0, for 10 min before the column was re-equilibrated to 5% acetonitrile in 100 mM ammonium formate, pH 4.0, over 5 min. The production of clavams was also monitored by imidazole-derivative analysis and processing. The following high performance liquid chromatography (His6) elution volume parameter (Kav) was calculated for each of the calibration proteins, and a calibration curve was constructed.

**Mass Spectrometric Analyses**—For molecular weight determination of ORF17, ESI-MS was performed using a VG Platform II spectrometer with Agilent 1100 series pump and auto-sampler. ORF17 samples were prepared in 1:1 water-acetonitrile containing 0.1% formic acid with a final protein concentration of 10 μM. Soft ESI-MS was performed on a quadrupole time-of-flight micro with collisional cooling (Micromass UK Ltd., Altrincham, UK). ORF17 samples were prepared in 50 mM ammonium acetate with a final protein concentration of 2 μM. A 20-fold excess of ATP was added to a final protein concentration of 2 μM in 50 mM ammonium acetate. The capillary voltage was set to 3 kV, and the sample cone voltage and extractor cone voltage were changed to 120 and 20 kV, respectively.

**ORF17 Activity Assays**—The rate of ATP hydrolysis was measured at 25 °C by coupling the production of ADP to pyruvate kinase and lactate dehydrogenase from rabbit muscle solution, and the assay was initiated with 1 mM potential carboxylate and/or amine substrates in a reaction mixture containing 100 mM potassium phosphate buffer (pH 7.5), 100 mM potassium chloride, 1 mM MgCl2, 1 mM phosphoenolpyruvate, 0.2 mM NADH, varying amounts of ATP (0.1–4.0 mM), 0.7 units of pyruvate kinase, and lactate dehydrogenase from rabbit muscle solution, and the assay was initiated by the addition of 80 μg of ORF17 in a final volume of 1 ml. The release of inorganic phosphate from the hydrolysis of ATP, by ORF17, was confirmed using the EnzChek® phosphate assay kit (Molecular Probes). ORF17 assays were carried out, in the same manner, using the pyruvate kinase/lactate dehydrogenase ATPase assay (45–47), but reaction mixtures contained 1 mM ATP with 80 μg of ORF17. The assay was initiated with 1 mM potential carboxylate and/or amine substrates in a final volume of 1 ml.
Coupled Assay Procedures—Recombinant CEAS, BLS, PAH, and CAS2 were prepared as reported (14, 15, 21, 24, 48). A typical coupled enzyme incubation mixture contained 150 mM Tris-HCl (pH 8.0), 12 mM CEAS, BLS, PAH, and CAS2 with appropriate cofactors (30 mM D/L-glyceraldehyde-3-phosphate, 1.5 mM thiamin diphosphate, 10 mM MgCl₂, 5 mM ATP, 0.5 mM MnCl₂, 10 mM FeSO₄, and 10 mM 2-OG) and substrates (30 mM D-arginine or 10 mM deoxyguanidinoproclavaminic acid), in a final volume of 100 μl. The assay mixtures were incubated at 37°C. The assay mixtures were incubated at 37°C. The assay mixtures were incubated at 37°C.
30 °C in a water bath for 1 h, and the reaction was then stopped by heating for 2 min at 100 °C. Protein was removed by centrifugation, and the resulting supernatant was stored at −80 °C before LC-MS analysis on a LUNA C18 reverse-phase column. Controls were conducted in the absence of enzyme(s).

Coupled CAS2/ORF17 Activity Assay—The CAS2 assay was modified from that reported by Lloyd et al. (24). The final incubation mixture contained 100 mM Tris-HCl (pH 7.5), 2 mM dithiothreitol, 10 mM 2-OG, 2 mM FeSO4, 50 µg of CAS2, and 2.5 mM synthetic racemic clavaminic acid (24) in a final volume of 50 µ. The assay mixture was incubated at 37 °C in a water bath. After 20 min, more FeSO4 and dithiothreitol (10 µl of a 10 mM stock of each) were added. After 40 min, the reaction mixture was centrifuged at 14,000 rpm for 2 min, to remove any precipitate, and the supernatant was filtered through a 10-kDa molecular mass cut-off membrane and centrifuged at 14,000 rpm for 10 min. 25 µl of the CAS2-free filtrate was then combined with a buffer solution of 10 mM MgCl2, 100 mM KCl, and 100 mM Tris-HCl, pH 7.5, containing 2 mM ATP, 2 mM glycine, or alternative potential substrates, and the assay was initiated by the addition of 80 µg of ORF17 in a final volume of 50 µl. The mixture was incubated at 37 °C in a water bath for 30 min before the reaction was stopped by the addition of acetone (120 µl). After centrifugation, the acetone was removed in vacuo (using a Speed-Vac SC110, Savant), and the mixture was then analyzed by LC-MS. This assay and substrate specificity assays were also conducted using the pyruvate kinase/lactate dehydrogenase ATPase assay (45–47), using either 25 µl of clavaminic acid stock, produced from CAS2 assays (as above), or alternative potential substrates. The assay was initiated with 1 mM potential carboxylate substrates in a final volume of 1 ml.

RESULTS
Expression and Purification of ORF17—Two possible start codons have been proposed for orf17 (4, 32). The sequence of orf17 identified by Mellado et al. (4) encodes a protein of 529 amino acids, whereas that proposed by Jensen et al. (32) lacks the first 92 residues (this difference is reflected in the different numbering schemes for each gene product). There are two further differences in the predicted polypeptide sequences (highlighted by bold italicized residues), 473VEKGVKKLR482 compared with 386VEKGDKLQR390 for Mellado et al. (4) and Jensen et al. (32), respectively. The two proposed versions of orf17 were cloned into the pET24a (+) vector. However, in E. coli BL21 (DE3), expression was only observed for the shorter of the two proposed forms of orf17. The DNA sequence encoding the longer form of orf17 (4) contains some rare codons and thus was transformed into the E. coli RosettaTM (DE3) strain, which possesses tRNAs corresponding to the pET24a (+) vector. However, in E. coli BL21 (DE3), expression was only observed for the shorter of the two proposed forms of orf17. The two proposed versions of orf17 were cloned into the pET24a (+) vector. However, in E. coli BL21 (DE3), expression was only observed for the shorter of the two proposed forms of orf17. The DNA sequence encoding the longer form of orf17 (4) contains some rare codons and thus was transformed into the E. coli RosettaTM (DE3) strain, which possesses tRNAs corresponding to these rare codons (49). However, no significant expression of the longer form of orf17 was observed under a variety of conditions, and thus, all further studies were carried out with the shorter form of orf17.

ORF17 (Jensen sequence), from orf17/pET24a (+), was produced as ~25% of total soluble protein in E. coli BL21(DE3), but initial attempts at purification were problematic. Thus, the orf17 gene was cloned into the pET28a (+) vector to produce ORF17 protein with an N-terminal His tag to aid purification. However, only insoluble protein was produced from this vector under standard growth conditions in E. coli BL21(DE3). Orf17 was then cloned into the pET24a (+) vector using alternative restriction sites to enable production of protein with a C-terminal His tag, which was produced at ~30% of total soluble protein by SDS-PAGE analysis (Fig. 3, lane 2).

The C-terminal His-tagged ORF17 was purified by affinity chromatography (Fig. 3, lane 3). However, imidazole concentrations over 50 mM resulted in precipitation of ORF17. Therefore, an imidazole gradient of 5–80 mM was used to elute ORF17 after washing with binding buffer. ORF17 elutes between imidazole concentrations of 20–50 mM, and all fractions were diluted with Tris-HCl, pH 7.9, to prevent precipitation during the concentration of desired fractions. The combined, concentrated fractions were then further purified using gel filtration to give protein of >95% purity by SDS-PAGE analysis (Fig. 3, lane 4). Edman sequencing revealed the protein to have the predicted N terminus (TTPLADTAKF) lacking the N-terminal methionine. ESI-MS analysis of the purified ORF17 was consistent with production of the predicted protein (observed 48,580 Da, cf. calculated 48,581 Da). Native PAGE, gel filtration chromatography, and soft ionization mass spectrometry all implied that C-terminal His-tagged ORF17 was monomeric.

Sequence and Structural Analysis of ORF17—PSI-BLAST (36) and 3D-PSSM (39) searches revealed that ORF17 shares homology with proteins from the ATP-grasp fold superfamily, of which carbamoyl phosphate synthetase, biotin carboxylase, d-alanine-d-alanine ligase, and glycinamide ribonucleotide synthetase have been structurally characterized by x-ray crystallography (40, 43, 50–52). Members of the ATP-grasp fold superfamily are thought to share a similar ATP-dependent carboxylate-amine or thiol ligation mechanism (Fig. 1d), wherein a carboxylate is activated as an acyl-phosphate intermediate, which subsequently reacts with the nucleophilic substrate (50). Conserved residues that are present throughout the ATP-grasp fold superfamily are also present in ORF17 (Fig. 2a). These include Glu301 and Asn303 of ORF17, which have been shown in the biotin carboxylase subfamily to be directly involved in binding Mg2+ within the ATP-binding domain. The conserved residues Lys130 and Lys169 of ORF17 have also been shown to be involved in nucleotide binding in the biotin carboxylase subfamily (50, 53). However, ORF17 displays poor sequence similarity to the glycine-rich phosphate-binding motif of most ATP-grasp fold enzymes (54) and is most similar to the phosphate-binding region of the d-alanine-d-alanine ligase subfamily (X-Gly-Ser-Ser-X-Gly) (55).

The 3D-PSSM sequence-based secondary structure search indicated that ORF17 is most closely related to E. coli biotin carboxylase (PSSM E_value 0.000401, SAWTED value 0.0731, and 95% certainty, Protein Data Bank code 1DV1). Thus, a structural model of ORF17 was built based on the E. coli biotin carboxylase ATP-complexed structure.

The sequence alignment and structural model of ORF17 underline the strong conservation of residues in the ATP-binding site (Fig. 2).
**N-Glycyl-clavaminic Acid Synthetase**

However, the limited conservation of residues in the predicted substrate-binding site suggests that ORF17 belongs to a new subfamily of ATP-grasp fold enzymes and does not utilize the same substrate as either the bovine carbamoylase or the d-alanine-d-alanine ligase subfamilies. Experimental evidence for the proposal that ORF17 is a member of the ATP-grasp fold superfamily came from soft ionization ESI-MS analyses in which one molecule of ATP was shown to bind to monomeric ORF17.

**ATP Hydrolase Activity of ORF17**—Some ATP-grasp fold enzymes exhibit ATP hydrolase (ATPase) activity in the presence of Mg\(^{2+}\), K\(^{+}\), and ATP in the absence of other substrates (56–58). The potential ATPase activity of ORF17 was tested using UV spectrophotometric analysis, by modification of a continuous assay employing pyruvate kinase and lactate dehydrogenase, monitoring the loss of NADH at 340 nm, corresponding to the production of ADP (45–47, 56, 58). ATPase activity was observed for ORF17 at low levels and only in the presence of Mg\(^{2+}\), K\(^{+}\), and ATP. In a separate assay, the release of inorganic phosphate by ORF17-induced ATP hydrolysis was also confirmed using the EnzChek\textsuperscript{®} phosphate assay kit.

**ORF17 Substrate Specificity**—Based on the proposal that ORF17 may function in a similar manner to other ATP-grasp fold enzymes, potential substrates were then assayed. ORF17-catalyzed ADP formation was examined in the presence of potential carboxylate substrates, including bicarbonate, acetate, formate, and acetyl-coenzyme A, and also in the presence of both carboxylate- and amine/thiol-containing substrates. Substrates were tested either individually or as a combination of any two of the following: bicarbonate, acetate, formate, acetyl-coenzyme A, coenzyme A, glycine, D/L-alanine, L-valine, L-leucine, L-aspartic acid, L-asparagine, N-\(\alpha\)-acetyl-glycine, N-glycyl-glycine, biotin, D/L-ornithine, N-\(\alpha\)-acyl-L-ornithine, L-lysine, and N-\(\alpha\)-acyl-L-lysine. However, no stimulation of ADP production was observed with any of these potential substrates.

**Assays with Clavaminic Acid Prepared via Enzymatic and Chemoenzymatic Routes**—Gene disruption studies by Jensen et al. (32) have indicated that orf17 is required for clavulanic acid biosynthesis, and therefore, intermediates in the clavaminic acid biosynthesis pathway were required for further studies on the enzyme. Access to these intermediates (in particular, clavaminic acid) via synthetic routes is a limiting factor in studying the pathway, due to the lack of, or lengthy, synthetic routes. To overcome the problem of the limited availability of candidate substrates, we investigated methods of producing these intermediates via in vitro reconstitution of the enzymatic pathway from primary metabolites or synthetically accessible intermediates.

Attempts to couple CEAS, BLS, PAH, and CAS2 to produce clavaminic acid in a "one-pot" procedure from L-arginine and D/L-glyceraldehyde-3-phosphate were carried out. The anticipated products for each of the enzymes were observed by LC-MS analyses of each of these four enzymes with the bacteria in the presence of appropriate cofactors and substrates. Coupled assays with CEAS and BLS produced the anticipated product, deoxyguanidinopropclavaminic acid, from L-arginine and D/L-glyceraldehyde-3-phosphate in the presence of thiamin diphosphate and ATP (Fig. 1 b). Combining CEAS, BLS, and CAS2 gave guanidinopropclavaminic acid, in the presence of thiamin diphosphate ATP, Fe(II), and 2-OG. The coupled assay of PAH and CAS2 only gave low yields of clavaminic acid from deoxyguanidinopropclavaminic acid, in the presence of Mn(II), Fe(II), and 2-OG; optimization of the relative Mn(II) and Fe(II) levels was required as Mn(II) inhibits CAS2 (24).

The coupled assays employing all four enzymes, using L-arginine and D/L-glyceraldehyde-3-phosphate as starting substrates, were only partially successful. The major product observed was proclavaminic acid, which was not efficiently processed further by CAS2, under the conditions tested. Attempts to couple BLS, PAH, and CAS2 using N\(^2\)-(2-carboxyethyl)arginine as the initial substrate were similarly unsuccessful. In this instance, the major product observed was deoxyproclavaminic acid, produced by the hydrolysis of the guanidino group of deoxyguanidinopropclavaminic acid as catalyzed by PAH (21, 28). This is consistent with the observation that deoxyproclavaminic acid has been reported to be a relatively poor substrate for CAS2 (24). The incubations of BLS and PAH with N\(^2\)-(2-carboxyethyl)arginine also produced deoxyproclavaminic acid as the major product.

The results of the coupled reactions are consistent with previous observations and support the accepted sequence of reactions in the biosynthesis of clavaminic acid (Fig. 1 b). However, they were not useful for providing sufficient clavaminic acid for ORF17 assays. Alternatively, clavaminic acid was produced by CAS2-mediated cyclization and desaturation of synthetic racemic proclavaminic acid (Fig. 4 a and b).

Clavaminic acid produced by incubation of CAS2 with proclavaminic acid was incubated with ORF17, ATP, and various potential acyl and formyl donors including glycine, D/L-alanine, L-valine, L-leucine, L-aspartic acid, L-asparagine, N-\(\alpha\)-acetyl-glycine, N-glycyl-glycine, acetyl coenzyme A, bicarbonate, acetate, acetic acid, formate, and proclavaminic acid. No acylated products were observed by LC-MS analyses except for the combination of ORF17, ATP, and glycine, where a mass corresponding to glycl-clavaminic acid was observed (Fig. 4 c) at the same retention time as clavaminic acid. Introduction of 1,2-[\(^{13}\)C]-labeled glycine with clavaminic acid in the presence of ORF17 and ATP resulted in a product with the expected retention time and a mass increase of 2 Da relative to the unlabeled material.

As glycine and clavaminic acid both contain carboxylate and amine functionalities, two possible products can be envisaged from ORF17-mediated carboxylate-amine ligation. Glycine could be attached via its carboxylate group to the C-9 amino group to give N-glycl-clavaminic acid (Fig. 5 a) or via its amino group to the C-3 carboxylate group of clavaminic acid to give carboxyl-linked glycl-clavaminic acid (C-glycl-clavaminic acid, Fig. 5 b).

Imidazole derivatization of clavaminic acid involves formation of fragments arising via opening of both rings of the bicycle followed by decarboxylation of the C-3 carboxylate (11, 35), which should not be observed if glycine was attached at the C-3 position. The absence of a mass at 324 Da [M+H]\(^+\) corresponding to the fragment of derivatized C-glycl-clavaminic acid (Fig. 5 a) or via its amino group to the C-3 carboxylate group of clavaminic acid to give carboxyl-linked glycl-clavaminic acid (C-glycl-clavaminic acid, Fig. 5 b).

Glycine functions as the carboxylate acceptor, and clavaminic acid functions as the amine donor. The product from the incubation of 1,2-[\(^{13}\)C]-labeled glycine with clavaminic acid was similarly imidazole-derivatized to give a fragment with the anticipated predicted mass for 1,2-[\(^{13}\)C]-labeled N-glycl-clavaminic acid (282 Da [M+H]\(^+\)). Insufficient material was produced to assign the absolute stereocchemistry of the product, and it is drawn as 3S,5S in the schemes. Given the nature of the ORF17 reaction, the determined stereochemistry of N-acytelyl-glycl-clavaminic acid and N-glycl-clavaminic acid isolated from non-clavulanic acid-producing mutants of *S. clavuligerus* (33), and the determined stereochemistry of clavaminic acid (10, 18–20), this is likely to be correct. However, since an epimerization must occur in the pathway, at the ORF17 stage or later stages, the assignment is provisional.

There was no evidence for the production of N-glycl-proclavaminic acid or N-glycl-dihydroclavaminic acid upon incubation of ORF17, ATP, and glycine with the crude products of the CAS assay. Separation...
of dihydroclavaminic acid and clavaminic acid on a preparative scale to allow each component to be assayed separately was not possible, due to the small amounts of unstable dihydroclavaminic acid available (22, 24). Thus, the CAS2 incubation mixture was filtered through a 10-kDa molecular mass cut off membrane to remove CAS2, and the filtrate was then incubated with ORF17, ATP, and glycine (Fig. 4).

N-Glycyl-clavaminic acid was observed with no evidence for the production of N-glycyl-dihydroclavaminic acid. The absence of CAS2 in this reaction mixture eliminates the possibility that N-glycyl-clavaminic acid is produced via N-glycyl-dihydroclavaminic acid, which might then be desaturated by CAS2 to give N-glycyl-clavaminic acid (Fig. 6). These results demonstrate that clavaminic acid is a substrate for ORF17. However, there remains the possibility, albeit unlikely, that N-glycyl-dihydroclavaminic acid could represent an alternative in vivo route to produce N-glycyl-clavaminic acid or that there is an alternative substrate for ORF17.

Comparison of the ATPase activity, observed using the pyruvate kinase/lactate dehydrogenase ATPase UV assay, with the rate of ATP hydrolysis in the presence of clavaminic acid and glycine, revealed a 20-fold increase in the rate of ATP hydrolysis. The initial amount of

FIGURE 4. LC-MS analyses of ORF17 incubations. a, LC-MS total ion count elution profile of proclavaminic acid, displaying the mass spectrum of peak at 1.80 min for proclavaminic acid m/z (ESI⁺): 203 Da [M+H]+. b, LC-MS total ion count elution profile of proclavaminic acid incubated with CAS2, displaying the mass spectrum of peak at 1.99 min for clavaminic acid, m/z (ESI⁺): 199 Da [M+H]+. c, LC-MS total ion count elution profile of proclavaminic acid incubated with CAS2 and then ORF17 (with ATP and glycine), displaying the mass spectrum of peak at 1.99 min for N-glycyl-clavaminic acid, m/z (ESI⁺): 256 Da [M+H]+. d, LC-MS total ion count elution profile of the imidazole-derivatized sample of proclavaminic acid incubated with CAS2 and then ORF17, displaying the mass spectrum of peak at 2.54 min for clavaminic acid m/z (ESI⁺), 223 Da [M+H]+, and the mass spectrum of peak at 3.26 min for N-glycyl-clavaminic acid m/z (ESI⁺), 280 Da [M+H]+, 212 Da [M-imidazole]+.
Clavaminic acid present can only be estimated based on CAS2 turnover of the proclavaminic acid; therefore, accurate kinetic data, in the presence of substrates, could not be obtained.

A range of potential alternative carboxylate substrates were then tested as ORF17 substrates, in the presence of ATP and clavaminic acid, using the pyruvate kinase/lactate dehydrogenase ATPase assay. The common proteinogenic α-amino acids (except l-cysteine, l-tryptophan, and l-tyrosine due to poor solubility under these conditions), bicarbonate, acetate, formate, acetyl coenzyme A, D-alanine, D-serine, glycolic acid, propanolic acid, N-glycyl-glycine, and N-acetyl-glycine failed to stimulate significant ORF17 activity in the presence of clavaminic acid, under standard assay conditions. Alternative clavam biosynthetic intermediates, other than clavaminic acid, were also tested as potential amine substrates in the presence of glycine in addition to the potential analogues of clavam biosynthetic intermediates mentioned earlier. D/L-Ornithine, N-α-acetyl-l-ornithine, l-lysine, and N-α-acetyl-l-lysine, N²-(2-carboxylethyl)arginine, deoxyguanidinoproclavaminic acid, guanidinoproclavaminic acid, deoxyproclavaminic acid, proclavaminic acid, and clavulanic acid all failed to stimulate detectable activity of ORF17 in the presence of Mg²⁺, K⁺, ATP, and glycine.

**DISCUSSION**

Gene disruption studies (32) have indicated that the orf17 gene product is essential for clavulanic acid production. Sequence analyses and modeling led to the proposal that ORF17 is a member of the ATP-grasp fold superfamily catalyzing a carboxylate-amine ligation reaction (Fig. 1d). This proposal was further supported by the use of soft ionization mass spectrometry that demonstrated binding of ATP to ORF17 and by the (low level) ORF17 catalyzed hydrolysis of ATP to ADP and phosphate.

A problem in defining complex biosynthetic pathways is the availability of potential enzyme substrates. This problem is severe in pathways where potential intermediates are labile or difficult to access by synthesis, as in the case of the latter stages of the clavulanic acid pathway. In this study, we addressed this problem through the use of enzymes from the earlier stages of the pathway leading to clavaminic acid to prepare candidate substrates for ORF17. We did not prepare enough clavaminic acid for conclusive assays from the primary metabolic precursors of the pathway or from early stage intermediates, by full in vitro reconstruction of the pathway leading to clavulanic acid. However, these results supported the proposed sequence of events in clavulanic acid biosynthesis leading to clavaminic acid. Sufficient clavaminic acid for ORF17 activity assays was prepared by CAS2-mediated conversion of synthetic proclavaminic acid to clavaminic acid. LC-MS analyses then demonstrated that the shorter of the two proposed forms of ORF17 (32) catalyzes the production of N-glycyl-clavaminic acid from ATP, glycine, and the semisynthetic (3S,5S)-clavaminic acid in the presence of Mg²⁺ and K⁺.

The only available intermediate from the clavulenic acid pathway that stimulated ORF17 activity was clavaminic acid. N²-(2-carboxylethyl)-arginine, deoxyguanidinoproclavaminic acid, deoxyproclavaminic acid, guanidinoproclavaminic acid, proclavaminic acid, and clavulanic acid did not stimulate activity under the assay conditions used. Furthermore, glycine, but not N-acetyl-glycine, was shown to be a substrate for ORF17. Although the results do not prove the identity of the natural substrate of ORF17 within cells, they demonstrate that ORF17 can act as an N-glycyl-clavaminic acid synthetase (GCAS). Although insufficient clavaminic acid was available for the assignment of the stereochemistry or detailed kinetic analyses, the results indicated that the mechanism of GCAS likely proceeds via an enzyme-bound O-glycyl-phosphate inter-
mediate analogous to the proposed D-alanyl phosphate intermediate for the D-alanine-D-alanine ligase subfamily (59, 60). GCAS also appeared in contrast to the D-alanine-D-alanine ligases, which, although they are specific for D-amino acids, do display some flexibility in their substrate specificity. For example, the Streptococcus faecalis D-alanine-D-alanine ligase was found to accept d-aminoacryluric acid in place of alanine as a carboxylic substrate but accepts a number of other D-amino acid substrates as the nucleophilic substrate, including serine, threonine, norvaline, and at high substrate concentrations, glycine (61).

N-Glycyl-clavamic acid and N-acetyl-glycyl-clavamic acid have both been observed in the supernatants obtained from orf15 and orf16 disruption mutants (32), but proof that they are actual intermediates in clavamic acid biosynthesis is still lacking. This proof is challenging to obtain since the glycyl component of the intermediate must be lost en route to clavamic acid (Fig. 1h). The observation that GCAS catalyzes conversion of clavamic acid to N-glycyl-clavamic acid, but was not observed to catalyze the formation of N-acetyl-glycyl-clavamic acid from N-acetyl-glycine and clavamic acid, suggested that the pathway may proceed via clavamic acid to N-glycyl-clavamic acid and then to N-acetyl-glycyl-clavamic acid.

Acknowledgments—We thank Dr. Z. H. Zhang for the cas2/pe21a expression vector, Dr. N. J. Oldham for assistance with soft ionization mass spectrometry analysis, and Dr. M. McDonough for assistance with multiple alignments and modeling.

REFERENCES

1. Brown, A. G., Butterworth, D., Cole, M., Hanscomb, G., Hood, J. D., Reading, C., and Rolinson, G. N. (1976) J. Antimicrob. (Tokyo) 29, 668–669
2. Baggaley, K. H., Brown, A. G., and Schofield, C. J. (1997) Nat. Prod. Rep. 14, 309–333
3. Townsend, C. A. (2002) Curr. Opin. Chem. Biol. 6, 583–589
4. Mellado, E., Lorenzana, L. M., Rodriguez-Saiz, M., Diez, B., Liras, P., and Barredo, J. L. (2002) Microbiology (Reading) 148, 1427–1438
5. Li, R., Khaleeli, N., and Townsend, C. A. (2000) J. Bacteriol. 182, 4087–4095
6. Hodgson, J. E., Fosberry, A. P., Rawlinson, N. S., Ross, H. N., Schofield, C. J., Earl, A. J., and Lawlor, E. J. (1995) Gene (Amst.) 166, 49–55
7. Jensen, S. E., Alexander, D. C., Paradkar, A. S., and Aidoo, K. A. (1993) in Microorganisms: Basic and Applied Molecular Genetics (Baltz, R. H., Hegeman, G. D., and Tyler, J. W.) 743–754
8. Galperin, M. Y., and Koonin, E. V. (1997) Nucleic Acids Res. 25, 565–573
9. N-Glycyl-clavaminic Acid Synthetase

10. Brown, A. G., Butterworth, D., Cole, M., Hanscomb, G., Hood, J. D., Reading, C., and Rolinson, G. N. (1976) J. Antimicrob. (Tokyo) 29, 668–669
11. Baggaley, K. H., Brown, A. G., and Schofield, C. J. (1997) Nat. Prod. Rep. 14, 309–333
12. Townsend, C. A. (2002) Curr. Opin. Chem. Biol. 6, 583–589
13. Mellado, E., Lorenzana, L. M., Rodriguez-Saiz, M., Diez, B., Liras, P., and Barredo, J. L. (2002) Microbiology (Reading) 148, 1427–1438
14. Li, R., Khaleeli, N., and Townsend, C. A. (2000) J. Bacteriol. 182, 4087–4095
15. Hodgson, J. E., Fosberry, A. P., Rawlinson, N. S., Ross, H. N., Schofield, C. J., Earl, A. J., and Lawlor, E. J. (1995) Gene (Amst.) 166, 49–55
16. Jensen, S. E., Alexander, D. C., Paradkar, A. S., and Aidoo, K. A. (1993) in Microorganisms: Basic and Applied Molecular Genetics (Baltz, R. H., Hegeman, G. D., and Tyler, J. W.) 743–754
17. Galperin, M. Y., and Koonin, E. V. (1997) Nucleic Acids Res. 25, 565–573
18. N-Glycyl-clavaminic Acid Synthetase

19. Brown, A. G., Butterworth, D., Cole, M., Hanscomb, G., Hood, J. D., Reading, C., and Rolinson, G. N. (1976) J. Antimicrob. (Tokyo) 29, 668–669
20. Baggaley, K. H., Brown, A. G., and Schofield, C. J. (1997) Nat. Prod. Rep. 14, 309–333
21. Townsend, C. A. (2002) Curr. Opin. Chem. Biol. 6, 583–589
22. Mellado, E., Lorenzana, L. M., Rodriguez-Saiz, M., Diez, B., Liras, P., and Barredo, J. L. (2002) Microbiology (Reading) 148, 1427–1438
23. Li, R., Khaleeli, N., and Townsend, C. A. (2000) J. Bacteriol. 182, 4087–4095
24. Hodgson, J. E., Fosberry, A. P., Rawlinson, N. S., Ross, H. N., Schofield, C. J., Earl, A. J., and Lawlor, E. J. (1995) Gene (Amst.) 166, 49–55
25. Jensen, S. E., Alexander, D. C., Paradkar, A. S., and Aidoo, K. A. (1993) in Microorganisms: Basic and Applied Molecular Genetics (Baltz, R. H., Hegeman, G. D., and Tyler, J. W.) 743–754
26. Galperin, M. Y., and Koonin, E. V. (1997) Nucleic Acids Res. 25, 565–573