Crystal Structure of Carbapenam Synthetase (CarA)

Matthew T. Miller‡, Barbara Gerratana§, Anthony Stapon§, Craig A. Townsend§, and Amy C. Rosenzweig‡,¶

‡Departments of Biochemistry, Molecular Biology, and Cell Biology and of Chemistry, Northwestern University, Evanston, Illinois 60208; §Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland 21218

¶To whom correspondence may be addressed. Tel.: 847-467-5301; Fax: 847-467-6489; E-mail: amyr@northwestern.edu

Running title: Crystal structure of CarA

The atomic coordinates and structure factors (codes 1Q15 and 1Q19) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

* This work was supported by funds from the David and Lucile Packard Foundation (to A. C. R.) and National Institutes of Health Grant AI14937 (to C. A. T.). M. T. M. was supported in part by National Institutes of Health training grant GM 008061.
Abbreviations used: CMPr, (2S,5S)-5-carboxymethylproline; β-LS, β-lactam synthetase; AS-B, class B asparagine synthetase; AMP-CPP, α,β-methyleneadenosine 5′-triphosphate; CEA, N²-(carboxyethyl)-L-arginine; CM A, N²-(carboxymethyl)-L-arginine.
SUMMARY

Carbapenam synthetase (CarA) is an ATP/Mg\(^{2+}\)-dependent enzyme that catalyzes formation of the \(\beta\)-lactam ring in (5R)-carbapenem-3-carboxylic acid biosynthesis. CarA is homologous to \(\beta\)-lactam synthetase (\(\beta\)-LS), which is involved in clavulanic acid biosynthesis. The catalytic cycles of CarA and \(\beta\)-LS mediate substrate adenylation followed by \(\beta\)-lactamization via a tetrahedral intermediate or transition state. Another member of this family of ATP/Mg\(^{2+}\)-dependent enzymes, asparagine synthetase (AS-B), catalyzes intermolecular, rather than intramolecular, amide bond formation in asparagine biosynthesis. The crystal structures of apo CarA and CarA complexed with the substrate, (2S,5S)-5-carboxymethylproline (CMPr), ATP analog \(\alpha,\beta\)-methyleneadenosine 5'-triphosphate (AMP-CPP), and a single Mg\(^{2+}\) ion have been determined. CarA forms a tetramer. Each monomer resembles \(\beta\)-LS and AS-B in overall fold, but key differences are observed. The N-terminal domain lacks the glutaminase active site found in AS-B, and an extended loop region not observed in \(\beta\)-LS or AS-B is present. Comparison of the C-terminal synthetase active site to that in \(\beta\)-LS reveals that the ATP binding site is highly conserved. By contrast, variations in the substrate binding pocket reflect the different substrates of the two enzymes. The Mg\(^{2+}\) coordination is also different. Several key residues in the active site are conserved between CarA and \(\beta\)-LS, supporting proposed roles in \(\beta\)-lactam formation. These data provide further insight into the structures of this class of enzymes and suggest that CarA might be a versatile target for protein engineering experiments aimed at developing improved production methodologies and new carbapenem antibiotics.
The carbapenem class of β-lactam antibiotics exhibits a broad spectrum of activity and is resistant to inactivation by β-lactamases (1). The core structure of carbapenems consists of a four-membered β-lactam ring fused to a five-membered ring containing only carbon atoms (2). Chemical modification of naturally-occurring carbapenems, which include olivanic acids from Streptomyces olivaceus and thienamycin from Streptomyces cattleya (2,3), has led to the development of successful drugs such as imipenem or meropenem, used to treat hospital acquired infections (1,4). Commercial fermentation of carbapenems is limited by low yields and chemical instability, however (2,4). Synthetic processes have been developed, but are expensive and not as efficient as the semisynthetic procedures used to generate penicillins and cephalosporins (4,5). A detailed understanding of carbapenem biosynthesis is thus crucial to developing improved production methodologies as well as new antibiotics.

Whereas knowledge of carbapenem biosynthesis by the streptomycetes remains limited, the genes encoding carbapenem biosynthetic enzymes have been cloned from two Gram negative bacteria, Erwinia carotovora (6) and Serratia marcescens (7). Both of these organisms produce the simplest carbapenem, (5R)-carbapen-2-em-3-carboxylic acid (8). Although this compound itself is not clinically useful, it contains the same nucleus as the thienamycin derivatives, suggesting that its biosynthesis could be engineered to generate other carbapenem antibiotics or semisynthetic intermediates. Five proteins, encoded by the carA-E genes, are involved in (5R)-carbapen-2-em-3-carboxylic acid biosynthesis, and three, encoded by carA-C and called CarA, CarB, and CarC, are absolutely essential (9,10). CarB is proposed to generate (2S,5S)-5-carboxymethylproline (CMPr)\(^1\) from glutamate semialdehyde/pyrroline 5-carboxylate and acetyl-CoA (2,11). Formation of the β-lactam ring to yield (3S,5S)-carbapenam-3-carboxylic acid is then catalyzed by CarA (Fig. 1) (10,12), followed by epimerization and desaturation by CarC to
produce (3S,5R)-carbapenam-3-carboxylic acid and (5R)-carbapen-2-em-3-carboxylic acid (5,10,11).

Both CarC and CarA are related to enzymes in the clavulanic acid biosynthetic pathway (2). CarC is weakly homologous to clavaminate synthase, a nonheme iron a-ketoglutarate-dependent enzyme that catalyzes hydroxylation, cyclization, and desaturation reactions in clavulanic acid biosynthesis (13,14). The crystal structure of CarC reveals a jellyroll fold characteristic of this enzyme family and details of Fe(II), a-ketoglutarate, and substrate binding (5). CarA is homologous to b-lactam synthetase (b-LS), an ATP/Mg²⁺-dependent enzyme that catalyzes formation of the b-lactam ring in clavulanic acid (Fig. 1) (15,16). Based on kinetic (17) and structural (18,19) data, the b-LS mechanism is proposed to involve substrate adenylation followed by b-lactamization via a tetrahedral intermediate or transition state. The chemistry performed by CarA is similar. Its ability to catalyze b-lactam formation in the presence of ATP and Mg²⁺ has been demonstrated in vitro, and activity assays with alternative substrates are consistent with formation of an adenylated intermediate (12).

The sequence homology and mechanistic parallels between CarA and b-LS extend to class B asparagine synthetase (AS-B), which converts aspartic acid to asparagine (20). Whereas CarA and b-LS catalyze intramolecular amide bond formation, the AS-B reaction involves intermolecular amide bond formation between adenylated aspartic acid and ammonia. AS-B contains two active sites, an N-terminal glutaminase site where ammonia is generated by hydrolysis of glutamine and a C-terminal synthetase site where aspartic acid is adenylated and then reacts with ammonia to form asparagine. The two active sites are connected by an extended tunnel (21). Although the AS-B N-terminal domain is conserved in b-LS, glutaminase activity is not observed (15), consistent with the absence of crucial catalytic residues and the tunnel
between domains (18). The differences between the b-LS and AS-B N-terminal domains together with variations in the C-terminal active sites reveal how the two enzymes accommodate b-lactam formation versus asparagine synthesis (18). To further investigate the evolutionary connections among this family of ATP/Mg\(^{2+}\)-dependent amide synthesizing enzymes, we have determined the X-ray structure of CarA in the absence of exogenous ligands (apo) and in the presence of substrate, CMPr, ATP analog, \(\alpha,\beta\)-methyleneadenosine 5'-triphosphate (AMP-CPP), and a single Mg\(^{2+}\) ion.

**EXPERIMENTAL PROCEDURES**

Protein overexpression and purification. CarA for initial crystallization trials and soaking experiments was expressed as described previously (12). To generate selenomethionine-substituted protein, the expression vector pET24a/carA (10) was transformed into Escherichia coli strain B834(DE3) (Novagen). Colonies from this transformation were used to inoculate two 5 ml tubes of LB medium containing 50 mg/ml kanamycin. After 16 hrs at 37 °C, 5 ml were used to inoculate a 100 ml culture in LB. This culture was incubated for six hours, and then the cells were split into two 500 ml cultures of M9 minimal medium supplemented with amino acids, vitamins, 40 mg/l selenomethionine, and 50 mg/l kanamycin (22). The cultures were grown in 2 l baffled Erlenmeyer flasks for 16 hrs and then transferred to a fermenter containing 10 l of the same medium supplemented with 200 ml 20% (w/v) glucose. At an O\(D_{600}\) of 0.9, 1 mM isopropyl-\(\beta\)-D-thiogalactopyranoside (IPTG), 200 ml 20% (w/v) glucose, and 2 ml antifoam (Sigma-Aldrich) were added. During the induction period, the impellor speed was increased to 750 rpm, the airflow rate was maximized, and a pH of 7 was maintained by manual addition of ammonium hydroxide. Three hours after induction, at an O\(D_{600}\) of 2.0, the cells were harvested.
by centrifugation at 6000 x g for 15 minutes. The cell paste was frozen in liquid nitrogen and stored at -80 °C.

The purification protocols for both native and selenomethionine-substituted CarA were based on described procedures (12). Thawed cell paste (~50 g) was resuspended in 200 ml of 100 mM Tris, pH 8.0, 1.8 mM EDTA, and 1 mM DTT. While stirring on ice, 0.75 mg/ml lysozyme (Research Organics) was added to the solution followed 10 min later by 0.1 mg/ml DNase I (Research Organics) and MgCl₂ to a final concentration of 1 mM. The solution was sonicated at 4 °C using a Sonic Dismembrator (VWR) with a 0.5 inch probe for 5 min of 1 sec pulses at power level 7. The sonicated suspension was then centrifuged for 20 min at 10,000 x g at 4 °C. After 35% and 65% ammonium sulfate fractionation, the pellet from the 65% cut was resuspended in 50 mM HEPES, pH 7.5, 1 mM DTT. This solution was transferred to a 10,000 molecular weight cut off SnakeSkin dialysis membrane (Pierce) and dialyzed against 1 L of 50 mM HEPES, pH 7.5, 1 mM DTT at 4 °C for 12 hrs with three changes of buffer. The dialyzed sample was then loaded onto a 53 ml High Load 16/10 Q Sepharose Fast Flow column (Amersham Biosciences). After washing with four column volumes of 50 mM HEPES, pH 7.5, 5% (v/v) glycerol, and 1 mM DTT, CarA was eluted with a 0-1 M NaCl 10 column volume gradient. CarA-containing fractions were identified by SDS-PAGE, pooled, applied in 5 ml aliquots to a 175 ml Superdex 200 column (Amersham Biosciences), and eluted at 1 ml/min with 50 mM HEPES, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, and 1 mM DTT. Purified CarA was concentrated to 13 mg/ml using an Amicon stirred cell with a YM-10 membrane followed by a Centriprep YM-10 (Millipore), and stored at 4 °C. The protein concentration was determined using the Bradford assay calibrated with a bovine serum albumin standard. The native molecular mass of purified CarA was estimated by gel filtration chromatography using a Sephacryl
S300HR column calibrated with the following protein standards (Amersham Biosciences): ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), albumin (67 kDa), ovalbumin (43 kDa), and chymotrypsinogen A (25 kDa). Each protein was eluted with 50 mM Tris, pH 7.5, 100 mM NaCl at 0.4 ml/min at 4 °C.

Crystallization and data collection. Initial crystallization trials for both native and selenomethionine-substituted CarA were conducted using the Hampton Research Crystal Screens HT™ and Index HT™ in 96 well sitting drop trays. Crystallization drops containing 1 ml protein and 1 ml of well solution were equilibrated versus 50 ml of each screen solution. The optimized precipitant solution for selenomethionine-substituted protein consisted of 30% (w/v) PEG 3350, 15% (v/v) glycerol, and 100 mM Na citrate, pH 5.3. For data collection, these crystals were transferred to a cryosolution composed of 30% (w/v) PEG 3350, 25 mM HEPES, pH 7.5, 50 mM Na citrate, pH 5.3, 15% (v/v) glycerol, and 250 mM NaCl. After a 1-5 min soak, the crystals were flash cooled in liquid nitrogen. Crystals of native CarA were obtained by the hanging drop method using 23-26 % (w/v) PEG 4000, 100 mM Na citrate, pH 5.6, 210-250 mM ammonium acetate, and 100 mM cesium chloride. Microseeding was required to produce sufficiently large crystals. Soaking was performed in a solution containing 26% (w/v) PEG 4000, 100 mM Na citrate, pH 5.6, 50 mM ammonium acetate, 20% (v/v) glycerol, 100 mM cesium chloride, 5 mM AMP-CPP, 5 mM MgCl₂, and 5 mM CMPr. CMPr was synthesized as described previously (12). After 6 hours, crystals were flash cooled in liquid nitrogen. All data were collected at −160 °C by using an ADSC Quantum 315 CCD detector at Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9-1. Data for multiwavelength anomalous dispersion (MAD) phasing were collected using the inverse beam technique. Both the selenomethionine-substituted and native CarA crystals belong to the space group P2₁ with unit cell dimensions of a = 61.1 Å,
b = 180.4 Å, c = 96.6 Å, \( b = 97.5^\circ \) and \( a = 103.0 \) Å, \( b = 61.9 \) Å \( c = 175.9 \) Å, \( b = 97.0^\circ \), respectively. Data sets were processed with MOSFLM (23) and SCALA (24).

Structure determination. Of 24 possible selenium positions, 20 were located using the program SOLVE, and an initial model was built with RESOLVE (25). The asymmetric unit contains four monomers labeled A, B, C, and D. The initial model from RESOLVE included 1,453 of a possible 2,012 residues and served as a starting point for manual model building with XtalView (26). Initially, only monomer A was built and refined with iterative cycles of simulated annealing and individual B-factor refinement using CNS (27). Once monomer A was complete, it was used to replace the partially modeled B, C and D monomers using the program LSQM AN (28), and iterative refinement and model adjustment were continued. The final model consists of residues A2-A441, A451-A501, B2-B441, B451-B501, C2-C441, C451-C501 and D2-D441, D468-D500, and 312 water molecules. The missing residues were not visible in the electron density map. Ramachandran plots generated with PROCHECK (29) indicate that the model exhibits good geometry with 90.7% of the residues in the most favored region and all others in the additionally allowed regions.

The structure of selenomethionine-substituted CarA was used as a starting model for refinement of the structure of native CarA in the presence of AMP-CPP, Mg\(^{2+}\), and CMPr. All four monomers in the asymmetric unit were located with molecular replacement using MOLREP (30). After rigid body refinement with CNS (27), \( F_o - F_c \) difference maps at the active sites revealed positive electron density that was modeled as AMP-CPP, Mg\(^{2+}\), and CMPr. The structure was refined by multiple rounds of simulated annealing and individual B-factor refinement with CNS (27) and model rebuilding with XtalView (26). The final model includes residues 2-501 for each of the four monomers and 399 water molecules. All four active sites
contain one molecule each of AMP-CPP, Mg\(^{2+}\), and CMPr. The model exhibits good geometry according to Ramachandran plots generated using PROCHECK (29) with 88.4\% of the residues in the most favored regions and all others in the additionally allowed regions. Buried surface area calculations were performed with CNS (27), and figures were generated with MOLSCRIPT (31), PyMOL (32), and Raster3D (33).

**RESULTS AND DISCUSSION**

Overall fold. CarA comprises two domains (Fig. 2a). The N-terminal domain includes residues 2-207, and the C-terminal domain consists of residues 208–501. The N-terminal domain is composed of two antiparallel six-stranded \(\beta\)-sheets that form a sandwich, flanked on each side by two \(\alpha\)-helices. This fold is quite similar to that of the \(\beta\)-LS (18) and AS-B (21) N-terminal domains (Fig. 3). In \(\beta\)-LS, one of the sheets is five-stranded and is covered by disordered loop regions. These regions correspond to an additional strand and two helices in CarA. In AS-B, the N-terminal domain houses the glutaminase active site, which includes the catalytically important N-terminal cysteine residue and other residues involved in glutamine binding. The cysteine is replaced with a phenylalanine in \(\beta\)-LS, and nine additional N-terminal residues fill the space corresponding to the AS-B glutaminase site. The CarA N-terminal residue, a serine, occupies approximately the same position as the AS-B cysteine, but several key structural elements that constitute the AS-B glutaminase site are missing. First, an extended loop comprising AS-B residues 49-59 is not conserved in CarA (Fig. 3). Second, residues 224-228 from the AS-B C-terminal domain extend toward the glutamine binding site whereas the equivalent residues in CarA are shifted back toward the second domain. Finally, specific residues that interact with glutamine in AS-B are not conserved in CarA. As a result, there is no
glutamine binding pocket in CarA. The N-terminal domain of CarA is further distinguished from those of the other two enzymes by the presence of an extended loop region encompassing residues 148-173 (Fig. 2). These residues interact with the C-terminal domain and are completely absent in b-LS and AS-B (Fig. 3). Part of the equivalent region in AS-B is occupied by the C-terminal residues of the second domain.

An extended linker region connects the N-terminal and C-terminal domains (Fig. 2a). The interface between the two domains is large with ~3800 Å² buried surface area. The C-terminal domain consists of a five-stranded parallel β-sheet and 14 α-helices, six on one side and eight on the other. Two additional short parallel β-strands are located on the exterior of the protein near the domain linker. The active site is situated in a cleft between four β-strands and five α-helices, in the same position as the b-LS and AS-B active sites. Structural comparisons of b-LS and AS-B revealed several features that allow the b-LS active site to accommodate the elongated N²-(carboxyethyl)-L-arginine (CEA) molecule (~12 Å) as opposed to the more compact aspartic acid (~5 Å) (18). One such feature that is also observed in CarA is the presence of a loop at residues 377-388 (Fig. 2a). The analogous residues in AS-B form a helix, constricting the active site. Several additional adjustments in the CarA active site are consistent with the intermediate size of CMPr (~8 Å). Residues 371-375 are shifted slightly toward the active site in CarA as compared to b-LS. In addition, the Arg 374 and Ile 375 side chains point toward the active site in contrast to their b-LS counterparts, Phe 377 and Asp 378, the side chains of which are oriented away from the active site. A loop comprising residues 442-450 covers the active site in the substrate structure (Fig. 2a), but is not observed in the apo structure. Part of this region, spanning residues 445-451, forms a 3₁₀ helix.
Subunit interactions. In both structures, the four monomers in the asymmetric unit are very similar and can be superimposed with root mean square (r.m.s.) deviations of 0.2-0.5 Å for Cα atom coordinates. The four monomers form a tetramer with a hole in the center (Fig. 2b). There are extensive contacts between the A and B subunits, also observed for the D and C subunits, and between the A and D subunits, repeated between B and C (Table 2). The buried surface area at these two types of interfaces exceeds 2000 Å², and is within the range expected for a stable complex (34). Although native gel analysis indicated that CarA is a dimer in solution (12), the molecular mass measured by calibrated gel filtration chromatography is 235 kDa. This value is consistent with the calculated molecular mass for a homotetramer of 224 kDa and the tetramer observed in the crystal. Both b-LS (18) and AS-B (21) form dimers, but not tetramers or higher order species.

The interactions between the A and B monomers primarily involve residues from one side of the N-terminal domain (Fig. 2b). Several key hydrogen bonds are formed between residues from the last strand in the N-terminal domain and the linker region. For example, the side chain of Asp A191 interacts with the side chains of Lys B181 and Tyr B373 and vice versa. In addition, the side chain of Glu A199 is hydrogen bonded to the backbone nitrogen of Thr B202. Other important interactions derive from the loop encompassing residues 148-173, including a salt bridge between Glu A156 and Lys B175. The involvement of the domain linker and the head-to-tail arrangement of the two monomers are similar to the b-LS dimer, but the N-terminal part of the linker forms the interactions in CarA whereas the C-terminal portion is important in b-LS (18). As a result, the monomers are shifted with respect to one another in the two enzymes.
The interactions between the A and D monomers include residues 63-78 from two helices in the N-terminal domain, residues from the domain linker region, and a few residues from the C-terminal domain (Fig. 2b). In particular, the carbonyl oxygen of Gly A70 is linked to the side chain of Ser D404, the side chains of Glu A73 and Glu A75 interact with Ser D208, and the carbonyl oxygen of Arg A91 is hydrogen bonded to the side chain of Asn D95. Two tyrosine residues, Tyr A77 and Tyr D226, are also within hydrogen bonding distance. Contacts between the A and C and between the B and D monomers are minimal and only involve Glu 49 and Arg 50 from all four monomers. A third residue, Asn A112, points toward the interface between A and C, but away from that between B and D, accounting for the size difference. The interface between B and D is smaller in the substrate structure because Arg D50 is directed toward the hole in tetramer center rather than toward the interacting monomer (Table 2).

The active site. Soaking crystals of apo CarA in solutions containing CMPr, AMP-CPP, and MgCl₂ resulted in an active site with a full complement of ligands (CarA/CMPr/AMP-CPP structure) (Fig. 4). The substrate is located in a hydrophobic pocket lined by residues Leu 319, Ile 323, Leu 349, and Ile 354 (Fig. 5a). The presence of a hydrophobic site was predicted from kinetic studies with alternate substrates (12). The CMPr molecule was initially positioned with its carboxymethyl group directed toward the AMP-CPP a-phosphate by analogy to the b-LS structures (18,19). Subsequent refinement and careful inspection of simulated annealing omit maps indicated that it is best modeled in the opposite orientation, however, with the carboxymethyl group, which is not as well ordered as the rest of the CMPr molecule, pointing away from the AMP-CPP. One of the carboxylate oxygen atoms is within hydrogen bonding distance of the side chain of Gln 371 in three of the four monomers (Fig. 5a). In monomers A and C, the side chain of Arg 374 is also within 4 Å. The b-LS counterpart to Arg 374 is Phe 377,
probably because an arginine in this position would be too close to the CEA guanidinium group. One α-carboxylate oxygen atom is hydrogen bonded to a water molecule (Fig. 5a), present in all four monomers, which is linked to the carbonyl oxygen of Gly 344, the backbone nitrogen of Asp 348, and one of the AMP-CPP α-phosphate oxygen atoms. This orientation of CMPr is clearly not a productive conformation for β-lactam formation. Notably, the second α-carboxylate oxygen atom of CMPr is within 3.5 Å of the AMP-CPP α-phosphate oxygen atom in all four monomers, suggesting that the CMPr α-carboxylate may be protonated in the pH 5.6 soaking solution. Such a pH effect could account for the presence of this aborted substrate complex in the crystal. Additional stabilization of the unproductive complex might derive from the interactions of Gln 371 and Arg 374 with the carboxymethyl group.

The AMP-CPP molecule is well ordered in all four monomers. The adenosine hydroxyl groups interact with carbonyl oxygen of Pro 244 and the amide nitrogens of Gly 344 and Tyr 345. The adenine nitrogens are hydrogen bonded to the backbone oxygen and nitrogen atoms of Ile 270. The corresponding residues in β-LS, Val 247, Gly 347, Tyr 348, and Met 273 (Fig. 3) participate in a similar set of interactions. A single Mg²⁺ ion is coordinated by terminal oxygen atoms from the α- and γ-phosphates, the oxygen atom bridging the β- and γ-phosphates, two water molecules, and the carbonyl oxygen atom of Ile 444 (Fig. 5a). The water molecules are not consistently present in all four monomers. The position of this Mg²⁺ ion is different from that observed in the structure of β-LS with CEA and AMP-CPP (β-LS/CEA/AMP-CPP) (18). In β-LS, the Mg²⁺ ion is located on the opposite side of the AMP-CPP phosphorus atoms and is coordinated by side chain oxygens from two aspartic acid residues as well as a water molecule and phosphate oxygen atoms (Fig. 5b). Notably, the water molecule, which forms a hydrogen bond with the β-carboxylate of CEA, is also observed in CarA, interacting with the substrate
despite the absence of the Mg$^{2+}$ ion. Although the equivalent aspartic acid residues are present in CarA, Asp 250 and Asp 348, they do not coordinate a Mg$^{2+}$ ion (Fig. 5b). By contrast, both possible Mg$^{2+}$ binding sites are occupied in structures of b-LS with ATP alone, with N$^2$-(carboxylmethyl)-L-arginine (CMA)-adenylate plus PP$i$, and with AMP plus the product deoxyguanidinoproclavaminic acid plus PP$i$ (19). In these three structures, one Mg$^{2+}$ ion, designated Mg2, is coordinated like in the CarA/CMPr/AMP-CPP structure and the other, designated Mg1, is coordinated as in the b-LS/CEA/AMP-CPP structure.

Since all the structural features involved in Mg$^{2+}$ coordination are conserved between CarA and b-LS, it is not obvious why distinct sites are occupied in the two structures with substrate and AMP-CPP. One difference is the pH at which crystallization and subsequent ligand soaking was conducted. Whereas all the b-LS crystals were grown and handled at pH 8.0, the CarA crystals were maintained at pH 5.6. The lower pH could affect the protonation state of the g-phosphate oxygens, one of which has a pK$_a$ of ~6 (35). Both Mg$^{2+}$ sites involve coordination by a terminal g-phosphate oxygen atom (Fig. 5b). If the g-phosphate oxygen atom near the two aspartic acid residues is protonated in CarA, Mg$^{2+}$ binding on the opposite side might be more favorable. It is likely that neither g-phosphate oxygen atom is protonated in the pH 8.0 b-LS/CEA/AMP-CPP structure, leaving both sides of the g-phosphate group available for coordination. In this situation, the site with the two aspartic acid residues might have higher affinity for the Mg$^{2+}$ ion. Consistent with this hypothesis, two lysine residues, Lys 423 and Lys 443, are hydrogen bonded to the g-phosphate oxygen atoms in b-LS (18,19), whereas Lys 443 has shifted away from the AMP-CPP in CarA (Fig. 5b) leaving just one lysine, Lys 421, within hydrogen bonding distance. This altered conformation of Lys 443 could reflect a decrease in negative charge due to protonation of one of the g-phosphate oxygen atoms. The presence of one
rather than two Mg\textsuperscript{2+} ions in the the b\textsuperscript{-}LS/CEA/AMP-CPP structure was attributed to the substitution of a carbon atom for the oxygen bridging the a\textsuperscript{-} and b\textsuperscript{-}phosphate groups (19), and the same explanation may apply to CarA.

A second difference between the CarA/CMPr/AMP-CPP and b\textsuperscript{-}LS/CEA/AMP-CPP structures is the ordering of a loop comprising residues 441-450 in CarA. The corresponding region remains disordered in the b\textsuperscript{-}LS/CEA/AMP-CPP structure (18), although it is visible in the structures with CMA\textsuperscript{-}adenylate and product (19). Since this loop provides a ligand to the Mg\textsuperscript{2+} ion, the carbonyl oxygen of Ile 444, its ordering could explain why this site is preferentially occupied in the CarA structure. Conversely and perhaps more likely, Mg\textsuperscript{2+} binding at this site could organize these residues. The loop is further linked to the active site by a salt bridge between Arg 441 and Glu 277, which in turn interacts with one of the solvent ligands to the Mg\textsuperscript{2+} ion. The corresponding residues in b\textsuperscript{-}LS, Arg 441 and Glu 280, only interact in one monomer of the structure with CMA\textsuperscript{-}adenylate. In all the other b\textsuperscript{-}LS structures, regardless of whether the following residues are disordered, Arg 441 adopts a different conformation.

Mechanistic implications. By analogy to b\textsuperscript{-}LS and AS-B and in accordance with kinetic data, the CarA mechanism has been proposed to involve substrate adenylation followed by b\textsuperscript{-}lactam formation (12). Similar to b\textsuperscript{-}LS (17) and AS-B (36), ATP binds first followed by CMPr, and the last product released is PP\textsubscript{i} (12). The arrangement of ligands in CarA is consistent with this scheme. The AMP-CPP is situated deep in the active site cleft covered by CMPr and residues 441-450. The interactions between AMP-CPP and CarA are very similar to those observed in the b\textsuperscript{-}LS structures, suggesting that the ATP binding site in this family of enzymes is highly conserved. In b\textsuperscript{-}LS, substrate binding is accompanied by a large shift of residue Tyr 326 as well as a lesser alteration in Tyr 348. In CarA, Ile 323 occupies the analogous position to b\textsuperscript{-}LS.
LS Tyr 326 (Fig. 5b) and does not change conformation upon CMPr and AMP-CPP binding. The CarA equivalent to b-LS Tyr 348, Tyr 345, also remains in the same position. The similarity in position between the apo and substrate structures combined with the substitution of Ile for Tyr suggest that these two residues may not function to secure the substrate in the active site, as suggested for b-LS (19). It is also possible, however, that these residues or others do change conformation upon substrate binding in the correct position.

Although CMPr is present in a nonproductive orientation, the structure provides insight into the details of substrate binding. Whereas the carboxylate of the carboxymethyl group interacts with two amino acid residues, Gln 371 and Arg 374, the a-carboxylate group interacts only with a water molecule. Therefore, in the proper orientation, the carboxymethyl group is not expected to form specific contacts with the enzyme. Similarly, the carboxyethyl chain of CEA does not interact with residues in the b-LS active site (18,19). The lack of specific interactions is consistent with the ability of CarA to catalyze cyclization of different diastereomers of CMPr and of various diacids (12). This apparent promiscuity taken together with the nonproductive substrate conformation observed in the crystal structure raises the question of how CarA places substrates correctly in the active site. Whereas the arginine side chain provides a means for positioning CEA in the b-LS active site via specific interactions, CMPr is more symmetrical, with the two sides of the proline ring differing only in one methylene group.

Once substrate and ATP are assembled in the active site, the next step is attack of ATP by the CMPr carboxymethyl group to yield an adenylated intermediate. The formation of such a species is supported by hydroxylamine-dependent AMP formation in the presence of diacid substrates lacking an amino group (12). Intramolecular amide bond formation then yields a tetrahedral intermediate or transition state, followed by collapse to (3S,5S)-carbapenam-3-
carboxylic acid, AMP, and PPₐ. Deprotonation of the ammonium ion is required to initiate β-lactam formation and by analogy to β-LS (19), may be facilitated by a catalytic dyad consisting of residues Glu 380 and Tyr 345 (Fig. 5b). These two residues form a strong hydrogen bond in all four monomers in both structures. Another residue proposed to be critical in β-LS (19), Lys 443, is also conserved in CarA (Fig. 5b). Although this residue points away from the AMP-CPP in the current structure, a change in side chain conformation would enable it to stabilize the oxyanion intermediate or transition state. The sequential and structural conservation of these three residues in the two enzymes underscores their potential mechanistic importance. Finally, the adenylation and β-lactamization steps are assisted by Mg²⁺, which both activates the a,b-bond in ATP and counterbalances the negative charge on the PP, and nascent AMP, as proposed previously (18,19). Although only one Mg²⁺ ion is present in the CarA/CMPr/AMP-CPP structure, all the ligands used by β-LS to bind two Mg²⁺ ions are present, suggesting that CarA can also bind two Mg²⁺ ions.

In sum, the structure of CarA reveals the overall fold, details of cofactor and substrate binding, and key residues in the active site. Comparison of the N-terminal domain to the corresponding domains in AS-B and β-LS demonstrates another manner in which the AS-B glutaminase site can evolve away. The function of this domain in CarA and β-LS remains unclear. Comparison of the CarA C-terminal active site to those in AS-B and β-LS indicates that the ATP binding site is highly conserved, but that the substrate binding pocket changes to accommodate different molecules. The positions of the Glu 380/Tyr 345 catalytic dyad and of Lys 443 are consistent with previously proposed functional roles for these residues. A major outstanding question is how CarA properly positions CMPr and other substrates in the active site. Further biochemical and structural studies are required to resolve this issue, but the current
data suggest that CarA might prove to be an even more versatile target than b-LS for protein engineering experiments.

REFERENCES

1. Neu, H. C. (1994) Curr. Opin. Inf. Dis. 7, S3-S10
2. McGowan, S. J., Bycroft, B. W., and Salmond, G. P. C. (1998) Trends Microbiol. 6, 203-208
3. McGowan, S. J., Holden, M. T. G., Bycroft, B. W., and Salmond, G. P. C. (1999) Antonie van Leeuwenhoek 75, 135-141
4. Coulton, S., and Hunt, E. (1996) Prog. Med. Chem. 33, 99-145
5. Clifton, I. J., Doan, L. X., Sleeman, M. C., Topf, M., Suzuki, H., Wilmouth, R. C., and Schofield, C. J. (2003) J. Biol. Chem. 278, 20843-20850
6. McGowan, S. J., Sebaihia, M., Porter, L. E., Stewart, G. S. A. B., Williams, P., Bycroft, B. W., and Salmond, G. P. C. (1996) Mol. Microbiol. 22, 415-426
7. Thomson, N. R., Crow, M. A., McGowan, S. J., Cox, A., and Salmond, G. P. C. (2000) Mol. Microbiol. 36, 539-556
8. Parker, W. L., Rathnum, M. L., Wells, J. S., Trejo, W. H., Principe, P. A., and Sykes, R. B. (1982) J. Antibiot. 35, 653-660
9. McGowan, S. J., Sebaihia, M., O’Leary, S., Hardie, K. R., Williams, P., Stewart, G. S. A. B., Bycroft, B. W., and Salmond, G. P. C. (1997) Mol. Microbiol. 26, 545-556
10. Li, R., Stapon, A., Blanchfield, J. T., and Townsend, C. A. (2000) J. Am. Chem. Soc. 122, 9296-9297
11. Stapon, A., Li, R., and Townsend, C. A. (2003) J. Am. Chem. Soc. 125, 8486-8493
12. Gerratana, B., Stapon, A., and Townsend, C. A. (2003) Biochem. **42**, 7836-7847
13. Salowe, S. P., Marsh, E. N., and Townsend, C. A. (1990) Biochem. **29**, 6499-6508
14. Zhang, Z., Ren, J., Stammers, D. K., Baldwin, J. E., Harlos, K., and Schofield, C. J. (2000) Nature Struct. Biol. **7**, 127-133
15. Bachmann, B. O., Li, R., and Townsend, C. A. (1998) Proc. Natl. Acad. Sci. USA **95**, 9082-9086
16. McNaughton, H. J., Thirkettle, J. E., Zhang, Z., Schofield, C. J., Jensen, S. E., Barton, B., and Greaves, P. (1998) Chem. Commun. **21**, 2325-2326
17. Bachmann, B. O., and Townsend, C. A. (2000) Biochem. **39**, 11187-11193
18. Miller, M. T., Bachmann, B. O., Townsend, C. A., and Rosenzweig, A. C. (2001) Nature Struct. Biol. **8**, 684-689
19. Miller, M. T., Bachmann, B. O., Townsend, C. A., and Rosenzweig, A. C. (2002) Proc. Natl. Acad. Sci. USA **99**, 14752-14757
20. Richards, N. G. J., and Schuster, S. M. (1998) Adv. Enzymol. Relat. Areas Mol. Biol. **72**, 145-198
21. Larsen, T. M., Boehlein, S. K., Schuster, S. M., Richards, N. G. J., Thoden, J. B., Holden, H. M., and Rayment, I. (1999) Biochem. **38**, 16146-16157
22. Doublié, S. (1997) Methods Enzymol. **276**, 523-530
23. Powell, H. R. (1999) Acta Cryst. **D55**, 1690-1695
24. Collaborative Computational Project Number 4. (1994) Acta Cryst. **D50**, 760-763
25. Terwilliger, T. C., and Berendzen, J. (1999) Acta Cryst. **D55**, 849-861
26. McRee, D. E. (1999) J. Struct. Biol. **125**, 156-165
27. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Cryst. D54, 905-921

28. Kleywegt, G. J., and Jones, T. A. (1994) CCP4/ESF-EACBM Newsletter on Protein Crystallography 31, 9-14

29. Laskowski, R. A. (1993) J. Appl. Cryst. 26, 283-291

30. Vagin, A., and Teplyakov, A. (1997) J. Appl. Cryst. 30, 1022-1025

31. Kraulis, P. J. (1991) J. Appl. Cryst. 24, 946-950

32. Delano, W. L. (2002) The PyMOL molecular graphics system, DeLano Scientific, San Carlos, CA

33. Merritt, E. A., and Bacon, D. J. (1997) Methods Enzymol. 277, 505-524

34. Jones, S., and Thornton, J. M. (1996) Proc. Natl. Acad. Sci. USA 93, 13-20

35. Haynie, D. T. (2001) Biological thermodynamics, Cambridge University Press, Cambridge, UK

36. Boehlein, S. K., Stewart, J. D., Walworth, E. S., Thirumoorthy, R., Richards, N. G. J., and Schuster, S. M. (1998) Biochem. 37, 13230-13238
**FIGURE LEGENDS**

**FIG. 1.** **Formation of** β-lactam rings by CarA and β-LS. CarA converts (2S,5S)-5-carboxymethylproline (CMPr) to (3S,5S)-carbapenam-3-carboxylic acid. Subsequent steps catalyzed by CarC yield (3S,5R)-carbapenam-3-carboxylic acid and (5R)-carbapen-2-em-3-carboxylic acid. β-LS cyclizes N²-(carboxyethyl)-L-arginine (CEA) to deoxyguanidinoproclavaminic acid, which is converted to clavulanic acid by other enzymes.

**FIG. 2.** **Overall structure of** CarA. a, Stereo view of one monomer. Regions that differ significantly from β-LS and AS-B are colored magenta and labeled by residue number. Bound CMPr and AMP-CPP are shown as ball-and-stick representations, and a Mg^{2+} ion is shown as a gray sphere. b, Stereo view of the tetramer. The four monomers are labeled A, B, C, and D.

**FIG. 3.** **Structure-based sequence alignment of** CarA, β-LS, and AS-B. Secondary structure elements are superimposed on each sequence. Helices are shown as yellow cylinders and β-strands are shown as blue arrows. Key residues in the CarA active site are highlighted magenta.

**FIG. 4.** **Stereo view of simulated annealing electron density omit map showing bound CMPr, AMP-CPP, and Mg^{2+}.** The map is contoured at 1σ.

**FIG. 5.** **The CarA active site.** a, Stereo view showing Mg^{2+} coordination, CMPr binding site, and key hydrogen bonding interactions. Water molecules are shown as red spheres. b, Stereo comparison of CarA (red) and β-LS (light blue). Mechanistically important residues are included.
and labeled for CarA with the corresponding b -LS residue in parentheses. Water molecules are shown as red spheres.
**TABLE 1**  
Crystallographic statistics

Data collection and MAD phasing

|     | 1 | 2 | 3 | Substrate |
|-----|---|---|---|-----------|
| Wavelength (Å) | 0.9789 | 0.9417 | 0.9793 | 1.0024 |
| Resolution range (Å) | 20.0 – 2.3 | 20.0 – 2.3 | 20.0 – 2.3 | 50.0 – 2.4 |
| Unique observations | 90,698 | 87,462 | 88,429 | 83,896 |
| Total observations | 628,203 | 583,124 | 596,458 | 294,337 |
| Completeness* (%) | 99.6 (99.9) | 97.0 (92.2) | 97.5 (97.5) | 99.2 (98.3) |
| \(R_{sym}\) | 0.091 (0.311) | 0.078 (0.275) | 0.083 (0.310) | 0.084 (0.368) |
| I/sigma | 7.4 (2.3) | 8.4 (2.7) | 8.0 (2.4) | 7.9 (2.0) |

**Figure of merit** | 0.41 |

**Refinement**

|     | Apo (1) | Substrate |
|-----|---------|-----------|
| Resolution (Å) | 20.0 – 2.3 | 50.0 – 2.4 |
| Number of reflections | 91,852 | 86,087 |
| R-factor* | 0.203 | 0.220 |
| R-free* | 0.249 | 0.273 |
| Number of protein, nonhydrogen atoms | 15,305 | 15,708 |
| Number of nonprotein atoms | 312 | 575 |
| R.m.s. bond lengths (Å) | 0.007 | 0.006 |
| R.m.s. bond angles (°) | 1.1 | 1.1 |
| Average B-value (Å²) | 29.3 | 33.7 |

*Values in parentheses are for the highest resolution shell: 1, 2, 3, 2.42 – 2.30 Å; Substrate, 2.53 – 2.40 Å.

\(R_{sym} = \sum |I_{obs} - I_{avg}|/\sum I_{obs}\), where the summation is over all reflections.

Figure of merit was obtained from the program SOLVE for 24 – 2.5 Å resolution.

For calculation of R-free, 10% of the reflections were reserved.
| Interface | Apo structure | Substrate structure |
|-----------|---------------|---------------------|
| A-B       | 2,276 Å²      | 2,297 Å²            |
| A-C       | 240 Å²        | 231 Å²              |
| A-D       | 2,775 Å²      | 2,739 Å²            |
| B-C       | 2,701 Å²      | 2,751 Å²            |
| B-D       | 144 Å²        | 63 Å²               |
| C-D       | 2,370 Å²      | 2,158 Å²            |

\(^a\) Calculations performed with CNS.
\(^b\) The monomer labels correspond to those in Fig. 2b.
FIG. 1
FIG. 2
FIG. 3
Crystal structure of carbapenam synthetase (CarA)
Matthew T. Miller, Barbara Gerratana, Anthony Stapon, Craig A. Townsend and Amy C. Rosenzweig

J. Biol. Chem. published online July 30, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M307901200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts