*Proc. Natl. Acad. Sci. USA*
Vol. 78, No. 8, pp. 4897–4901, August 1981
Biochemistry

# *ampC* cephalosporinase of *Escherichia coli* K-12 has a different evolutionary origin from that of β-lactamases of the penicillinase type

(protein sequence/signal peptide/nucleotide sequence/sequence comparisons/serine enzymes)

BENGTÅKE JAURIN AND THOMAS GRUNDSTRÖM

Department of Microbiology, University of Umeå, S-901 87 Umeå, Sweden

*Communicated by Jack Leonard Strominger, May 13, 1981*

**ABSTRACT** **A 1536-nucleotide-long sequence that carries the *ampC* β-lactamase gene of the *Escherichia coli* K-12 chromosome has been determined. This gene codes for a protein of 377 amino acids, of which the first 19 amino acids form a signal peptide. The molecular weight of the mature enzyme was determined to be 39,600. The *ampC* β-lactamase with a substrate specificity for cephalosporins showed no significant sequence homologies with β-lactamases of the penicillinase type or with D-alanine carboxypeptidases. However, because the region around serine-80 of the *ampC* β-lactamase has extensive homology with an active-site fragment of the *Pseudomonas aeruginosa* cephalosporinase, we suggest that the *ampC* cephalosporinase as well as related cephalosporinases form a distinct group of serine β-lactamases that have an evolutionary origin different from that of the serine penicillinases and thus constitute a new class of β-lactamases.**

β-Lactamases of chromosomal or plasmid origin have been found in a large number of Gram-positive and Gram-negative bacteria (1). These enzymes have been classified according to such properties as substrate profile, isoelectric point, and molecular weight. The protein sequence has been determined for β-lactamases from *Staphylococcus aureus* PC1 (2), *Bacillus licheniformis* 749/C (3, 4), and *Escherichia coli*/R6K, R-TEM (5). Nearly the entire sequence of the *B. cereus* 569/H β-lactamase has also been elucidated (6, 7). In addition, the amino acid sequence of the β-lactamase encoded by the plasmid pBR322 has been deduced from its nucleotide sequence (8). The two plasmid-mediated TEM β-lactamases from Gram-negative species differ only by one amino acid residue (5, 8).

These β-lactamases of known sequence all show substrate specificity for penicillins and have therefore been termed "penicillinases" (1). The molecular weight of these enzymes is around 29,000 (7). They show significant sequence homologies with each other (7) as well as with regions of D-alanine carboxypeptidases from *B. stearothermophilus* and *B. subtilis* (9). By the use of substrate analogues, the active site has been determined for three penicillinases and two carboxypeptidases (9–13). They have been referred to as "serine enzymes" because the reagents react with a serine residue. β-Lactamases of the metalloenzyme type have also been identified. From incomplete sequence data it is suggested that this class has a different evolutionary origin from that of the serine penicillinases of known sequence (7).

The chromosomally encoded β-lactamases of Gram-negative enterobacteria in general are basic proteins with a substrate specificity for cephalosporins (14). One such cephalosporinase is encoded by the *ampC* gene of *E. coli* K-12. This gene, which is located at 93.8 min on the *E. coli* linkage map (15), was isolated from a gene bank containing *E. coli* chromosomal DNA (16, 17). By subcloning, *ampC* was localized to a 1370-base-pair (bp)-long DNA fragment (17, 18). We have reported (19) a characterization of the regulatory region that precedes *ampC*. The expression from *ampC* was decreased by the presence of a terminator that caused a marked attenuation of transcription. Mutations in both the promoter and the attenuator for *ampC* can confer increased *ampC* β-lactamase production (19).

In this paper the entire sequence of the *ampC* gene is presented. We conclude that the *ampC* enzyme is a serine cephalosporinase that shows no sequence homologies to the internally related group of serine penicillinases and serine D-alanine carboxypeptidases.

## MATERIALS AND METHODS

**DNA Techniques.** The pBR322 derivatives plasmids pNU5 and pNU6, which carry the *ampC* gene in opposite orientations, were used as a source of DNA (17). A crude plasmid DNA preparation was made from chloramphenicol-treated cells (17) and further purified by two consecutive CsCl gradient centrifugations of 6 hr each in a Beckman VTi 65 rotor. Restriction enzyme fragments were prepared from polyacrylamide gels (20), and 5′- or 3′-end-labeled as described (19). The labeled DNA fragments were either strand separated (21) by using 5% or 10% gels (bisacrylamide/acrylamide, 1:30) or recleaved with a restriction enzyme. The DNA sequence was determined by the methods of Maxam and Gilbert (22). The sequence was analyzed by using the computer programs of Staden (23), as modified by P. Gustafsson and P. Hagblom (this laboratory, personal communication).

**Purification of β-Lactamase.** The *ampC* gene has been cloned onto the plasmid pKN402 which carries a temperature-sensitive replication control (24). Such plasmid-containing cells were harvested after 4 hr of uncontrolled plasmid replication at 37°C. The amount of β-lactamase in these cells was approximately 3% of the total cell protein. The purification procedure for the β-lactamase was that described by Lindström *et al.* (25) with the following modification: after chromatography on SE-cellulose, the extract was applied to a Sephadex G-75 column. A purity of >99% was demonstrated by $NaDodSO_4$/10–17.5% polyacrylamide gradient gel electrophoresis (19) and by crossed immunoelectrophoresis (26) with antibodies raised against an *E. coli* cell extract.

**$NH_2$-Terminal Sequence Determination.** Purified β-lactamase (1.5 mg; 38 nmol) was dissolved in 0.3 ml of trifluoroacetic acid. The $NH_2$-terminal amino acids were determined by using a Beckman 890 C protein sequencer; the phenylthiohydantoin derivatives were identified by high-pressure liquid chromatography as described by Wiman *et al.* (27).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: bp, base pair(s).

**Computer Analysis.** The amino acid sequence of the *ampC* β-lactamase was compared to the known complete β-lactamase sequences of *Staphylococcus aureus* PC1 (2), *B. licheniformis* 749/C (3, 4), and *E. coli*/R6K, R-TEM (5). Two programs were used. The first one, ALIGN, allows the introduction of gaps into the sequence (28, 29). A gap penalty of 6 and a bias parameter of 6 were used. Alignment scores were calculated for the two sequences in question. The second program, RELATE, is designed for comparing proteins of different length (29, 30).

## RESULTS AND DISCUSSION

**Strategy of the DNA Sequence Determination.** We have determined the DNA sequence of the chromosomal β-lactamase gene, *ampC*, of *E. coli* K-12. The gene for the *ampC* β-lactamase had previously been localized to within the DNA segment between the rightmost *Pst* I site in Fig. 1 and the insertion point for the transposable element γδ, located at about 220 bp to the left of the *Xma* I site in Fig. 1 (18). By analysis of both strands, the sequence of 1536 bp encompassing the entire *ampC* gene was obtained (Fig. 2).

**Identification of the Coding Sequence for the *ampC* β-Lactamase.** The molecular weight of the *ampC* β-lactamase has been estimated to be about 36,000 from $NaDodSO_4$/polyacrylamide gel electrophoresis (18, 19). Assuming a mean molecular weight of 110 for the amino acids, the protein would consist of about 330 amino acids.

The sequence of the $NH_2$-terminal amino acids was determined to be $NH_2$-Ala-Pro-Gln-Gln-Ile-Asn-Asp-Ile-Val-His-Arg-Thr-. This stretch of amino acids corresponds exactly to the 12 codons from base +117 to base +152 (Fig. 2). The first stop codon downstream in this reading frame is the ochre codon, TAA, at bases +1191 to +1193. This reading frame contains 358 codons, and no other reading frame was open for more than 87 codons. The amino acid composition of the mature β-lactamase deduced from the nucleotide sequence is displayed in Table 1 and compared with the composition of the purified enzyme (25). We have recalculated the number of residues per molecule from the composition found by Lindström *et al.* (25) on the basis of a molecular weight of 39,600 as found by us. The amino acid composition of the purified protein corresponds well with the composition deduced from the sequence. We therefore conclude that the mature *ampC* β-lactamase of *E. coli* K-12 is encoded by these 358 codons and would have a molecular weight of 39,600. This value is about 40% greater than the molecular weights (about 29,000) found for the four penicillinases of known sequence (7). The number of basic residues (lysine and arginine) exceeds the number of acidic residues (aspartic acid and glutamic acid), which agrees with the high isoelectric point (pI, 9.9) previously reported for the *ampC* β-lactamase (33).

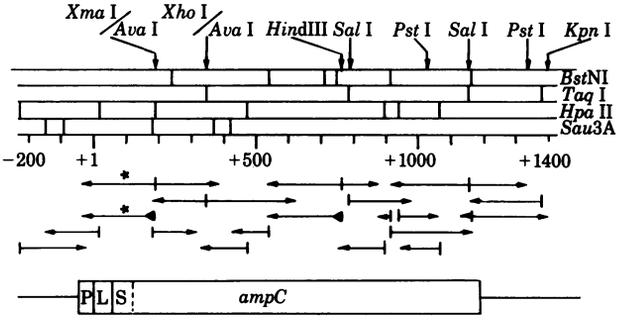


FIG. 1. Restriction enzyme map and sequencing strategy for the *ampC* gene of *E. coli* K-12. The numbers represent length in bp and correspond to the DNA sequence shown in Fig. 2. The top strip represents the map for enzymes that cut, at most, twice in the *ampC* gene. Each horizontal strip represents the cleavage map for a different restriction enzyme as indicated at the end of the strip. Arrows with vertical lines represent sequencing readings from 5'-labeled ends; arrows with filled triangles at their tail indicate 3'-labeling. The entire sequence was determined from both strands. Asterisks indicate previously published sequences of the regulatory region (19). The box at the bottom shows the location of the *ampC* gene. P, L, and S, promoter region, leader region, and signal peptide, respectively.

**Primary Gene Product Is a Precursor.** The codon ATG at bases +60 to +62 is preceded by a ribosome binding site (31) (-T-A-T-G-G-A-) at bases +43 to +48 (Fig. 2). This is the only possible translational start point within the open reading frame preceding the $NH_2$ terminus of the β-lactamase. It has the features common to all translation initiation sequences (32). Five of six bases of the β-lactamase ribosome binding site are complementary to the 3'-end of 16S RNA (3' A-U-U-C-C-U-) (31), and the binding site is located at a distance of 11 bases from the initiation codon. This distance is close to the average distance found between ribosome binding sites and initiation codons (32). The translational start codon is thus 19 codons upstream from the $NH_2$-terminal alanine of the mature *ampC* β-lactamase. We have previously reported that the β-lactamase is synthesized as a precursor. Both a coupled *in vitro* transcription-translation system (19) and minicells (18) make a pre-β-lactamase with a molecular weight about 2000 greater than that of the mature enzyme.

The *ampC* β-lactamase of *E. coli* is secreted into the periplasmic space (25). Proteins exported through membranes have been shown to be synthesized with an $NH_2$-terminal signal peptide extension (34). Such signal sequences have been shown to have a stretch of hydrophobic amino acids in the middle, one or more positively charged residues near the $NH_2$ terminus, and an amino acid with a side chain containing at most one carbon at the COOH terminus (35). In the signal peptide of the pre-β-lactamase, a positively charged lysine residue is found at position three, and six of the eight residues from positions 6–13 are hydrophobic. Furthermore, the last amino acid, alanine, contains one carbon in its side chain and has been found in this position in the majority of signal peptides of known sequence in *E. coli* (35). Thus, the *E. coli ampC* β-lactamase is made with a 19-amino acid-long $NH_2$-terminal extension which possesses all the general features common to signal peptides.

**Regulatory Regions for Transcription of *ampC*.** The sequences at nucleotide positions −13 to −8 (-T-A-C-A-A-T-) and −35 to −30 (-T-T-G-T-C-A-) show a five-of-six-bp homology with the conserved −10 (-T-A-T-A-A-T-) and −35 (-T-T-G-A-C-A-) regions of promoters, respectively (36). Even in the less-well-conserved regions surrounding the −35 and the −10 regions, sequence homologies are observed. We have recently shown that this region in fact represents the promoter for *ampC* (19). The adenine at position +1 was shown by RNA sequence determination to be the first base of β-lactamase mRNA. *In vitro* transcription studies revealed that *ampC* was controlled by attenuation of transcription (19). The terminator structure and the points of termination are indicated in Fig. 2.

The only other β-lactamase control region whose sequence has been determined is that of the TEM-1 β-lactamase gene (8). The localization of the promoter and the regulation of transcription of this operon have not been studied in detail. However, a sequence very much resembling a promoter is located close to the potential ribosome binding site (8). No dyad symmetry element similar to that found in the *ampC* leader precedes the site of initiation of translation of this β-lactamase.

Nineteen bases downstream from the translational stop codon is the center of an eight-base-long dyad symmetry. The sequence is followed by a series of thymidine residues on the

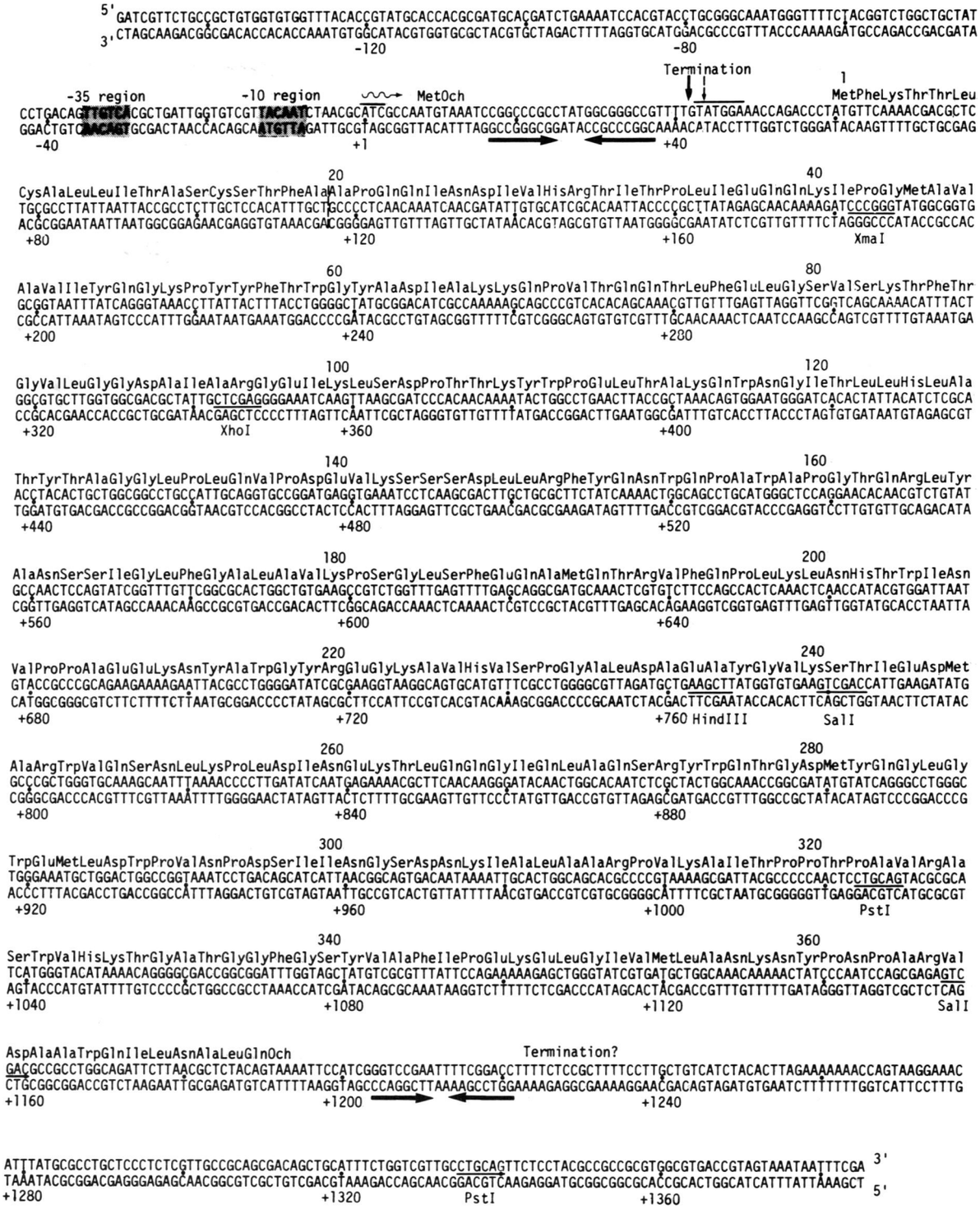


FIG. 2. DNA sequence of the *ampC* gene from *E. coli* K-12. Every 20th bp is marked with a dot between the strands. The position of the first base in the β-lactamase mRNA was chosen as position +1 and the count was written below every second dot. The three-letter abbreviations for the amino acids of the *ampC* β-lactamase appear directly over their three-base codons and are numbered (every 20 amino acids) starting from the first methionine. The positions of restriction enzyme sites (Fig. 1) are marked with a horizontal line between the strands, and the names of the enzymes are written below the strands. The start of transcription is marked by +1 and the wavy arrow. The major and minor termination points of the attenuator (19) are indicated by vertical solid and dashed arrows, respectively. The regions of dyad symmetry in the attenuator and the possible terminator are marked by horizontal arrows. Solid lines designate possible ribosome binding sites (31, 32). The boundary between the signal peptide and the mature β-lactamase is marked by a vertical dashed line. The −35 and −10 regions of the β-lactamase promoter are marked by stippled boxes.

noncoding strand. Such features are common to rho-independent terminators (37). We have no indications of the existence of a second structural gene downstream from *ampC* in the β-lactamase operon. We therefore think that the *ampC* β-lactamase mRNA may terminate around nucleotide positions +1225 to +1230 (Fig. 2).

**Codon Distribution in *ampC*.** Table 2 summarizes the distribution of codons used in the assembly of the *E. coli* β-lac-

Table 1. Amino acid composition of *ampC* β-lactamase

| Residue | Protein hydrolyzate* | DNA sequence |
|---|---|---|
| Glycine | 28.9 | 29 |
| Alanine | 36.3 | 37 |
| Valine | 21.5 | 22 |
| Leucine | 30.1 | 31 |
| Isoleucine | 20.3 | 22 |
| Serine | 19.3 | 18 |
| Threonine | 23.6 | 23 |
| Aspartic acid | 31.1 | 14 |
| Asparagine | | 16 |
| Glutamic acid | 41.1 | 15 |
| Glutamine | | 25 |
| Phenylalanine | 9.6 | 9 |
| Tyrosine | 14.2 | 15 |
| Tryptophan | — | 13 |
| Cysteine | 0 | 0 |
| Methionine | 5.6 | 6 |
| Proline | 24.6 | 26 |
| Lysine | 21.7 | 21 |
| Histidine | 5.3 | 5 |
| Arginine | 11.7 | 11 |
| Total | | 358 |

Results are shown as residues per molecule.

* Data are from Lindström *et al.* (25). The results presented are the means of composition determinations made from enzyme purified from two different *E. coli* K-12 strains. We have recalculated the number of residues per molecule on the basis of the molecular weight of 39,600 found by us. Aspartic acid and glutamic acid were not distinguished from their respective amino derivatives. Tryptophan was degraded during the analysis.

tamase. We find a higher preference for codons for the major iso-accepting species of tRNAs than that found in most *E. coli* genes of known sequence. The frequency of CGY (Y = pyrimidine) for arginine is 82%, GGY for glycine is 72%, ATY for isoleucine is 91%, GAA for glutamic acid is 60%, and AAA for lysine is 68%. Preference for these codons has also been found in the *lacI* gene (38) and to an even greater extent in some sequenced r-protein genes (39). One exception is CUG, which is preferentially used for leucine in the *lacI* and these r-protein genes (61% and 91%, respectively); the usage in *ampC* is only 29%.

**ampC β-Lactamase Is Distinct from Previously Analyzed β-Lactamases.** At present, the complete or partial amino acid sequences of four penicillinases are known from direct analysis. Although of different origin, these penicillinases show extensive sequence homologies with each other throughout the whole length of the polypeptide chain (7). When these enzymes were aligned (7) the same amino acid was found in all four proteins at 20% of the positions; at only 15% of the positions were four different amino acids found.

A computer search for sequence homology between the *ampC* β-lactamase and the three β-lactamases with known complete amino acid sequence was kindly performed by W. Barker and M. Dayhoff at the National Biomedical Research Foundation (Washington, DC). The program ALIGN (28, 29), with MDM250 scoring matrix, a bias of 6, and a gap penalty of 6, gave alignment scores (Z values) of 0.67, 1.89, and 0.99 when the sequence was compared to that of the *S. aureus*, *B. licheniformis*, and TEM-1 β-lactamases, respectively. These values are significantly lower than 3.0, considered to be the limit of significance for homology (29). Likewise, no significant homologies were found by using the program RELATE (29, 30). For all three comparisons the homology score with this program was less than 0.11 SD above the random score. We therefore conclude that the *ampC* β-lactamase is distinct from previously analyzed β-lactamases.

**Sequence Around Serine-80 Shows Homologies to the Active Site Region of *Pseudomonas aeruginosa* β-Lactamase.** Serine-44 of the *B. cereus* β-lactamase I and the corresponding peptides from the TEM-1 β-lactamase and the *S. aureus* β-lactamase can be covalently coupled to active-site-directed substrate analogues (10–12). Furthermore, serine-36 of the *B. subtilis* and *B. stearothermophilus* D-alanine carboxypeptidases can be covalently coupled to active-site-directed substrate analogues (9, 13). Significant homology exists in the active-site region between these penicillinases and the D-alanine carboxypeptidases (7, 9). Therefore, these two groups of enzymes seem to be evolutionary related (7, 9). None of the serine residues in the *ampC* β-lactamase is within a sequence with significant homologies to the active site region of these enzymes.

A β-lactamase with a preference for cephalosporins from *P. aeruginosa* binds an active-site-directed analogue to the serine in the peptide Ile-Gly-Ser (40). Additional studies have expanded the sequence to 14 residues (S. Waley, personal communication). Eleven of the 14 amino acids are identical when this peptide is compared to the segment from residues 70 to 83 of the *ampC* product. Each of the three differences (Gln-72 → Pro; Gln-73 → Glu; Leu-78 → Ile) can be explained by a single base substitution. These findings strongly suggest that

Table 2. Use of codons in *ampC* β-lactamase gene

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Phe | TTT | 7 | Ser | TCT | 3 | Tyr | TAT | 10 | Cys | TGT | 0 |
| Phe | TTC | 4 | Ser | TCC | 3 | Tyr | TAC | 5 | Cys | TGC | 2 |
| Leu | TTA | 7 | Ser | TCA | 2 | | TAA | 1 | | TGA | 0 |
| Leu | TTG | 5 | Ser | TCG | 3 | | TAG | 0 | Trp | TGG | 13 |
| Leu | CTT | 6 | Pro | CCT | 7 | His | CAT | 5 | Arg | CGT | 2 |
| Leu | CTC | 4 | Pro | CCC | 7 | His | CAC | 0 | Arg | CGC | 7 |
| Leu | CTA | 2 | Pro | CCA | 6 | Gln | CAA | 14 | Arg | CGA | 1 |
| Leu | CTG | 10 | Pro | CCG | 6 | Gln | CAG | 11 | Arg | CGG | 0 |
| Ile | ATT | 12 | Thr | ACT | 4 | Asn | AAT | 8 | Ser | AGT | 3 |
| Ile | ATC | 9 | Thr | ACC | 8 | Asn | AAC | 8 | Ser | AGC | 6 |
| Ile | ATA | 2 | Thr | ACA | 9 | Lys | AAA | 15 | Arg | AGA | 1 |
| Met | ATG | 7 | Thr | ACG | 6 | Lys | AAG | 7 | Arg | AGG | 0 |
| Val | GTT | 1 | Ala | GCT | 10 | Asp | GAT | 7 | Gly | GGT | 10 |
| Val | GTC | 5 | Ala | GCC | 9 | Asp | GAC | 7 | Gly | GGC | 11 |
| Val | GTA | 6 | Ala | GCA | 12 | Glu | GAA | 9 | Gly | GGA | 3 |
| Val | GTG | 10 | Ala | GCG | 9 | Glu | GAG | 6 | Gly | GGG | 5 |

The numbers represent how many times the codons are used in the structural gene for *ampC* β-lactamase, including its signal peptide. TAA is the terminator used.

serine-80 is the active-site residue of the *ampC* β-lactamase.

The protein sequence homology around the active site that is common to all β-lactamases mentioned above is limited to a phenylalanine four residues on the $NH_2$-terminal side of the serine and a lysine three residues from the serine toward the COOH terminus. Only the lysine is common also to the D-alanine carboxypeptidases.

***ampC* β-Lactamase Is a Member of a Third Structural Class of β-Lactamases.** Ambler (7) has suggested that the β-lactamases have a polyphyletic origin. The previously reported serine β-lactamases with extensive sequence homologies with each other and with a preference for penicillin substrates are grouped as class A enzymes. The *B. cereus* β-lactamase II is a zinc-requiring enzyme. Preliminary partial sequence analysis suggests it to be structurally unrelated to the class A enzymes. In addition to this class B enzyme, other β-lactamases have been found that do not fit into either of these groups (7).

Interestingly, DNA probes from *ampC* of *E. coli* K-12 hybridize to fragments of the same size from the chromosome of many Gram-negative enterobacteria (41). This shows that the chromosomally encoded β-lactamases of many species of Enterobacteriaceae have extensive sequence homologies. They therefore constitute a third group of β-lactamases (class C) that are serine enzymes but probably have an evolutionary origin different from that of serine penicillinases.

We are indebted to Professor Staffan Normark for his encouragement and support of this study and for help in the preparation of the manuscript. We thank Torbjörn Nilsson and Per A. Petersson for help with the amino acid sequence determination. We also thank S. Waley and coworkers (Oxford University) for communicating to us the additional sequence data on the *Pseudomonas aeruginosa* β-lactamase prior to publication. The expert assistance of Drs. W. C. Barker and M. Dayhoff with the computer analysis is gratefully acknowledged. We also thank Monica Persson for technical assistance and Christopher Korch for carefully reading the manuscript. This work was supported by grants from the Swedish Natural Science Research Council (Dnr 3373) and from the Swedish Medical Research Council (Dnr 5428).

1. Citri, N. (1971) in *The Enzymes*, ed. Boyer, P. D. (Academic, New York), 3rd Ed., Vol. 4, pp. 23–46.
2. Ambler, R. P. (1975) *Biochem. J.* **151**, 197–218.
3. Meadway, R. J. (1969) *Biochem. J.* **115**, 12P–13P.
4. Yamamoto, S. & Lampen, J. O. (1976) *J. Biol. Chem.* **251**, 4095–4101.
5. Ambler, R. P. & Scott, G. K. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3732–3736.
6. Thatcher, D. R. (1975) *Biochem. J.* **147**, 313–326.
7. Ambler, R. P. (1980) *Philos. Trans. R. Soc. London Ser. B* **289**, 321–331.
8. Sutcliffe, J. G. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3737–3741.
9. Waxman, D. J. & Strominger, J. L. (1980) *J. Biol. Chem.* **255**, 3964–3976.
10. Knott-Hunziker, V., Waley, S. G., Orlek, B. S. & Sammes, P. G. (1979) *FEBS Lett.* **99**, 59–61.
11. Fisher, J., Belasco, J. G., Khosla, S. & Knowles, J. R. (1980) *Biochemistry* **19**, 2895–2901.
12. Cartwright, S. J. & Coulson, A. F. W. (1980) *Philos. Trans. R. Soc. London Ser. B* **289**, 370–372.
13. Yocum, R. R., Waxman, D. J., Rasmussen, J. R. & Strominger, J. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2730–2734.
14. Sykes, R. B. & Matthew, M. (1976) *J. Antimicrob. Chemother.* **2**, 115–157.
15. Bachmann, B. J. & Low, K. B. (1980) *Microb. Rev.* **44**, 1–56.
16. Clarke, L. & Carbon, J. (1976) *Cell* **9**, 91–99.
17. Edlund, T., Grundström, T. & Normark, S. (1979) *Mol. Gen. Genet.* **173**, 115–125.
18. Grundström, T., Jaurin, B., Edlund, T. & Normark, S. (1980) *J. Bacteriol.* **143**, 1127–1134.
19. Jaurin, B., Grundström, T., Edlund, T. & Normark, S. (1981) *Nature (London)* **290**, 221–225.
20. Akusjärvi, G. & Pettersson, U. (1978) *Virology* **91**, 477–480.
21. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. (1979) *Nature (London)* **280**, 288–294.
22. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
23. Staden, R. (1979) *Nucleic Acids Res.* **6**, 2601–2610.
24. Uhlin, B. E., Molin, S., Gustafsson, P. & Nordström, K. (1979) *Gene* **6**, 91–106.
25. Lindström, E. B., Boman, H. G. & Steele, B. B. (1970) *J. Bacteriol.* **101**, 218–231.
26. Weeke, B. (1973) *Scand. J. Immunol.* Suppl. 1, **2**, 47–56.
27. Wiman, K., Trägårdh, L., Rask, L. & Peterson, P. A. (1979) *Eur. J. Biochem.* **95**, 265–273.
28. Needleman, S. B. & Wunsch, C. D. (1970) *J. Mol. Biol.* **48**, 443–453.
29. Dayhoff, M. O. (1976) in *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 2, pp. 1–8.
30. Fitch, W. M. (1966) *J. Mol. Biol.* **16**, 9–16.
31. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
32. Steitz, J. (1979) in *Biological Regulation and Development*, ed. Goldberger, R. F. (Plenum, New York), Vol. 1, pp. 349–399.
33. Jaurin, B. & Normark, S. (1979) *J. Bacteriol.* **138**, 896–902.
34. Blobel, G. & Dobberstein, B. (1975) *J. Cell. Biol.* **67**, 835–851.
35. Inouye, M. & Halegoua, S. (1980) *Crit. Rev. Biochem.* **7**, 339–371.
36. Siebenlist, U., Simpson, R. B. & Gilbert, W. (1980) *Cell* **20**, 269–281.
37. Rosenberg, M. & Court, D. (1979) *Annu. Rev. Genet.* **13**, 319–353.
38. Farabaugh, P. J. (1978) *Nature (London)* **274**, 765–769.
39. Post, L. E., Strycharz, G. D., Nomura, M., Lewis, H. & Dennis, P. P. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1697–1701.
40. Knott-Hunziker, V., Redhead, K., Petursson, S. & Waley, S. G. (1980) *FEBS Lett.* **121**, 8–10.
41. Jaurin, B., Grundström, T., Bergström, S. & Normark, S. (1981) in *Proceedings on Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids*, eds. Clones, R. & Levy, S. (Plenum, New York), in press.