Nucleotide Sequence and Phylogeny of a Chloramphenicol Acetyltransferase Encoded by the Plasmid pSCS7 from *Staphylococcus aureus*

STEFAN SCHWARZ* AND MARISA CARDOSO

Institut für Bakteriologie und Immunologie der Justus Liebig Universität Giessen, Frankfurterstrasse 107, 6300 Giessen, Federal Republic of Germany

The nucleotide sequence of the chloramphenicol acetyltransferase gene (*cat*) and its regulatory region, encoded by the plasmid pSCS7 from *Staphylococcus aureus*, was determined. The structural *cat* gene encoded a protein of 209 amino acids, which represented one monomer of the enzyme chloramphenicol acetyltransferase (CAT). Comparisons between the amino acid sequences of the pSCS7-encoded CAT from *S. aureus* and the previously sequenced CAT variants from *S. aureus*, *Staphylococcus intermedius*, *Staphylococcus haemolyticus*, *Bacillus pumilis*, *Clostridium difficile*, *Clostridium perfringens*, *Escherichia coli*, *Shigella flexneri*, and *Proteus mirabilis* were performed. An alignment of CAT amino acid sequences demonstrated the presence of 34 conserved amino acids among all CAT variants. These conserved residues were considered for their possible roles in the structure and function of CAT. On the basis of the alignment, a phylogenetic tree was constructed. It demonstrated relatively large evolutionary distances between the CAT variants of enteric bacteria, *Clostridium*, *Bacillus*, and *Staphylococcus* species.

The determinants for Cm^r in staphylococci have been found to be exclusively located on small plasmids up to 5.2 kb in size (17, 25–28). These plasmids have been classified on the basis of restriction endonuclease analyses and DNA-DNA hybridizations to belong to three families (10). The prototype plasmids of these families were the plasmids pC221 (2, 22), pC194 (13), and pC223 (20), all from *Staphylococcus aureus* of human origin. All of these staphylococcal Cm^r plasmids encode an inducible chloramphenicol acetyltransferase (CAT). So far, molecular analyses have been performed on the *cat* genes of the plasmids pC221 (2, 22), pUB112 from *S. aureus* (3), pSCS1 from *Staphylococcus intermedius* (28), pC194 from *S. aureus* (13), and the recently described pSCS5 from *Staphylococcus haemolyticus* (25). These analyses revealed a high degree of homology among the *cat* genes of Cm^r plasmids of one family, but there are distinct differences between those from different families. So far, no sequence data have been reported for the *cat* gene of the prototype Cm^r plasmid pC223.

In this regard, the *cat* gene from the Cm^r plasmid pSCS7 from *S. aureus* of bovine origin was investigated. This plasmid was previously identified to be a pC223 analog (5). The predicted amino acid sequence of CAT from pSCS7 was used for comparisons with the CAT variants of gram-positive and gram-negative bacteria. A phylogenetic tree based on the alignment of the CAT amino acid sequences was constructed to provide insight into the origin and the evolutionary development of the pSCS7-encoded CAT and other bacterial CAT variants.

MATERIALS AND METHODS

Bacterial strains, plasmids, and phages. For cloning experiments, the previously analyzed plasmid pSCS7 (5) was subjected to partial and complete digestion with *Mbo*I. Plasmid pBluescript II SK$^+$ (Stratagene, Heidelberg, Fed-

eral Republic of Germany) served as the cloning vector. *Escherichia coli* TG1 [*supE hsd* Δ5 *thi* Δ(*lac-proAB*) F'(*traD36 proAB*$^+$ *lacI*^{q} *lacZ* ΔM15); Stratagene] was used as the recipient for plasmid pBluescript II SK$^+$ and its recombinant derivatives. Growth conditions were as described previously (25, 28). The production of single-stranded DNA with the helper phage M13K07 (Stratagene) was performed by the recommendations of Vieira and Messing (35).

Reagents and enzymes. Restriction endonucleases were purchased from Boehringer (Mannheim, Federal Republic of Germany), T$_4$ ligase and alkaline phosphatase were from Stratagene, and T$_7$ DNA polymerase was from Pharmacia (Freiburg, Federal Republic of Germany). [α-^{35}S]dATP was a product of Amersham Corp. (Braunschweig, Federal Republic of Germany), and ampicillin was from Sigma (Deisenhofen, Federal Republic of Germany). Sequencing reagents were purchased from Roth (Karlsruhe, Federal Republic of Germany) or GIBCO/BRL (Paisley, Scotland).

DNA preparation and nucleotide sequence determination. pSCS7 DNA was prepared by using the Qiagen system (Diagen, Düsseldorf, Federal Republic of Germany), modified for staphylococci as described previously (27). Cloning of *Mbo*I-digested pSCS7 DNA into pBluescript II SK$^+$ was performed as described by Sambrook et al. (23). Transformation of *E. coli* TG1 with recombinant pBluescript II SK$^+$ was done by the recommendations of Dagert and Ehrlich (7). Sequence analyses were performed on both strands by using single-stranded phage DNA and double-stranded plasmid DNA as templates (23). The dideoxy chain termination method of Sanger et al. (24) was applied by using T$_7$ DNA polymerase (33) and [α-^{35}S]dATP. T$_3$, T$_7$, KS, and reverse primer (Stratagene) served for initial sequence determinations. Another four 16-mer oligonucleotide primers, synthesized with an Applied Biosystems 380B DNA synthesizer, were used to complete the *cat* gene sequence.

Nucleotide sequence accession number. The pSCS7 *cat*
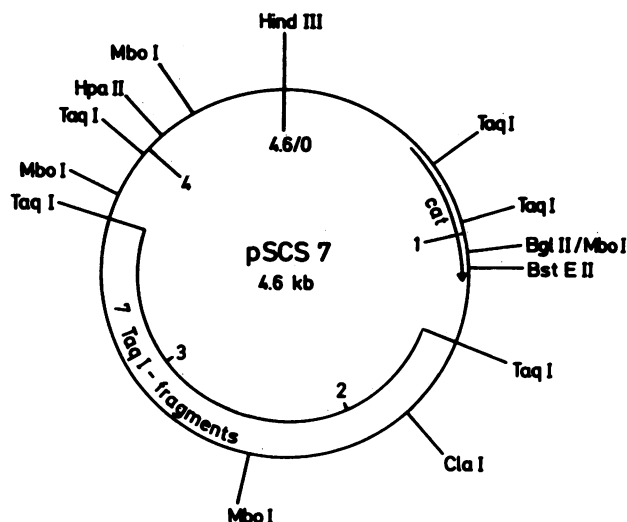
---

* Corresponding author.

FIG. 1. Circular restriction endonuclease cleavage map of the Cm^r plasmid pSCS7 from *S. aureus*. The arrow indicates the extent and the direction of translation of the *cat* gene. The positions of some of the *Taq*I sites on this plasmid were not determined, but the region containing seven unordered *Taq*I fragments is marked.

sequence had been submitted to GenBank and assigned accession number M 58516.

Computer analysis. Alignment of the CAT amino acid sequences, determination of the percentages of CAT identity, and construction of the cladogram were performed by using the alignment program of Hein (12).

The strategy of this alignment program was to use first pairwise alignments. For each pair of aligned sequences, one ancestral sequence was reconstructed. Thus, the total number of pairwise comparisons in our study was 65 = (11 × 10)/2 (distance tree construction) + (11 − 1) (number of ancestral sequences reconstructed, including one abitrary root) (12). The distances between two sequences were calculated on the basis of amino acid substitutions. Amino acid substitutions were weighted by the minimum mutation matrix of Dayhoff (8). The gap penalty used was $g_k = 8 + 3k$, which means that an insertion-deletion of length k was weighted $8 + 3k$. The percentages of CAT identity, as well as the branch lengths, were calculated on the basis of this alignment.

RESULTS

Mapping and cloning of the pSCS7 *cat* gene. To localize the *cat* gene, pSCS7 was digested with *Mbo*I. The resulting fragments were cloned in both orientations into the single *Bam*HI site of the pBluescript II SK$^+$ vector. The *E. coli* TG1 clones which were transformed with these recombinant plasmids were tested for Cm^r. Only clones that harbored an *Mbo*I insert of about 2,630 bp, resulting from partial pSCS7 digestion, exhibited Cm^r. All clones which contained *Mbo*I inserts of only 1,250 or 1,380 bp showed no Cm^r independent of the orientation of the inserts. This suggested that the *Mbo*I cleavage site which separated these two fragments is located within the *cat* gene (Fig. 1).

Nucleotide sequence of the pSCS7 encoded *cat* gene. The two previously defined *Mbo*I fragments were used for sequence analyses. These two fragments could be distinguished by digestion with *Hind*III and *Cla*I. The single
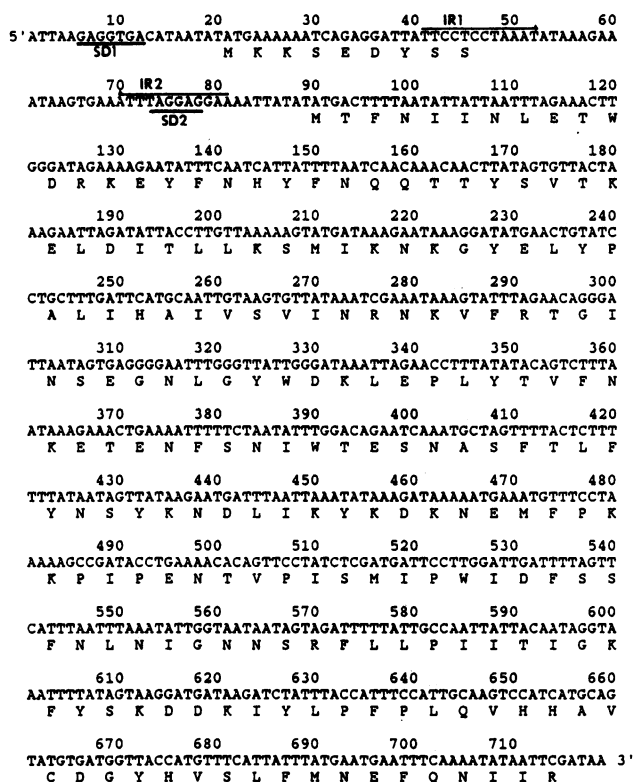

FIG. 2. Nucleotide sequence of a 719-bp region of pSCS7 containing the structural *cat* gene as well as its regulatory region, presented as the noncoding strand. The two possible ribosomal binding sites (SD1, SD2) as well as the two inverted repeats (IR1, IR2) are indicated. The amino acid sequences of the two ORFs, which were predicted from the respective nucleotide sequences, are displayed as a single-letter code.

*Hind*III site of pSCS7 occurred in the 1,380-bp fragment, and the single *Cla*I site of pSCS7 was located in the 1,250-bp fragment. Moreover, these digests served to determine the orientations of the inserts, since cleavage sites for *Hind*III and *Cla*I appeared only once in the cloning vector. Sequence analyses started at those ends of the respective inserts which were purportedly within the *cat* gene. First, commercially available 17-mer primers were used. The sequence data achieved with these primers were then used to choose another four 16-mer oligonucleotide primers which enabled the completion of the *cat* gene sequence (Fig. 2).

Analysis of the nucleotide sequence revealed the presence of two open reading frames (ORFs). ORF1 encoded a peptide of 9 amino acids (positions 21 to 47) and was preceded by a potential ribosome-binding site 1 (Shine-Dalgarno-sequence 1 [SD1]; positions 6 to 12). The final 10 nucleotides of ORF1, including the stop codon, were found to be a part of inverted repeat 1 (IR1; positions 41 to 52), while the SD2 sequence (positions 74 to 79) proved to be located within IR2 (positions 70 to 81). The start codon for ORF2 was found 10 bp 3' from the SD2 sequence. ORF2 (positions 90 to 716) encoded a protein of 209 amino acids which was considered to represent one monomer of the enzyme CAT. Some 97 bp of the CAT-coding sequence as well as the stop codon were located within the 1,380-bp fragment, while the remaining 530 bp of the *cat* gene was detected in the 1,250-bp fragment.

Comparisons with CAT enzymes of other bacteria. The pSCS7-encoded CAT monomer was slightly smaller than the previously determined CAT monomers in staphylococci, which ranged between 215 and 219 amino acids (2, 3, 13, 22, 25, 28). Comparisons revealed that the *cat* gene of pSCS7 exhibited 90% nucleotide sequence homology and 85% amino acid identity to the *cat* gene of pSCS5. A slightly lower homology of 84% in the nucleotide sequences and 80% in the predicted amino acid sequences were observed between the *cat* genes of pSCS7 and pC221. However, distinct differences could be detected by comparing the *cat* genes of pSCS7 and pC194. These comparisons revealed 66% nucleotide sequence homology and 57% amino acid identity.

An alignment of the amino acid sequences of known bacterial CAT variants is shown in Fig. 3. This alignment revealed the presence of 34 completely conserved amino acids in all CAT variants. The amino acid identities shown in Table 1 were calculated on the basis of this alignment. Thus, the pSCS7-encoded CAT showed the highest identity to the other staphylococcal CAT variants (57 to 85%), followed by the CAT variants of the gram-positive, endospore-forming bacteria *Clostridium difficile* (47%), *Clostridium perfringens* (47%), and *Bacillus pumilis* (42%). However, the pSCS7-encoded CAT shared less than 40% identity with the CAT variants of gram-negative bacteria, such as type III CAT from *Shigella flexneri* (36%), type I CAT from *E. coli* (38%), and CAT from *Proteus mirabilis* (39%). Type I CAT was indistinguishable in its amino acid sequence from the CAT variant designated Tn*204 cam* (18) and that encoded by the transposon Tn*9* (1), both from *E. coli*.

Phylogeny of CATs. A phylogenetic tree was constructed for the amino acid sequences of the 12 CAT variants. The respective cladogram is shown in Fig. 4. The branching order followed the identity calculations. The numbers on the branches represent relative evolutionary distances calculated according to the values from the distance matrix.

This phylogenetic tree suggested the presence of at least four main groups of CAT enzymes. Group 1 included the staphylococcal CAT variants encoded by the plasmids pC194, pC221, pUB112, pSCS1, pSCS5, and pSCS7. Within the staphylococcal CAT variants, three subgroups could be distinguished in the phylogenetic tree. Interestingly, the Cm^r plasmids which encoded these CAT variants have been classified on the basis of restriction endonuclease analysis, DNA-DNA hybridization, and heteroduplex analyses to belong to three different families of staphylococcal Cm^r plasmids (10). Thus, the differences observed in the restriction maps of the Cm^r plasmids could be strongly confirmed by the identity calculations of the respective CAT amino acid sequences.

The second group was represented by the CAT variants of gram-negative bacteria, such as type III CAT, type I CAT, and CAT from *P. mirabilis*. Within this group, the best-studied CAT enzyme, type III CAT, appeared to have branched off relatively early in evolution, while the other two enzymes of this group exhibited higher degrees of sequence identity and shorter evolutionary distances to one another. Relatively low degrees of amino acid identity to one another as well as to all other CAT variants could be observed for the CAT enzymes from *B. pumilis* (39 to 43%) and *Clostridium* species (44 to 47%). This suggested that these CAT variants might be representatives of individual evolutionary directions.

DISCUSSION

Cm^r is observed in a wide variety of bacterial species. It is usually due to an enzymatic inactivation of the drug by acetylation (31). The enzyme responsible for the acetylation was often found to be encoded by extrachromosomal, mobile DNA elements (17, 29). In staphylococci, Cm^r is the only known resistance marker, which is, to the best of our knowledge, exclusively located on plasmids (17, 25–28). In gram-negative bacteria, the determinants for Cm^r have also been found to be associated with plasmids or transposons (1, 18, 19, 29). However, chromosomally encoded Cm^r, which produces a permeability barrier, has also been found in gram-negative bacteria (4, 34). Despite the widespread occurrence and the clinical importance of Cm^r, only 13 *cat* genes have been sequenced. Of these, five were isolated from gram-negative bacteria, including *E. coli* (1, 18, 30), *S. flexneri* (19), and *P. mirabilis* (6). Another five *cat* genes have been found in *S. aureus* (2, 3, 13), *S. intermedius* (28), and *S. haemolyticus* (25). *cat* genes also were isolated from *C. difficile* (36), *C. perfringens* (32), and *B. pumilis* (11).

We sequenced the *cat* gene of the plasmid pSCS7 from *S. aureus*, which was previously identified to be an analog of the prototype plasmid pC223 (5). To date, pC223 has been analyzed for its replication properties (9), but the nucleotide sequence of its *cat* gene is not available from a data library. However, hybridization studies, heteroduplex analyses, and differences in the restriction maps suggested that pC223 and its analogs might carry *cat* genes which differ from those of the prototype plasmids pC221 and pC194 (10). Comparisons of the nucleotide sequence and the predicted amino acid sequence of the pSCS7-encoded CAT confirmed the expected differences. Despite the differences in the *cat* structural genes, the expression of all staphylococcal *cat* genes, including that of pSCS7, is inducible (2, 3, 13, 22, 25–28) and is regulated by translational attenuation (3, 13, 16, 28). The role of the small ORF for the 9-amino-acid peptide in CAT induction had been studied in detail for the *S. aureus* plasmid pUB112 (3) and *cat-86* from *B. pumilis* (16). The nucleotide sequence of the 9-amino-acid peptide proved to be identical in pC221, pUB112, pSCS1, pSCS5, and pSCS7. However, in pC194 this peptide consisted only of 6 amino acids which also differed from those of the 9-amino-acid peptide (13). All the differences observed among *cat* genes of closely related bacterial species such as *S. aureus*, *S. intermedius*, and *S. haemolyticus* encouraged further comparisons with *cat* genes of gram-negative bacteria and those of gram-positive, endospore-forming bacteria. Independent of the bacteria from which CAT was isolated, these enzymes revealed some remarkable structural similarities. All known CAT variants proved to have a trimeric structure composed of identical subunits (14, 25, 29). Furthermore, the sizes of the CAT monomers varied in a narrow range between 209 and 220 amino acids (1, 6, 19, 25, 29, 30).

For a better understanding of the relationships between these functionally homologous enzymes, we performed an alignment of their amino acid sequences. The progressive alignment showed that a total of 34 amino acids were highly conserved among all CAT variants. This observation suggests that the conserved residues might play a role in the structure and/or the function of the CAT enzymes. So far, structurally or functionally important amino acids have only been identified by X-ray crystallography of type III CAT (14, 15). Of the 20 amino acids which were considered to play a role in connecting the CAT monomers in type III CAT, only three (Met-1, Asn-143, Asn-161) were found to be con-

```
CAT pSCS5     MT-----FNIINLETWDRKEYFNHYFNQQT-TYSVTKEFDITLLKSMIKNKGYELYPALI
CAT pSCS7     MT-----FNIINLETWDRKEYFNHYFNQQT-TYSVTKELDITLLKSMIKNKGYELYPALI
CAT pUB112    MT-----FNIIKLENWDRKEYFEHYFNQQT-TYSITKEIDITLFKDMIKKKGYEIYPSLI
CAT pSCS1     MT-----FNIIKLENWDRKEYFEHYFNQQT-TYSITKEIDITLFKDMIKKKGYEIYPSLI
CAT pC221     MT-----FNIIKLENWDRKEYFEHYFNQQT-TYSITKEIDITLFKDMIKKKGYEIYPSLI
CAT pC194     MN-----FNKIDLDNWKRKEIFNHYLNQQT-TFSITTEIDISVLYRNIKQEGYKFYPAFI
CAT D         MV-----FEKIDKNSWNRKEYFDHYFASVPCTYSMTVKVDITQ----IKEKGMKLYPAML
CAT P         MV-----FEKIDKNSWNRKEYFDHYFASVPCTYSMSLKVDITQ----IKEKGMKLYPAML
CAT 86        M------FKQIDENYL-RKEHFHHYMTLTRCSYSLVINLDITKLHAILKEKKLKVYPVQI
TypeI CAT     MEKKITGYTTVDISQWHRKEHFEAFQSVAQCTYNQTVQLDITAFLKTVKKNKHKFYPAFI
CAT Pm        MDTKRVGILVVDLSQWGRKEHFEAFQSFAQCTFSQTVQLDITSLLKTVKQNGYKFYPTFI
TypeIII CAT   MN-----YTKFDVKNWVRREHFEFYRHRLPCGFSLTSKIDITTLKKSLDDSAYKFYPVMI
                 *            *  *  *                      **              **


CAT pSCS5     YTIVNIINQNKVFRTGINSSGNLGFWDKLNPLYTCLINETEKFLNIWIESNVSFSSFYN
CAT pSCS7     HAIVSVINRNKVFRTGINSEGNLGYWDKLEPLYTVFNKETENFSNIWTESNASFTLFYN
CAT pUB112    YAIMEVVNKNKVFRTGINSENKLGYWDKLNPLYTVFNKQTEKFTNIWTESDNNFTSFYN
CAT pSCS1     YAIMEVVNKNKVFRTGINSENKLGYWDKLNPLYTVFNKQTEKFTNIWTESDNNFTSFYN
CAT pC221     YAIMEVVNKNKVFRTGINSENKLGYWDKLNPLYTVFNKQTEKFTNIWTESDNNFTSFYN
CAT pC194     FLVTRVINSNTAFRTGVNSDGELGYWDKLEPLYTIFDGVSKTFSGIWTPVKNDFKEFYD
CAT D         YYIAMIVNRHSEFRTAINQDGELGIYDEMIPSYTIFHNDTETFSSLWTECKSDFKSFLA
CAT P         YYIAMIVNRHSEFRTAINQDGELGIYDEMIPSYTIFHNDTETFSSLWTECKSDFKSFLA
CAT 86        YLLARAVQKIPEFRMDQVND-ELGYWEILHPSYTILNKETKTFSSIWTPFDENFAQFYK
TypeI CAT     HILARLMNAHPEFRMAMK-DGELVIWDSVHPCYTVFHEQTETFSSLWSEYHDDFRQFLH
CAT Pm        YIISLLVNKHAEFRMAMK-DGELVIWDSVNPGYNIFHEQTETFSSLWSYYHKDINRFLK
TypeIII CAT   YLIAQAVNQFDELRMAIKDD-ELIVWDSVDPQFTVFHQETETFSALSCPYSSDIDQFMV
                 *           *          *               *                *


CAT pSCS5     SYKSDLLEYKDTNEMFPKKPIPENTGPISMIPWIDFSSFNLNIGNNSRFLLPIITIGKF
CAT pSCS7     SYKNDLIKYKDKNEMFPKKPIPENTVPISMIPWIDFSSFNLNIGNNSRFLLPIITIGKF
CAT pUB112    NYKNDLFEYKDKEEMFPKKPIPENTIPISMIPWIDFSSFNLNIGNNSSFLLPIITIGKF
CAT pSCS1     NYKNDLFEYKDKEEMFPKKPIPENTIPISMIPWIDFSSFNLNIGNNSSFLLPIITIGKF
CAT pC221     NYKNDLLEYKDKEEMFPKKPIPENTIPISMIPWIDFSSFNLNIGNNSNFLLPIITIGKF
CAT pC194     LYLSDVEKYNGSGKLFPKTPIPENAFSLSIIPWTSFTGFNLNINMNSNYLLPIITAGKF
CAT D         DYESDTQRYGNNHRMEGKPNAPENIFNVSMIPWSTFDGFNLNLQKGYDYLIPIFTMGKI
CAT P         DYESDTQRYGNNHRMEGKPNAPENIFNVSMIPWSTFDGFNLNLQKGYDYLIPIFTMGKI
CAT 86        SCVADIETFSKSSNLFPKPHMPENMFNISSLPWIDFTSFNLNVSTDEAYLLPIFTIGKF
TypeI CAT     IYSQDVACYGENLAYFPKGFI-ENMFFVSANPWVSFTSFDLNVANMDNFFAPVFTMGKY
CAT Pm        TYSEDIAQYGDDLAYFPKEFI-ENMFFVSANPWVSFTSFNLNMANINNFFAPVFTIGKY
TypeIII CAT   NYLSVMERYKSDTKLFPQGVTPENHLNISALPWVNFDSFNLNVANFTDYFAPIITMAKY
                      **       *  **    *  *  **           *    *   *


CAT pSCS5     YSKNNKIYLPVSLQVHHAVCDGYHVSL-------FMSEFQNIVDSVNEWDLNLK-------
CAT pSCS7     YSKDDKIYLPFPLQVHHAVCDGYHVSL-------FMNEFQNI--------IR---------
CAT pUB112    YSENNKIYIPVALQLHHAVCDGYHASL-------FINEFQDI--------IN-K--VDDWI
CAT pSCS1     YSENNKIYIPVALQLHHAVCDGYHASL-------FINEFQDI--------IK-K--VDDWI
CAT pC221     YSENNKIYIPVALQLHHAVCDGYHASL-------FMNEFQDI--------IH-K--VDDWI
CAT pC194     INKGNSIYLPLSLQVHHSVCDGYHAGL-------FMNSIQEL--------SDRP--NDWLL
CAT D         IKKDNKIILPLAIQVHHAVCDGFHICR-------FVNELQEL--------II-V--TQVCL
CAT P         IKKDNKIILPLAIQVHHAVCDGFHICR-------FVNELQEL--------II-V--TQVCL
CAT 86        KVEEGKIILPVAIQVHHAVCDGYHAGQYVEYLRWLIEHCDEW--------LN-D--SLHIT
TypeI CAT     YTQGDKVLMPLAIQVHHAVCDGFHVGR-------MLNELQQY--------CD-EWQGG--A
CAT Pm        YTQGDKVLMPLAIQVHHAVCDGFHVGR-------LLNEIQQY--------CD-E--GC--K
TypeIII CAT   QQEGDRLLLPLSVQVHHAVCDGFHVAR-------FINRLQEL--------CN-S--KL--K
                 *    *  **  ****   *
```

FIG. 3. Alignment of the amino acid sequence of the pSCS7-encoded CAT with the sequences of CAT variants from other bacteria according to the alignment program of Hein (12). Abbreviations are as follows: CAT pSCS5, CAT encoded by the plasmid pSCS5 from canine *S. haemolyticus* (25); CAT pSCS7, CAT encoded by the plasmid pSCS7 from bovine *S. aureus*; CAT pUB112, CAT encoded by the plasmid pUB112 from human *S. aureus* (3); CAT pSCS1, CAT encoded by the plasmid pSCS1 from canine *S. intermedius* (28); CAT pC221, CAT encoded by the plasmid pC221 from human *S. aureus* (2, 22); CAT pC194, CAT encoded by the plasmid pC194 from human *S. aureus* (13); CAT D, chromosomally encoded CAT from a clinical isolate of *C. difficile* (36); CAT P, CAT encoded by the plasmid pIP401 from *C. perfringens* (32); CAT 86, chromosomally encoded CAT from *B. pumilis* NCIB 8600 (11); Type I CAT, CAT encoded by the plasmid JR66b from *E. coli* (30); CAT Pm, chromosomally encoded CAT from *P. mirabilis* (6); Type III CAT, CAT encoded by the plasmid R387 from *S. flexneri* (19). Asterisks indicate conserved amino acids in all CAT sequences.

served. In type III CAT, 26 amino acids were involved in the folding of the CAT monomers. However, only seven of them, namely, Arg-18, Asp-40, Arg-74, Glu-142, Leu-160, Gln-192, and Asp-199, proved to be conserved, as shown in Fig. 3. A total of 17 amino acids were found to be involved in the formation of the chloramphenicol-binding pocket in

type III CAT (14). Five of these residues were conserved in all CAT variants: Phe-103, Ser-148, Leu-160, Thr-174, and His-195. Two residues exhibited direct hydrogen bonds to chloramphenicol (14). These were His-195 and Tyr-25, of which the latter was replaced by Phe in type I CAT and CAT from *P. mirabilis*. The conserved Thr-174 also contributed to

TABLE 1. Amino acid sequence identities of CAT variants found in different bacterial species

Variant	% Amino acid sequence identity[a]											
	CAT 86	CAT D	CAT P	Type I CAT	CAT Pm	CAT pC194	CAT pC221	Type III CAT	CAT pUB112	CAT pSCS1	CAT pSCS5	CAT pSCS7
CAT 86	100	43	43	39	39	42	42	40	43	42	41	42
CAT D		100	99	45	44	46	44	44	44	44	47	47
CAT P			100	45	44	46	44	44	44	44	47	47
Type I CAT				100	76	38	37	47	37	37	35	38
CAT Pm					100	40	39	47	39	39	38	39
CAT pC194						100	54	41	53	53	54	57
CAT pC221							100	35	98	98	78	80
Type III CAT								100	36	36	36	36
CAT pUB112									100	99	77	80
CAT pSCS1										100	77	80
CAT pSCS5											100	85
CAT pSCS7												100

[a] Values were calculated as described by Hein (12) from the alignment in Fig. 3; CAT abbreviations are as defined in the legend to Fig. 3.

substrate binding via a bridging water molecule (14). Of the two residues (Leu-29, Phe-135) which were located at the entrance to the chloramphenicol-binding pocket in type III CAT, Leu-29 appeared to be relatively variable. It was altered in all staphylococcal CAT variants to Gln, while it was replaced in the other CAT enzymes by Val, Ala, or Thr. However, Phe-135 was found to be exchanged by Glu, solely in the clostridial CAT variants. The staphylococcal CAT variants exhibited other amino acid alterations in the chloramphenicol-binding pocket, such as the absence of Cys-31 and the replacement of Phe-24 by His, Gln-92 by Leu,
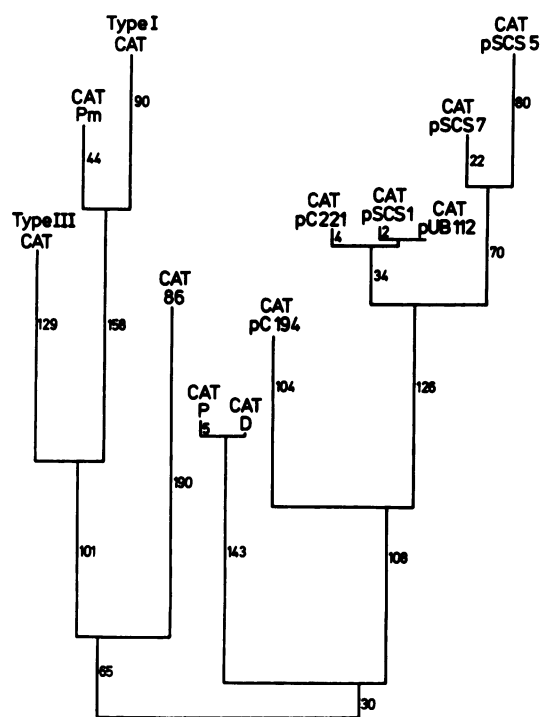

FIG. 4. Cladogram of CAT variants from gram-negative and gram-positive bacteria determined by the method of Hein (12). Branch lengths were determined on the basis of the amino acid alignment; the branch length values indicate relative phylogenetic distances.

Ala-105 by Asn or Gly (pC194), Asn-146 by Pro or Ser (pC194), Val-162 by Ile, and finally, Tyr-168 by Phe.

The comparisons of CAT enzymes of different bacteria suggested that amino acid exchanges also occurred in the regions obviously associated with structural and functional properties. These alterations might be expected to be a consequence of the extrachromosomal nature of most *cat* genes. Such plasmid- or transposon-encoded genes were usually not essential for the survival of the bacteria under normal conditions. Therefore, they could undergo more radical evolutionary changes than chromosomal genes without affecting the viability of the bacteria (17). These changes might result in a variety of phenotypically homologous, but genotypically unique, enzymes. The origins of these enzymes remain unclear, since neither a common ancestor for CAT has been found (29) nor distinct sequence homology to other proteins has become obvious (29). Moreover, a naturally occurring substrate which would be able to replace chloramphenicol as an acetyl acceptor and which could provide insight into the origin of CAT enzymes also has not been detected (29). Therefore, the phylogenetic tree shown in Fig. 4 represents an unrooted tree. Nevertheless, it shows that, among the CAT enzymes, a certain diversity can be observed.

We carefully suggested a possible classification into four main groups. The relatively long evolutionary distances shown in the cladogram suggest that the different directions presumably diverged relatively early in evolution. This might be confirmed by changes in the substrate spectrum and the properties of the CAT variants. Thus, type I CAT also mediated resistance to the steroid antibiotic fusidic acid (18) and was able to bind rosaniline dyes (21). These properties could not be observed either in the CAT variants of gram-positive or in the other CAT variants of gram-negative bacteria.

Furthermore, the expression of all staphylococcal CAT variants and that of the CAT from *B. pumilis* was found to be inducible, whereas the *cat* gene expression in gram-negative bacteria, *C. difficile*, and *C. perfringens* occurred constitutively. Within the staphylococcal group, the early branching off of the pC194-encoded CAT was associated with distinct differences not only in the structural *cat* gene but also in the respective regulatory region. The three CAT enzymes of the pC221 subgroup were separated by very short evolutionary distances. This might indicate that the CAT variants of

pSCS1 and pUB112 are very recent derivatives of the pC221-encoded CAT. The alignment of the amino acid sequences revealed only four amino acid substitutions among the three members of the pC221 CAT family. The differences in the amino acid sequences between the CAT variants of pSCS5 from canine *S. haemolyticus* and pSCS7 from bovine *S. aureus* might be a consequence of adaption to the different host bacteria.

In conclusion, CAT encoded by the *S. aureus* plasmid pSCS7 corresponded more closely to the staphylococcal CAT variants than to the CAT variants of other bacteria. As suggested from the restriction map, the differences in the nucleotide sequence and the predicted amino acid sequence allowed its assignment to a third staphylococcal CAT family, which differed from those represented by the CAT variants from pC221 and pC194. This classification was confirmed by the phylogenetic analyses of the known CAT enzymes.

REFERENCES

1. **Alton, N. K., and D. Vapnek.** 1979. Nucleotide sequence analysis of the chloramphenicol resistance transposon (Tn9). Nature (London) **282**:864–869.
2. **Brenner, D. G., and W. V. Shaw.** 1985. The use of synthetic oligonucleotides with the universal templates for rapid DNA sequencing: results with staphylococcal replicon pC221. EMBO J. **4**:561–568.
3. **Brückner, R., and H. Matzura.** 1985. Regulation of the inducible chloramphenicol acetyltransferase gene of the *Staphylococcus aureus* plasmid pUB112. EMBO J. **4**:2295–2300.
4. **Burns, J. L., P. M. Mendelman, J. Levy, T. L. Stull, and A. L. Smith.** 1985. A permeability barrier as a mechanism of chloramphenicol resistance in *Haemophilus influenzae*. Antimicrob. Agents Chemother. **27**:46–54.
5. **Cardoso, M., and S. Schwarz.** Submitted for publication.
6. **Charles, I. G., J. W. Keyte, and W. V. Shaw.** 1985. Nucleotide sequence analysis of the *cat* gene of *Proteus mirabilis*: comparisons with the type I (Tn9) *cat* gene. J. Bacteriol. **164**:123–129.
7. **Dagert, M., and S. D. Ehrlich.** 1979. Prolonged incubation in calcium chloride improves the competence of *E. coli* cells. Gene **6**:23–28.
8. **Dayhoff, M. O.** 1978. Atlas of protein sequences and structure. vol. 5, suppl. 3. National Biomedical Research Foundation, Washington, D.C.
9. **Ehret, M., and H. Matzura.** 1988. Replication control of the *Staphylococcus aureus* chloramphenicol resistance plasmids pC223 and pUB112 in *Bacillus subtilis*. Nucleic Acids Res. **16**:2045–2962.
10. **Gillespie, M. T., and R. A. Skurray.** 1988. Structural relationships among chloramphenicol-resistance plasmids of *Staphylococcus aureus*. FEMS Microbiol. Lett. **51**:205–210.
11. **Harwood, C. R., D. M. Williams, and P. S. Lovett.** 1983. Nucleotide sequence of a *Bacillus pumilis* gene specifying chloramphenicol acetyltransferase. Gene **24**:163–169.
12. **Hein, J.** 1990. Unified approach to alignment and phylogenies. Methods Enzymol. **183**:626–645.
13. **Horinouchi, S., and B. Weisblum.** 1982. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bacteriol. **150**:815–825.
14. **Leslie, A. G. W.** 1990. Refined crystal structure of type III chloramphenicol acetyltransferase at 1.75 A resolution. J. Mol. Biol. **213**:167–186.
15. **Leslie, A. G. W., P. E. C. Moody, and W. V. Shaw.** 1985. Structure of chloramphenicol acetyltransferase at 1.75 A resolution. Proc. Natl. Acad. Sci. USA **85**:4133–4137.
16. **Lovett, P. S.** 1990. Translation attenuation as the regulator of inducible *cat* genes. J. Bacteriol. **172**:1–6.
17. **Lyon, B. R., and R. Skurray.** 1987. Antimicrobial resistance of *Staphylococcus aureus*: genetic basis. Microbiol. Rev. **51**:88–134.
18. **Marcoli, R., S. Iida, and T. A. Bickle.** 1980. The DNA sequence of an IS1-flanked transposon coding for resistance to chloramphenicol and fusidic acid. FEBS Lett. **110**:11–14.
19. **Murray, I. A., A. R. Hawkins, J. W. Keyte, and W. V. Shaw.** 1988. Nucleotide sequence analysis and overexpression of the gene encoding a type III chloramphenicol acetyltransferase. Biochem. J. **252**:173–179.
20. **Novick, R. P.** 1976. Plasmid-protein relaxation complexes in *Staphylococcus aureus*. J. Bacteriol. **127**:1177–1187.
21. **Proctor, G. N., and R. H. Rownd.** 1982. Rosanilins: indicator dyes for chloramphenicol resistant enterobacteria containing chloramphenicol acetyltransferase. J. Bacteriol. **150**:1375–1382.
22. **Projan, S. J., J. Kornblum, S. L. Moghazeh, I. Edelman, M. L. Gennaro, and R. P. Novick.** 1985. Comparative sequence and functional analysis of pT181 and pC221, cognate plasmid replicons from *Staphylococcus aureus*. Mol. Gen. Genet. **199**:452–464.
23. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Sanger, F., S. Nicklen, and A. Coulsen.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
25. **Schwarz, S., and M. Cardoso.** 1991. Molecular cloning, purification, and properties of a plasmid-encoded chloramphenicol acetyltransferase from *Staphylococcus haemolyticus*. Antimicrob. Agents Chemother. **35**:1277–1283.
26. **Schwarz, S., M. Cardoso, and H. Blobel.** 1989. Plasmid-mediated chloramphenicol resistance in *Staphylococcus hyicus*. J. Gen. Microbiol. **135**:5329–3336.
27. **Schwarz, S., M. Cardoso, and H. Blobel.** 1990. Detection of a novel chloramphenicol resistance plasmid from "equine" *Staphylococcus sciuri*. J. Vet. Med. Ser. B **37**:674–679.
28. **Schwarz, S., U. Spies, and M. Cardoso.** 1991. Cloning and sequence analysis of a plasmid-encoded chloramphenicol acetyltransferase gene from *Staphylococcus intermedius*. J. Gen. Microbiol. **137**:977–981.
29. **Shaw, W. V.** 1983. Chloramphenicol acetyltransferase: enzymology and molecular biology. Crit. Rev. Biochem. **14**:1–46.
30. **Shaw, W. V., L. C. Packman, B. D. Burleigh, A. Dell, H. R. Morris, and B. S. Hartley.** 1979. Primary structure of a chloramphenicol acetyltransferase specified by bacterial plasmids. Nature (London) **282**:870–872.
31. **Shaw, W. V., and J. Unowsky.** 1968. Mechanism of R factor-mediated chloramphenicol resistance. J. Bacteriol. **95**:1976–1978.
32. **Steffen, C., and H. Matzura.** 1989. Nucleotide sequence analysis of a chloramphenicol-acetyltransferase coding gene from *Clostridium perfringens*. Gene **75**:349–354.
33. **Tabor, S., and G. C. Richardson.** 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA **84**:4767–4771.
34. **Toro, C. S., S. R. Lobos, I. Calderson, M. Rodriguez, and G. C. Mora.** 1990. Clinical isolate of a porinless *Salmonella typhi* resistant to high levels of chloramphenicol. Antimicrob. Agents Chemother. **34**:1715–1719.
35. **Vieira, J., and J. Messing.** 1987. Production of single stranded plasmid DNA. Methods Enzymol. **153**:3–11.
36. **Wren, B. W., P. Mullany, C. Clayton, and S. Tabaqchali.** 1989. Nucleotide sequence of a chloramphenicol acetyltransferase gene from *Clostridium difficile*. Nucleic Acids Res. **17**:4877.