Volume 16 Number 5 1988 Nucleic Acids Research

Nucleotide sequence of pS194, a streptomycin-resistance plasmid from *Staphylococcus aureus*

Steven J.Projan, Soraya Moghazeh and Richard P.Novick

Department of Plasmid Biology, Public Health Research Institute, 455 First Avenue, New York, NY 10016, USA

Received January 12, 1988; Accepted January 27, 1988 Accession no.X06627

ABSTRACT

pS194 is a naturally occurring *Staphylococcus aureus* plasmid encoding streptomycin resistance. The plasmid has a copy number of about 25 per cell, and belongs to the *inc5* incompatibility group. The nucleotide sequence of pS194 has been determined and consists of 4397 base pairs including four open reading frames potentially encoding proteins of greater than 100 amino acids. All four of these reading frames are on the same coding strand. The first reading frame, *repE*, encodes a 38 kd protein specifically required for pS194 replication. The second open reading frame, *str*, encodes a 34 kd polypeptide required for streptomycin resistance, probably a streptomycin adenylyltransferase. The third potential polypeptide, *rlx*, would be 37 kd and is probably required for relaxation complex formation and plasmid mobilization by conjugative plasmids. The fourth, orfD, overlapping the *rlx* reading frame, is potentially 27 kd, and may also be involved in mobilization.

INTRODUCTION

The streptomycin resistance plasmid pS194 was originally obtained from a clinical isolate of *Staphylococcus aureus* by S. Iordanescu (1). It is virtually indistinguishable from the three other known *S. aureus* plasmids conferring streptomycin resistance, namely pS169, pS177 and pUB109 (2) and all of these plasmids belong to the inc5 incompatibility group (2).

On the basis of a variety of observations, it is thought that pS194 belongs to a family of closely related *S. aureus* plasmids of which pT181 is the prototype (3). The known members of this family share considerable base sequence homology (2), have similar overall functional organization (Projan, et al., in preparation) and replicate via an asymmetric rolling circle mechanism in which lagging strand initiation occurs only after leading strand replication is complete or nearly complete (4). They regulate copy number by means of short antisense transcripts that inhibit the translation of a trans-acting leading strand initiation factor (5).

Previous studies have identified the leading and lagging strand replication origins of pS194 and have demonstrated the existence of a trans-acting leading strand initiation factor, RepE, for this plasmid (Projan, et al., in preparation). A single regulatory countertranscript has been identified (5) and the streptomycin resistance determinant has been localized by the finding that interruption of the plasmid's unique HindIII site inactivates resistance (6).

All of the Sm^r plasmids of the pS194 type form relaxation complexes (7) which are involved

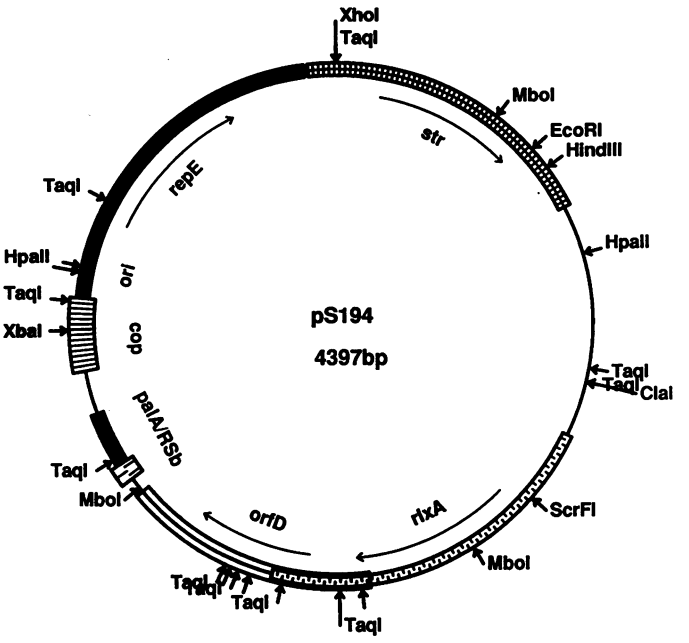


Figure 1. Physcial map of pS194. Restriction sites indicated are predicted from the nucleotide sequence. *repE* is the gene for the plasmid encoded initiator protein, *str* is the streptomycin resistance determinant, *rlxA* is a gene required for relaxation complex formation, *orfD* encodes a putative protein required for plasmid mobilization by conjugative plasmids, *palA*/RS_B is the lagging strand origin of replication, *cop* is the copy control region, *ori* is the leading strand origin of replication. The arrows indicate the direction of transcription.

in conjugative mobilization (Archer and Projan, in preparation); the region of the plasmid involved in relaxation complex formation has also been localized (Projan, et al., in preparation).

In this paper, we present the pS194 nucleotide sequence which confirms that the plasmid is closely related to other members of the pT181 family and has a similar functional organization.

MATERIALS AND METHODS

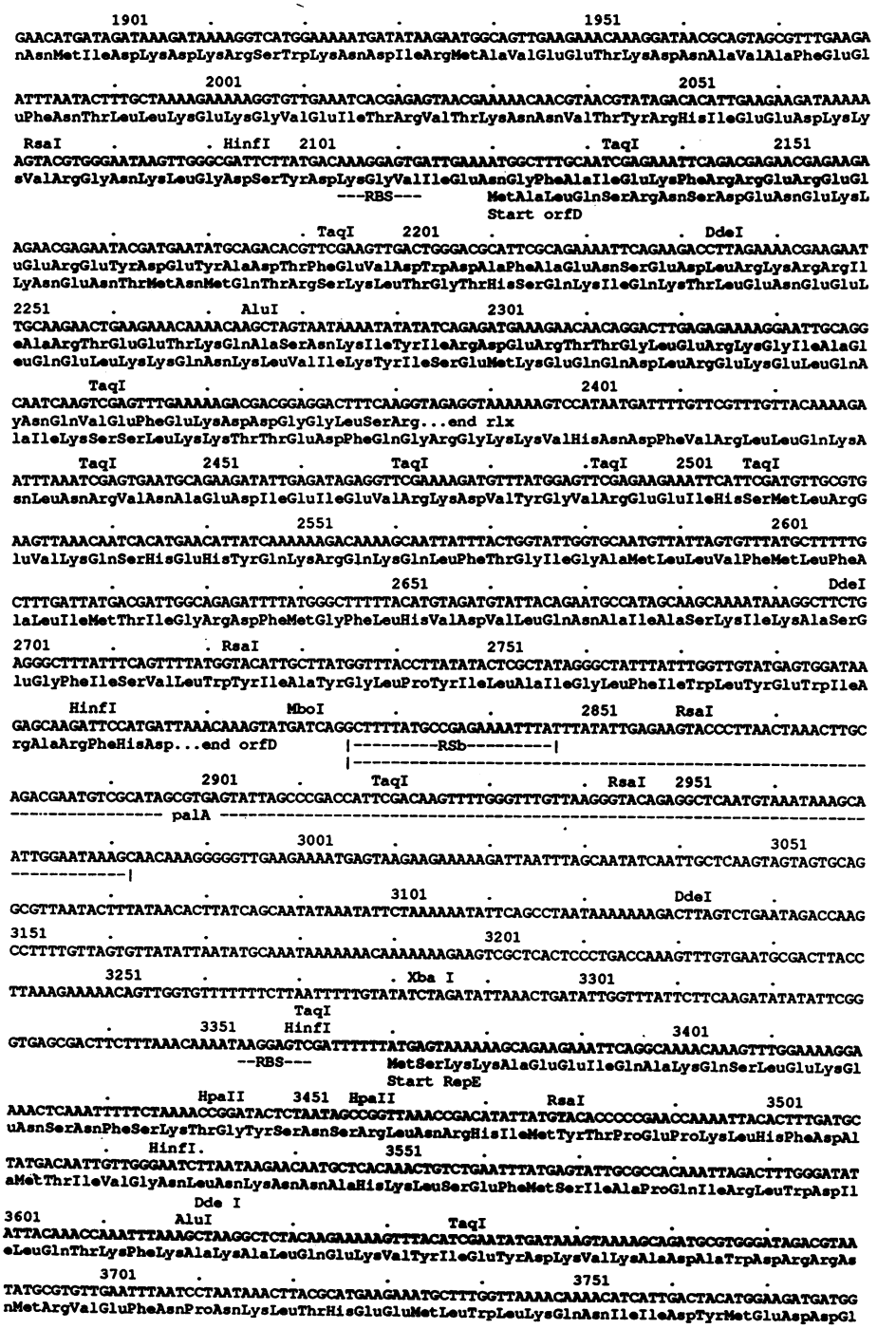
Bacterial strains and plasmids.

The staphylococcal strains used are derivatives of strain NTCC 8325. (8).

Media and growth conditions.

CY broth was used for liquid cultures of *S. aureus*, and VY broth (9) was used for liquid cultures of *B. subtilis*. Cultures were shaken at 37°C (unless otherwise indicated) and monitored turbidometrically with a Klett-Summerson photoelectric colorimeter with a green filter. GL agar (*S. aureus*) (9) or TBAB agar (*B. subtilis*) (Difco) was supplemented with antibiotics as indicated.

XhoI
TaqI
HinfI
MboI
HindIII
AluI
EcoRI
HpaII
Cla I
HinfI
Dde I
---RBS---
Start rlx
ScrFI


1901 1951
GAACATGATAGATAAAGATAAAAGGTCATGGAAAAATGATATAAGAATGGCAGTTGAAGAAACAAAGGATAACGCAGTAGCGTTTGAAGA
nAsnMetIleAspLysAspLysArgSerTrpLysAsnAspIleArgMetAlaValGluGluThrLysAspAsnAlaValAlaPheGluGl
2001 2051
ATTTAATACTTTGCTAAAAGAAAAAGGTGTTGAAATCACGAGAGTAACGAAAAACAACGTAACGTATAGACACATTGAAGAAGATAAAAA
uPheAsnThrLeuLeuLysGluLysGlyValGluIleThrArgValThrLysAsnAsnValThrTyrArgHisIleGluGluAspLysLy
RsaI HinfI 2101 TaqI 2151
AGTACGTGGGAATAAGTTGGGCGATTCTTATGACAAAGGAGTGATTGAAAATGGCTTTGCAATCGAGAAATTCAGACGAGAACGAGAAGA
sValArgGlyAsnLysLeuGlyAspSerTyrAspLysGlyValIleGluAsnGlyPheAlaIleGluLysPheArgArgGluArgGluGl
---RBS--- MetAlaLeuGlnSerArgAsnSerAspGluAsnGluLysL
Start orfD
TaqI 2201 DdeI
AGAACGAGAATACGATGAATATGCAGACACGTTCGAAGTTGACTGGGACGCATTCGCAGAAAATTCAGAAGACCTTAGAAAACGAAGAAT
uGluArgGluTyrAspGluTyrAlaAspThrPheGluValAspTrpAspAlaPheAlaGluAsnSerGluAspLeuArgLysArgArgIl
LyAsnGluAsnThrMetAsnMetGlnThrArgSerLysLeuThrGlyThrHisSerGlnLysIleGlnLysThrLeuGluAsnGluGluL
2251 AluI 2301
TGCAAGAACTGAAGAAACAAAACAAGCTAGTAATAAAATATATCAGAGATGAAAGAACAACAGGACTTGAGAGAAAAGGAATTGCAGG
eAlaArgThrGluGluThrLysGlnAlaSerAsnLysIleTyrIleArgAspGluArgThrThrGlyLeuGluArgLysGlyIleAlaGl
euGlnGluLeuLysLysGlnAsnLysLeuValIleLysTyrIleSerGluMetLysGluGlnGlnAspLeuArgGluLysGluLeuGlnA
TaqI 2401
CAATCAAGTCGAGTTTGAAAAAGACGACGGAGGACTTTCAAGGTAGAGGTAAAAAAGTCCATAATGATTTTGTTCGTTTGTTACAAAAGA
yAsnGlnValGluPheGluLysAspAspGlyGlyLeuSerArg...end rlx
laIleLysSerSerLeuLysLysThrThrGluAspPheGlnGlyArgGlyLysLysValHisAsnAspPheValArgLeuLeuGlnLysA
TaqI 2451 TaqI 2501 TaqI
ATTTAAATCGAGTGAATGCAGAAGATATTGAGATAGAGGTTCGAAAAGATGTTTATGGAGTTCGAGAAGAAATTCATTCGATGTTGCGTG
snLeuAsnArgValAsnAlaGluAspIleGluIleGluValArgLysAspValTyrGlyValArgGluGluIleHisSerMetLeuArgG
2551 2601
AAGTTAAACAATCACATGAACATTATCAAAAAAGACAAAAGCAATTATTTACTGGTATTGGTGCAATGTTATTAGTGTTTATGCTTTTG
luValLysGlnSerHisGluHisTyrGlnLysArgGlnLysGlnLeuPheThrGlyIleGlyAlaMetLeuLeuValPheMetLeuPheA
2651 DdeI
CTTTGATTATGACGATTGGCAGAGATTTTATGGGCTTTTTACATGTAGATGTATTACAGAATGCCATAGCAAGCAAAATAAAGGCTTCTG
laLeuIleMetThrIleGlyArgAspPheMetGlyPheLeuHisValAspValLeuGlnAsnAlaIleAlaSerLysIleLysAlaSerG
2701 RsaI 2751
AGGGCTTTATTTCAGTTTTATGGTACATTGCTTATGGTTTACCTTATATACTCGCTATAGGGCTATTTATTTGGTTGTATGAGTGGATAA
luGlyPheIleSerValLeuTrpTyrIleAlaTyrGlyLeuProTyrIleLeuAlaIleGlyLeuPheIleTrpLeuTyrGluTrpIleA
HinfI MboI 2851 RsaI
GAGCAAGATTCCATGATTAAACAAAGTATGATCAGGCTTTATGCCGAGAAAATTTATTTATATTGAGAAGTACCCTTAACTAAACTTGC
rgAlaArgPheHisAsp...end orfD |---------RSb---------|
|----------
2901 TaqI RsaI 2951
AGACGAATGTCGCATAGCGTGAGTATTAGCCCGACCATTCGACAAGTTTTGGGTTTGTTAAGGGTACAGAGGCTCAATGTAAATAAAGCA
---------- palA ----------
3001 3051
ATTGGAATAAAGCAACAAAGGGGGTTGAAGAAAATGAGTAAGAAGAAAAAGATTAATTTAGCAATATCAATTGCTCAAGTAGTAGTGCAG
----------|
3101 DdeI
GCGTTAATACTTTATAACACTTATCAGCAATATAAATATTCTAAAAAATATTCAGCCTAATAAAAAAAGACTTAGTCTGAATAGACCAAG
3151 3201
CCTTTTGTTAGTGTTATATTAATATGCAAATAAAAAAACAAAAAAAGAAGTCGCTCACTCCCTGACCAAAGTTTGTGAATGCGACTTACC
3251 Xba I 3301
TTAAAGAAAAACAGTTGGTGTTTTTTTCTTAATTTTGTATATCTAGATATTAAACTGATATTGGTTTATTCTTCAAGATATATATTCGG
TaqI
3351 HinfI 3401
GTGAGCGACTTCTTTAAACAAAATAAGGAGTCGATTTTTTATGAGTAAAAAAGCAGAAGAAATTCAGGCAAAACAAAGTTTGGAAAAGGA
--RBS--- MetSerLysLysAlaGluGluIleGlnAlaLysGlnSerLeuGluLysGl
Start RepE
HpaII 3451 HpaII RsaI 3501
AAACTCAAATTTTTCTAAAACCGGATACTCTAATAGCCGGTTAAACCGACATATTATGTACACCCCGAACCAAAATTACACTTTGATGC
uAsnSerAsnPheSerLysThrGlyTyrSerAsnSerArgLeuAsnArgHisIleMetTyrThrProGluProLysLeuHisPheAspAl
HinfI. 3551
TATGACAATTGTTGGGAATCTTAATAAGAACAATGCTCACAAACTGTCTGAATTTATGAGTATTGCGCCACAAATTAGACTTTGGGATAT
aMetThrIleValGlyAsnLeuAsnLysAsnAsnAlaHisLysLeuSerGluPheMetSerIleAlaProGlnIleArgLeuTrpAspIl
Dde I
3601 AluI TaqI
ATTACAAACCAAATTTAAAGCTAAGGCTCTACAAGAAAAAGTTTACATCGAATATGATAAAGTAAAAGCAGATGCGTGGGATAGACGTAA
eLeuGlnThrLysPheLysAlaLysAlaLeuGlnGluLysValTyrIleGluTyrAspLysValLysAlaAspAlaTrpAspArgArgAs
3701 3751
TATGCGTGTTGAATTTAATCCTAATAAACTTACGCATGAAGAAATGCTTTGGTTAAAACAAAACATCATTGACTACATGGAAGATGATGG
nMetArgValGluPheAsnProAsnLysLeuThrHisGluGluMetLeuTrpLeuLysGlnAsnIleIleAspTyrMetGluAspAspGl

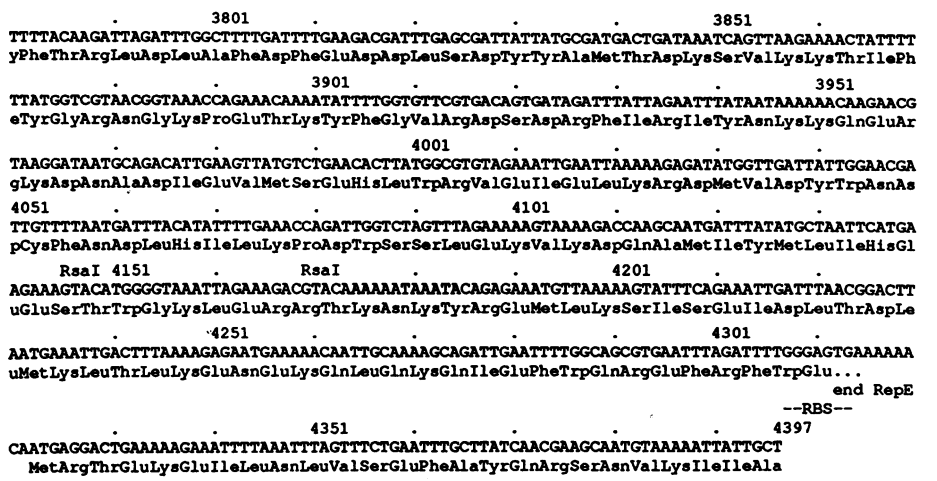


Figure 2. Complete nucleotide sequence of pS194 numbered from the unique XhoI site. Open reading frames of greater than 100 amino acid residues with their potential ribosome binding sites (RBS) are indicated. *palA*, RS_B, and t_1 loci are described in text.

Tetracylcine (Tc), chloramphenicol (Cm), streptomycin (Sm) and erythromycin (Em) were used at 5 μg/ml unless otherwise indicated.

Protoplast transformation was by the method of Chang and Cohen (10) as modified for *S. aureus* (11).

Isolation and analysis of plasmid DNA.

DNA was prepared and analyzed as described (12, 13). Copy numbers were determined by fluorimetric densitometry of ethidium-bromide stained agarose gels run with sheared whole-cell lysates of exponentially growing cultures (13).

Restriction mapping and cloning.

Restriction enzymes were purchased from Boehringer-Mannheim Biochemicals and New England Biolabs. For molecular cloning, specific fragments were extracted from polyacrylamide gels, phenol extracted, ethanol precipitated, and stored lyophilized prior to use (14). For ligations, samples were combined in approximately equimolar ratios and incubated with T4 DNA ligase (Collaborative Research) using a DNA concentration of at least 10 μg/ml and a ligase concentration of 40 units/ml for fragments with complementary ends and 100 units/ml for fragments with blunt ends. Incubation was at room temperature for 1 to 4 hours. Ligated samples were used to transform *S. aureus* protoplasts or *B. subtilis* or *E. coli* competent cells with selection for the appropriate antibiotic resistance marker. Transformants were screened for plasmid content and those carrying plasmids of the required size were used to prepare plasmid DNA for restriction analysis and, if necessary, sequencing.

DNA sequencing.

Determination of nucleotide sequences utilized the dideoxy chain-termination method (15) with isolated restriction fragments cloned into either mp18 or mp19 M13 vectors (16).

RESULTS AND DISCUSSION

A partial restriction map of pS194 determined by digestion with restriction enzymes either singly or in combination was used to begin the sequencing of pS194. As the sequencing of the plasmid proceeded, refinements in the map were made. Two different strategies were employed. Initially, the restriction fragments generated by digestion with MboI, TaqI or HinfI, were each cloned into M13mp18 or mp19. These were sequenced and a nearly complete, single-stranded sequence was assembled. A complete double-stranded sequence was generated by cloning to M13mp18 linear plasmid monomers cleaved by HindIII or ClaI digestion; oligodeoxynucleotides corresponding to regions where sequence information was incomplete were synthesized and used as sequencing primers. This has been shown to be an efficient method for generating a complete double-stranded sequence (17). A map of pS194 identifying the regions involved in replication, streptomycin resistance, relaxation, and copy control together with positions of the four reading frames of significant size is shown in Fig. 1. Fig. 2 shows the sequence of pS194 and the four reading frames described above.

The Replication Region

Cloning experiments have localized the pS194 leading strand replication origin to the 277 bp TaqI-F fragment, and have shown that the reading frame designated *repE* encodes a *trans* acting factor required for leading strand initiation. The TaqI-F fragment can support the replication of a plasmid carrying it providing a second plasmid carrying the *repE* reading frame is also present. As is the case with other related plasmids, the replication origin is located within the initiator protein coding sequence. As with pT181 and pC221, the copy number control region of pS194 is located just 5' to the *rep* protein coding sequence and it is similar to the copy control regions of these other two plasmids; deletions in this region cause an increase in copy number of 10 fold or more (Projan, unpublished), and we have identified a small antisense RNA species (countertranscript) that is encoded in this region and has a potential secondary structure similar to that predicted for the pC221 and pT181 countertranscripts (3). The start and end for this countertranscript are predicted by analogy with the other members of the family to be at positions 3280 and 3195 as indicated.

Most small *S. aureus* plasmids analyzed to date have in common a sequence called RS_B (18). The RS_B sequence was originally identified as a site of cointegrate formation between heterologous plasmids. It has subsequently been noted that this sequence is always found as part of an approximately 150 bp axis of dyad symmetry, *palA* (4). The *palA* elements of 8 different plasmids, including pS194, have pairwise base sequence homologies ranging from 35-88%; the pS194 *palA* is most closely related (88% homologous) to that of pC194, an otherwise totally unrelated plasmid, and is

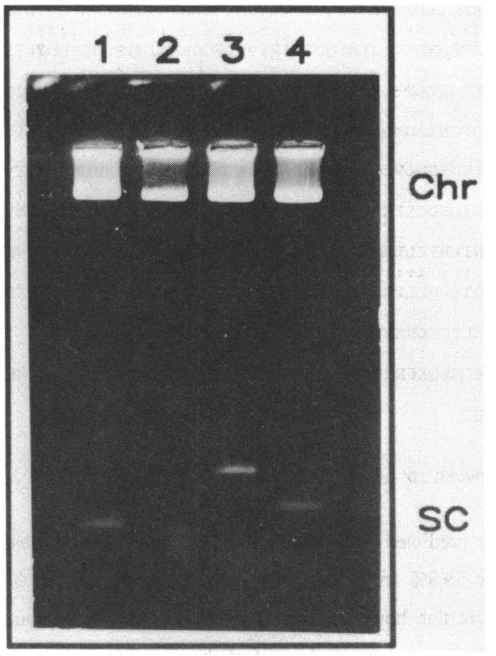


Figure 3. Agarose gel electropherogram of whole cell lysates. Lane 1: RN5343 (pE5); lane 2: RN5838 (pRN5202); lane 3: RN5170 (pRN5405); lane 4: RN5169 (pRN5404). Chromosome = chromosomal DNA, SC = supercoiled plasmid DNA.

least closely related (41% homology) to that of pIM13, which is also unrelated. Several of the *palA* elements have been found to be interchangeable and we have verified this for pS194by cloning a HinfI(2798-3360) fragment containing pS194 *palA* to pRN5202, a *palA* deletion derivative of pE5 (19). The pS194 *palA* element restored normal plasmid copy number and stability to the pE5 *palA* deletion derivative (see pRN5170, lane 3, Fig. 3), in the same manner as a restriction fragment carrying the pUB110 lagging strand origin (A. Gruss, personal communication; S. Projan, unpublished data) (see pRN5169, lane 4, Fig. 3).

Streptomycin Resistance

It has previously been observed that interuption of the HindIII site of pS194 inactivates streptomycin resistance (6). The HindIII site is located within a reading frame specifying a 34kd peptide which is therefore assumed to encode a streptomycin resistance protein. The most probable start for this protein is an ATG codon at nucleotide positions 4323 to 4325, which would place the Shine-Dalgarno sequence at 4308 to 4314, just within the C-terminus of the *repE* reading frame. The *str* transcript would then originate within the *repE* reading frame or even further upstream, suggesting that the expression of *repE* and *str* may be coupled. As shown in Fig. 5, the predicted Str protein

```
pS194    1' MRTEKEILNLVSEFAYQRSNVKIIALEGSRTNENIKKDKFQDYDFAFFVSDIEYFTHEES
            ::.:::   ::   : :   ..:. ::::: : :: :: :::::  .:::::: :    .
pJH1     1" MRSEKEMMDLVLSLAEQDERIRIVTLEGSRANINIPKDEFQDYDITYFVSDIEPFISNDD

        61' WLSLFGELLFIQKPEDMELFPPDLDYGYSYIMYFKDGIKMDITLINLKDLNRYFSDSDGL
            ::  :: ..  ::::::::::::. . :.::.: : :  : :.::. : .:  :    : :
        61" WLNQFGNIIMMQKPEDMELFPPE-EKGFSYLMLFDDYNKIDLTLLPLEELDNYL-KGDKL

       121' VKILVDKDNLVTQEIVPDDSNYWLKKPTEREFYDCCNEFWSVSTYVAKGVFRREILFALD
            .:.:.:::  .  .::: :  : ..::. ::. :::::::: :. :: :: :.:::::.:
       119" IKVLIDKDCRIKRDIVPTDIDYHVRKPSAREYDDCCNEFWNVTPYVIKGLCRKEILFAID

       181' HFNNILRPELLRMISWYIGFNRGFDFSLGKNYKFINKYLTDKEFNMLLATFEMNGYRKTY
            :::.:.: ::::::::: .:   ::  :.:::::::: .:...  .  :: :. : :  .
       179" HFNQIVRHELLRMISWKVGIETGFKLSVGKNYKFIERYISEDLWEKLLSTYRMDSYENIW

       241' QSFKLCCELFKYYSNKVSCLGNYNYPNYEKNIENFIRNNYEN
               :: .::.  :  :     : :: :..::  . :  :
       239" EALFLCHQLFRAVSGEVAERLHYAYPEYDRNITKYTRDMYKKYTGKTGCLDSTYAADIEE

       299" RREQ
```

Figure 4. Similarity between *str* of pS194 and the *aadE* reading frame of *S. faecalis* plasmid pJH1.

is homologous with the predicted AadE protein of the *Streptococcus faecalis* plasmid pJH1 (20). These two proteins have 48.9% amino acid identity and an additional 15% of conservative replacements. These proteins are not, however, detectably homologous to known Str proteins from Gram-negative bacteria. Just downstream from the end of the *str* reading frame is a sequence that resembles a ρ-independent terminator. This region, called t_1, has perfect homology with the corresponding region of both pT181 and pC221, immediately downstream from the reisistance determinants of these two plasmids.

Relaxation Complexes

Supercoiled pS194 DNA is relaxable by SDS, indicating that it exists largely as a protein-DNA relaxation complex (7).

pS194 and other relaxable plasmids such as pC221 are mobilized by conjugative *S. aureus* plasmids such as pG01 (21), whereas non-relaxable plasmids such as pT181 and pE194 are not (G. Archer, personal communication). The site of pC221 relaxation has been mapped and a plasmid gene required for relaxation has been identified.

On the basis of sequence homology, it is suggested that the pS194 relaxation site and gene are correspondingly located (see Fig. 1). A fourth open reading frame, *orfD*, overlaps the *rlx* gene (Figs. 1 and 2). It is 56% homologous to the corresponding open reading frame of pC221 at the amino acid level; the latter has recently been implicated in conjugative mobilization and it is considered likely that the same is true for pS194 *orfD*.

ACKNOWLEDGMENTS

We thank Ms. D. Everett for expert secretarial assistance and Dr. Ellen Murphy and Dr. Serban Iordanescu for useful and stimulating discussions. This work was supported by Public Health Ser-

vice grant GM-33278 to S.J.P. and grant DMB-8503619 from the National Science Foundation to R.P.N. P.H.R.I. computer facilities used for sequence analysis were supported by National Science Foundation grant DMB-8502189 and by Public Health Service grant RR-02990.

REFERENCES

1. Iordanescu, S. (1976) Arch. Roum. Pathol. Exp. Microbiol. 35, 111-118.
2. Iordanescu., S., Surdeanu, M., Della Latta, P. and Novick, R. (1978) Plasmid 1, 468-479.
3. Projan, S. J., Kornblum, J., Moghazeh, S. L., Edelman, I., Gennaro, M. L. and Novick, R. P. (1985) Mol. Gen. Genet. 199, 452-464.
4. Gruss, A., Ross, H.F. and Novick, R.P. (1987) Proc. Natl. Acad. Sci. USA 84, 2165-2169.
5. Novick, R.P., Projan, S.J., Kumar, C.C., Carleton, S., Gruss, A., Highlander, S.K. and Kornblum, J. (1985) in *Plasmids in Bacteria*, , vol 30, Plenum Press, New York, p. 299-320.
6. Gryczan, T., Shivakumar, A.G. and Dubnau, D. (1980) J. Bacteriol. 141, 246-253.
7. Novick, R. (1976) J. Bacteriol. 127, 1177-1187.
8. Novick, R. P. (1967) Virology 33, 155-166.
9. Novick, R.P. and Brodsky, R.J. (1972) J. Mol. Biol. 68, 285-302.
10. Chang, S. and Cohen, S.N. (1979) Mol. Gen. Genet. 168, 111-115.
11. Murphy, E. (1983) Plasmid 10, 260-269.
12. Novick, R.P., Murphy, E., Gryczan, T.J., Baron, E. and Edelman, I. (1979) Plasmid 2, 109-129.
13. Projan, S.J., Carleton, S. and Novick, R.P. (1983) Plasmid. 9, 182-190.
14. Maxam, A. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
15. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
16. Messing, J. (1983) Methods in Enzymol. 101, 20-78.
17. Brenner, D.G. and Shaw, W.V. (1985) EMBO J. 4, 561-568.
18. Novick, R. P, Projan, S. J., Rosenblum, W. and Edelman, I. (1984) Mol. Gen. Genet. 195, 374-377.
19. Projan, S.J., Monod, M., Narayanan, C.S. and Dubnau, D. (1987) J. Bacteriol. 169, 5131-5139.
20. Caillaud, F., Trieu-Cuot, P., Carlier, C. and Courvalin, P. (1987) Mol. Gen. Genet. 207, 509-513.
21. Archer, G.L. and Johnston, J.L. (1983) Antimicrob. Ag. Chem. 24, 70-77.