Molecular Characterization of the Complete 23F Capsular Polysaccharide Locus of *Streptococcus pneumoniae*

MARIO RAMIREZ AND ALEXANDER TOMasz*

The Rockefeller University, New York, New York 10021

Received 13 May 1998/Accepted 21 July 1998

The complete DNA sequence of the capsular locus 23F of *Streptococcus pneumoniae* is presented. The 18.6-kb cps23F locus is composed of 18 open reading frames flanked at the 5′ and 3′ ends by the genes *dexB* and *aIA*, an arrangement similar to those of some of the other identified cps loci.

Capsular polysaccharides deposited on the outermost surfaces of most clinical strains of *Streptococcus pneumoniae* are major virulence factors involved with evasion of the host immune system (1, 8). The capsule is the main target of protective antibodies which are specific for the particular polysaccharide (1, 32). Of the 90 different capsular types identified by immunological (17) and chemical (39) techniques, genetic characterization of the capsular biosynthetic genes (cps) is available only for a few capsular loci, namely, 1, 3, 14, 19F, and 19B (4, 9, 10, 14, 15, 19, 25, 27). Understanding the genetic determinants of the capsular polysaccharides should help in unraveling their pathway of biosynthesis which, in turn, could lead to the design of inhibitors capable of blocking the expression of these important pneumococcal virulence factors.

We describe here the complete DNA sequence of the cps locus of a type 23F *S. pneumoniae* Him18, a recent clinical isolate originating in Mexico. Analysis of a large number of serotype 23F isolates from diverse geographic areas and with a wide variety of isolation dates and chromosomal backgrounds confirmed that all genes identified by sequencing (as well as their structural organization) were conserved in each of the 23F isolates tested.

**Bacterial strains, plasmids, and growth conditions.** *S. pneumoniae* strains were obtained from the American Type Culture Collection (11, 12) or were from the Rockefeller University collection. A semisynthetic medium (C+Y) (20) or tryptic soy agar plates (Difco, Detroit, Mich.) supplemented with 3% sterile sheep blood and incubation at 37°C were used for growth. *Escherichia coli* HB101 (16) was obtained from Promega Corp. (Madison, Wis.) and used as the host strain for all recombinant plasmids generated in this study. Plasmid pSP64 (22) was used as the cloning vector. *E. coli* cells were grown in liquid medium (LB; Difco), SOC (16) with aeration, or solid medium (LA; Difco) incubated at 37°C. Selection of transformants was on LA supplemented with 50 μg of ampicillin/ml.

**PFGE.** Chromosomal DNA fragments, generated by *Smal* digestion, were separated and analyzed as described previously (36).

**Southern blot hybridization.** DNA fragments separated by PFGE or conventional gels were transferred to nylon membranes (Hybond-N+; Amersham, Little Chalfont, Buckinghamshire, United Kingdom) with the Vacuum Gene System (Pharmacia LKB Biotech, Uppsala, Sweden) in accordance with the manufacturer’s instructions. Blots were hybridized to specific DNA probes labelled with the ECL Direct Labelling System (Amersham) in accordance with the manufacturer’s recommendations. Hybridization conditions included a sodium chloride concentration of 0.5 M as recommended by the manufacturer.

**Probe preparation.** For the preparation of capsule 19F-specific probes, primers were designed from the published sequence of this capsular type (accession no. U09239) (24, 30) so that the generated probes resembled the ones used by Morona et al. (24). PCR was performed with strain OP5248 (serotype 19F) (13) as template and the GeneAmp system (Perkin-Elmer, Branchburg, NJ) basically as described previously (28). Amplified products were visualized by agarose gel electrophoresis, and fragments of the expected size were purified by using the Wizard PCR Preps Purification System (Promega) in accordance with the manufacturer’s recommendations. Preparation of serotype 23F-specific probes was based on the sequence determined in this study from the penicillin-resistant serotype 23F clinical isolate Him18 following the same procedure as described above.

**Long PCR.** Amplification of a region extending from the gene homologous to *cps19F* to the one homologous to *cps19F* of strain Him18 (serotype 23F) was accomplished by long PCR (5) with the GeneAmp XL system (Perkin-Elmer). The primers used were XL19Bd1 (AGGGGGTGCAGAACCATTGTC) and XL19Lr1 (AGCCAAGCAAAGCCAGGTCC). The reaction was performed in a DNA Thermal Cycler 480 (Perkin-Elmer) with a hot start (33), a magnesium acetate concentration of 1.1 mM, and primer and template concentrations of 1 ng/μl. The program used was the following: denaturation at 94°C for 2 min, 16 extension cycles of 94°C for 30 s and 68°C for 11 min, followed by 15 extension cycles of 94°C for 30 s and 68°C for 11 min, with an autoextension of 30 s per cycle.

**Cloning of the 23F-specific region.** Long PCR products were purified with the Wizard PCR Preps Purification System in accordance with the manufacturer’s recommendations. The products were digested with either *HindIII* or *PstI* (New England Biolabs, Beverly, Mass.), ligated to pSP64, and transformed into *E. coli* HB101 by standard procedures (34). Colonies resistant to ampicillin were screened for the presence of fragments of correct size by using the procedure described by Corton and Gustafsson (7) with the following modifications. The primers used were pSP64f (GATAAGGTCTCTCCTCA GTAAGC) and pSP64r (CGCCAGCGTTTCCACGTC). No glycerol or cresol red was added to the PCR mixture, and LB with ampicillin was used instead of YT plus kanamycin. Clones with inserts of approximately 0.5, 2.4, and 1.3 kb after *HindIII* digestion and 0.6, 5.7, and 5.1 kb after digestion with *PstI* were isolated.

**Nucleotide sequencing.** Plasmid purified with the Wizard Plus Midiprep DNA Purification System (Promega) was used
for automated sequencing. Sequencing of PCR products was performed after purification with the Wizard PCR Preps Purification System (Promega). All sequencing was done by primer walking with the TaqFS fluorescent dye terminator sequencing method run on a Perkin-Elmer Applied Biosystems Division model 377 automated sequencer.

Nucleotide and amino acid sequence analyses were performed with either the DNASTAR (Lasergene, Madison, Wis.) sequence analysis package, the GeneDoc program (29), or servers accessible via the Internet. Homology searches were performed with the Gapped BLAST or PSI-BLAST algorithms (2) or the COG system (37) available at the National Center for Biotechnology Information site, preliminary characterization and analysis of the open reading frames (ORFs) was performed with the GENEQUIZ (35) server, and sequence alignments were performed with the Clustal W (38) algorithm as implemented in DNASTAR or at the Clustal W server at the European Bioinformatics Institute. Promoter analysis was done using the NNPP (31) program available through the Baylor College of Medicine server.

**PCR.** Based on the sequences of the clones and the published sequence information for dexB, cps19fA, cps19fB, cps19fL, cps19fM, cps19fN, cps19fO, and aliA genes, primers were designed to obtain a series of overlapping PCR products covering the entire region between the dexB and aliA genes.

**Conventional gel electrophoresis.** Total chromosomal DNA was prepared with a modified version of the procedure described by Marmur (21). Restriction was performed with one of the following: EcoRV, HaeIII, HincII, or a mixture of SpeI and HindIII restriction enzymes. All enzymes were obtained from New England Biolabs. Agarose gel electrophoresis was performed in accordance with standard procedures (34).

**Nucleotide sequence accession number.** The nucleotide sequences and predicted amino acid translations described in this communication have been submitted to GenBank and are available under accession no. AF057294.

**Sequence of the complete type 23F capsular locus.** In a previous study (24), it was shown that a strain of serogroup 23 carried genes with high homology to dexB, cps19fA, cps19fB, cps19fL, cps19fM, cps19fN, cps19fO, and aliA. Using PCR, we established that the organization of the homologous genes in the 23F locus was similar to the one found on serotype 19F, in agreement to what has since been reported by others (6). We therefore designed primers to amplify the region between the cps19fB and cps19fL homologues, which presumably contained sequence information specific for type 23F. A fragment of approximately 14 kb was obtained and subcloned into pSP64. The sequence of the clones was determined by primer walking. In order to confirm the accuracy of the obtained sequence and to close the gaps in the locus due to the cloning strategy, a series of overlapping PCR products covering the entire 23F capsular locus was generated and sequenced.

**Computer analysis of the region located between the dexB and aliA homologues.** Computer analysis of the region located between the dexB and aliA homologues revealed the existence of 19 ORFs (Fig. 1).

The first ORF is transcribed in the orientation opposite that of dexB, and its product has a high degree of similarity to the C-terminal portion of IS1202 transposase (23) up to a deletion of one base in a stretch of T residues at position 630 (accession no. AF057294) that causes a frameshift. Moreover, a detailed examination of the DNA sequence in this region revealed a high degree of identity (95%) at the DNA level between the region encompassing nucleotides 495 to 1892 (accession no. 5274 NOTES J. BACTERIOL. 1998;180:5274–5280)
AF057294) and the previously reported coding sequence for IS1202 transposase (23). The insertion of a C residue at position 135 relative to the IS1202 transposase coding sequence (accession no. U04047) creates a frameshift that generates a premature stop codon 53 bp downstream, preventing the synthesis of a correct polypeptide.

Although regions with homology to the repeats previously recognized to be associated with IS1202 (23) could be found immediately upstream of cps23fA, these did not present the same organization described before, and the corresponding regions downstream of the dexB homologue could not be identified. Taken together, these findings suggest that the insertion sequence associated with the capsule 23F genes is no longer functional but is an inactivated form of IS1202.

The location, properties, significant similarities with known proteins, and proposed functions for each of the remaining ORFs associated with serotype 23F are summarized in Table 1.

### Analysis of other isolates expressing serotype 23F

In order to assess possible variations at the 23F locus, 11 strains from diverse genetic backgrounds (as defined by SmaI PFGE restriction profile) and different geographic origins (Brazil, United States, South Korea, and Bulgaria) were analyzed for their hybridization pattern to probes specific for the cps23f genes after enzymatic digestion. All ORFs identified in strain Him18 (whose sequence was determined in this study) and other pneumococcal capsular determinants, individual genes of strain Him18 were used to probe Southern blots of restricted chromosomal DNA from strains belonging to 21 different serotypes and two non-typeable strains. The results obtained are summarized in Fig. 2.

The properties and proposed functions of the 18 ORFs of the 23F capsular locus are summarized in Table 1. In most cases, the predicted functions of the homologous genes were themselves based on DNA or amino acid sequence homology. In some of the cases, proposed functions were based on genetic or biochemical evidence.

### Structure of 23F capsular determinant of the Spanish/USA clone

The structure of the 23F capsular determinant of the widespread multiresistant Spanish/USA clone (26) was analyzed by using representatives of this clone recovered from different geographic sites but sharing a similar SmaI PFGE restriction pattern. We found that these isolates constituted a very homogeneous group (data not shown). Differences detected between the capsular determinant of this group and strain Him18 (whose sequence was determined in this study) could be explained solely by differences in the region between dexB and cps23fA.

### Diversity of pneumococcal capsular loci

To examine the relationship between the genes found in the capsule 23F locus and other pneumococcal capsular determinants, individual genes of strain Him18 were used to probe Southern blots of restricted chromosomal DNA from strains belonging to 21 different serotypes and two non-typeable strains. The results obtained are summarized in Fig. 2.

The properties and proposed functions of the 18 ORFs of the 23F capsular locus are summarized in Table 1. In most cases, the predicted functions of the homologous genes were themselves based on DNA or amino acid sequence homology. In some of the cases, proposed functions were based on genetic or biochemical evidence.

Upstream of cps23fA, a putative promoter could be identified. Although promoters could also be detected in intergenic region I and II by using the NNPP software (31) (Fig. 1 and results not shown), the fact that these transcripts lack a convincing Rho independent terminator makes them less biologically plausible. Downstream of cps23fR, a putative Rho independent terminator could be identified (−23.4 Kcal). Taken together, these findings suggest that the entire locus could be transcribed in a single 18.6-kb transcript.

The variable G+C composition of the locus and the location of intergenic spaces suggest a modular organization in which different parts of the locus could have been acquired indepen-
dently (Fig. 1). It is noteworthy that the four genes responsible for the synthesis and activation of rhamnose have the same organization and have a very high sequence identity to genes identified in the serotypes 19F and 1 loci (24, 27).

The 115-nucleotide sequence found upstream of \( \text{cap1A} \) and conserved in all pneumococcal capsular loci (27) was also found associated with the 23F locus (89.6% identity).

Analysis of isolates expressing the 23F serotype. The fact that isolates from unrelated genetic backgrounds (as defined by \( \text{SmaI} \) PFGE restriction profile) had homologues to all the genes identified in strain Him18 is in agreement with the hypothesis that these genes are essential for capsular biosynthesis. The observed differences could be explained by two phenomena: the loss in some strains of a \( \text{HaeIII} \) site near the 5' end of the \( \text{cps23fO} \) coding region and the variability in the regions flanking the \( \text{cps} \) genes. The highly variable nature of these regions is apparent from the \( \text{cps} \) loci sequenced to date (3, 6, 19, 23).

**FIG. 2.** Cross-reactivity of 23F genes with \( S. \) pneumoniae isolates with different capsular types. A strong signal, a weak signal, and no signal are indicated by +, ±, and −, respectively. All positive signals are shaded.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Serogroup/ type</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
<th>Q</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR639</td>
<td>1</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D39S</td>
<td>2</td>
<td>+</td>
<td>+</td>
<td>*</td>
<td>*</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATCC6303</td>
<td>3</td>
<td>+</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34034</td>
<td>4</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR314</td>
<td>5</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR861</td>
<td>6B</td>
<td>+</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATCC6326</td>
<td>6B</td>
<td>+</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATCC10368</td>
<td>9V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K40</td>
<td>9V</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR35</td>
<td>10F</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP96</td>
<td>10A</td>
<td>+</td>
<td>*</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K41</td>
<td>11A</td>
<td>+</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD5229</td>
<td>14</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATCC6314</td>
<td>14</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP108</td>
<td>15B</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>±</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OP5248</td>
<td>19F</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8249</td>
<td>19A</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>±</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR459</td>
<td>22F</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>±</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATCC6323</td>
<td>23F</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>±</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATCC10346</td>
<td>23A</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATCC10364</td>
<td>23B</td>
<td>-</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATCC6328</td>
<td>28F</td>
<td>-</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR96</td>
<td>34</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR472</td>
<td>35B</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP148</td>
<td>36</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1138-80</td>
<td>NT</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1139-80</td>
<td>NT</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
cps23F/A were present in all strains with the exception of one of the nontypable isolates (Fig. 2). Sequences homologous to cps23F/B were also detected in all isolates hybridizing to cps23F/A. This observation is in apparent contradiction to results of a previous study (18), which found that strains expressing serotypes 9V, 11A, and 11B did not hybridize to the cps14B probe. However, the discrepancies may be explained by the nature of the probes used in the two studies. The probe used in the study described here spans nucleotides 99 to 389 of the coding region of cps23F/B while the probe used in the previous analysis (18) covered the region from nucleotide 260 of the cps14B coding region to nucleotide 26 of the cps14C coding region. The 5' regions of the cpsB genes sequenced to date show a high sequence identity that rapidly degenerates as one passes the middle of the gene. Therefore, if the structure of the cpsB genes of these serotypes follows the same principles, but with an even higher divergence, the construction of the probes would explain the differences in results.

Genes homologous to at least some of the cps32F/C to -E genes were also found in 11 other serotypes in agreement with their proposed functions (Table 1).

Of the 21 different serotypes analyzed, 7 contained rhamnose. All strains with rhamnose containing polysaccharides had genes homologous to cps23F/O to -Q similar to what was previously reported (24). It is interesting that isolates expressing serotypes 1 and 34, although having no rhamnose constituent, nevertheless carried a set of genes homologous to cps23F/O to -Q. Dispensability of these genes for capsule 1 expression and their presence in a set of 19 different isolates of type 1 pneumococci were interpreted as evidence of a clonal origin of type 1 strains through recombination with DNA from an unknown origin (27). A similar phenomenon could explain the presence of these genes in serotype 34, which also does not have rhamnose. All strains with rhamnose containing polysaccharides had genes homologous to cps23F/A and 23/23F involved in the synthesis of the capsular polysaccharide of Streptococcus pneumoniae type 3. Mol. Gen. Genet. 239:188–195.


