JOURNAL OF BACTERIOLOGY, Oct. 1994, p. 6375–6383
0021-9193/94/$04.00+0

Vol. 176, No. 20

Molecular Characterization of *cap3A*, a Gene from the Operon Required for the Synthesis of the Capsule of *Streptococcus pneumoniae* Type 3: Sequencing of Mutations Responsible for the Unencapsulated Phenotype and Localization of the Capsular Cluster on the Pneumococcal Chromosome

CARLOS ARRECUBIETA, RUBENS LÓPEZ,* AND ERNESTO GARCÍA

Departamento de Microbiología Molecular, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Madrid, Spain

Received 17 May 1994/Accepted 27 July 1994

The complete nucleotide sequence of the *cap3A* gene of *Streptococcus pneumoniae*, which is directly responsible for the transformation of some unencapsulated, serotype 3 mutants to the encapsulated phenotype, has been determined. This gene encodes a protein of 394 amino acids with a predicted M_r of 44,646. Twelve independent *cap3A* mutations have been mapped by genetic transformation, and three of them have been sequenced. Sequence comparisons revealed that *cap3A* was very similar (74.4%) to the *hasB* gene of *Streptococcus pyogenes*, which encodes a UDP-glucose dehydrogenase (UDP-GlcDH) that catalyzes the conversion of UDP-glucose to UDP-glucuronic acid, the donor substances in the pneumococcal type 3 capsular polysaccharide. Furthermore, a PCR-generated *cap3A*$^+$ gene restored encapsulation in our *cap3A* mutants as well as in a mutant previously characterized as deficient in UDP-GlcDH (R. Austrian, H. P. Bernheimer, E. E. B. Smith, and G. T. Mills, J. Exp. Med. 110:585–602, 1959). These results support the conclusion that *cap3A* codes for UDP-GlcDH. We have also identified a region upstream of *cap3A* that should contain common genes necessary for the production of capsule of any type. Pulsed-field gel electrophoresis and Southern blotting showed that the capsular genes specific for serotype 3 are located near the genes encoding PBP 2X and PBP 1A in the *S. pneumoniae* chromosome, whereas copies of the common genes (or part of them) appear to be present in different locations in the genome.

Streptococcus pneumoniae is a human pathogen causing high levels of morbidity and mortality throughout world. Austrian has pointed out that in the United States more than one million cases of pneumococcal pneumonia may occur each year, with a fatality rate of 5 to 7%•(1), and penicillin-resistant pneumococci are now found worldwide and exhibit MICs of penicillin that are up to 1,000-fold greater than those of typical penicillin-sensitive strains (25). Although several pneumococcal proteins appear to play an important role in the pathogenesis of the disease (31), the polysaccharide capsule is essential for virulence. Currently, 84 different serological serotypes have been identified (3). Avery et al. (4) demonstrated that the so-called "transforming principle" actually encoding the encapsulation gene(s) of this bacterium was the DNA, but no genes responsible for capsule production in *S. pneumoniae* had been cloned until recently (19). Early genetic analyses of the pneumococcus revealed that capsular genes are clustered (30), and this type of genetic organization has been demonstrated for *Escherichia coli* (33), *Haemophilus influenzae* (26) and *Neisseria meningitidis* (18), for which a common association of genes encoding components for polysaccharide translocation, sugar biosynthesis and polymerization, and postpolymerization modifications has been postulated. More recently, studies of capsule production by several gram-positive pathogens have confirmed the clustering of the capsular genes, although their whole organization remains to be elucidated (14, 34).

We have recently cloned in *E. coli* some genes required for capsule production by *S. pneumoniae* type 3 (19). Some of the difficulties found in cloning these capsular genes might be ascribed to the toxicity of the corresponding products for the host, particularly when high-copy-number plasmids are used. This type of limitation led us to a partial characterization of a gene (*cap3-1*, now renamed *cap3A*) involved in capsule biosynthesis as the result of a cloning artifact (see below). In the present study, the complete DNA sequence of the *cap3A* has been determined, and three mutations responsible for the unencapsulated (rough) phenotype of the pneumococcus have been characterized for the first time. In addition, the capsular genes of *S. pneumoniae* have been localized in the pneumococcal chromosome.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The pneumococcal strains used are described in Table 1. Serogrouping (or serotyping) was carried out by the Quellung reaction as described by Lund and Henrichsen (29), with the use of 46 antisera provided by the Staten Seruminstitut (Copenhagen, Denmark). We also employed *Streptococcus oralis* NCTC 11427 (type strain) in the blotting experiments. The *E. coli* strains TG1 and DH1 (35) were used as hosts for recombinant plasmids and grown in LB medium (35). Plasmid pLGL1 and its derivatives have been described in a previous publication (15). We also employed pXF405 (*recP*) (32) as a probe in some blotting experiments. *S. pneumoniae* was grown

* Corresponding author. Mailing address: Departamento de Microbiología Molecular, Centro de Investigaciones Biológicas (CSIC), Velázquez 144, 28006 Madrid, Spain. Phone: (1) 5611800. Fax: (1) 5627518. Electronic mail address: CIBEG47@CC.CSIC.ES.

TABLE 1. Pneumococcal strains used in this study

Strain(s)	Relevant genotype or phenotype	Source or reference
R6	Lyt$^+$ S2$^-$	Rockefeller University
406[a]	Lyt$^+$ S3$^+$	J. Casal
M23	Lyt$^-$ S3$^+$	19
M24	Lyt$^-$ S3$^-$	19
M32	Δ*lytA32* S2$^-$	28
A66R2[b]	Lyt$^+$ S3$^-$	7
NS3-1[c]	Lyt$^-$ S3$^+$	This work
NR3-4 to NR3-9[d]	Lyt$^-$ S3$^-$	This work
NR3-11[d]	Lyt$^-$ S3$^-$	This work
NR13 to NR3-16[d]	Lyt$^-$ S3$^-$	This work
49450[a]	Lyt$^+$ S3$^+$	J. Casal
12928[a]	Lyt$^+$ S6$^+$	A. Fenoll
13805[a]	Lyt$^+$ S9$^+$	A. Fenoll
13783[a]	Lyt$^+$ S14$^+$	A. Fenoll
8249[a]	Lyt$^+$ S19A$^+$	A. Tomasz
12950[a]	Lyt$^+$ S23$^+$	A. Fenoll

[a] Clinical isolate of the indicated serogroup or serotype.
[b] This strain has also been designated 662 (12) and $S\text{-}_{III_4}$ (2).
[c] Strain constructed by transformation of M32 with DNA from strain 406.
[d] Spontaneous S3$^-$ mutants of strain NS3-1.

in liquid C medium (39) containing 0.8% yeast extract with or without 0.08% bovine serum albumin without shaking or on reconstituted tryptose blood agar base plates (Difco Laboratories) supplemented with 5% (vol/vol) defibrinated sheep blood.

Isolation of unencapsulated (S3$^-$) mutants. Spontaneous rough mutants (other than M24) were isolated by repeated subculturing of the NS3-1 strain in C medium–0.08% yeast extract and on blood agar plates, as previously described in detail (19). Each mutant came from a different colony of NS3-1. The Quellung reaction was used to ascertain the absence of a detectable capsule in each mutant.

DNA manipulation, plasmid construction, and genetic transformation. The preparation of pneumococcal DNA and the transformation procedure for *S. pneumoniae* have been described elsewhere (39). When a nonisogenic pneumococcal strain, i.e., strain A66R2, was used as a recipient, agglutinins for noncapsulated cells (anti-R) were used to enhance the number of capsulated transformants (9). In some experiments, competent pneumococcal cells were transformed with restriction fragments separated by low-melting-point agarose gel electrophoresis as previously described (5). Plasmid DNA was prepared by the alkaline method, including equilibrium centrifugation in CsCl-ethidium bromide gradients, and transformation of *E. coli* was carried out by the RbCl method (35). Restriction endonucleases and T4 DNA ligase were purchased from Amersham Searle, New England Biolabs, or Boehringer Mannheim. All of these enzymes were used according to the recommendations of the suppliers. Restriction fragments and plasmids were analyzed by agarose gel electrophoresis and purified by using the Geneclean kit (Bio 101, La Jolla, Calif.).

DNA hybridization. DNA probes were labeled with the PolarPlex Chemiluminescent Blotting Kit (Millipore). Southern blots and hybridization were carried out according to the manufacturer's instructions.

PCR amplification. PCR amplifications were performed with 2 U of *Taq* polymerase (Perkin-Elmer Cetus), 1 μg of chromosomal DNA, 1 μM each synthetic oligonucleotide primer, 250 μM each deoxynucleoside triphosphate, and 5 mM $MgCl_2$ in the buffer recommended by the manufacturers. Amplification was achieved with a cycle of 3 min of denaturation at 95°C, 2 min of annealing at 69°C, and 5 min of polymerase extension at 72°C repeated 40 times, with a final 6-min extension at 72°C and slow cooling to 4°C with a Pharmacia-LKB Gene ATAQ controller. The following synthetic oligonucleotide primers, identified from the sequence analysis of *cap3A*, were used: PCAP3 (5′-CGAGGTCATTG GTCAGCAGAAAGTA-3′), PCAP6 (5′-CTCTGCTTCATT AAAAGCAACAACC-3′), and PCAP7 (5′-CCGCCCAACG AATAAGAATAAAGAC-3′).

DNA sequence determination and analysis. Both strands of DNA were sequenced by the dideoxy-chain termination method (36), using α-^{35}S-dATP (1,000 Ci/mmol) (Amersham International, Bucks, United Kingdom) and the Sequenase kit from U.S. Biochemicals. Oligonucleotide primers were purchased from Bio-Synthesis Dal, S. A. (Madrid, Spain). Direct PCR sequencing was carried out with the *fmol* DNA Sequencing System from Promega. DNA and protein sequence comparisons were done by using software from the Wisconsin Genetics Computer Group (Madison, Wis.). Sequence similarity searches were done by using the EMBL/GenBank, SWISS-PROT, and PIR databases.

Pulsed-field gel electrophoresis (PFGE) of chromosomal DNA. DNA embedded in agarose plugs were prepared from *S. pneumoniae* and *S. oralis* as previously described (20). DNAs were digested with either *Sma*I, *Sac*II, or *Apa*I before electrophoresis. Gels were cast and run in 75 mM Tris–25 mM boric acid–0.1 mM EDTA at 9°C in a contour-clamped homogenous electric field DRII apparatus (Bio-Rad) at 200 V with the switch time ramped from 1 to 40.2 s over a 24-h running time. After the gel was stained for 20 min in an ethidium bromide solution (1 μg/ml), it was photographed. The DNA fragments were then transferred to Immobilon-S membranes (Millipore) according to the manufacturer's instructions and hybridized as described above.

Nucleotide sequence accession number. The DNA sequence of the *S. pneumoniae cap3A* gene is available from the GenBank/EMBL/DDBJ databases under accession number Z12159.

RESULTS

Cloning and DNA sequence of the *cap3A* gene. We have previously reported that a 10-kb *Eco*RI fragment of DNA from strain 406 was capable of transforming an isogenic S3$^-$ mutant to the S3$^+$ phenotype (19). DNA fragments of about 10 kb obtained by treatment of DNA from the clinical strain 406 with *Eco*RI were ligated to *Eco*RI-digested pBR325 and used to transform *E. coli* DH1. Plasmids were isolated from chloramphenicol-sensitive, ampicillin-resistant transformants and used to transform M24 and other rough (S3$^-$) pneumococcal mutants to the smooth (S3$^+$) phenotype. The recombinant plasmid pKER1 (Fig. 1) was able to transform several, but not all, of the rough mutants to the smooth phenotype (Table 2). It has been impossible, so far, to clone the entire *Eco*RI fragment in *E. coli*, since pKER1 turned out to have a deletion of about 2.3 kb at the right end of the 2.5-kb *Sca*I fragment previously cloned (19). To accomplish the sequencing of the *Eco*RI fragment, we constructed a new recombinant plasmid, pKER2, by cloning a *Sca*I-*Eco*RI fragment into pUC18 (Fig. 1). To do that, the 10-kb *Eco*RI fragments from strain 406 were digested with *Sca*I, ligated to pUC18 previously treated with *Sma*I and *Eco*RI, and used to transform *E. coli* DH1. Plasmids were isolated from ampicillin-resistant transformants and used to transform NR3-6 and other S3$^-$ pneumococcal mutants to the S3$^+$ phenotype. The combined use of pKER1 and pKER2 allowed us to completely sequence the *cap3A* gene (Fig. 2). It should be pointed out that in a previous attempt to character-

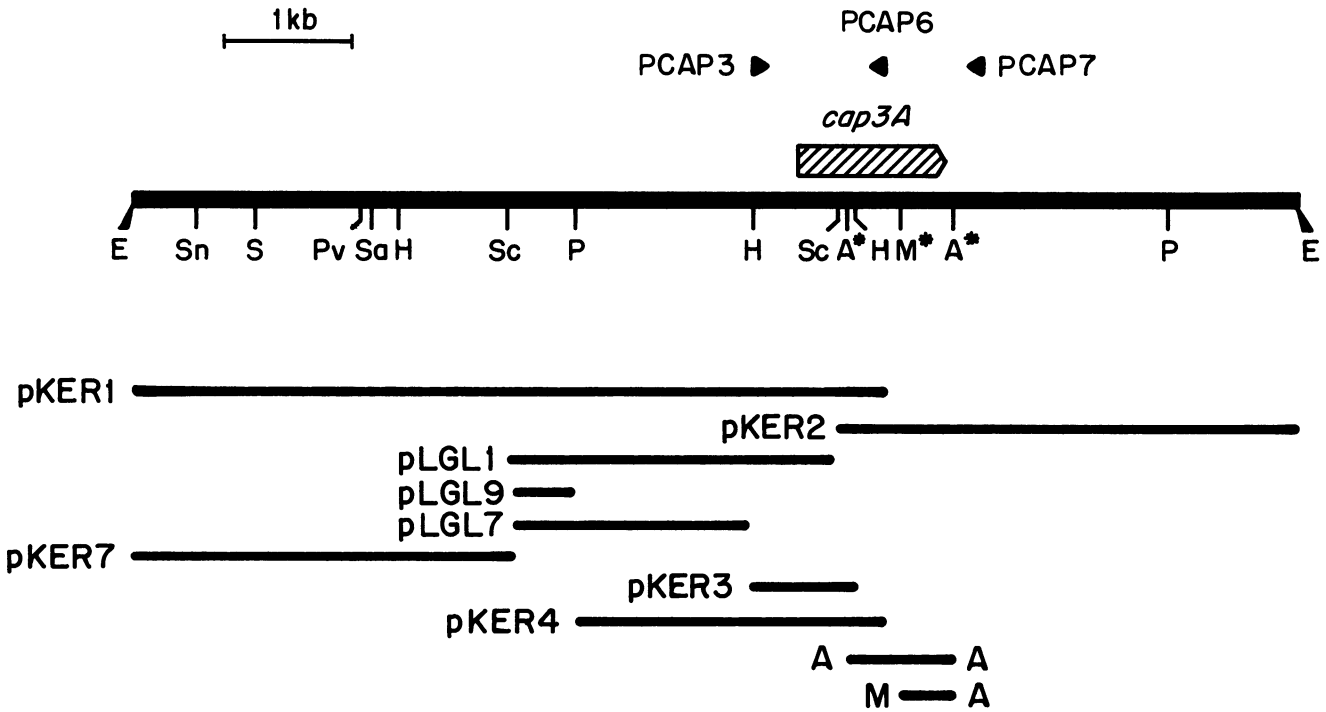


FIG. 1. Physical map of the 10-kb *Eco*RI fragment of strain 406 DNA, showing the location and direction of transcription of the *cap3A* gene (hatched arrow) as well as some of the plasmids that have been constructed. Arrowheads indicate the locations and directions (5′→3′) of the synthetic oligonucleotides used for PCR amplification. A, *Acc*I; E, *Eco*RI; H, *Hin*dIII, M, *Mun*I; P, *Pst*I; Pv, *Pvu*I; S, *Sal*I; Sa, *Sac*II; Sc, *Sca*I; Sn, *Sna*BI. An asterisk indicates that only some of the *Mun*I and *Acc*I sites present in this fragment are represented.

ize this gene, we sequenced from plasmid pLGL1 an open reading frame (ORF) of 414 nucleotides coding for a putative protein of 138 amino acids (19) that was suggested to be an UDP-glucose dehydrogenase (UDP-GlcDH) on the basis of comparative analyses of the primary structures of the product of this ORF and of the well characterized GDP-glucose dehydrogenase of *Pseudomonas aeruginosa*. Comparative analyses of the plasmids pKER1, pKER2, and pLGL1 have led us to conclude that the last plasmid suffered a cloning artifact leading to the insertion of a fragment of pneumococcal DNA (probably located far from the cluster of capsular genes) at one of the ends of the 2.5-kb *Sca*I fragment. An ORF of 1,182 bp, a potential Shine-Dalgarno sequence, and consensus promoter −35 and −10 sites were identified (Fig. 2). This ORF codes for a protein of 394 amino acids (predicted M_r, 44,646). No obvious transcription terminator-like sequences were found.

TABLE 2. Mapping of *cap3A* mutations by transformation with different donor DNAs

Mutant	Capsule restoration[a] after transformation with:					
	pLGL1	pKER1	pKER2	*Mun*I-*Acc*I-treated *cap3A*	*Acc*I-*Acc*I-treated *cap3A*	*cap3A*
M24	+	+	−	ND	−	+
A66R2	−	−	+	+	+	+
NR3-4	−	−	+	+	+	+
NR3-5	−	−	+	ND	+	+
NR3-6	−	−	+	+	+	+
NR3-7	−	+	+	ND	ND	+
NR3-8	−	+	+	ND	ND	+
NR3-9	−	+	+	ND	ND	+
NR3-11	−	−	+	ND	ND	+
NR3-13	−	+	+	ND	ND	+
NR3-14	−	+	+	ND	ND	+
NR3-15	−	−	+	ND	ND	+
NR3-16	−	+	+	ND	ND	+

[a] +, capsule restored; −, capsule not restored; ND, not determined. Strains were transformed with the indicated plasmid or with *cap3A*$^+$ that was amplified by using strain 406 DNA and oligonucleotides PCAP3 and PCAP7, digested (or not) with the restriction enzymes indicated, electrophoresed in low-melting-point agarose, and purified.

Sequence similarity comparisons. Sequence similarity searches with the Cap3A protein and the sequences compiled in the international databases revealed 57% identity and 74% similarity over 394 residues with the HasB protein of *Streptococcus pyogenes* (Fig. 3). The *hasB* gene encodes the UDP-GlcDH that converts the UDP-glucose to UDP-glucuronic acid (14). Figure 3 also shows that Cap3A was very similar to two other proteins of unknown function present in *E. coli* and *Salmonella enterica*.

***cap3A* codes for UDP-GlcDH.** Chromosomal DNA from the encapsulated strain 406 was amplified by using oligonucleotides PCAP3 and PCAP7, located 5′ and 3′ of *cap3A*, respectively (Fig. 1). Analysis of the PCR-generated products by agarose gel electrophoresis showed a band of the expected size (1.7 kb), and direct nucleotide sequencing of this fragment revealed an ORF identical to that reported in Fig. 2 (data not shown), confirming that no cloning artifacts were introduced during the isolation of pKER1 and pKER2. The PCR-generated *cap3A*$^+$ gene was also used to transform to the encapsulated phenotype the *cap3A* mutants described above (Table 2).

Most interestingly, transformation of a previously characterized mutant suggested that *cap3A* codes for UDP-GlcDH. The nonencapsulated mutant A66R2 was shown to be deficient in the production of UDP-GlcDH (2) because of mutations in the

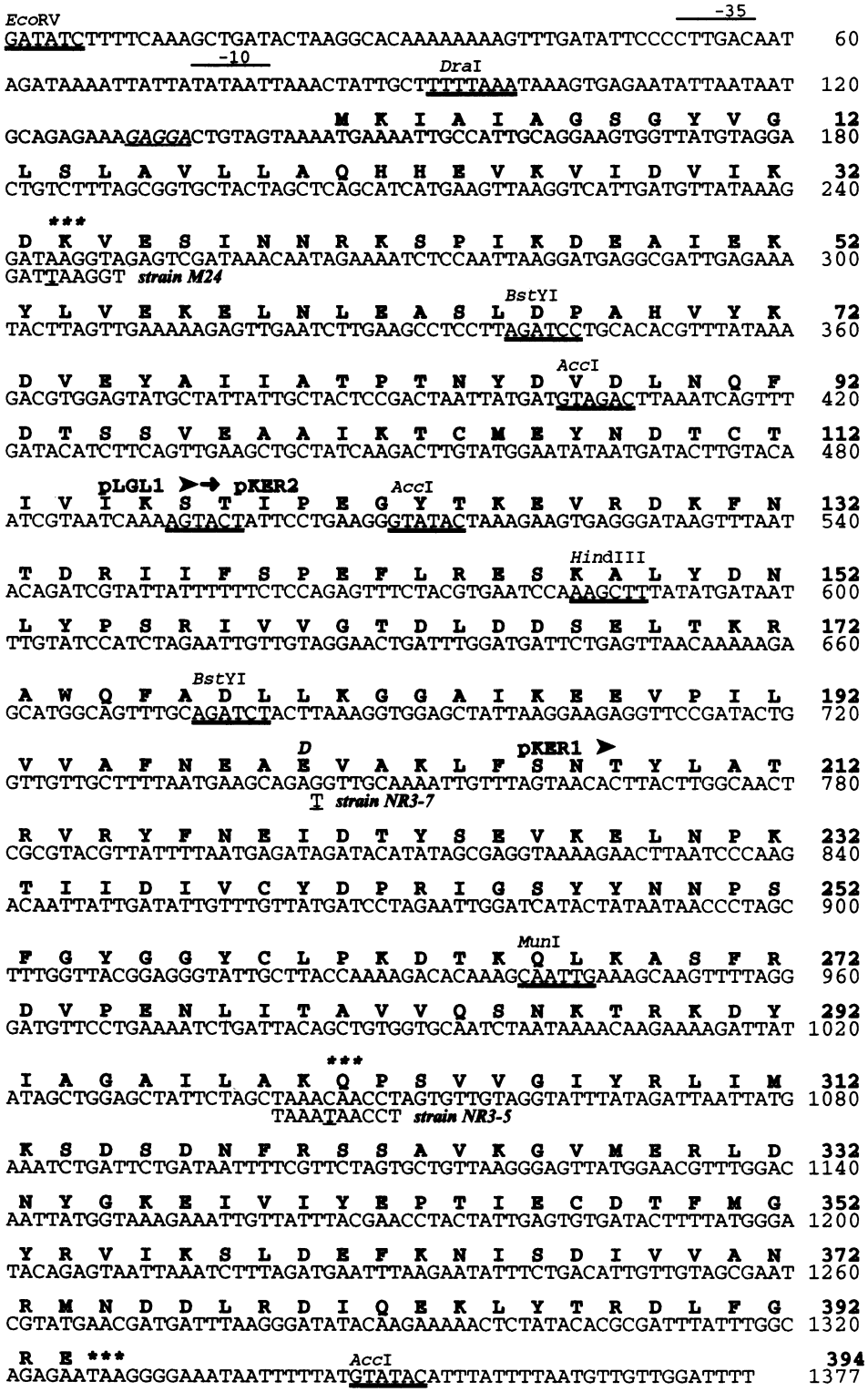
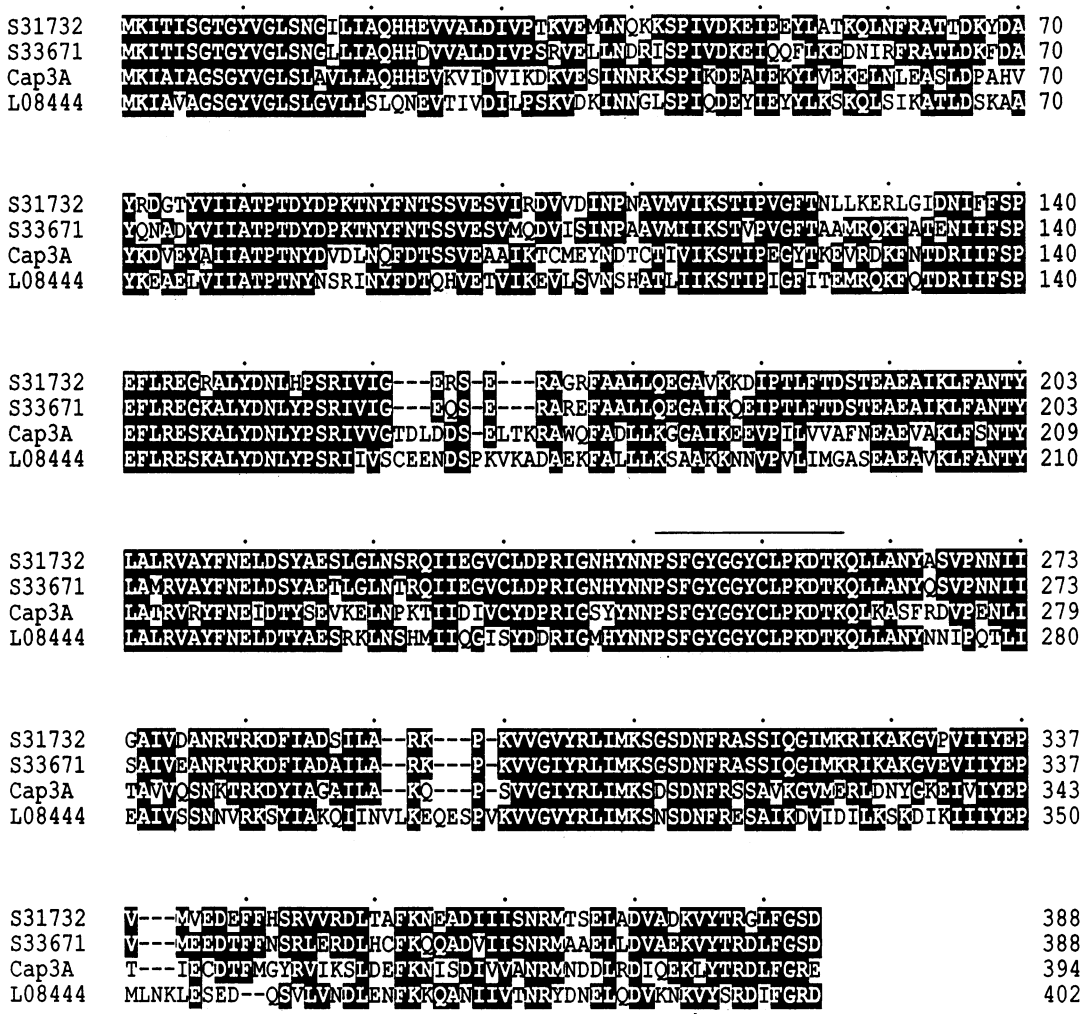


FIG. 2. Nucleotide sequence of the *cap3A* gene. The deduced amino acid sequence is shown above the nucleotide sequence in the one-letter code. The −10 and −35 regions of a possible promoter for the *cap3A* gene are shown, and the putative ribosome binding site is underlined and italicized. The mutations responsible for the unencapsulated phenotype of M24, NR3-5, and NR3-7 are in italics and underlined below the main sequence. Termination codons are indicated by asterisks. Important restriction sites are double underlined. The ends of plasmids pLGL1 and pKER1 and the beginning of pKER2 are indicated by ➤ and →, respectively.



	HasB Identity/similarity (%)	S31732 Identity/similarity (%)	S33671 Identity/similarity (%)
Cap3A	56.9/74.4	56.1/72.6	56.4/73.7
HasB		56.2/75.6	57.1/75.1
S31732			81.2/91.0

FIG. 3. Computer-generated alignment (PILEUP) of Cap3A protein with the UDP-GlcDH (HasB) of *S. pyogenes* (accession number L08444) and two unidentified proteins from *E. coli* (accession number S31732) and *S. enterica* (accession number S33671). Residues on black boxes indicate amino acids identical in at least two of the four proteins shown. The percentages of identical and similar amino acid residues resulting from pairwise comparisons (BESTFIT) are shown at the bottom. The region marked with a thin line above the sequences corresponds to the proposed active-site peptide of the bovine UDP-GlcDH (17).

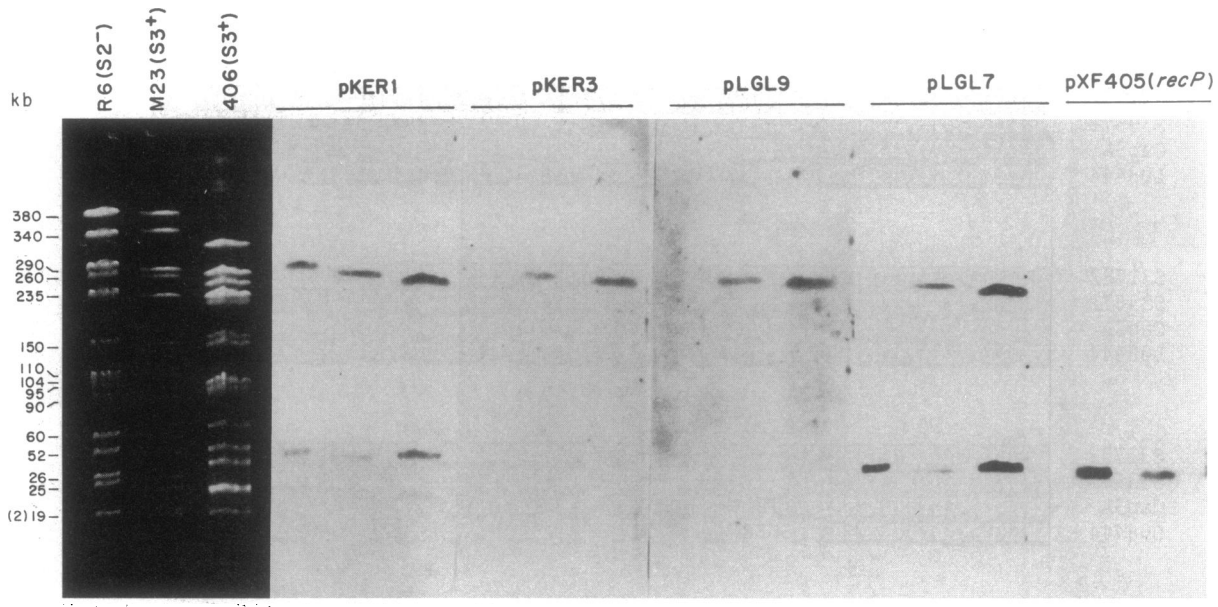


FIG. 4. PFGE of the DNAs prepared from strains R6 ($S2^-$), M23 ($S3^+$), and 406 ($S3^+$) digested with *Sma*I (left) and blotted and hybridized with biotin-labeled pKER1, pKER3, pLGL7, or pLGL9. Plasmid pXF405 (*recP*) was also used to identify unambiguously the 52-kb *Sma*I fragment.

locus designated *capD* (8). Transformation of A66R2 with the PCR-generated *cap3A*$^+$ gene of strain 406, or with several plasmids containing part of this gene, restored encapsulation (Table 2). In addition, comparative analyses of these transformation experiments revealed that the mutation in A66R2 is located in a 0.4-kb *Mun*I-*Acc*I fragment of *cap3A*. Serological analyses confirmed that the type 3 serotype has been restored in all the smooth transformants reported in Table 2 (not shown).

The fact that *cap3A* restored encapsulation in the UDP-GlcDH mutant as well as in the mutants isolated in this work, together with the high sequence similarity of Cap3A to the UDP-GlcDH (HasB protein) of *S. pyogenes* (Fig. 3), strongly suggests that *cap3A* encodes the pneumococcal UDP-GlcDH of serotype 3.

Mapping and characterization of *cap3A* mutations. The availability of two capsulated $S3^+$ strains (M23 and NS3-1) allowed the isolation of 12 isogenic, spontaneous rough mutants. The mutations were mapped by genetic transformation of these unencapsulated strains to the mucoid colony phenotype by using, as donor DNA, pKER1, pKER2, different restriction fragments of *cap3A*$^+$ and its surrounding regions, and the PCR-amplified *cap3A*$^+$ gene (Fig. 1 and Table 2). It should be mentioned that the mutation present in the M24 strain had been located in a 0.24-kb *Dra*I-*Bst*YI fragment (19). The remaining 11 mutants isolated in this work fall into two groups: one consisting of strains NR3-4, NR3-5, NR3-6, NR3-11, and NR3-15 and the other consisting of the rest of the mutants (Table 2). For three of these mutants, the mutations affecting capsular formation have been characterized (Fig. 2). The unencapsulated M24 strain resulted from the insertion of a single base (T) at nucleotide 100 from the ATG initiation codon of *cap3A*, causing the appearance of a termination codon (TAA) in the 5′-end part of the sequence. Strain NR3-7 showed a G-to-T transversion that would change Glu to Asp in the amino acid chain. This conservative amino acid change was located in a quite conserved region in the four proteins compared in Fig. 3. On the other hand, strain NR3-5 showed a C-to-T transition that resulted in a TAA termination codon at position 901 of the nucleotide sequence of the structural gene.

Localization of the capsular genes in the chromosome of *S. pneumoniae*. The recent development of the physical map of *S. pneumoniae* R6 (20) provides a useful framework for gene localization. Dot blot analyses of DNAs prepared from strains of *S. pneumoniae* expressing different capsular antigens provided preliminary information on those regions that may contain genes common to different pneumococcal serogroups and those that are specific for the synthesis of the type 3 capsule (19). PFGE and Southern blotting experiments (Fig. 4 and 5) using pKER1, pKER2, or several of their derivatives as probes showed that *cap3A* and its downstream region (up to the end of the 10-kb *Eco*RI fragment) correspond to type 3-specific capsular genes, whereas most, if not all, of the region located upstream from *cap3A* contains genes common to all the pneumococcal strains tested that belong to the most common serogroups (or serotypes) (16, 38, 40). In contrast, *S. oralis* DNA, which has 56% overall homology with *S. pneumoniae* DNA (24), did not hybridize with any of the probes used. These experiments also showed that type 3-specific capsular genes reside in a 290-kb *Sma*I fragment, but, in addition to this band, fragments of about 52 and 25 kb were also labeled when common genes were used as probes. It should be mentioned that great polymorphism was exhibited by the various pneumococcal strains used, in agreement with a previous report by Lefevre et al. (27). Using the restriction endonucleases *Sac*II and *Apa*I, we found that the type 3-specific genes are located in fragments 12 (60 kb) and 13 (59 kb), respectively, of the physical map of strain R6 (Fig. 6). In addition, the results of the PFGE experiments using *Sma*I, *Sac*II, and *Apa*I clearly showed that the 52-kb *Sma*I fragment that hybridized with probes containing common genes corresponds to *Sma*I fragment 12, where the *recP* gene is located (20).

DISCUSSION

The capsular polysaccharide of *S. pneumoniae* has been shown to be the most important virulence factor of this species (23). Nevertheless, experimental approaches using recombinant DNA technology to study the capsule production in this

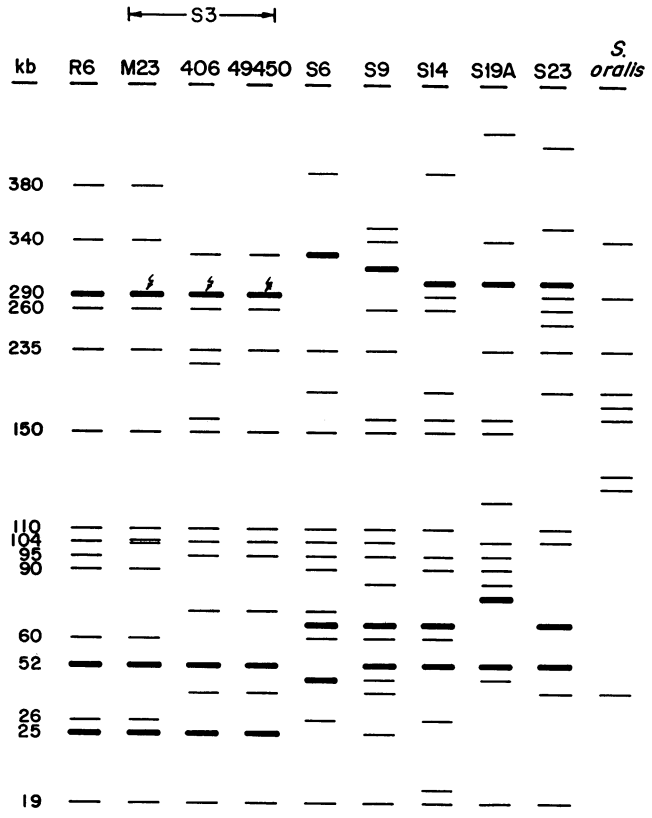


FIG. 5. Schematic representation of a hybridization experiment using *Sma*I fragments of DNAs, separated by PFGE, prepared from several pneumococcal isolates expressing different capsular antigens. *Sma*I-digested *S. oralis* DNA was also used. Thin lines correspond to fragments that did not hybridize with any of the probes tested (pKER1, pKER2, and pKER7); thick lines indicate those fragments that hybridized with pKER1 and pKER7. pKER2, which contains the type 3-specific genes, hybridized only with those thick bands indicated by arrows. The sizes of the *Sma*I fragments of R6 DNA (20) are indicated at the left.

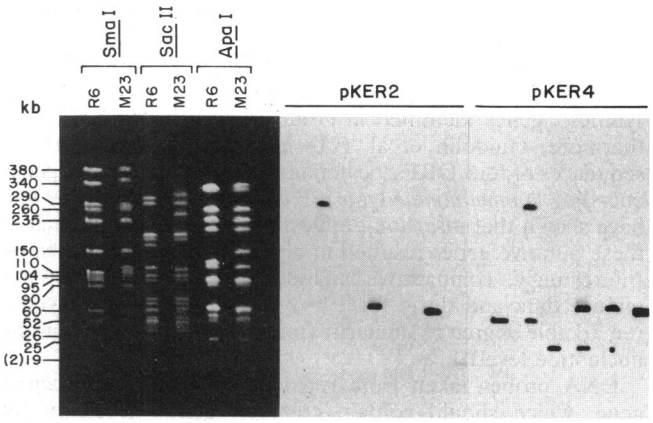


FIG. 6. PFGE of the DNAs obtained from strains R6 and M23 digested with *Sma*I, *Sac*II, or *Apa*I and blotted and hybridized with pKER2 or pKER4.

important pathogen, which would complement early biochemical studies (reviewed in reference 30), have not been reported until recently (19). Furthermore, this approach should also contribute to the gathering of knowledge on the organization of capsular genes, a fundamental step for studying the mechanisms of regulation of capsular polysaccharide biosynthesis. Unfortunately, we have been unable, so far, to clone an intact *cap3A* gene, which has hampered the biochemical characterization of Cap3A. Toxicity during expression in *E. coli* appears to have hindered the cloning of the *hasB* gene of *S. pyogenes* (14). Nevertheless, the determination of the complete sequence of the *cap3A* gene reported here confirms and extends previous suggestions on the remarkable similarity of the protein encoded by this gene to several dehydrogenases (19). In the case of HasB of *S. pyogenes*, the similarity to Cap3A extends on the whole ORF, and local sequence homology indicated that a motif involved in NAD^+ binding, characteristic of many dehydrogenases (41), was present at the N termini of both enzymes. The finding that *cap3A* was able to transform a previously characterized UDP-GlcDH-defective mutant (strain A66R2) (2) to the encapsulated phenotype reinforces the assumption that *cap3A* does encode this enzyme. We also observed that two previously unidentified ORFs of *E. coli* and *S. enterica* located near the *rfb* cluster, which is responsible for the synthesis of the O antigen (6), were also very similar to Cap3A, not only in amino acid sequence (Fig. 3) but also at the nucleotide level (not shown). These proteins probably represent the UDP-GlcDHs of those species. This assumption is based on two features: (i) the amino acid compositions of these proteins matched that which was reported some years ago for the UDP-GlcDH of *E. coli* (37), and (ii) the amino acid sequence ASVGFGGSCFZZGK (where Z corresponds to glutamic acid or glutamine), reported to be located at the active center of the bovine UDP-GlcDH (17), is one of the most conserved of all the aligned sequences shown in Fig. 3. To the best of our knowledge, the gene encoding UDP-GlcDH has not been mapped yet either in *E. coli* or in *S. enterica*.

Classical genetic studies of pneumococcal capsular type 3 strains had suggested that all the mutations altering the formation of type 3 polysaccharide were located in the gene coding for UDP-GlcDH (10). We isolated 12 independent rough mutants, and transformation experiments showed, again, that all the mutations map in this gene (Table 2). The molecular characterization of three of these mutations (Fig. 2), together with the finding that a PCR-generated *cap3A* gene restored encapsulation in these mutants and also in a previously characterized UDP-GlcDH mutant (2), provided direct experimental evidence of the involvement of *cap3A* in capsule formation (Table 2). Since the S3 polysaccharide contains glucose and glucuronic acid, one might have expected to find rough mutants that were defective in UDP-glucose pyrophosphorylase and in the polysaccharide polymerase as well as those defective in UDP-GlcDH. It is tempting to assume that mutations in capsular genes other than those affecting UDP-GlcDH activity might not be compatible with viability in the pneumococcus, although other explanations cannot be ruled out at this time.

Detailed molecular analyses of gram-negative bacteria have demonstrated that genes participating in polysaccharide synthesis, polysaccharide transport, and modification of capsular polysaccharides are clustered together in different regions of the polysaccharide regulon (18, 26, 33). Clustering of the capsular genes of several gram-positive bacteria has also been suggested recently (13, 14, 34), although it is unclear whether a common organization of capsulation loci similar to that demonstrated for gram-negative bacteria exists. Experimental support of the notion that the region upstream of *cap3A*

contains common genes necessary for production of capsule of any type comes from our preliminary sequence data (unpublished observations) showing the presence of four ORFs within this region that have remarkable similarities to polysaccharide synthesis genes identified in group B streptococci (34). Furthermore, Guidolin et al. (21) have recently reported the sequences of four ORFs located upstream of the specific genes encoding *S. pneumoniae* type 19F capsular polysaccharide and have shown that insertion-duplication mutagenesis of each of these putative genes resulted in nonencapsulated phenotypes. Interestingly, comparative analyses of our preliminary sequence data and those for 19F pneumococci (21) showed a remarkable degree of similarity (more than 80% identity at the nucleotide level).

DNA probes taken from regions upstream of the *cap3A* gene, which should contain common genes necessary for production of capsule of any type, showed homology with DNA fragments of several serotypes (Fig. 4, 5, and 6), located at different positions on the physical map of the pneumococcus (20), indicating the presence of several copies of these genes (or part of them) in the genome. A similar conclusion had previously been reached from genetic transformation experiments (9): when a rough mutant of type 3, deficient only in UDP-GlcDH, was exposed to DNA from a smooth strain of type 1, it was observed that the majority of encapsulated transformants were of the donor type and presumably resulted from the replacement of the whole type 3 gene cluster by the type 1 gene cluster. However, a minority of the transformants were of the binary type; that is, they reacted with both type 3 and type 1 sera. Transformation experiments carried out with such binary transformants suggested that the supernumerary (i.e., type 1) gene cluster had been inserted in the genome of the receptor strain somewhere other than the usual capsular-polysaccharide-determining region. On the other hand, duplications of capsular genes outside well-described capsular clusters have also been reported for *E. coli* (15) and *N. meningitidis* (22).

It is worth noting that combined PFGE and Southern blot analyses of pneumococcal DNA digested with the restriction enzymes *Apa*I, *Sac*II, and *Sma*I (Fig. 4, 5, and 6) revealed that the genes involved in capsular formation are located close to the genes encoding PBP 2X and PBP 1A, that is, in the same 59-kb *Apa*I fragment (fragment 14 on the map), demonstrating that these genes are less than 2.6% apart on the low-resolution physical map of the pneumococcus (20).

It has recently been concluded that the increase in the frequency of invasive infections caused by strains of pneumococci that are resistant to penicillin and other antimicrobial drugs calls for the development of effective polysaccharide-protein conjugated vaccines against *S. pneumoniae* (11). Molecular studies that result in the identification of potential vaccine antigens and in detailed knowledge of the structure and organization of the genes responsible for the synthesis of the main factors that contribute to pathogenesis of the pneumococcus may help in a better examination of the mechanisms by which this microorganism colonizes, invades, and frequently kills the host.

ACKNOWLEDGMENTS

We thank P. García and J. L. García for helpful suggestions and for critical reading of the manuscript and B. Murdock for his revision of the English. We also thank M. Carrasco, E. Cano, and D. Vicioso for technical assistance. The artwork of A. Hurtado is gratefully acknowledged.

C.A. was a recipient of a graduate fellowship from Euskal Herriko Jaurlaritza. This research was supported by grants from Comisión Interministerial de Ciencia y Tecnología (SAL91-0898-C02-01) and Programa Sectorial de Promoción del Conocimiento (PB 90-0069).

ADDENDUM IN PROOF

After this article was submitted, mutants deficient in type 3 capsule production have been identified and characterized by J. P. Dillard and J. Yother (Mol. Microbiol. **12**:959–972, 1994).

REFERENCES

1. **Austrian, R.** 1981. Pneumococcus. The first one hundred years. Rev. Infect. Dis. **3**:183–189.
2. **Austrian, R., H. P. Bernheimer, E. E. B. Smith, and G. T. Mills.** 1959. Simultaneous production of two capsular polysaccharides by pneumococcus. II. The genetic and biochemical bases of binary capsulation. J. Exp. Med. **110**:585–602.
3. **Austrian, R., C. Boettger, M. Dole, L. Fairly, and M. Fried.** 1985. *Streptococcus pneumoniae* type 16A, a hitherto undescribed pneumococcal type. J. Clin. Microbiol. **22**:127–128.
4. **Avery, O. T., C. M. MacLeod, and M. McCarthy.** 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. I. Induction of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III. J. Exp. Med. **79**:137–157.
5. **Barany, F., and A. Tomasz.** 1980. Genetic transformation of *Streptococcus pneumoniae* by heterologous plasmid deoxyribonucleic acid. J. Bacteriol. **144**:698–709.
6. **Bastin, D. A., G. Stevenson, P. K. Brown, A. Haase, and P. R. Reeves.** 1993. Repeat unit polysaccharides of bacteria: a model for polymerization resembling that of ribosomes and fatty acid synthetase, with a novel mechanism for determining chain length. Mol. Microbiol. **7**:725–734.
7. **Bernheimer, H. P.** 1979. Lysogenic pneumococci and their bacteriophages. J. Bacteriol. **138**:618–624.
8. **Bernheimer, H. P., and I. E. Wermundsen.** 1972. Homology in capsular transformation reactions in *Pneumococcus*. Mol. Gen. Genet. **116**:68–83.
9. **Bernheimer, H. P., I. E. Wermundsen, and R. Austrian.** 1967. Qualitative differences in the behavior of pneumococcal deoxyribonucleic acids transforming to the same capsular type. J. Bacteriol. **93**:320–333.
10. **Bernheimer, H. P., I. E. Wermundsen, and R. Austrian.** 1968. Mutation in pneumococcus type III affecting multiple cistrons concerned with the synthesis of capsular polysaccharide. J. Bacteriol. **96**:1099–1102.
11. **Broome, C. V., and R. F. Breiman.** 1991. Pneumococcal vaccine: past, present, and future. N. Engl. J. Med. **325**:1506–1508.
12. **Dawson, M. H.** 1928. The interconvertibility of "R" and "S" forms of pneumococcus. J. Exp. Med. **47**:577–591.
13. **DeAngelis, P. L., J. Papaconstantinou, and P. H. Weigel.** 1993. Isolation of a *Streptococcus pyogenes* gene locus that directs hyaluronan biosynthesis in acapsular mutants and in heterologous bacteria. J. Biol. Chem. **268**:14568–14571.
14. **Dougherty, B., and I. van de Rijn.** 1993. Molecular characterization of *hasB* from an operon required for hyaluronic acid synthesis in group A streptococci. Demonstration of UDP-glucose dehydrogenase activity. J. Biol. Chem. **268**:7118–7124.
15. **Drake, C. R., G. J. Boulnois, and I. S. Roberts.** 1993. The *Escherichia coli serA*-linked capsule locus and its flanking sequences are polymorphic, genetic evidence for the existence of more than two groups of capsule gene clusters. J. Gen. Microbiol. **139**:1707–1714.
16. **Fenoll, A., C. Martín Bourgon, R. Muñoz, D. Vicioso, and J. Casal.** 1991. Serotype distribution and antimicrobial resistance of *Streptococcus pneumoniae* isolates causing systemic infections in Spain, 1979–1989. Rev. Infect. Dis. **13**:56–60.
17. **Franzen, B., C. Carrubba, D. S. Feingold, J. Ashcom, and J. S. Franzen.** 1981. Amino acid sequence of the tryptic peptide containing the catalytic-site thiol group of bovine liver uridine diphosphate glucose dehydrogenase. Biochem. J. **199**:599–602.
18. **Frosch, M., C. Weisgerber, and T. F. Meyer.** 1989. Molecular characterization and expression in *Escherichia coli* of the gene complex encoding the polysaccharide capsule of *Neisseria menin-*

gitidis group B. Proc. Natl. Acad. Sci. USA **86**:1669–1673.
19. **García, E., P. García, and R. López.** 1993. Cloning and sequencing of a gene involved in the synthesis of the capsular polysaccharide of *Streptococcus pneumoniae* type 3. Mol. Gen. Genet. **239**:188–195.
20. **Gasc, A. M., L. Kauc, P. Barraillé, M. Sicard, and S. Goodgal.** 1991. Gene localization, size, and physical map of the chromosome of *Streptococcus pneumoniae*. J. Bacteriol. **173**:7361–7367.
21. **Guidolin, A., J. K. Morona, R. Morona, D. Hansman, and J. C. Paton.** 1994. Nucleotide sequence analysis of genes essential for the production of *Streptococcus pneumoniae* type 19F capsular polysaccharide, abstr. M47, p. 24. Abstr. 4th Int. ASM Conf. Strep. Genet. 1994. American Society for Microbiology, Washington, D.C.
22. **Hammerschmidt, S., C. Birkholz, U. Zähringer, B. D. Robertson, J. van Putten, O. Ebeling, and M. Frosch.** 1994. Contribution of genes from the capsule gene complex (*cps*) to lipooligosaccharide biosynthesis and serum resistance in *Neisseria meningitidis*. Mol. Microbiol. **11**:885–896.
23. **Johnston, R. B., Jr.** 1991. Pathogenesis of pneumococcal pneumonia. Rev. Infect. Dis. **13**(Suppl. 6):S509–S517.
24. **Kilpper-Bälz, R., P. Wenzig, and K. H. Schleifer.** 1985. Molecular relationships and classification of some viridans streptococci as *Streptococcus oralis* and amended description of *Streptococcus oralis* (Bridge and Sneath 1982). Int. J. Syst. Bacteriol. **35**:482–488.
25. **Klugman, K. P.** 1990. Pneumococcal resistance to antibiotics. Clin. Microbiol. Rev. **3**:171–196.
26. **Kroll, J. S., S. Zamze, B. Loynds, and E. R. R. Moxon.** 1989. Common organization of chromosomal loci for the production of different capsular polysaccharides in *Haemophilus influenzae*. J. Bacteriol. **171**:3343–3347.
27. **Lefevre, J. C., G. Faucon, A. M. Sicard, and A. M. Gasc.** 1993. DNA fingerprinting of *Streptococcus pneumoniae* strains by pulsed-field gel electrophoresis. J. Clin. Microbiol. **31**:2724–2728.
28. **López, R., J. M. Sánchez-Puelles, E. García, J. L. García, C. Ronda, and P. García.** 1986. Isolation, characterization and physiological properties of an autolytic defective mutant of *Streptococcus pneumoniae*. Mol. Gen. Genet. **204**:237–242.
29. **Lund, E., and J. Henrichsen.** 1978. Laboratory diagnosis, serology and epidemiology of *Streptococcus pneumoniae*. Methods Microbiol. **12**:241–262.
30. **Mäkelä, P. H., and B. A. D. Stocker.** 1969. Genetics of polysaccharide biosynthesis. Annu. Rev. Genet. **3**:291–322.
31. **Paton, J., P. Andrew, J. Boulnois, and T. Mitchell.** 1993. Molecular analysis of the pathogenicity of *Streptococcus pneumoniae*. Annu. Rev. Microbiol. **47**:89–115.
32. **Rhee, D. K., and D. A. Morrison.** 1988. Genetic transformation in *Streptococcus pneumoniae*: molecular cloning and characterization of *recP*, a gene required for genetic recombination. J. Bacteriol. **170**:630–637.
33. **Roberts, I. S., R. Mountford, R. Hodge, K. B. Jann, and G. J. Boulnois.** 1988. Common organization of gene clusters for production of different capsular polysaccharides (K antigens) in *Escherichia coli*. J. Bacteriol. **170**:1305–1310.
34. **Rubens, C. E., L. M. Heggen, R. F. Haft, and M. R. Wessels.** 1993. Identification of *cpsD*, a gene essential for type III capsule expression in group B streptococci. Mol. Microbiol. **8**:843–855.
35. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
36. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
37. **Schiller, J. G., F. Lamy, R. Frazier, and D. S. Feingold.** 1976. UDP-glucose dehydrogenase from *Escherichia coli*. Purification and subunit structure. Biochim. Biophys. Acta **453**:418–425.
38. **Shapiro, E. D., and R. Austrian.** 1994. Serotypes responsible for invasive *Streptococcus pneumoniae* infections among children in Connecticut. J. Infect. Dis. **169**:212–214.
39. **Tomasz, A.** 1970. Cellular metabolism in genetic transformation of pneumococci: requirement for protein synthesis during induction of competence. J. Bacteriol. **101**:861–871.
40. **Vinther Nielsen, S., and J. Henrichsen.** 1992. Capsular type of *Streptococcus pneumoniae* isolated from blood and CSF during 1982–1987. Clin. Infect. Dis. **15**:794–798.
41. **Wierenga, R. K., P. Terpstra, and W. G. J. Hol.** 1986. Prediction of the occurrence of the ADP-binding βαβ-fold in proteins, using an amino acid sequence fingerprint. J. Mol. Biol. **187**:101–107.