INFECTION AND IMMUNITY, Sept. 1996, p. 3576–3583 Vol. 64, No. 9
0019-9567/96/$04.00+0


# Variation in Repeat Number within the Alpha C Protein of Group B Streptococci Alters Antigenicity and Protective Epitopes

CLAUDIA GRAVEKAMP,[1]* DEBRA S. HORENSKY,[1] JAMES L. MICHEL,[1] AND LAWRENCE C. MADOFF[1,2]

*Channing Laboratory, Department of Medicine, Brigham and Women's Hospital,[1] and Division of Infectious Diseases, Beth Israel Hospital,[2] Harvard Medical School, Boston, Massachusetts*

Received 20 June 1996/Accepted 21 June 1996

**Variable expression of repeating units of the protective alpha C proteins among clinical isolates of group B streptococci (GBS) may have implications for vaccine development. In this study, alpha C protein genes containing various numbers of repeats (1, 2, 9, and 16) were cloned in a T7 overexpression vector in *Escherichia coli*. Expression was induced by isopropyl-β-D-thiogalactopyranoside, and proteins were purified by anion-exchange, gel filtration, or affinity chromatography or by isoelectric focusing. Rabbits were immunized with purified 1-, 2-, 9-, or 16-repeat proteins. All proteins appeared to be highly immunogenic. Enzyme-linked immunosorbent assay inhibition with 9-repeat protein as the coating antigen and 9-repeat-antigen-elicited antiserum showed that a 200-fold-higher concentration of 1-repeat antigen than of 9- or 16-repeat antigen was required for 50% inhibition of antibody-antigen binding. The concentration of 2-repeat antigen required for 50% inhibition was intermediate relative to the concentrations of 1- and 9-repeat antigens. These results suggested that antibodies to 9-repeat antigen recognized predominantly a conformational epitope(s) contained in proteins with higher numbers of repeats (9 or 16) but lost considerable binding affinities for an epitope(s) contained in alpha C proteins with fewer repeats (1 or 2). Similar results were obtained with antiserum to 16-repeat antigen. However, antibodies to 1- and 2-repeat antigens recognized 1-, 2-, 9-, and 16-repeat antigens with equal binding affinities. This finding suggested that 1- and 2-repeat-elicited antibodies recognized an epitope(s) on individual repeats. Loss of repeating units from the alpha C proteins may result in decreased protection because the loss of epitopes (including conformational epitopes) gives the microorganisms the opportunity to escape host antibodies. If 1- and 2-repeat-elicited antibodies bind all alpha C proteins with equal affinity, regardless of their repeat number, they may prevent GBS strains with fewer repeats from escaping host immunity. Protection data obtained with antisera to the proteins with different repeat numbers support this hypothesis: mouse pups challenged with GBS strain A909 were better protected when immunized with 1- or 2-repeat-elicited antiserum (76 and 75%, respectively) than when immunized with 9- or 16-repeat-elicited antiserum (41 and 48%, respectively).**

Group B streptococci (GBS) are the leading cause of meningitis, pneumonia, and sepsis in neonates (3). The type-specific capsular polysaccharides expressed on the surface of GBS are considered protective antigens (16). In addition, the cell surface-associated C proteins of GBS are known to be protective in experimental animals. Lancefield et al. (15) showed that antibodies to C proteins raised in rabbits protected mice challenged with GBS expressing C proteins. Later, the biochemical and immunological characterization of two protective C-protein antigens, alpha and beta, was reported (21). Alpha C proteins are present in approximately 50% of all GBS isolates, and beta C proteins are present in 10% (10a). With respect to the beta C protein, we found that active immunization of mice with the purified protein protected offspring against lethal infection with a beta-positive GBS strain (18). With respect to the alpha proteins, a monoclonal antibody (MAb), 4G8, that conferred passive protection against lethal GBS infections to mice challenged with a GBS strain expressing alpha C proteins was produced (19). Western blots (immunoblots) of the cloned alpha C protein gene (*bca*) product demonstrated a regularly laddered pattern of heterogeneous polypeptides similar to that in the native protein (22). The nucleotide sequence of the *bca* gene revealed nine identical tandem repeats, each 246 bp in length, and a single partial repeat of 33 bp; the repeat region was flanked by an N-terminal and a C-terminal region (22). This regularly laddered pattern of the alpha C proteins has also been found among clinical GBS isolates. However, alpha C proteins of variable sizes (62.5 to 167 kDa) were expressed by the latter isolates (17). It was found that the sizes of the alpha C proteins corresponded to the number of repeats (20). Recently, the occurrence of alpha repeat deletions during transmission of GBS from mother to neonate was described (18a). To study this phenomenon in more detail, an animal model in which GBS mutants with lower repeat numbers could be selected for by antibody pressure was developed (18a). It was demonstrated that many of these mutants contained deletions of tandem repeat elements, with the smallest mutants containing only one full and one partial repeat.

The importance of this variable expression of repeat numbers to virulence and immunity and thus to vaccine development is not known. It has been shown that, in the presence of the alpha-specific MAb 4G8, opsonophagocytic killing of clinical GBS isolates with variable numbers of repeats correlates directly with the increasing molecular mass of the alpha antigen (17). Tandemly repeated sequences have also been found in antigens of *Rickettsia rickettsii* (1), *Trypanosoma cruzi* (10),

* Corresponding author. Mailing address: Channing Laboratory, 180 Longwood Ave., Boston, MA 02115. Phone: (617) 432-3822. Fax: (617) 731-1541.

*Trypanosoma brucei* (26, 27), *Leishmania major* (33), *Plasmodium* spp. (12), *Streptococcus pneumoniae* (36), group A streptococci (9, 24, 32), *Bordetella pertussis* (25), *Staphylococcus aureus* (28), *Ureaplasma urealyticum* (38), and *Blastomyces dermatitidis* (14). Moreover, phenotypic diversity by variation in the repeat number of cell surface proteins has been demonstrated among clinical isolates of *Rickettsia* spp., *Trypanosoma* spp., *U. urealyticum*, group A streptococci, and *S. pneumoniae* (1, 7, 9, 10, 34, 37, 38). It has been suggested that the repeating units of the alpha antigen represent a mechanism for phenotypic and genotypic variability and provide natural sites for gene rearrangements that can generate antigenic diversity (22). It is possible that antigens containing different numbers of repeats elicit different repertoires of antibodies. Wessels et al. (35) described the loss of binding affinity of antibodies accompanying the loss of repeating subunits for capsular polysaccharides of GBS. They concluded that these repeating subunits contained a conformational epitope. The presence of conformational epitopes has also been suggested for the *rompA* protein gene repeat region of *Rickettsia* spp. (8), the *MB* antigen of *U. urealyticum* (38), and M proteins of group A streptococci (11). Loss of repeating units from the alpha C protein could result in (i) loss of binding affinity because of lower valency of the protein (i.e., fewer antibody-binding sites), (ii) loss of a possible conformational epitope(s), and/or (iii) poorer presentation of an epitope(s) at the cell surface when expressed by a smaller protein.

To study the relationship between the overall relative affinity of binding of antibody and the number of repeats, we created constructs of the alpha C protein gene with variable numbers of repeats (1, 2, 9, and 16), purified their gene products, immunized rabbits with the purified repeat-containing proteins, and performed enzyme-linked immunosorbent assay (ELISA) inhibition experiments, combining antisera to purified 1-, 2-, 9-, and 16-repeat proteins with the purified proteins as inhibiting antigens. We then conducted studies with mice to determine the protective efficacies of rabbit antisera elicited by the different repeat-containing proteins.

## MATERIALS AND METHODS

**Bacterial strains.** *Escherichia coli* strains HMS174(DE3) (6) and BL21(DE3) (29) containing the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible T7 RNA polymerase gene (30) were used as recipients for the newly constructed plasmids. Plasmid pT7-7 (kindly provided by S. Tabor [31]), which contains the T7 promoter and a strong ribosome-binding site optimized for translation beginning at an *Nde*I site, was used as a vector. The *bca* gene was obtained from plasmid pJMS23, which has been described previously (22). A909, the prototype of an $Ia^+$ $\alpha^+$ $\beta^+$ GBS strain which contains an alpha C protein gene with nine tandem repeats and a partial repeat, was used for challenge in mouse protection studies (22).

**DNA procedures and electroporation.** Plasmid DNA preparation, restriction endonuclease digestion, and agarose gel electrophoresis were performed by standard techniques (2). Large-scale plasmid DNA was prepared with the Magic Maxiprep Purification System (Promega, Madison, Wis.). *E. coli* cells were transformed by electroporation (Gene Pulser; Biorad, Melville, N.Y.) according to the manufacturer's instructions.

**Plasmids.** Plasmid pT7LM2 (see Fig. 1) was created for this project to optimize alpha C protein expression under the control of a T7 promoter. In order to place the ATG start site of the gene in optimal proximity to the T7 promoter and ribosome-binding site of the pT7-7 vector, it was necessary to create a new *Nde*I site at the start codon. The *bca* gene was therefore modified by PCR mutagenesis to contain such a site (CA ↓ T ATG).

An upstream fragment of the gene (287 bp) was removed from the plasmid by complete digestion with *Nde*I and partial digestion with *Alw*NI (to preserve a second *Alw*NI site further downstream). The desired fragment (5.8 kb) was purified from an agarose gel and treated with calf alkaline phosphatase according to the instructions of the manufacturer (Boehringer Mannheim Biochemicals). A new fragment was synthesized by PCR with an intact gene as the template. The forward primer (5′ GTA AAG GAG AAT CA ↓ T ATG TTT AGA AGG) was chosen to introduce the new *Nde*I site (underlined) and to allow sufficient 5′ overhang to be cut by *Nde*I. The reverse primer (5′ AGG TCG CTG CAC TCC) was chosen to span the region of interest and to include the *Alw*NI site within the product. PCR was conducted under standard conditions with use of *Taq* polymerase (Perkin-Elmer, Norwalk, Conn.) in a thermal cycler for 20 cycles. The product was digested with *Nde*I and *Alw*NI and purified with a Centricon 100 cartridge (Millipore, Bedford, Mass.). Finally, the 197-bp fragment was ligated into the vector-*bca* fragment above with T4 ligase. The sites of ligation and the fidelity of PCR were confirmed by restriction mapping and sequencing of the region of pT7LM2 containing the PCR product and of flanking regions with standard chain termination techniques (2).

**IPTG induction and protein expression.** *E. coli* HMS174(DE3) and BL21 (DE3) containing plasmids with the different repeat constructs were grown in tryptone medium (Difco, Detroit, Mich.) plus ampicillin (100 μg/ml) to an optical density at 650 nm of 0.6. Expression of the repeats was induced by IPTG (0.4 mM) for 3 h and then stopped on ice.

**Preparation of cytoplasmic fractions.** For the preparation of cytoplasmic fractions, the IPTG-induced *E. coli* cells were centrifuged and the pellet was resuspended in phosphate-buffered saline (PBS) (1% of the culture volume). The cells were then mechanically lysed with a French pressure cell (AMINCO, SLM Instruments Inc., Urbana, Ill.) at 750 lb/in$^2$ and subjected to ultracentrifugation at 140,000 × *g* for 1 h. The supernatant containing aberrant alpha C protein was used for subsequent purification.

**Purification of recombinant alpha C proteins. (i) Nine- and 16-repeat proteins.** Proteins were purified by anion-exchange column and fast protein liquid chromatography (FPLC) with a 0 to 0.5 M NaCl gradient in Tris buffer (50 mM, pH 7.8). A 5-ml volume of the cytoplasmic fraction was diluted (1:1) with Tris buffer (50 mM, pH 7.8) and loaded onto a Resource Q column (Pharmacia, Piscataway, N.J.); 5-ml fractions were collected (2 ml/min) and examined for alpha protein by dot blot analysis with alpha-specific MAb 4G8 (19). Alpha C protein was eluted at 0.12 M NaCl. Positive fractions were analyzed qualitatively as well as quantitatively by comparison of a sodium dodecyl sulfate (SDS) gel stained with Coomassie brilliant blue and an SDS gel that was subsequently subjected to Western blotting. Fractions with the highest quality and largest amount of alpha C protein were pooled and treated first with DNase and RNase (325 μg/25 ml of pooled fractions) in the presence of $CaCl_2$ and $MgCl_2$ (1 mM, final concentration) and then with trypsin (substrate-to-enzyme ratio, 4:1) (Sigma, St. Louis, Mo.) at 37°C for 4 h. The trypsin resistance of alpha C proteins (4) was useful in this purification process. After dialysis in distilled water overnight at 4°C and lyophilization, the remaining proteins were dissolved in 500 μl of Tris buffer (50 mM, pH 7.8) and subjected to gel filtration with Superose-12 column FPLC (Pharmacia). Positive fractions were identified and analyzed as described above.

**(ii) Two-repeat protein.** For the 2-repeat protein, the same procedure was followed as for the 9- and 16-repeat proteins, but an additional step of preparative isoelectric focusing (Rotofor Cell; Biorad) was required. A solution of ampholytes (Bio-lyte; Biorad) with a pH range of 3 to 5 was used (the isoelectric point of alpha C protein is 4.5). Positive fractions were identified and analyzed as described above.

**(iii) One-repeat protein.** For the purification of the 1-repeat protein, we used affinity chromatography. Protein A-agarose beads of an Immunopure immunoglobulin G orientation column (Pierce, Rockford, Ill.) were cross-linked according to the manufacturer's instructions with rabbit antiserum to 2-repeat proteins. The cytoplasmic fraction of *E. coli* expressing 1-repeat protein was applied to the column, and the alpha C protein was eluted by 0.1 M glycine (pH 2.8). The eluted protein was then subjected to anion-exchange chromatography (Resource Q column). Positive fractions were identified and analyzed as described above.

**Production of rabbit antisera.** Two New Zealand White female rabbits (6 to 8 weeks old) were subcutaneously immunized with each antigen (1, 2, 9, and 16 repeats). The first dose (100 μg/0.5 ml of PBS) was emulsified in 0.5 ml of complete Freund's adjuvant, and the subsequent doses (days 21 and 42; 100 μg) were emulsified in 0.5 ml of incomplete Freund's adjuvant. Serum was collected before each immunization and again 2 weeks after the last dose.

**ELISA.** Antibody in rabbit antisera to 1-, 2-, 9-, and 16-repeat alpha C proteins was measured by ELISA as described previously (17). The various purified repeat proteins were used as the coating antigen. Optimal coating concentrations (in micrograms per milliliter), determined by checkerboard titration, were as follows: 1-repeat antigen, 1; 2-repeat antigen, 0.25; 9-repeat antigen, 0.125; and 16-repeat antigen, 0.081. Twofold dilutions of the rabbit antisera were made (starting dilution, 1:100). Alkaline phosphatase-conjugated antibody to rabbit immunoglobulin G (whole molecule; 1:2,000) was used as the second antibody (Cappel/Organon Teknika, Durham, N.C.). The titer was defined as the highest dilution with an $A_{405}$ of >0.2 after development with substrate (*p*-nitrophenyl phosphate; Sigma) for 30 min.

**ELISA inhibition.** ELISA inhibition was performed as described previously (17). Serial dilutions of inhibiting antigen were made and preincubated with the first antibody for 1 h at 37°C. The antigen-antibody complex was transferred to the antigen-coated microtiter plate and incubated for 1 h at 37°C, a step that allowed free antibody to bind the coating antigen. After the plates were washed, the second antibody was added to the microtiter wells and the plates were incubated for 1 h at 37°C. The plates were again washed and were then developed with the substrate. The $A_{405}$ was measured with an automated ELISA reader (Dynatech, Chantilly, Va.). ELISA inhibition was measured by the following formula: % inhibition = $\{[A_{405(\text{uninhibited control})} - A_{405(\text{sample})}]/A_{405(\text{uninhibited control})}\} \times 100$.

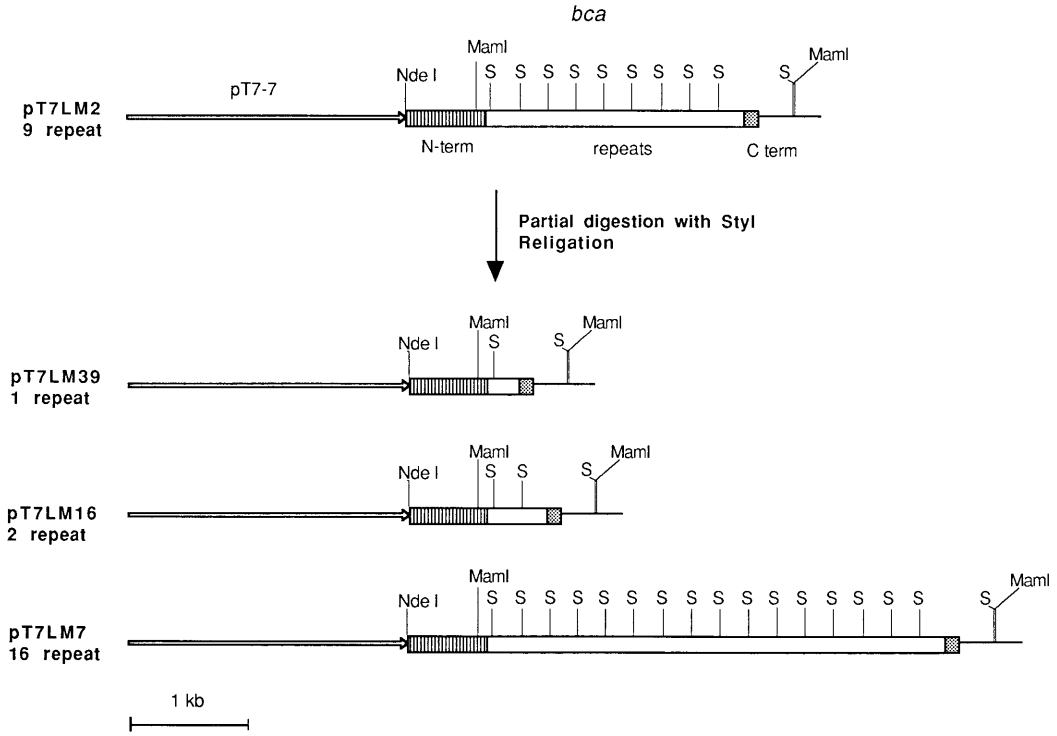


FIG. 1. Schematic view of multimeric repeat constructs of the alpha C protein gene. pT7LM2 consists of the T7 expression plasmid pT7 ligated directly to the start codon of the alpha C protein gene *bca*. Each repeat contains a single *Sty*I restriction site near its center. After partial digestion of pT7LM2 containing the *bca* gene (9 repeats) with *Sty*I, a mixture of pT7LM2 plasmids containing different numbers of repeats was obtained. Multimeric constructs were obtained by ligation of these *Sty*I-digested pT7LM2 plasmids. Each construct contained a repeat region flanked by identified N- and C-terminal domains (N-term and C-term, respectively).

To study the effect of the repeat number of alpha C proteins on the antibody repertoire elicited in rabbits, ELISA inhibition was performed with the purified 1-, 2-, 9-, and 16-repeat proteins (as inhibiting and coating antigens) and the obtained antisera to these antigens. The optimal concentrations of the 1-, 2-, 9-, and 16-repeat coating antigens were 1, 0.25, 0.125, and 0.081 μg/ml, respectively. Serial twofold dilutions of inhibiting antigen were made (starting at 5 μg/ml); rabbit antisera to the different repeats were used as the first antibody (1:8,000), and alkaline phosphatase-conjugated antibody to rabbit immunoglobulin G (whole molecule) was used as the second antibody (1:2,000) (Cappel).

**Mouse protection studies.** The mouse protection method was adapted from that described by Rodewald et al. (25a). In our study, groups of four CD-1 outbred pregnant mice were given 0.5 ml of undiluted rabbit antiserum by intraperitoneal injection (3 or 4 days before delivery). Pups were challenged intraperitoneally with $3 \times 10^4$ to $5 \times 10^4$ CFU of GBS strain A909 ($10^3$ times the 50% lethal dose determined in a pilot study) within 48 h of birth. Survival was assessed 48 h after challenge, and survival data for groups whose dams had been immunized with different antisera were compared by Fisher's exact test.

## RESULTS

**Creation of multimeric tandem repeats.** Alpha C protein gene constructs were generated by modification of the parent plasmid pT7LM2 containing the alpha C protein gene *bca* immediately downstream of the phage T7 promoter. Each 246-bp repeat of the *bca* gene contains a single *Sty*I restriction site (C↓CTTGG) near its center (22). One additional *Sty*I site exists outside the repeat area near the C terminus of the gene. Since the *Sty*I site within each repeat is not a palindrome, *Sty*I-digested DNA fragments can religate in only one orientation. The C-terminus *Sty*I consists of a sequence different from those internal to the repeat (C↓CT AGG); this situation forced cloning at this site. Plasmid DNA (500 ng) was partially digested with *Sty*I (Boehringer Mannheim Biochemicals; 1 U) in the appropriate buffer (total volume, 15 μl). Digestion was allowed to continue for 15 min at 37°C and was then stopped by immersion of the digest mixture first in ice and then in a 65°C bath for 5 min. The mixture was purified on a Sepharose CL6B mini-column in an Eppendorf tube (23). The *Sty*I partial-digest mixture was religated with 20 μl (8,000 U) of T4 DNA ligase (New England Biolabs) in ligase buffer overnight at 16°C. A 1-μl volume of the ligation mixture was electroporated into 35 μl of *E. coli* DH5α rendered competent by standard techniques (2). *E. coli* DH5α (*recA*) was used for plasmid maintenance in order to minimize spontaneous recombination. A schematic view of the multimeric repeat variants is shown in Fig. 1.

Transformants were selected on ampicillin-containing plates and screened for the size of *Mam*I fragments, which span the repeat region of the *bca* gene. Transformants containing variable numbers of repeats were transformed into expression strain BL21(DE3) or HMS174(DE3). A representative sample of alpha variants containing various numbers of tandem repeats was obtained (Fig. 2). These multimeric variants displayed insertions and deletions of tandem repeats. For example, the size of the *Mam*I fragment from pT7LM2 was 1.55 kb (Fig. 2, lane 4). Thus, the insert contained six identical repeats of 246 bp as well as flanking sequences. Characteristics of these variants are summarized in Table 1.

**Expression of multimeric tandem-repeat variants.** Immunoblots screened with MAb 4G8 (data not shown) indicated that the sizes of multimeric variants ranged from 36 to 150 kDa. All of the alpha C protein variant clones expressed protein immunoreactive with the MAb. In addition, each of the variants displayed the characteristic ladder appearance on SDS-polyacrylamide gel electrophoresis (SDS-PAGE). The size of the *Bsa*BI fragment containing the repeat region of the gene in each variant correlated with the size of the largest expressed protein product ($r^2 = 0.97$). The number of bands on the

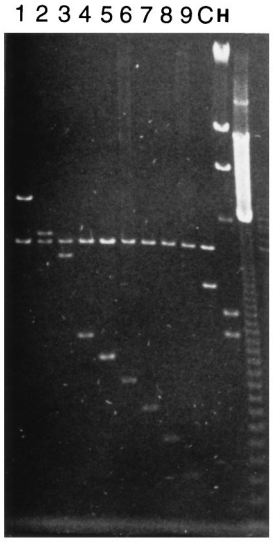


FIG. 2. Restriction analysis of DNA from constructs containing various numbers of tandem repeats. DNA cut with *Mam*I (which excises the repeat region of the gene) exhibits a vector band of 3.2 kb and inserts ranging from 5.5 kb (lane 1; 16 repeats) to 0.6 kb (lane 9; 1 repeat). A digest of pT7LM2 (lane C) and lambda DNA digested with *Hin*dIII markers (lane H) are also shown. Markers (123 bp) are shown on the far right.

immunoblot increased with the number of tandem repeats. The level of expression of the protein (the intensity of staining of the immunoblot) increased after induction with IPTG (data not shown). In order to study the effect of the number of repeats on the immune response of rabbits, clones containing 1, 2, 9, and 16 repeats were selected for purification.

**Purification of alpha C protein.** Anion-exchange chromatography (Resource Q) eliminated many but not all high-molecular-weight contaminants of cytoplasmic fractions containing 2-, 9-, or 16-repeat proteins (data not shown). Therefore, alpha-positive fractions were pooled and treated with trypsin. However, contamination by some low-molecular-weight proteins was still detected (data not shown). This contamination was eliminated by gel filtration chromatography (Superose-12), which resulted in highly purified 9- and 16-repeat proteins (Fig. 3). It is noteworthy that these purified high-molecular-weight antigens contained predominantly the high-molecular-weight species but also small amounts of the low-molecular-weight components. The laddered patterns of the purified 9- and 16-repeat proteins in Western blots were identical to that of the native antigen (17). Each of the individual bands visualized with Coomassie brilliant blue reacted with alpha-specific MAb 4G8 in Western blots (Fig. 3). An additional step of isoelectric focusing was successfully applied to separate the

TABLE 1. Characteristics of GBS alpha C protein gene expression constructs containing various numbers of identical tandem repeats

| Plasmid | *Mam*I fragment size (kb) | No. of repeats[a] | Molecular mass of protein (kDa)[b] |
|---|---|---|---|
| pT7LM2 | 2.8 | 9 | 105 |
| pT7LM7 | 5.5 | 16 | 160 |
| pT7LM20 | 1.55 | 6 | 79 |
| pT7LM16 | 0.79 | 2 | 47 |
| pT7LM39 | 0.59 | 1 | 37 |

[a] Calculated on the basis of the size of the *Mam*I fragment.
[b] Molecular mass of full-size protein, i.e., the most prominent band on the immunoblot.

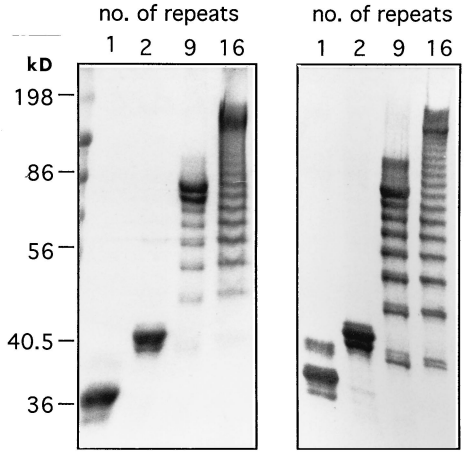


FIG. 3. Results of SDS-PAGE with Coomassie blue staining (left) or with subsequent Western blotting (right) of purified GBS alpha C proteins expressing 1, 2, 9, and 16 repeats (as indicated above the lanes). MAb 4G8 was used to visualize the laddered pattern of the alpha C proteins in the Western blot.

low-molecular-weight contaminating proteins from the 2-repeat proteins (Fig. 3). Because of the lower yield of purified 2-repeat antigen than of 9- and 16-repeat antigens (0.2 versus 1.5 mg/liter of *E. coli* culture), we decided to adopt another strategy for purification of the 1-repeat protein. Application of affinity chromatography (protein A cross-linked with 2-repeat-protein-elicited antibodies) followed by ion-exchange chromatography (Resource Q) resulted in highly purified 1-repeat protein, as shown in Fig. 3. The yield of purified 1-repeat antigen was 0.7 mg/liter of *E. coli* culture.

**Immunogenicity of 1-, 2-, 9-, and 16-repeat proteins in rabbits as determined by ELISA.** All four proteins were highly immunogenic in rabbits. ELISA titers ranged from $1 \times 10^5$ to $2 \times 10^5$, which is an 800- to 2,000-fold increase over preimmunization titers.

**Antibody responses in immunized rabbits.** ELISA inhibition was used to study the relationship between the overall relative binding affinities of antibodies and the number of repeats expressed by alpha C proteins (Fig. 4A). By overall relative binding affinity, we mean the complex interaction of a population of antibodies with a particular antigen as quantified by ELISA inhibition. Binding may be affected both by the relative concentrations of antibodies to distinct epitopes and by the avidity of individual antibodies within this population of antibodies. When 9-repeat alpha C protein was used as the coating antigen in combination with 9-repeat antiserum, a 200-fold-higher concentration of 1-repeat antigen than of 9- or 16-repeat antigen was required for 50% inhibition of antigen-antibody binding (Fig. 4A). The concentration of 2-repeat antigen required for 50% inhibition was intermediate relative to the necessary concentrations of 1- and 9-repeat antigens. (The combination of 16-repeat coating antigen and 16-repeat antiserum gave results similar to those obtained with the 9-repeat combination and therefore are not shown.) However, when the 2-repeat protein was used as the coating antigen in combination with 2-repeat antiserum, the difference in the concentrations of the 1-, 2-, 9-, and 16-repeat antigens required for 50% inhibition of antigen-antibody binding closely paralleled the difference in repeat number; that is, the concentrations of the 2-, 9-, and 16-repeat antigens required for 50% inhibition were 2-, 6-, and 15-fold lower, respectively, than the required concentrations of 1-repeat antigen (Fig. 4B). The combination of 1-repeat coating antigen and 1-repeat anti-

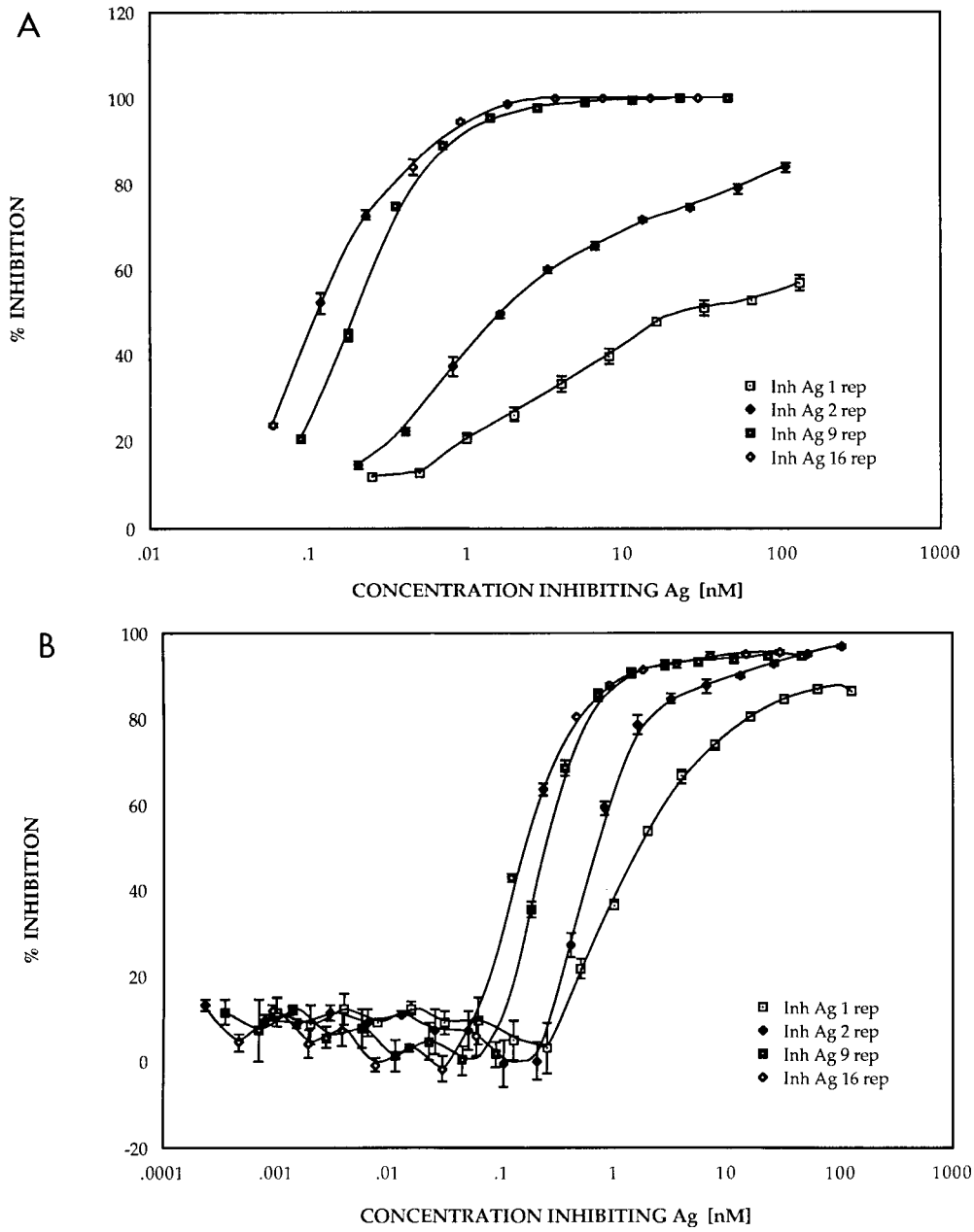


FIG. 4. ELISA inhibition showing relative binding affinities of antibodies elicited to alpha C proteins expressing 9 repeats (A) and 2 repeats (B). In each experiment, antisera from two immunized rabbits were pooled. Plates were coated with the corresponding antigen (Ag) (9- and 2-repeat antigens, respectively), and 1-, 2-, 9-, and 16-repeat proteins were used as the inhibiting antigen (Inh Ag 1, 2, 9, and 16 rep, respectively).

serum gave results similar to those obtained with the 2-repeat protein (data not shown).

**Fractionation of lower-molecular-weight from higher-molecular-weight bands of 9-repeat alpha C protein.** Since both cloned and native alpha C proteins are composed of proteins of various molecular sizes, we deemed it necessary to demonstrate that the effect of repeat number on antibody binding (Fig. 4A) was due primarily to the parent, or full-size, molecule and not to the less abundant, lower-molecular-weight constituents (Fig. 3). We therefore separated the 9-repeat alpha C protein into a higher-molecular-weight pool (predominantly the six top bands) and a lower-molecular-weight pool (predominantly the four bottom bands), using sequential fractions of the Superose-12 column, from the last step of our purification method (not shown). We compared these pools, normalized for protein content, for their ability to inhibit antibody binding by the 9-repeat-protein-elicited antiserum. Fifty percent inhibition was obtained at a concentration of 0.1 μg/ml for the high-molecular-weight pool and at 0.3 μg/ml for the low-molecular-weight pool. Thus, as expected, the overall inhibition shown for the 9-repeat alpha C protein in Fig. 4A was attributed mainly to the high-molecular-weight component, while the smaller bands behaved more like the low-repeat-number alpha C protein.

**Mouse protection.** In order to study the effect of variation in repeat number on binding to a protective epitope(s), we tested

TABLE 2. Protection of mice by rabbit antisera to alpha C proteins with various numbers of identical tandem repeats

| Antiserum[a] | No. of pups alive/total | % Survival[b] |
|---|---|---|
| 1-repeat | 38/50 | 76 |
| 2-repeat | 40/53 | 75 |
| 9-repeat | 21/51 | 41 |
| 16-repeat | 16/33 | 48 |
| Ia-TT | 71/72 | 99 |
| Preimmune | 11/84 | 13 |

[a] Rabbit antiserum to alpha C proteins with 1, 2, 9, and 16 repeats was elicited. Preimmune serum was used as a negative control, and antiserum to tetanus toxoid (TT) protein conjugated with polysaccharide type III (Ia-TT) was used as a positive control.

[b] Significant differences in protection levels for different antisera were identified by Fisher's exact test. The results were as follows: for 1- versus 9- and 16-repeat antisera, $P < 0.0001$; for 2- versus 9- and 16-repeat antisera, $P < 0.0001$; and for Ia-TT versus 1-, 2-, 9-, and 16-repeat antisera, $P < 0.0001$.

the protective efficacy of rabbit antisera to 1-, 2-, 9-, and 16-repeat alpha C proteins in mice. One- and 2-repeat-elicited antibodies appeared to be more protective (survival rates of 76 and 75%, respectively) than 9- and 16-repeat-elicited antibodies (survival rates of 41 and 48%, respectively) (Table 2).

## DISCUSSION

Alpha C protein is a potentially valuable component in the construction of GBS conjugate vaccines. In previous studies, it has been shown that beta C protein conjugated to the type III polysaccharide of GBS protects neonatal mice from GBS challenge by functioning both as a carrier protein to augment the immunogenicity of the polysaccharide and as a protective immunogen (18). However, the beta C protein is present in only 10% of clinical GBS isolates. Alpha C proteins are present in approximately 50% of all clinical isolates and in 90% of isolates bearing non-type III polysaccharides (10a). The alpha C protein, like surface antigens of many pathogenic microorganisms, contains a series of tandem repeats. It has been found that the alpha C proteins vary in their repeat number in naturally occurring GBS. The role of these variable repeat regions in protective immunity is not well understood.

In this study, we tested the effects of the alpha C protein repeat number on (i) immunogenicity in rabbits, (ii) the repertoire of antibodies elicited in rabbits, and (iii) the protective efficacy of these antibodies in mice. The immunogenicity of protective antigens is crucial to vaccine development. We found that 1-, 2-, 9-, and 16-repeat alpha C proteins were all highly immunogenic in rabbits. In contrast, Khan et al. found that the antibody response of mice to the peptide consisting of amino acids 115 to 131 (aa115-131 peptide) of the P28 glutathione *S*-transferase protective protein of *Schistosoma mansoni* increases dramatically as the copy number increases from one to two, four, and eight repeats (13). Likewise, Broekhuijsen and colleagues found similar effects for the VP1 region of foot-and-mouth disease virus (one, two, and four repeats) in immunogenicity studies with mice (5). Neither the aa115-131 peptide nor the VP1 region is found in nature as a repeating protein; rather, each is expressed as a single copy on the surface of the microorganism. Thus, constructs with more copies of the aa115-131 or VP1 region were artificially created for immunization and protection studies. In addition to the number of repeats, the size of the repeat region is important for immunogenicity. One repeat of the alpha C protein is 82 amino acids long (22), while the aa115-131 region is 16 amino acids long (13) and the VP1 region is 25 amino acids long (5). The difference in size between one repeat of the alpha C protein and one repeat of the VP1 or aa115-131 region may explain the differences in immunogenicity when copy numbers are increased.

The repertoires of antibodies elicited by the repeat-variant alpha C protein constructs exhibited important differences. The alpha C proteins with the most repeats (9 or 16) appeared to elicit a repertoire of antibodies distinct from that elicited by alpha C proteins with fewer repeats (1 or 2). Our ELISA inhibition results suggest that antibodies elicited by 9- and 16-repeat alpha C proteins recognize predominantly a conformational epitope(s) expressed by proteins containing more repeats (9 or 16) but lose their overall binding affinity for an epitope(s) expressed by proteins containing fewer repeats (1 or 2). Evidence in support of this conclusion includes the fact that antibodies elicited by 9- and 16-repeat proteins showed much higher overall binding affinity for the 9- and 16-repeat antigens than for the 1- and 2-repeat antigens than was expected simply on the basis of their repeat number. In other words, if 9-repeat antibodies were directed to an epitope(s) on individual repeats, the overall binding affinity for 9- or 16-repeat antigens would be expected to be 9 and 16 times higher, respectively, than that for 1-repeat antigen. However, the overall binding affinity for 9- and 16-repeat antigen was in fact 200 times higher than that for 1-repeat antigen (Fig. 4A). The results were the same for 16-repeat-protein-elicited antibodies.

In contrast, immunization of rabbits with 1- or 2-repeat proteins elicited antibodies that recognized an epitope(s) that did not depend on repeat number. This conclusion is supported by the equal binding affinities of the 1- and 2-repeat antisera for 1-, 2-, 9-, and 16-repeat proteins (Fig. 4B). Hypothetically, conformational epitopes might arise as a result of changes in the tertiary structure of the proteins evoked by the increased length (and repeat number) of the protein; these changes could be due to repeat-repeat interactions, to alterations in the angles of peptide bonds, or to interaction of the repeat region(s) with other domains (i.e., the N-terminal and C-terminal regions). Alternatively, exposure of a hidden epitope(s) could be evoked by the decreased length (and repeat number) of the protein. As mentioned in the introduction, each repeat construct of the alpha C protein contained a partial repeat of 33 nucleotides encoding 11 amino acids. Therefore, all repeat constructs, including the 1-repeat construct, contained a joining segment, and new primary (linear) epitopes would not be introduced by addition of more segments.

Our results support the use of 1- and 2-repeat alpha C proteins as vaccines because these proteins elicit antibodies that bind all alpha C proteins with equal overall affinity. Antibodies elicited by 9- and 16-repeat alpha C proteins have lower overall affinities of binding to lower-repeat-number proteins. These findings suggest that immunization with 9- or 16-repeat alpha C protein permits the escape of lower-repeat-number GBS mutants from antibody-mediated immunity. It has been shown previously that the number of repeats varies in clinical GBS isolates (17). Recently, it was established in an animal model that GBS mutants with fewer repeats in their alpha C proteins arise and can escape antibody-mediated immunity (18a). In these experiments, rabbit antiserum raised to purified 9-repeat alpha C protein was injected intraperitoneally into mice, which were then challenged with GBS. Most of the splenic GBS isolates that survived expressed one repeat or none at all. It could be hypothesized that antiserum raised to the lower-repeat-number alpha C proteins binds equally well to these lower-repeat-number mutants and thus blocks their

escape. Mouse protection experiments have supported this hypothesis: immunization with 9- or 16-repeat-protein-elicited antisera provided protection to only 41 and 48% of mice, respectively, whereas immunization with 1- or 2-repeat-protein-elicited antisera produced significantly higher levels of protection (76 and 75%, respectively).

Since many pathogenic microorganisms other than GBS express repeat-containing antigens, these organisms may also exhibit antigenic variation based on differences in repeat number. One well-studied example is the M6 protein of group A streptococci. Jones et al. demonstrated variation in opsonic epitopes when spontaneous mutants of group A streptococci lost internal M-protein repeats (11). They hypothesized that the loss of opsonic epitopes resulted from changes in conformational epitopes, which in turn resulted from loss of repeats. In the study presented here, we found strong evidence for such alteration in conformational epitopes in the alpha C protein of GBS. Using purified alpha C proteins differing only in the number of tandem repeats, we demonstrated that variation in repeat number gives rise to immunologically important changes in the antigenic structure and protective epitopes of this protein. We showed that 1- and 2-repeat alpha C proteins are more protective than 9- and 16-repeat proteins. However, when the repeat number varies, other repeat-containing cell surface proteins may show other forms of antigenic variation that may affect protection differently. In other words, variation in repeat number may lead to increased or decreased protection, depending on the protein. Either way, repeat number may be an important consideration in the design of vaccines directed at repeat-containing antigens.

## ACKNOWLEDGMENTS

We thank Dennis Kasper for strong support and practical advice, Mike Wessels for fruitful discussions, Fred Ausubel for help in designing multimeric repeat constructs, Elizabeth Gong for excellent technical support, and Michael Lee for substantial help with the production of purified alpha C proteins.

This research was supported by NIH grants AI00981 and AI28500 and by Public Health Service contract N01-A1-25152 ("Prevention of group B streptococcal infection in neonatal and infant populations").

## REFERENCES

1. **Anderson, B. E., G. A. McDonald, D. C. Jones, and R. L. Regnery.** 1990. A protective protein antigen of *Rickettsia rickettsii* has tandemly repeated, near-identical sequences. Infect. Immun. **58:**2760–2769.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. John Wiley & Sons, New York.
3. **Baker, C., and M. S. Edwards.** 1990. Group B streptococcal infections, p. 742–811. *In* J. Remington and J. O. Klein (ed.), Infectious diseases of the fetus and newborn infants. The W. B. Saunders Co., Philadelphia.
4. **Bevanger, L., and J. A. Maeland.** 1979. Complete and incomplete Ibc protein fraction in group B streptococci. Acta Pathol. Microbiol. Scand. Sect. B **93:**51–54.
5. **Broekhuijsen, M. P., M. M. J. van Rijn, A. J. M. Blom, P. H. Pouwels, B. E. Enger-Valk, F. Brown, and M. J. Francis.** 1987. Fusion proteins with multiple copies of the major antigenic determinant of foot-and-mouth disease virus protect both the natural host and laboratory animals. J. Gen. Virol. **68:**3137–3143.
6. **Campbell, J. L., C. C. Richardson, and F. Studier.** 1978. Genetic recombination and complementation between bacteriophage T7 and cloned fragments of T7 DNA. Proc. Natl. Acad. Sci. USA **75:**2276–2280.
7. **Crain, M. J., W. D. Waltman II, J. S. Turner, J. Yother, D. F. Talkington, L. S. McDaniel, B. M. Gray, and D. E. Briles.** 1990. Pneumococcal surface protein A (PspA) is serologically highly variable and is expressed by all clinically important capsular serotypes of *Streptococcus pneumoniae*. Infect. Immun. **58:**3293–3299.
8. **Gilmore, R. D.** 1993. Comparison of the rompA gene repeat regions of Rickettsiae reveals species-specific arrangements of individual repeating units. Gene **125:**97–102.
9. **Hollingshead, S. K., V. A. Fischetti, and J. R. Scott.** 1987. Size variation in group A streptococcal M protein is generated by homologous recombination between intragenic repeats. Mol. Gen. Genet. **207:**196–203.
10. **Ibanez, C. F., J. L. Affranchino, R. A. Macina, M. B. Reyes, S. Leguizamon, M. E. Camargo, L. Aslund, U. Pettersson, and A. C. C. Frasch.** 1988. Multiple *Trypanosoma cruzi* antigens containing tandemly repeated amino acid sequence motifs. Mol. Biochem. Parasitol. **30:**27–33.

10a. **Johnson, D. R., and P. Ferrieri.** 1984. Group B streptococcal Ibc protein antigen: distribution of two determinants in wild-type strains of common serotypes. J. Clin. Microbiol. **19:**506–510.

11. **Jones, K. F., S. K. Hollingshead, J. R. Scott, and V. A. Fischetti.** 1988. Spontaneous M6 protein size mutants of group A streptococci display variation in antigenic and opsonogenic epitopes. Proc. Natl. Acad. Sci. USA **85:**8271–8275.
12. **Kemp, D. J., R. L. Coppel, H. D. Stahl, A. E. Bianco, L. M. Corcoran, P. McIntyre, C. J. Langford, J. M. Favaloro, P. E. Crewther, G. V. Brown, G. F. Mitchel, J. G. Cullvanor, and R. F. Anders.** 1986. Genes for antigens of *Plasmodium falciparum*. Parasitology **91:**83–108.
13. **Khan, C. M. A., B. Villareal-Ramos, R. J. Pierce, R. D. de Hormaeche, H. McNeill, T. Ali, S. Chatfield, A. Capron, G. Dougan, and C. E. Hormaeche.** 1994. Construction, expression, and immunogenicity of multiple tandem copies of the *Schistosoma mansoni* peptide 115-131 of the P28 glutathione *S*-transferase expressed as a C-terminal fusion to tetanus toxin fragment C in a live aro-attenuated vaccine strain of *Salmonella*. J. Immunol. **153:**5634–5642.
14. **Klein, B. S., L. H. Hogan, and J. M. Jones.** 1993. Immunologic recognition of a 25-amino acid repeat arrayed in tandem on a major antigen of *Blastomyces dermatitidis*. J. Clin. Invest. **92:**330–337.
15. **Lancefield, R. C., M. McCarty, and W. N. Everly.** 1975. Multiple mouse-protective antibodies directed against group B streptococci. Special reference to antibodies effective against protein antigens. J. Exp. Med. **142:**165–179.
16. **Levy, N. J., A. Nicholson-Weller, C. J. Baker, and D. L. Kasper.** 1984. Potentiation of virulence by group B streptococcal polysaccharides. J. Infect. Dis. **149:**851–860.
17. **Madoff, L. C., S. Hori, J. L. Michel, C. J. Baker, and D. L. Kasper.** 1991. Phenotypic diversity in the alpha C protein of group B streptococci. Infect. Immun. **59:**2638–2644.
18. **Madoff, L. C., J. L. Michel, E. W. Gong, A. K. Rodewald, and D. L. Kasper.** 1992. Protection of neonatal mice from group B streptococcal infection by maternal immunization with beta C protein. Infect. Immun. **60:**4989–4994.

18a. **Madoff, L. C., J. L. Michel, E. W. Gong, D. E. Kling, and D. L. Kasper.** 1996. Group B streptococci escape host immunity by deletion of tandem repeat elements of the alpha C protein. Proc. Natl. Acad. Sci. USA **93:**4131–4136.

19. **Madoff, L. C., J. L. Michel, and D. L. Kasper.** 1991. A monoclonal antibody identifies a protective C-protein alpha-antigen epitope in group B streptococci. Infect. Immun. **59:**204–210.
20. **Michel, J. L., B. D. Beseth, L. C. Madoff, S. K. Olken, D. L. Kasper, and F. M. Ausubel.** 1994. Genotypic diversity and evidence for two distinct classes of the C protein alpha antigen of group B *Streptococcus*, p. 331–332. *In* A. Totolian (ed.), Pathogenic streptococci: present and future. Lancer Publications, St. Petersburg.
21. **Michel, J. L., L. C. Madoff, D. K. Kling, D. L. Kasper, and F. M. Ausubel.** 1991. The C proteins of group B streptococcus, p. 214–218. *In* G. M. Dunny, P. C. Cleary, and L. L. McKay (ed.), Genetics and molecular biology of streptococci, lactococci, and enterococci. American Society for Microbiology, Washington, D.C.
22. **Michel, J. L., L. C. Madoff, K. Olson, D. E. Kling, D. L. Kasper, and F. M. Ausubel.** 1992. Large, identical, tandem repeating units in the C protein alpha antigen gene, bca, of group B streptococci. Proc. Natl. Acad. Sci. USA **89:**10060–10064.
23. **Nickoloff, J. A.** 1994. Sepharose spin column chromatography: a fast, nontoxic replacement for phenol:chloroform/ethanol precipitation. Mol. Biotech. **1:**105–108.
24. **Rakonjac, J. V., J. C. Robbins, and V. A. Fischetti.** 1995. DNA sequence of the serum opacity of group A streptococci: identification of a fibronectin-binding repeat domain. Infect. Immun. **63:**622–631.
25. **Relman, D. A., M. Domenighini, E. Tuomanen, R. Rappuoli, and S. Falkow.** 1989. Filamentous hemagglutinin of *Bordetella pertussis*, nucleotide sequence and crucial role in adherence. Proc. Natl. Acad. Sci. USA **86:**2637–2641.

25a. **Rodewald, A. K., A. B. Onderdonk, H. B. Warren, and D. L. Kasper.** 1992. Neonatal mouse model of group B streptococcal infection. J. Infect. Dis. **166:**635–639.

26. **Roditi, I., M. Carrington, and M. Turner.** 1987. Expression of a polypeptide containing a dipeptide repeat is confined to the insect stage of *Trypanosoma brucei*. Nature (London) **325:**272–274.
27. **Schneider, A., A. Hemphill, T. Wyler, and T. Seebeck.** 1988. Large microtubule-associated protein of *T. brucei* has tandemly repeated, near-identical sequences. Science **241:**459–462.
28. **Signas, C., G. Raucci, K. Jonsson, P. E. Lindgren, G. M. Anantharamaiah, M. Hook, and M. Lindberg.** 1989. Nucleotide sequence of the gene for a fibronectin-binding protein from *Staphylococcus aureus*: use of this peptide sequence in the synthesis of biologically active peptides. Proc. Natl. Acad. Sci. USA **86:**699–703.

29. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189:**113–130.
30. **Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185:**60–89.
31. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82:**1074–1078.
32. **Timoney, J. F., J. Walker, M. Zhou, and J. Ding.** 1995. Cloning and sequence analysis of a protective M-like protein gene from *Streptococcus equi* subsp. *zooepidemicus*. Infect. Immun. **63:**1440–1445.
33. **Wallis, A. E., and W. R. McMaster.** 1987. Identification of *Leishmania* genes encoding proteins containing tandemly repeating peptides. J. Exp. Med. **166:**1814–1824.
34. **Watson, H. L., D. K. Blalock, and G. H. Cassell.** 1990. Variable antigens of *Ureaplasma urealyticum* containing both serovar-specific and serovar-cross-reactive epitopes. Infect. Immun. **58:**3679–3688.
35. **Wessels, M. R., A. Munoz, and D. L. Kasper.** 1987. A model of high-affinity antibody binding to type III group B streptococcus capsular polysaccharide. Proc. Natl. Acad. Sci. USA **84:**9170–9174.
36. **Yother, J., and D. E. Briles.** 1992. Structural properties and evolutionary relationships of PspA, a surface protein of *Streptococcus pneumoniae*, as revealed by sequence analysis. J. Bacteriol. **174:**601–609.
37. **Zeng, X., H. L. Watson, K. B. Waites, and G. H. Cassell.** 1992. Serotype diversity and antigen size variation among invasive isolates of *Ureaplasma urealyticum* from neonates. Infect. Immun. **60:**3472–3474.
38. **Zheng, X., L. J. Teng, H. L. Watson, J. I. Glass, A. Blanchard, and G. H. Cassell.** 1995. Small repeating units within the *Ureaplasma urealyticum* MB antigen gene encode serovar specificity and are associated with antigen size variation. Infect. Immun. **63:**891–898.

*Editor:* V. A. Fischetti