MOLECULAR EPIDEMIOLOGY

Distribution and expression of bca, the gene encoding the c alpha protein, by Streptococcus agalactiae

J. A. MAELAND, O. G. BRAKSTAD*, L. BEVANGER and S. KROKSTAD

Department of Microbiology, School of Medicine, Norwegian University of Science and Technology, N-7006, Trondheim and *SINTEF Applied Chemistry, N-7465, Trondheim, Norway

A total of 52 clinical isolates of group B streptococci (GBS) was tested for expression of the c protein cα by a fluorescent antibody test (FAT) and by PCR amplification of a 202-bp stretch within the repeat unit of the bca gene encoding the cα protein. The strains were categorised as follows: cα FAT positive and PCR positive with amplification products of multiple sizes (category A, n = 12); FAT negative and with PCR products of multiple sizes (category B, n = 11); FAT negative and with a single PCR product of c.200 bp (category C, n = 5); negative in both tests (category D, n = 24). A single amplification product of minimum size and additional products of larger sizes corresponded to one and more bca repeats, respectively. Five of the 11 category B strains showed expression of low M₉ cα in whole cell-based Western blotting. The results showed that a proportion of the GBS isolates harboured bca gene elements that either were not expressed or they expressed cα molecular variants which could not be detected by the whole cell-based FAT. This genotype/phenotype discrepancy should be considered in relation to GBS typing, including the selection of antibody reagents and the technical approach to cα protein detection.

Introduction

Streptococcus agalactiae (group B streptococci; GBS) can be classified into nine serotypes on the basis of the capsular polysaccharide antigens, i.e., Ia, Ib and II–VIII. Any of the capsular serotypes may be further subtyped on the basis of surface-anchored and strain-variable protein antigens, including the c proteins c alpha (cα) and c beta (cβ), and the R proteins R1–R4 [1, 2]. The combination of capsular polysaccharide and protein antigen typing allows subdivision of GBS into a large number of serovariants [1], which is important in epidemiological studies of GBS infections. Several of the surface-localised GBS protein antigens induce protective immunity in animal models, including the trypsin-resistant cα [3]. This protein is encoded by the bca gene and is expressed by 45–50% of GBS strains [1, 4, 5]. In the prototype GBS strain A909 (Ia/cα/cβ), bca has an open reading frame of 3060 nucleotides with a central domain of nine identical, tandemly arranged repeats [4]. Each repeat contains 246 bp and encodes a peptide of 82 amino acids [4]. The bca gene from different GBS strains varies in the number of repeat regions, resulting in strain variability of the maximal cα M₉ [6–8]. The repeat domain is subjected to mutational deletions of repeat units down to a single repeat [6]. Such changes result in altered antigenic specificity of cα and probably affect the mechanism by which the bacteria may evade opsonophagocytic killing [7, 8].

A recent study showed that the gene encoding the cβ protein could be detected by PCR in some GBS strains which were negative in a fluorescent antibody test (FAT) designed to detect the cβ protein [9]. Lack of expression of the bca gene encoding cα protein has also been suggested [10]. As cα is an important GBS serotype marker, in addition to its function as a virulence factor [11] and target of protective antibodies [3, 11], the present study compared bca detection by PCR and cα protein detection by the FAT.

Materials and methods

Bacterial strains

A total of 52 GBS strains was examined. The strains were clinical isolates from neonates or adults, mostly
blood culture isolates, which were submitted by various Norwegian hospitals for serotyping. The isolates examined belonged to the capsular antigen types Ia (n = 12), Ib (n = 7), II (n = 18) and III (n = 15). Both strains which were cα protein positive in antibody-based testing and strains which were negative were included, but in other respects the isolates were selected arbitrarily from our strain collection. The prototype GBS strains A909 (serotype 1a/cα/cβ) and 335 (1a/cα/cβ) were also included. Bacteria were preserved in Greaves’s medium [12] at −80°C and were cultured on blood agar plates, in Todd-Hewitt broth or in glucose 1% w/v broth at 37°C for 18 h.

Anticα antibodies

The generation and characterisation of the anti-cα monoclonal antibody (MAb) has been described previously [13]. The MAb is of the lgG1/kappa isotype and showed an ELISA titre of 128 000 when the ascitic fluid was examined [13]. The MAb is cα protein-specific, but the structure and localisation of the epitope have not been determined. Rabbit anti-cα serum was raised against immunoadsorbent-purified cα protein as described previously [13].

FAT

Slides of GBS were prepared and tested by an indirect fluorescent antibody test (FAT) as described previously [13, 14]. Briefly, smears were air-dried, heat-fixed, incubated with the MAb (diluted 1 in 200), washed, incubated with FITC-conjugated anti-mouse immunoglobulin (1 in 40; Dako A/S, Glostrup, Denmark), washed and read with a Nikon epifluorescence microscope. Only linear peripheral fluorescence (2+ or 3+ reaction) was judged positive.

Immunoblotting

Bacteria were lysed by hot SDS 1% w/v and components were separated by SDS-PAGE with polyacrylamide 12% gels and transferred on to PVDF membranes (BioRad, Richmond, CA, USA), pore size 0.2 μm, according to Laemmli [15] and as described previously [13]. Membrane strips were processed for probing with the anti-cα polyclonal antibody (1 in 1000) or MAb (1 in 1000) and IgG antibody binding was detected with peroxidase-labelled antisera (Dako) as described previously [13]. Colour was developed with o-phenylenediammine/H2O2 [13].

Oligonucleotide primers and probe

Two PCR primer sets based upon the published sequence of the bca gene of GBS strain A909 were synthesised [4]. The sequences of the primer set alpha primer 1 (21 nucleotides) and alpha primer 2 (23 nucleotides) are shown in Fig. 2. These primers targeted sites within the bca repeat unit and were used in a repeat unit PCR (RU-PCR). For the N-terminal repeat, the PCR product corresponded to nucleotides 785–985 of the A909 bca gene [4], a region which comprised 82% of the repeat. The second primer set, alpha primer 3 (5’T-GCG AGA GTA CAG GAA GGG CT-3’) and alpha primer 4 (5’T-GTG TCA CAG CAA TAA ATG GTG A-3’), were identical to a primer set described previously [6] and targeted within the N-terminus and downstream from the C-terminus, respectively. This set was used in a repeat area PCR (RA-PCR) to amplify the whole repeat region of bca in addition to the flanking areas. A 32-mer alpha probe corresponding to nucleotides 888–919 of the N-terminal bca repeat unit was synthesised and labelled with alkaline phosphatase. The sequence of the probe is shown in Fig. 2. The primers and the probe were synthesised and preserved as described previously [16].

PCR amplification

Glucose broth cultures (37°C, 20 h) of GBS (10 μl) were diluted 1 in 10 in sterile saline, pelleted by centrifugation (10 000 g, 5 min) and washed three times with saline (100 μl). Washed bacteria were suspended in 100 μl of lysis buffer (50 mM KCl; 10 mM Tris-HCl, pH 8.3; 2.5 mM MgCl2; Nonidet P40 0.45% v/v; Tween 20 0.45% v/v; gelatin 0.1 mg/ml) containing proteinase K 6 μg. The mixture was heated at 56°C for 1 h and heat-inactivated (95°C, 5 min) to obtain the bacterial lysates for PCR. For the RU-PCR, 2 μl of lysisate were mixed with 46 μl of a PCR reaction mixture - 2 μl of each primer (0.6 μM), 2.5 μl of dNTPs (1 mM of ATP, GTP, TTP and CTP), 5 μl (10×) PCR buffer (Perkin-Elmer Cetus, Norwalk, CT, USA), 1.5 mM MgCl2, sterile water 34.5 μl and mineral oil 50 μl. The mixture was heated (94°C) and 1 unit of Taq DNA polymerase (Perkin-Elmer) was added. A total of 30 PCR cycles was carried out (denaturing 94°C, 1 min; annealing 56°C, 1 min; DNA extension 72°C, 2 min) with a terminating time delay of 7 min at 72°C. In the RA-PCR, with the alpha 3/alpha 4 primer set, similar conditions were used except that 5 μl dNTPs (1 mM of each nucleotide), annealing temperature of 63°C (1 min) and 35 cycles were used. For some GBS strains, elevated alpha 3/alpha 4 primer concentrations - 4 μl (1.2 μM) of each primer - were used. PCR products were detected by electrophoresis in agarose 2% w/v gels as described previously [16].

Southern blot analysis

PCR products separated by electrophoresis were transferred to nylon membranes (Gene-Screen Plus; Biotechnology Systems, Boston, MA, USA) after DNA denaturation and neutralisation as described previously [16], except that depurination was omitted. Membranes were prehybridised (5 × SSC, SDS 0.5% w/v, bovine serum albumin 1 mg/ml, Ficoll 1 mg/ml, polyvinylpyrrolidine 1 mg/ml) at 60°C (90 min) and hybridised in the same solution with 20 μl of the internal DNA
probe at 60°C for 20 min. The E-link Plus Oligonucleotide kit (Genosys Biotechnologies, Houston, TX, USA) was used for labelling the probe with alkaline phosphatase. A chemiluminescent substrate included in the E-link kit was used for signal development, according to the manufacturer’s instructions.

Cloning and DNA sequencing

The pGem-T Easy Vector System (Promega, Madison, WI, USA) was used for the cloning of the RU-PCR product from two GBS strains, according to the manufacturer’s instructions (Promega Technical Manual, pGEM-T and pGEM-T Easy Vector Systems, 1997), including composition of solutions, buffers, the LB and the SOC medium. Briefly, 1 ml of PCR product was mixed with 1 ml of (10×) T4 ligase buffer (300 mM Tris-HCl, pH 7.8; 100 mM MgCl2; 100 mM DTT; 10 mM ATP), pGEM-T Easy Vector (50 ng) 1 ml, T4 DNA ligase (3 Weiss units = 1 ml) to a final volume of 20 μl. The extension products were purified by ethanol precipitation to remove excess terminators. The DNA sequences of the purified products were analysed in an ABI 373 DNA sequencer. The nucleotide sequence was aligned with the nucleotide sequence of accession no. M97256 from the Gene Bank database at the National Centre of Biotechnology Information.

Results

Of the 52 GBS strains tested, 12 strains (23%) were both C4 FAT and bca repeat unit PCR (RU-PCR) positive (Table 1). All the 12 FAT-positive strains generated RU-PCR products with three-to-five visible bands on gel electrophoresis. These strains represented category A isolates according to the definition of categories of FAT/RU-PCR reactivity patterns presented in Table 1. The sizes of the various PCR products were very close to the theoretical estimates of 202, 447, 693 and 939 bp, respectively, when assuming that products spanning a variable number of tandem bca repeats were generated.

The 40 FAT-negative strains could be subdivided into the categories B, C and D (Table 1), all categories being illustrated by the examples shown in Fig. 1a. Of the 40 strains, 16 were bca RU-PCR positive, 11 of 16 with two or more bands of amplification products (category B) and 5 of 11 with only one band (category C). Both category B and C strains occurred among GBS of different capsular-antigen types (Table 1). No FAT-negative/RU-PCR negative strains (category D) occurred among the type Ia strains, the GBS serotype which expresses C4 protein with the highest frequency [1]. Category D strains occurred among type Ib (14%), type II (61%) and type III (80%) strains, known to express C4 protein less frequently [1]. In Southern blotting, the internal probe for the bca repeat unit hybridised corresponding to the position of the RU-PCR products of different sizes (Fig. 1b), except for the category C strain 15388/81 (Fig. 1b, lane 6), an urogenital isolate. The product of this strain and the category C strain 45/95 (Fig. 1b, lane 5), a blood culture isolate from a 43-year-old woman, were

Table 1. Correlation of FAT for C4 protein detection and PCR (RU-PCR) for bca gene detection in 52 GBS strains

<table>
<thead>
<tr>
<th>Number of strains of capsular type</th>
<th>C4 FAT / bca category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>la</td>
</tr>
<tr>
<td>A</td>
<td>FAT+ / PCR+ / multiple*</td>
</tr>
<tr>
<td>B</td>
<td>FAT+ / PCR+ / multiple*</td>
</tr>
<tr>
<td>C</td>
<td>FAT+ / PCR+ / single†</td>
</tr>
<tr>
<td>D</td>
<td>FAT+ / PCR*</td>
</tr>
<tr>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>

*Multiple bands observed after RU-PCR amplification.
†A single band (202 bp) observed after RU-PCR amplification.
sequenced. The strain 45/95 product showed 100% homology with the corresponding segment of bca of the prototype strain A909. However, the strain 15388/81 product showed only 84% homology because of multiple point mutations with clustering of mutations in the 818–824-bp area and also occurred with high frequency (28% of nucleotides) in the region targeted by the DNA probe (Fig. 2). The latter observation probably explains why the GBS 15388/81 PCR product failed to bind the probe under the conditions of stringency used. The probable amino-acid sequence indicated an exchange of 10 residues within the amplified region of an expressed 15388/81 c' protein, eight of which were non-synonymous.

All the 16 category B or C strains were analysed by the RA-PCR designed to generate amplification products, which included the whole bca repeat region and parts of the regions flanking the repeat area. A single product was generated by the RA-PCR, shown in Fig. 3 for one strain of each of the categories A, B, C and D. The category C strains generated products of 0.8 kb, corresponding to a single repeat, and the category B strains generated products in the calculated range of c. 1.0 to slightly above 2.0 kb, corresponding to two-to-seven repeats.

To further test if the c'' FAT-negative category B and C strains expressed c'' protein in low quantities, whole

---

**Fig. 1.** Gel electrophoresis of products of the RU-PCR for bca gene amplification (a) and hybridisation of the products to an internal DNA probe targeting bca repeats (b). The isolates tested were the GBS strains: 1, 335 (serotype Ia/c''); 2, A909 (Ia/c''C226); 3, 49/95 (II); 4, 11/93 (Ib/c''C226); 5, 45/95 (Ia); 6, 15388/81 (III); 7, 14/95 (Ib/c''). GBS categories are shown at the bottom and DNA standard is shown to the left.

---

**Fig. 2.** Alignment of nucleotide sequence of the RU-PCR product of the GBS strain 15388/81 (Fig. 1, lane 6) and the corresponding DNA segment of the N-terminal repeat of bca of strain the A909 (Gene Bank database, accession no. M97256). Stretches corresponding to the primers 1 and 2 and the DNA probe are underlined. Numbering of nucleotides is according to ref. [4]. Nucleotide substitutions are indicated for strain 15388/81 and consensus areas indicated by hyphen.
cells of all strains were tested by immunoblotting against the polyclonal and monoclonal anti-c\^\text{\textalpha} antibodies. Both the antibodies revealed c\^\text{\textalpha} expression by 5 of the 11 category B strains. Typically, these strains produced c\^\text{\textalpha} proteins which resulted in only one or a few weakly stained bands in the M\_r range 52–100 kDa, as shown by the examples in Fig. 4 (lanes 2 and 3). In the RA-PCR, the category B strains expressing c\^\text{\textalpha} generated amplimers which indicated two-to-seven bca repeats. None of the category C or D strains expressed protein that could be detected by the whole cell-based immunoassay.

Discussion

The c\^\text{\textalpha} protein of GBS has important biological functions [3, 11, 17] and it is an important serotype marker together with other strain-variable GBS antigens [1]. Whereas serotyping depends on expression of the serotype marker, typing can also be performed by methods designed to identify the gene encoding the marker. This implies a risk of discrepancy between gene- and gene product-based typing methods. Such a discrepancy was demonstrated for the c\^\text{\textbeta} protein of GBS [9], has been suggested by other investigators for the c\^\text{\textalpha} protein [10] and is a major finding of the present study. Of 52 GBS strains tested in the present study, 16 strains (31%) showed c\^\text{\textalpha} protein/bca gene discrepancy when the isolates were tested by a c\^\text{\textalpha} FAT and a PCR (RU-PCR) which amplified a region of the bca repeat unit. The frequency of the discrepancy varied with the capsular antigen type of the bacteria (Table 1) as does c\^\text{\textalpha} protein expression [1]. Thus, all the c\^\text{\textalpha} FAT-negative type Ia strains, the serotype with the highest frequency of c\^\text{\textalpha} expression [1], showed a discrepancy. It is possible that GBS of the serotype Ia always harbour the bca gene. The discrepancy occurred less frequently in GBS serotypes which express c\^\text{\textalpha} protein more rarely.

The bca gene shows strain variability in the number of repeats and this coincides with variability of the M\_r and the number of peptides of c\^\text{\textalpha} proteins [6]. This bca characteristic was also observed in the RU-PCR and RA-PCR used in the present study. The majority of RU-PCR positive strains showed ladder formation of PCR products, probably because the extension proceeded over a variable number of repeats during the amplification of bcas which contained several repeats, and the size of the RA-PCR product, which comprised the whole of the bca repeat region, coincided with the number of RU-PCR product lines.

Although none of the category C strains, with only one full-length bca repeat, showed c\^\text{\textalpha} protein expression, 5 of 11 category B strains, FAT-negative and RU-PCR positive with more than one band, demonstrated c\^\text{\textalpha} expression in immunoblotting.

Comprehensive interpretations of these findings may be that some of the bca gene-containing strains expressed c\^\text{\textalpha} protein in very low quantities, or that the antibodies...
employed, raised against a c\textsuperscript{n} protein with many repeats [13], were unable to detect c\textsuperscript{n} proteins with fewer repeats [7]. Also, steric hindrance due to the positioning of low M, c\textsuperscript{n} at the bacterial cell surface is a possible explanation for the negative FAT in strains containing bca [8]. Alternatively, the bca gene may be inactive in a proportion of the GBS that harbour the gene, because of mutational changes in a regulatory region. With the FAT-negative/RU-PCR positive strains, the protein was not detected in fluid culture supernates (not shown).

Other investigators have demonstrated that mutational deletions of c\textsuperscript{n} repeats may occur and that this provides selective advantages to the mutants [6, 8]. The present study demonstrated that the category C GBS strain 15388/81 had numerous point mutations within the repeat region amplified by the RU-PCR. Although the strain 15388/81 bca may be exceptional in this regard, these findings suggest the possibility that epitope specificity also may be altered by point mutations within a c\textsuperscript{n} protein region which is the target of protective antibodies [18].

This study has demonstrated that antibody-based methods to detect c\textsuperscript{n} protein expression as in traditional GBS serotyping and methods to detect bca, the gene encoding the c\textsuperscript{n} protein, will give discrepant results in a sizeable proportion of GBS strains. This is an important consideration in the selection of GBS typing methods.

We are grateful to Grete Iversen and Randi Valsøe Lyng for technical assistance.

References