JOURNAL OF BACTERIOLOGY, Oct. 2000, p. 5919–5921 Vol. 182, No. 20
0021-9193/00/$04.00+0

The *Streptococcus pneumoniae* Beta-Galactosidase Is a Surface Protein

DOROTHEA ZÄHNER AND REGINE HAKENBECK*

Department of Microbiology, University of Kaiserslautern, D-67663 Kaiserslautern, Germany

Received 30 May 2000/Accepted 19 July 2000

The β-galactosidase gene of *Streptococcus pneumoniae*, *bgaA*, encodes a putative 2,235-amino-acid protein with the two amino acid motifs characteristic of the glycosyl hydrolase family of proteins. In addition, an N-terminal signal sequence and a C-terminal LPXTG motif typical of surface-associated proteins of gram-positive bacteria are present. Trypsin treatment of cells resulted in solubilization of the enzyme, documenting that it is associated with the cell envelope. In order to obtain defined mutants suitable for *lacZ* reporter experiments, the *bgaA* gene was disrupted, resulting in a complete absence of endogenous β-galactosidase activity. The results are consistent with β-galactosidase being a surface protein that seems not to be involved in lactose metabolism but that may play a role during pathogenesis.

Streptococcus pneumoniae is an important human pathogen, causing invasive diseases such as pneumonia, bacteremia, and meningitis. In order to analyze gene expression in this organism, the availability of reporter constructs is highly desirable. The *Escherichia coli* β-galactosidase gene *lacZ* has been used in several studies. It has long been known that *S. pneumoniae* produces a β-galactosidase that can be purified from the growth medium (8, 11), necessitating the isolation of mutants devoid of this enzyme activity for gene expression studies. However, the β-galactosidase-negative *S. pneumoniae* strains described so far were spontaneously obtained and were not further characterized (6, 20). A β-galactosidase activity of *S. pneumoniae* has been isolated from culture supernatants. The objectives of the present study were the identification of the gene encoding the β-galactosidase from *S. pneumoniae* and the construction of a genetically defined β-galactosidase-negative mutant suitable for work with *lacZ* reporter constructs in *S. pneumoniae*.

Identification of the *bgaA* gene. For sequence database searches, the BLAST program was used (2). A BLAST homology search of the unfinished *S. pneumoniae* capsular type 4 strain genome, obtained from The Institute for Genomic Research at http://www.tigr.org, revealed a 365-residue peptide with 26% identical amino acids compared to the β-galactosidase of *Streptococcus thermophilus*. The peptide represented an internal region of a putative 2,235-amino acid protein, the product of a 6,704-bp open reading frame. The region covered the two motifs characteristic of glycosyl hydrolase family 2 (9), both of which showed some anomalies in the *S. pneumoniae* protein: the highly conserved residues Y in motif I and H in motif II were both replaced by an N (Fig. 1). The presence of these alterations was verified by direct sequencing between codons 285 and 716, using PCR products obtained from chromosomal DNA of *S. pneumoniae* strain R6, a nonencapsulated derivative of Rockefeller University strain R36A (3). For direct sequencing, a BigDye terminator cycle sequencing kit (Perkin-Elmer, Warrington, England) was used.

The deduced 2,235-amino-acid sequence (247.3 kDa) revealed several features not typical of described β-galactosidases, all of which are only approximately 1,000 amino acids long. The conserved β-galactosidase motifs are located in the first half of the protein. A region of about 100 amino acid residues at the N terminus contains a putative signal peptide (Fig. 1). The structure of this signal peptide is similar to those of the consensus sequence of signal peptides from gram-positive bacteria, suggesting that the enzyme is exported (17, 19). The second half of the protein has no similarity to other proteins, except for the extreme C terminus, which contains an LPXTG motif preceding a hydrophobic domain, features of gram-positive bacterial surface proteins (15). Similar C termini are present in the *Streptococcus pyogenes* M protein (16) and several pneumococcal surface proteins, e.g., neuraminidase A (5), hyaluronidase (4), and *N*-acetylglucosaminidase (7) (Fig. 1).

Disruption of the *bgaA* gene. In order to test whether *bgaA* encodes the *S. pneumoniae* β-galactosidase, the erythromycin resistance determinant *ermA* from the *Enterococcus faecalis* plasmid pAMβ1 (14) was inserted into *bgaA*, resulting in disruption of the reading frame. A single *Mun*I site was introduced into an internal *bgaA* fragment by site-directed mutagenesis (10), and the mutagenized fragment was cloned into vector pCR2.1 (Invitrogen, Leiden, The Netherlands). The *ermA* gene was amplified with oligonucleotide primers 5′-AGAGTGTGTTGATAGTGCAGTATC and 5′-TTATTTC CTCCCGTTAAATAATAG from pJDC9 (14), cloned into pCR2.1, and reisolated after *Eco*RI restriction; the *ermA* gene could now be cloned into the *Mun*I site of the *bgaA* fragment to give plasmid pBER. A 1.6-kb DNA fragment containing the *erm* cassette and flanking *bgaA* regions was amplified from pBER by PCR and used as donor DNA in transformation experiments with *S. pneumoniae* R6 as a recipient. Transformation was performed essentially as described previously (18). With 1 μg of erythromycin per ml for selection, Ery^r colonies were readily obtained, and integration of the Ery^r marker in the *bgaA* gene in individual mutants was verified by PCR.

β-Galactosidase activity in *bgaA* mutants. The Ery^r transformant R6 *bgaA*::*erm* showed no β-galactosidase activity when tested on D-agar plates (1) containing 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and catalase (17,000 U ml^{-1}; Sigma-Aldrich, Munich, Germany), whereas colonies of the parental R6 strain appeared light blue. In addition, β-galactosidase activity was determined in liquid cultures using C medium (12) supplemented with 0.2% yeast extract at 37°C

* Corresponding author. Mailing address: Department of Microbiology, University of Kaiserslautern, Paul Ehrlich Strasse, D-67663 Kaiserslautern, Germany. Phone: 49-631-205-2353. Fax: 49-631-205-3799. E-mail: hakenb@rhrk.uni-kl.de.

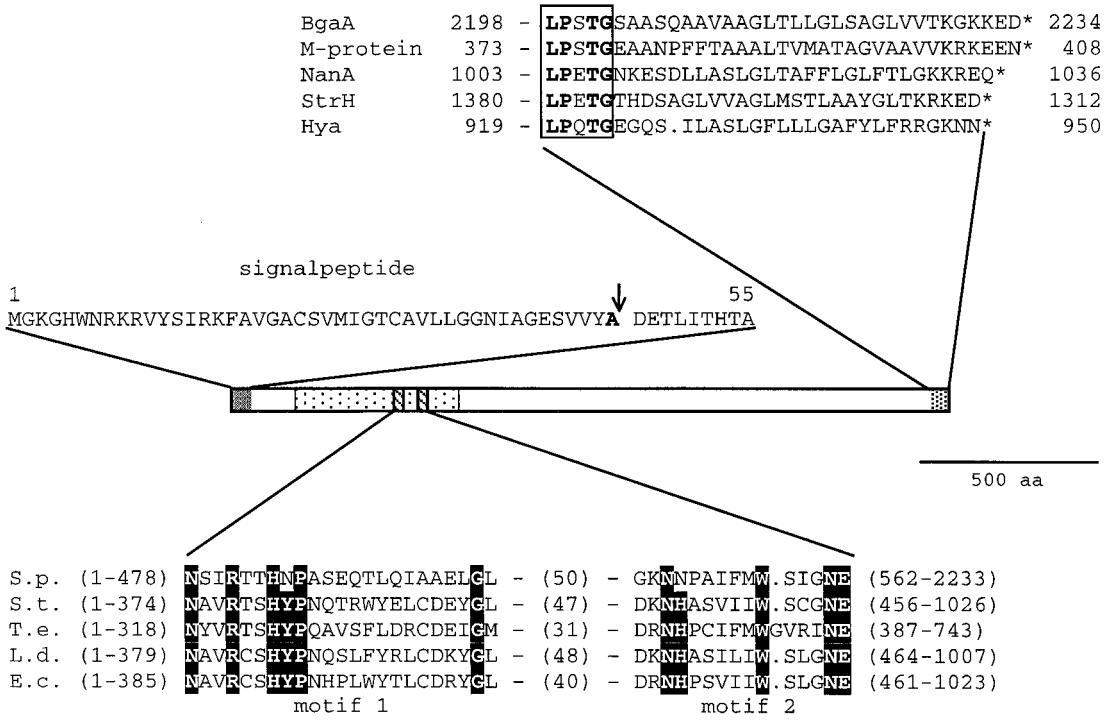

FIG. 1. Schematic representation of the β-galactosidase of *S. pneumoniae* (S.p.). The amino acid (aa) sequences are given for three important regions: the N-terminal signal peptide (residues 1 to 55; the arrow marks the putative signal peptidase cleavage site); the two conserved motifs of glycosyl hydrolase family 2 (hatched boxes) within the region homologous to other β-galactosidases (stippled box; approximately residues 90 to 600); and 35 C-terminal residues with the cell wall-anchoring LPXTG motif (stippled box at right end). The putative active site is aligned with β-galactosidases of the following organisms: S.t., *S. thermophilus* (SWISS-PROT P23989); T.e., *Thermoanaerobacter ethanolicus* (SWISS-PROT P77989); L.d., *Lactobacillus delbrueckii* (SWISS-PROT P33486); and *E. coli* (GenBank AJ002684). Strictly conserved amino acid residues within the motifs are highlighted. The C-terminal region is compared to those of other proteins containing an LPXTG motif: M-protein, *S. pyogenes* M protein (PIR S30283); *S. pneumoniae* proteins NanA (neuraminidase A; SWISS-PROT 59959), StrH (β-*N*-acetylglucosaminidase; SWISS-PROT P49610), and Hya (hyaluronidase; SWISS-PROT Q54873). Dots indicate gaps in the alignment; the asterisk marks the C terminus.

without aeration. Samples were removed at different cell densities. After centrifugation (5 min at 12,000 × *g*), cells were resuspended in 0.1 M sodium phosphate buffer–0.1% Triton X-100 and lysed during 5 min of incubation at 37°C. β-Galactosidase activity was determined essentially as described by Miller (13) with *o*-nitrophenyl-β-D-galactopyranoside as a substrate. None of the transformants showed any activity compared to the 60 Miller units of activity of strain R6. These results demonstrate that the *bgaA* gene encodes the β-galactosidase of *S. pneumoniae* and that it is solely responsible for the endogenous β-galactosidase activity of this species.

Localization of β-galactosidase. Since the β-galactosidase of *S. pneumoniae* has been isolated from culture supernatants (11), enzyme activity in the growth medium was determined and compared to cell-associated activity, i.e., in cell lysates (Table 1). Ninety-five percent of the activity was found in cell lysates. In order to distinguish between intracellular and surface localization of β-galactosidase, intact cells were harvested by centrifugation, washed once with 0.1 M sodium phosphate buffer (pH 7.5), resuspended in the same buffer containing trypsin (1 μg/ml; Promega, Madison, Wis.), and incubated for 5 min at 37°C. After centrifugation, 57% of the β-galactosidase activity was found in the supernatant. In a control experiment, *bgaA* mutant cells that expressed the *E. coli lacZ* gene (R6 *bgaA*::*erm uppS-lacZ*) were used (22). This mutant produces *E. coli* β-galactosidase in the cytoplasm and, after trypsin treatment, only traces of the enzyme could be detected in the solubilized fraction; these results demonstrate that trypsin does not induce cellular lysis (Table 1). In wild-type cells, the overall β-galactosidase activity after trypsin treatment was reduced to 51%, probably due to a limited resistance of the pneumococcal protein to trypsin.

Growth on lactose. Most β-galactosidases play a major role in lactose metabolism and can be induced by lactose. When *S. pneumoniae* R6 *bgaA*::*erm* was grown in C medium lacking

TABLE 1. Cellular localization of *S. pneumoniae* β-galactosidase

S. pneumoniae strain	β-Galactosidase activity (U) in the indicated sample:			
	Without trypsin[a]		With trypsin[b]	
	Cell lysate	Medium	Cell lysate	Supernatant
R6 *bgaA*::*erm*	0	0	0	0
R6	62	3.3	12	21
R6 *bgaA*::*erm uppS-lacZ*[c]	20	0.5	20	0.4

[a] Exponentially growing cells were centrifuged, the supernatant (growth medium) was removed, and cells were washed once with 0.1 M sodium phosphate buffer (pH 7.8). They were resuspended in the same buffer containing 0.1% Triton X-100 and lysed during 5 min of incubation at 37°C. β-Galactosidase activity in the growth medium and the cell lysate was determined as described in the text.

[b] Cells were harvested and washed as described in footnote *a* but were resuspended in buffer containing 1 μg of trypsin per ml. After incubation for 5 min at 37°C, cells were centrifuged immediately, and β-galactosidase activity in the supernatant and the cell pellet after lysis with Triton X-100 was determined.

[c] *S. pneumoniae* R6 *bgaA*::*erm uppS-lacZ* is a *lacZ* reporter derivative expressing LacZ of *E. coli* under the control of the promoter of the *uppS* gene (22).

glucose and sucrose (the only defined carbon sources in this medium) and without yeast extract, but with lactose, no effect on generation time was detected compared to that of the parental strain R6. Strain R6 contained similar β-galactosidase activities independent of the carbon source, and the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) also had no effect on β-galactosidase expression (data not shown). These results confirm that BgaA is the only β-galactosidase in *S. pneumoniae* and strongly suggest that BgaA is not involved in lactose metabolism. This conclusion is further confirmed by the presence in the pneumococcal genome of two ORFs encoding proteins highly similar to the 6-phospho-β-galactosidases of *Lactococcus lactis* and *Staphylococcus aureus*, both of which are preceded by genes encoding protein IIB/C of a putative lactose-specific phosphotransferase system.

Concluding remarks. For several surface proteins of gram-positive bacteria, proteolytic cleavage and subsequent release into the environment have been described (15). The present characterization of the *S. pneumoniae* β-galactosidase as a surface protein is in agreement with a putative role of the enzyme in the interaction with host cells, rather than an involvement in lactose metabolism. Surface proteins such as neuraminidase A, hyaluronidase, and *N*-acetylglucosaminidase of *S. pneumoniae*, all of which contain an LPXTG motif, have been described as virulence factors (4, 5, 7), and it is conceivable that β-galactosidase is a virulence factor as well. The presence of antibodies against β-galactosidase in convalescent-phase serum from a patient with a history of pneumococcal infection is in agreement with this assumption (24). The unusual high specificity of the enzyme for β-1,4-glycosidic bonds and a 10-times-higher specificity for Galβ1-4GlcNAc than for lactose (23) make it an ideal candidate for attacking polysaccharides conjugated to surface components of eukaryotic cells. Indeed, this property has been exploited for the analysis of complex polysaccharides (21). Curiously, no homologues of the pneumococcal enzyme with a signal peptide and the surface-anchoring LPXTG motif were detected in the genome databases of *Streptococcus mutans* and *S. pyogenes*. The availability of defined *S. pneumoniae bgaA* mutants will help to clarify the role of β-galactosidase in vivo and provides a suitable genetic background for expression studies using *lacZ* reporter constructs.

Nucleotide sequence accession number. The DNA sequence of *S. pneumoniae bgaA* has been deposited in GenBank under accession number AF282987.

This work was supported by the Deutsche Forschungsgemeinschaft (grants Ha 1011/7-1 and Ha 1011/7-2) and by European Commission contract no. BI04-CT98-0424.

REFERENCES

1. **Alloing, G., C. Granadel, D. A. Morrison, and J.-P. Claverys.** 1996. Competence pheromone, oligopeptide permease, and induction of competence in *Streptococcus pneumoniae*. Mol. Microbiol. **21:**471–478.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
3. **Avery, O. T., C. M. MacLeod, and M. McCarty.** 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. J. Exp. Med. **79:**137–158.
4. **Berry, A. M., R. A. Lock, S. M. Thomas, D. P. Rajan, D. Hansman, and J. C. Paton.** 1994. Cloning and nucleotide sequence of the *Streptococcus pneumoniae* hyaluronidase gene and purification of the enzyme from recombinant *Escherichia coli*. Infect. Immun. **62:**1101–1108.
5. **Càmara, M., G. J. Boulnois, P. W. Andrew, and T. J. Mitchell.** 1994. A neuraminidase from *Streptococcus pneumoniae* has the features of a surface protein. Infect. Immun. **62:**3688–3695.
6. **Campbell, E. A., S. Y. Choi, and H. R. Masure.** 1998. A competence regulon in *Streptococcus pneumoniae* revealed by genomic analysis. Mol. Microbiol. **27:**929–939.
7. **Clarke, V. A., N. Platt, and T. D. Butters.** 1995. Cloning and expression of the beta-N-acetylglucosaminidase gene from *Streptococcus pneumoniae*. Generation of truncated enzymes with modified aglycon specificity. J. Biol. Chem. **270:**8805–8814.
8. **Glasgow, L. R., J. C. Paulson, and R. L. Hill.** 1977. Systematic purification of five glycosidases from *Streptococcus* (*Diplococcus*) *pneumoniae*. J. Biol. Chem. **252:**8615–8623.
9. **Henrissat, B.** 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. **280:**309–316.
10. **Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease.** 1996. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene **77:**51–59.
11. **Hughes, R. C., and R. W. Jeanloz.** 1964. The extracellular glycosidases of *Diplococcus pneumoniae*. I. Purification of a neuraminidase and a beta-galactosidase active on the alpha-1-acid glycoprotein of human plasma. Biochemistry **10:**1535–1548.
12. **Lacks, S. A., and R. D. Hotchkiss.** 1960. A study of the genetic material determining an enzyme activity in pneumococcus. Biochim. Biophys. Acta **39:**508–517.
13. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
14. **Morrison, D. A., M. C. Trombe, M. K. Hayden, G. A. Waszak, and J.-D. Chen.** 1984. Isolation of transformation-deficient *Streptococcus pneumoniae* mutants defective in control of competence, using insertion-duplication mutagenesis with the erythromycin resistance determinant of pAMβ1. J. Bacteriol. **159:**870–876.
15. **Navarre, W. W., and O. Schneewind.** 1999. Surface proteins of gram-positive bacteria and the mechanism of their targeting to the cell wall envelope. Microbiol. Mol. Biol. Rev. **63:**174–229.
16. **Pancholi, V., and V. A. Fischetti.** 1989. Identification of an endogenous membrane anchor-cleaving enzyme for group A streptococcal protein. J. Exp. Med. **170:**2119–2133.
17. **Silhavy, T. J., S. A. Benson, and S. D. Emr.** 1983. Mechanisms of protein localization. Microbiol. Rev. **4:**313–334.
18. **Tiraby, J.-G., and M. S. Fox.** 1974. Marker discrimination and mutagen-induced alterations in pneumococcal transformation. Genetics **77:**449–458.
19. **Von Heijne, G.** 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14:**4683–4690.
20. **Wang, L. F., and R. H. Doi.** 1986. Nucleotide sequence and organization of *Bacillus subtilis* RNA polymerase major sigma (sigma 43) operon. Nucleic Acids Res. **14:**4293–4307.
21. **Yang, Y., and R. Orlando.** 1996. Simplifying the exoglycoside digestion/MALDI-MS procedure for sequencing N-linked carbohydrate side chains. Anal. Chem. **68:**570–572.
22. **Zähner, D.** 1999. Identifizierung von Zielgenen des signaltransduzierenden Zwei-Komponenten-Systems *cia* von *Streptococcus pneumoniae*. Ph.D. thesis. Universität Kaiserslautern, Kaiserslautern, Germany.
23. **Zeleny, R., F. Altmann, and W. Praznik.** 1997. A capillary electrophoretic study on the specificity of β-galactosidases from *Aspergillus oryzae*, *Escherichia coli*, *Streptococcus pneumoniae*, and *Canavalia ensiformis* (jack bean). Anal. Biochem. **246:**96–101.
24. **Zysk, G., R. J. M. Bongaerts, E. ten Thoren, G. Bethe, R. Hakenbeck, and H.-P. Heinz.** 2000. Detection of 23 immunogenic pneumococcal proteins using convalescent-phase serum. Infect. Immun. **68:**3740–3743.