A Proline-Rich Region with a Highly Periodic Sequence in Streptococcal beta Protein Adopts the Polyproline II Structure and Is Exposed on the Bacterial Surface.

Areschoug, Thomas; Linse, Sara; Stålhammar-Carlemalm, Margaretha; Hedén, Lars-Olof; Lindahl, Gunnar

Published in: Journal of Bacteriology

DOI: 10.1128/JB.184.22.6376-6393.2002

2002

Link to publication

Citation for published version (APA):

Total number of authors: 5

General rights
Unless other specific re-use rights are stated the following general rights apply:
• Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Proline-rich regions have been identified in many surface proteins of pathogenic streptococci and staphylococci. These regions have been suggested to be located in cell wall-spanning domains and/or to be required for surface expression of the protein. Because little is known about these regions, which are found in extensively studied and biologically important surface proteins, we characterized the proline-rich region in one such protein, the β protein of group B streptococci. The proline-rich region in β, designated the XPZ region, has a proline at every third position, and the sequence is highly periodic in other respects. Immunochemical analysis showed that the XPZ region was not associated with the cell wall but was exposed on the bacterial surface. Moreover, characterization of a β mutant lacking the XPZ region demonstrated that this region was not required for surface expression of the β protein. Comparison of the XPZ region in different β proteins showed that it varied in size but always retained the typical sequence periodicity. Circular dichroism spectroscopy indicated that the XPZ region had the structure of a proline rich sequence with unique properties that may allow protein Adopts the Polyproline II Structure and Is Exposed on the Bacterial Surface

Thomas Areschoug,1 Sara Linse,2 Margaretha Stålhammar-Carlemalm,1 Lars-Olof Hedén,3 and Gunnar Lindahl1*

Department of Medical Microbiology, Dermatology and Infection1, Department of Biophysical Chemistry,2 and Department of Cell and Organism Biology,3 Lund University, Lund, Sweden

Received 22 April 2002/Accepted 23 June 2002

Proline-rich regions have been identified in many surface proteins of pathogenic streptococci and staphylococci. These regions have been suggested to be located in cell wall-spanning domains and/or to be required for surface expression of the protein. Because little is known about these regions, which are found in extensively studied and biologically important surface proteins, we characterized the proline-rich region in one such protein, the β protein of group B streptococci. The proline-rich region in β, designated the XPZ region, has a proline at every third position, and the sequence is highly periodic in other respects. Immunochemical analysis showed that the XPZ region was not associated with the cell wall but was exposed on the bacterial surface. Moreover, characterization of a β mutant lacking the XPZ region demonstrated that this region was not required for surface expression of the β protein. Comparison of the XPZ region in different β proteins showed that it varied in size but always retained the typical sequence periodicity. Circular dichroism spectroscopy indicated that the XPZ region had the structure of a polyproline II helix, an extended and solvent-exposed structure with exactly three residues per turn. Because of the three-residue sequence periodicity in the XPZ region, it is expected to be amphipathic and to have different nonpolar and polar surfaces. This study identified a proline-rich structure with unique properties that is exposed on the surface of an important human pathogen.
two unrelated GBS strains (19, 24). Comparison of the two deduced protein sequences shows that they have XPZ regions containing 30 and 40 XPZ motifs, respectively, but apart from that difference the amino acid sequences are identical except for a single residue. These limited data suggest that the protein varies in size due to size variability in the XPZ region.

We show here that the XPZ region is exposed on the bacterial surface and is not required for cell wall anchoring of the protein. Moreover, we present spectroscopic evidence that the XPZ region adopts the conformation of a left-handed polyproline II (PPII) helix, an extended and solvent-exposed structure (55, 59). The length and sequence of the XPZ region varied in different proteins, but the sequence periodicity was strictly conserved. Together, these data indicate that the proline-rich XPZ region of the protein represents a novel type of surface structure in bacteria.

MATERIALS AND METHODS

Bacterial strains and media. The GBS strains used were of capsular serotypes Ia, Ib, and II. Strain SB35 was used to sequence the gene (see text). Strains A909 and H36B (51) were obtained from C. Schalen (Lund University, Lund, Sweden), and strain 70339 (9) was from L. Bevanger (University of...
The results plasmid, designated pLZ, was ligated into pLZ-UP which had been digested with EcoRI and BamHI, resulting in plasmid pLZ-UP.

The region in bac downstream of the XPZ region was amplified with synthetic oligonucleotides 5'-AAATTGGAATTCGTCGAAAGCTATTATTCGGAATG-3' and 5'-AAATTTGGACTATCCTTTGATTTTC-3'. The resulting PCR fragment was digested with EcoRI and BamHI, recognition sequences for which had been introduced through the primers. This fragment was ligated into plasmid pLZ12Spec (21) which had been digested with EcoRI and BamHI, resulting in plasmid pLZ-XP.

The region in bac downstream of the XPZ coding region was amplified with synthetic oligonucleotides 5'-AAATTGGATCCGTTAATAAAAGTGGG ACAAGCA-3' and 5'-AAATTTAAGCTTTTATGATTTTCCTCCCATACCAAT-3'. The PCR fragment was digested with EcoRI and BamHI, recognition sequences for which had been introduced through the primers. This fragment was ligated into plasmid pLZ-UP which had been digested with BamHI and HinIII, generating the Δbac/plLZΔbac strain.

Synthetic peptides. A 24-residue linear peptide, designated (XPZ)3, with the sequence SPKAPEAPRVPESPKTPEAPHVPE was purchased from Neosystem (Strasbourg, France). The (XPZ)3 peptide corresponds to amino acid residues 819 to 842 in the XPZ region of the β protein. The synthetic peptide was used in raising specific antibodies.

Proteins and antisera. The β protein expressed by strain SB35 was purified as described previously (34). Briefly, a suspension of washed bacteria was incubated at an elevated pH, which causes selective release of almost pure β protein (35), and the protein was purified to homogeneity by a combination of ion-exchange and molecular sieve chromatographies. β proteins expressed by other GBS strains were recovered after incubation of the bacteria at an elevated pH (35) and used without further purification. Human polyclonal IgA was from Cappel-Organon Teknika (Turnhout, Belgium). Protein G was from Amersham Pharmacia Biotech (Upplands, Sweden). Antibodies against the XPZ region were raised by immunizing a rabbit with the (XPZ)3-cys peptide conjugated to the carrier protein bovine serum albumin via the C-terminal cysteine (Neosystem, Strasbourg, France). The antisera was raised by immunizing a rabbit subcutaneously on the back with 200 μg of conjugate, with complete Freund's adjuvant for the first immunization and incomplete Freund's adjuvant for three boosters, which were given at monthly intervals. An antisera directed against the N-terminal B6 fragment of β, corresponding to the first 434 amino acid residues of the mature protein, has been described (19).
in all XPZ regions. As indicated in Fig. 1B, the different protein sequences may all be derived from the shortest sequence by the addition of blocks containing multiples of six amino acid residues, i.e., by the addition of multiples of six codons at the DNA level.

To analyze whether the XPZ region also varies in size within a bacterial strain, strains A909 and H36B were grown in 10 ml of broth and diluted 100-fold for 10 consecutive days, corresponding to ~66 generations of growth. For each of the final cultures, 15 colonies were picked, and the length of the XPZ region was analyzed by PCR. No size variation was seen in this analysis, which would have detected a change in length corresponding to six codons. Thus, the XPZ region appears to be stable within a strain under laboratory conditions.

The XPZ region is not required for surface expression of β protein. For some surface proteins of streptococci and staphylococci, it has been suggested that the proline-rich region threads through the peptidoglycan layer (16, 18, 45), suggesting that this region is important for surface localization of the protein. Moreover, it has been reported that a proline-rich region in the P1 adhesin of Streptococcus mutans is required for surface expression (12). This situation made it of interest to analyze whether the XPZ region was needed for surface expression of β. For this purpose, we constructed a bac gene with an in-frame deletion covering the entire XPZ region, and a plasmid carrying this construct (plasmid pLZΔxpz) was used to transform a GBS strain (the Δbac strain) in which the bac gene had been deleted, generating the Δbac/pLZΔxpz strain. Surface expression of β was analyzed in this strain and in a strain expressing the wild-type β protein (the Δbac/pLZbac strain), with the Δbac strain as the negative control (Fig. 2A).

For this analysis, we used an antiserum directed against the N-terminal B6 fragment of β, which does not include the XPZ region (Fig. 1B). This antiserum reacted equally well with the strain expressing the wild-type β protein and that expressing the XPZ deletion variant of β (designated βΔXPZ) but did not react with the β-negative Δbac strain (Fig. 2A). Moreover, the binding of IgA-Fc was only slightly reduced for the strain expressing βΔXPZ compared to the strain expressing the wild-type protein (Fig. 2B). These data indicate that the XPZ region is not required for surface expression of the β protein and that it is not required for IgA-Fc binding to surface-expressed β.

The XPZ region is exposed on the bacterial cell surface. An antiserum directed against a synthetic peptide derived from the XPZ region (Fig. 1B) was used to analyze whether this region is exposed on the bacterial surface. The specificity of this antiserum was first studied by Western blot analysis of purified preparations of the β protein and the β variant lacking the XPZ region, the βΔXPZ protein (Fig. 3A). Both proteins reacted with antibodies to the N-terminal B6 fragment of β, but the anti-XPZ serum reacted only with the wild-type β protein. Thus, the anti-XPZ serum could be used to specifically detect the XPZ region.

The anti-XPZ serum reacted with bacteria expressing the β protein but not with the isogenic β-negative Δbac mutant (Fig. 3B), indicating that the XPZ region is at least partially exposed on the bacterial surface and is not buried in the cell wall. As expected, the anti-XPZ serum did not react with a GBS strain expressing the βΔXPZ protein. Although these data indicated that the XPZ region is expressed on the surface of GBS, it seemed possible that the result was due to β protein that had been released from the bacteria, followed by (nonspecific) readsoption to the surface. A control experiment was therefore performed in which a suspension of the β-negative Δbac mutant (10⁹ CFU/ml) was incubated for 30 min at 23°C in the

![FIG. 2. A mutant of the β protein lacking the XPZ region is expressed on the bacterial surface. (A) Surface expression of β was analyzed in the Δbac strain, in which the chromosomal bac gene (encoding β) has been deleted, in the Δbac strain transcomplemented with a plasmid carrying the intact bac gene (Δbac/pLZbac), and in the Δbac strain transcomplemented with a plasmid encoding a β protein (βΔXPZ) lacking the XPZ region (Δbac/pLZΔxpz). The analysis was performed with an antiserum directed against the N-terminal B6 fragment of β, which does not include the XPZ region (19) (Fig. 1). This antiserum did not react with the β-negative strain but reacted equally well with the strain expressing the wild-type β protein and that expressing βΔXPZ, implying that the XPZ region is not required for surface expression of the β protein. (B) Binding of radiolabeled IgA to the three strains described for panel A. The data were expressed as percent of the binding of IgA to the wild-type strain.](image-url)
With that of antiserum against the XPZ region. These two antisera were first analyzed for reactivity with pure β immobilized in microtiter wells, i.e., under conditions in which all parts of the β protein should be equally accessible to antibodies. Under these conditions, the titer of the anti-XPZ serum was ~2-fold lower than that of the anti-B6 serum (Fig. 4A). When the same two antisera were analyzed for reactivity with β present on the bacterial surface, the titer of the anti-XPZ serum was also ~2-fold lower (Fig. 4B). This result shows that the N-terminal region and the XPZ region are equally accessible to antibodies whether β is present in pure form or is exposed on the surface of GBS. Finally, electron microscopic analysis with immunogold techniques showed that β is surface exposed on virtually all bacteria (data not shown). Together, these data indicate that the XPZ region is fully surface exposed on a GBS strain expressing the β protein.

**CD spectroscopy indicates that the XPZ region has PPII helix structure.** A sequence of four or more prolines in a row adopts the conformation of a left-handed PPII helix, an extended structure with exactly three residues per turn (29, 55, 59). This conformation is also known to be adopted by many regions that are rich in proline but contain other amino acids, including some regions in which every third residue is proline (59).

To analyze whether the XPZ region adopts a PPII structure, we characterized a 24-residue synthetic peptide derived from the XPZ region. This peptide, designated (XPZ)6, was studied by circular dichroism (CD) spectroscopy (Fig. 5). At 5°C, the CD spectrum of (XPZ)6 showed a deep negative peak at 200 nm (−33 deg mol$^{-1}$ m$^{-1}$). The CD spectrum of polyproline was characterized by a large negative band (−59 deg mol$^{-1}$ m$^{-1}$) at 205 nm and a small positive band (+4.6 deg mol$^{-1}$ m$^{-1}$) at 228 nm (23). However, many other peptides with the PPII helical structure display spectra very similar to the one we observed here for (XPZ)6, with a minimum at around 200 and the positive peak at 228 nm missing (2, 39, 48).

The spectral shifts relative to polyproline are often attributed to slight deviation from the ideal PPII helical structure, the presence of cis-proline residues, or formation of higher-order helical forms (4, 43). Raising the temperature to 25, 45, 65, and 85°C causes a gradual decrease in intensity of the negative band, a shift towards a higher wavelength, and the appearance of a negative shoulder at 220 to 230 nm. Similar temperature effects are seen for many other PPII structures and are due to unfolding to an unordered state. In summary, we may conclude that (XPZ)6 adopts a left-handed PPII helix that is disrupted at elevated temperature.

Because of the intriguing sequence of alternating positive and negative charges in the Z position, CD spectra were recorded as a function of salt and pH to evaluate the electrostatic contribution to the structure. Spectra recorded at 0.005, 0.15, and 0.7 M ionic strength superimposed completely (data not shown), indicating that screening of the electrostatic interactions does not affect the structure of (XPZ)6. Another way to perturb potentially stabilizing contributions from attractive electrostatic interactions between opposite charges is to go to extreme pHs. One may expect that the Z position carries alternating positive and uncharged residues at pH 1.5 and alternating negative and uncharged side chains at pH 12. The CD spectra obtained at pH 1.5 and pH 12 were, however, very
similar to that at pH 7, but with a small decrease in the intensity at 200 nm (by 8% at pH 1.5 and by 12% at pH 12) (data not shown). The lack of substantial salt and pH effects shows that ionic interactions are only minor determinants of the PPII structure of \((XPZ)_{8}\), which is hence governed by other forces.

**DISCUSSION**

The data reported here indicate that the proline-rich XPZ region of the streptococcal \(\beta\)H9252 protein is exposed on the bacterial surface and adopts the conformation of a left-handed PPII helix. Although the exact function of the XPZ region is not yet known, these findings shed new light on the properties of proline-rich regions in surface proteins of streptococci and staphylococci. In particular, our data indicate that the XPZ region is not required for cell wall anchoring and/or surface expression of the \(\beta\)H9252 protein. Rather, the surface exposure of this region and the knowledge that proline-rich regions commonly participate in protein-protein interactions (59) suggest that the XPZ region could promote interactions with host proteins or with other bacterial surface proteins. The XPZ region could also have a structural role, e.g., it might act as a scaffold that allows surface exposure of other regions of the \(\beta\)H9252 protein. However, our data do not support this hypothesis, because deletion of the XPZ region had little or no effect on surface exposure of the N-terminal IgA-binding part of \(\beta\)H9252.

The conclusion that the XPZ region adopts the conformation of a PPII helix was based on CD analysis of the \((XPZ)_{8}\) peptide, which exhibits the characteristic sequence periodicity. Because this periodicity is conserved throughout the XPZ region, it seems likely that the intact XPZ region is a PPII helix. This extended and solvent-exposed conformation, which has been studied extensively due to its importance for protein-protein interactions (1, 29), may even occur in regions lacking proline (1). The importance of this conformation is underlined by its presence in each of the three strands of collagen (7, 59) and by the finding that class II major histocompatibility complex molecules bind peptides in a PPII-like conformation (22, 42).

Because there are exactly three residues per turn in a PPII helix (29, 59) and because the XPZ region exhibits three-
residue periodicity, the helix corresponding to the XPZ region is expected to be amphipathic and to have three surfaces with very different properties, corresponding to the X, P, and Z residues. The X surface is mainly hydrophobic, although with a fair number of hydroxyl groups. The proline rings will line up to form an extended hydrophobic surface, while the Z residues will form a hydrophilic stretch of alternating charges. This separation of surface properties may allow the XPZ region to interact simultaneously with molecules having binding sites of very different characters. Possibly, the alternating charges of the third residue could also contribute to dimerization of two staggered XPZ regions by a zipper-like mechanism (46).

The XPZ region varies in size between different β proteins and is present in all β proteins analyzed, and the sequence periodicity is strictly maintained in XPZ regions of different lengths. These data suggest that the XPZ region is important for the function of the β protein, and they are in good agreement with recent studies that employed PCR to demonstrate size variation in the XPZ region (8, 30). The mechanisms that cause size variation in the XPZ region are not known, and it is not clear whether the mechanisms commonly used to explain size variability in repetitive sequences (10, 14) can also explain size variation in the XPZ region, because this region has a periodicity rather than a repetitive sequence. Possibly, the size variability arises through an error-prone process (41), causing duplications and deletions of multiples of six codons, followed by a selection that eliminates mutants that do not retain the typical periodicity.

In this context, it is of interest that three of the four proline codons are used for the central residue in the XPZ motif, suggesting that the XPZ region is indeed subject to a high mutation rate and to a selection process responsible for conservation of the periodicity (33). The selection of XPZ variants that retain the sequence periodicity may reflect a structural requirement, while the sequence variability could reflect immunological selection of antigenically different variants, a hypothesis implying that sequence variability in the XPZ region allows GBS to escape host immunity.

In summary, we have identified a PPII helix with a highly periodic sequence that is exposed on the surface of GBS. To our knowledge, a surface structure of this type has not previously been identified in bacteria. Finally, we note that “natural” antibodies directed against self-antigens preferentially recognize epitopes rich in proline and were suggested to provide a first line of defense against pathogens expressing proline-rich surface proteins (56). Conversely, it seems possible that surface-exposed proline-rich regions in pathogens may contribute to the emergence of natural antibodies or even to the development of autoimmunity. This argument is particularly relevant with regard to the XPZ region studied here, because the PPII helix structure of this region may favor presentation by major histocompatibility complex class II molecules (22, 42).

ACKNOWLEDGMENTS

We are grateful to Lars Stenberg and Charlotte Bergquist for contributions in the early stages of this study and to Erik Lindahl for advice. We also thank the students in the Prokaryotic Molecular Genetics course at Lund University for help with sequence analysis.

This work was supported by grants from the Swedish Research Council (grant 09490 to G.L. and grant 11552 to S.L.), the Medical Faculty of Lund University, the Royal Physiographic Society in Lund, the Swedish Society for Medical Research, and the Golje, Groschinsky, Hedberg, Hierta, Jerring, Kock, Lundström, and Österlund foundations.

REFERENCES


