Volume 235, number 1,2, 262-266 FEB 06150 August 1988

# Complete amino acid sequence of protein B

J. Rühlmann, B. Wittmann-Liebold*, D. Jürgens and F.J. Fehrenbach

*Robert Koch-Institut des Bundesgesundheitsamtes, Department of Microbiology, Nordufer 20, D-1000 Berlin 65* and **Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, Ihnestr. 73, D-1000 Berlin 33, FRG*

Received 3 June 1988

The complete amino acid sequence of protein B (=CAMP factor) of *Streptococcus agalactiae* has been determined. The sequence data were obtained mainly by manual sequencing of peptides derived from digestion with lysyl-peptidase, clostripain and *Staphylococcus aureus* protease and by solid phase sequencing of cyanogen bromide fragments. The protein contains 226 amino acids and has an $M_r$ of 25 263. The sequence was compared with sequences of other Fc-binding proteins and partial sequence homology was found between protein B and the Fc-binding region of protein A.

Protein B; CAMP factor; Fc-binding protein; Amino acid sequence; (*Streptococcus agalactiae*)

## 1. INTRODUCTION

Protein B (=CAMP factor) [1] is produced by most strains of *Streptococcus agalactiae* (group B streptococci) and was described first by Christie et al. [2] in 1944. The authors demonstrated that the 'CAMP factor' lyses in a synergistic way sphingomyelinase-treated sheep erythrocytes. Protein B has been purified recently and characterized as polypeptide with an apparent molecular mass of 25 kDa and an isoelectric point of 8.9 [3].

The interaction of protein B with membranes has been studied in detail mainly with the effort to characterize its membrane damaging potential [4–6]. Protein B is released from growing cultures [7] together with four satellite proteins, which are very similar in their isoelectric points, molecular masses and hydrophobicity [3]. Interestingly, protein B is the only one active in synergistic hemolysis.

Additionally a recent study has shown that protein B belongs to the group of bacterial Fc-binding proteins [1] such as protein A of *Staphylococcus aureus* and the Fc-receptor proteins of a variety of streptococci [8]. Unlike protein A [9] and the Fc-binding proteins of streptococci of group A, C and G [10] protein B remains not cell wall associated, but is released to a large extend into the supernatant during mid to late logarithmic growth phase [11].

In respect to the possible role of protein B in streptococcal pathogenicity the newly described Fc-binding property may be more important than the well known co-cytolytic activity of the CAMP factor which seems to be rather an epiphenomenon with little relevance in vivo.

The complete amino acid sequence of protein B, presented in this paper, has been determined to obtain information on the structure/function relationship of the protein and to compare the sequence with those of other Fc-binding proteins.

*Correspondence address:* J. Rühlmann, Robert Koch-Institut des Bundesgesundheitsamtes, Dept of Microbiology, Nordufer 20, D-1000 Berlin 65, FRG

## 2. MATERIALS AND METHODS

### 2.1. *Purification of protein B*

Protein B was purified from the culture supernatant of group B streptococci as described by Jürgens et al. [3].

### 2.2. *Sequence determination*

Sequence studies were performed with peptides prepared by digestion of intact protein B or subdigestion of protein B fragments with the following enzymes: clostripain (Sigma), lysyl-peptidase (Wako, Tokyo), *Staphylococcus aureus* pro-



*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/88/$3.50 

tease (Miles) and thermolysin (Boehringer-Mannheim). In addition, peptides obtained by cleavage with cyanogen bromide (CNBr) were used for sequence analysis.

Peptides resulting from CNBr cleavage were isolated by gel filtration on a TSK 2000 SW column (7.5 × 600 mm) at a flow rate of 0.2 ml/min. Peptides were applied to the column in 6 M guanidinium hydrochloride in 0.1% trifluoroacetic acid and detected by their absorbance at 206 nm. All other peptides were isolated by reversed-phase HPLC on a Nucleosil 300 5 $\mu$m C18 column (4 × 250 mm) using different gradients of 0–80% acetonitrile in 0.1% trifluoroacetic acid. Peptides were eluted at a flow rate of 0.5 ml/min and detected by their absorbance at 230 nm. For amino acid analysis peptides were hydrolysed in 5.7 M HCl at 110°C for 20 h and the amino acid composition was determined on a HPLC system using *o*-phthaldialdehyde as derivatizing reagent [12].

Peptides were sequenced manually by the DABITC/PITC (4′-*N,N*′-dimethylaminoazobenzene-4′-isothiocyanate/phenylisothiocyanate) double-coupling method [13]. Larger peptides were also sequenced automatically with a solid-phase sequencer (LKB 4020) as described [14]. For determination of the C-terminal amino acids intact protein B was digested with carboxypeptidase P (Boehringer-Mannheim) and the released amino acids were quantitated by amino acid analysis.

### 2.3. *Computer analysis*

Hydrophobicity analysis was performed with the program HYDROP (M. Dzionara and B. Wittmann-Liebold, unpublished) using the parameters of Kyte and Doolittle [15]. The amino acid sequence of protein B was compared with the sequences of Fc-binding proteins (protein A [16], protein G [17]) with the program ALIGN [18] on a VAX/VMX computer.

## 3. RESULTS AND DISCUSSION

### 3.1. *Sequence determination*

The complete amino acid sequence of protein B is presented in fig. 1. The sequence was derived mainly by sequence analysis of the CNBr fragments and the fragments obtained by cleavage with lysyl-peptidase. These peptides were aligned by sequencing overlapping peptides from *S. aureus* protease and clostripain digestion.

Sequencing of the N-terminal region was difficult due to the accumulation of asparagine and glutamine residues and the absence of cleavage sites for more specific proteases, but was finally achieved by sequence analysis of smaller peptides produced by subdigestion of the N-terminal CNBr fragment with thermolysin.

The C-terminal sequence of protein B was derived from sequence analysis of the C-terminal *S. aureus* protease peptide and from results of carboxypeptidase P digestion.

### 3.2. *Characteristics of the sequence*

Protein B contains 226 amino acids and has a calculated $M_r$ of 25263. This value is in good agreement with the apparent molecular masses of 25 kDa and 23.5 kDa determined by Jürgens [3] and Bernheimer [5], respectively. The amino acid composition as derived from the sequence is shown in table 1. With the exception of cysteine all common amino acids were present and no unusual amino acids were found. The relative frequency of the amino acids detected either by amino acid analysis [5,19] or derived from the sequence is in good agreement. The small quantitative discrepancies found for individual amino acids may be ex-

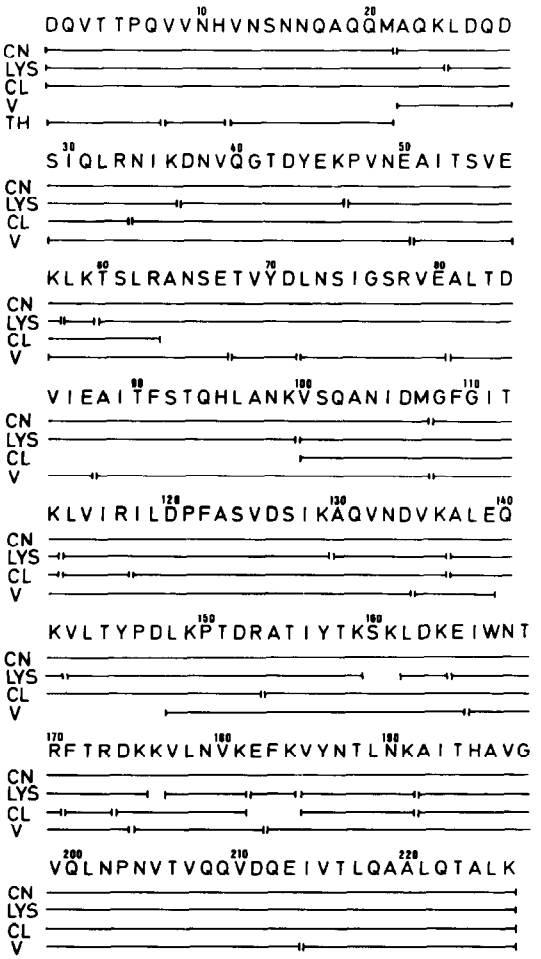


Fig. 1. Amino acid sequence of protein B. CN, LYS, CL, V and TH indicate peptides derived from cleavage with cyanogen bromide, lysyl-peptidase, clostripain, *Staphylococcus aureus* V8 protease and thermolysin, respectively.

Table 1

Amino acid composition of protein B

| Amino acid | Residues in protein B | | |
|---|---|---|---|
| | Calculated from | | From sequence |
| | [19] | [5] | |
| Asx | 28 | 33 | 34 |
| Glx | 29 | 29 | 29 |
| Val | 29 | 24 | 25 |
| Thr | 18 | 16 | 21 |
| Lys | 23 | 18 | 20 |
| Leu | 19 | 17 | 19 |
| Ala | 20 | 20 | 17 |
| Ile | 13 | 18 | 15 |
| Ser | 10 | 11 | 12 |
| Arg | 6 | 4 | 7 |
| Pro | 6 | 6 | 6 |
| Phe | 5 | 6 | 5 |
| Gly | 12 | 10 | 5 |
| Tyr | 4 | 8 | 5 |
| His | 3 | 2 | 3 |
| Met | 1 | 4 | 2 |
| Trp | n.d. | 2 | 1 |

Calculated values were obtained from [19] and [5] assuming an $M_r$ of 25263

plained by differences in the degree of purification or by using different producer strains for protein B production.

Protein B contains 25% charged residues and has a net charge of +2.5 assuming that histidine carriers a charge of +0.5 [20]. The polarity index according to Vanderkooi and Capaldi [21] is 50%, which is close to the mean value of soluble proteins (46 ± 6%). The hydrophobic residues are distributed equally throughout the protein. This is also documented by the hydrophobicity profile (fig.2) which shows the absence of pronounced hydrophobic domains. The sequence data reported here do not support the view that the 9 kDa CNBr fragment (pos. 22–107) in contrast to the 13.5 kDa CNBr fragment (pos. 108–226) may represent the 'membrane-binding domain' of protein B [22]. Extended hydrophobic regions, possibly responsible for membrane binding, have not yet been detected in protein B and were also found to be absent in other bacterial cytolytic toxins [23,24]. However, hydrophobic domains may result from protein folding in a way not predictable from sequence data. On the other hand binding of the above mentioned membrane damaging proteins to a target may be mediated by restricted hydrophobic regions not clustered in an individual domain. The Fc-binding regions of proteins A and G consist of several extended repeating sequences. Protein A contains five homologous Fc-binding domains with sizes ranging from 50 to 61 amino acids [25], whereas protein G contains two or three homologous Fc-binding domains, each with a size of 55 amino acids [17,26]. Although the Fc-binding domains of these proteins are approximately of the same size, sequence comparison does not reveal any homology.

In contrast to the extended repeats (50–61 amino acids) found in proteins A and G, protein B contains two repeating sequences of 5 and 11 amino acids only.

$^{48}$Val-Asn-Glu-Ala-Ile-Thr
$^{79}$Val-....-Glu-Ale-Leu-Thr
$^{85}$Val-Ile-Glu-Ala-Ile-Thr
$^{158}$Thr-Lys-Ser-Lys-Leu-Asp-Lys-Gln-Ile-Trp-Asn
$^{169}$Thr-Arg-Phe-Thr-Arg-Asp-Lys-Lys-Val-Leu-Asn
$^{180}$Val-Lys-Glu-Phe-Lys-Val-Tyr-Asn-Thr-Leu-Asn

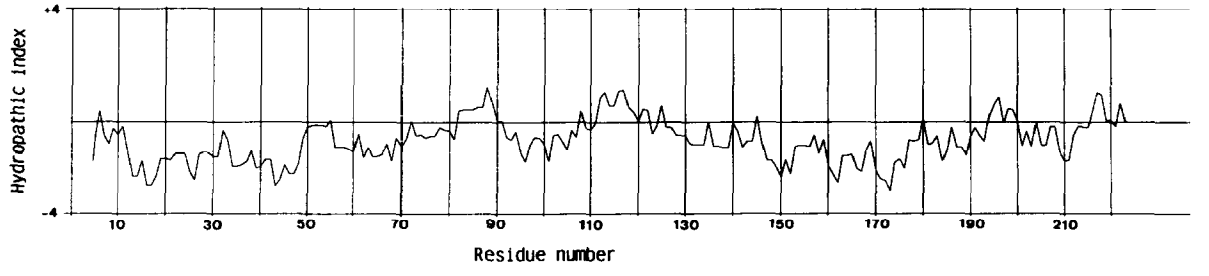


Fig. 2. Hydrophobicity analysis of protein B using the HYDROP program (see text). Hydrophobic regions appear as positive peaks of hydropathicity.

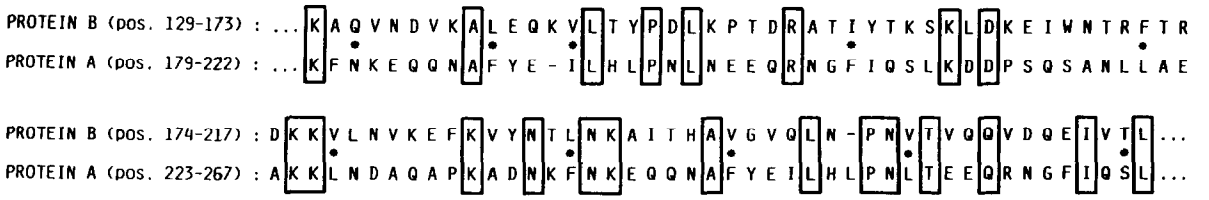


Fig. 3. Comparison of partial sequences of proteins B and A. Identical residues in both sequences have been boxed, conservative replacements have been marked with asterisks (*). Two gaps have been introduced for better alignment.

However, it is not yet known whether the repeats in protein B are of any importance for Fc binding. The complete amino acid sequence of protein B was compared with the sequences of the Fc-binding domains of proteins A and G in order to identify regions which may be responsible for Fc binding of protein B. According to the statistical analysis (program ALIGN [18]) no homology has been found between protein B and the Fc-binding domains of protein G. However, comparison of a partial sequence of protein B (pos. 129–217) with a segment of the Fc-binding domain of protein A (pos. 179–267) [16] revealed 22 identical residues and 10 conservative replacements. For this comparison two gaps were introduced for better alignment (fig.3). The 89-residue string of protein B (pos. 129–217), which exhibits 36% homology with segment 179–267 of protein A, is located in the 13,5 kDa CNBr fragment. This 13.5 kDa peptide has been shown to bind to immunoglobulins whereas no binding was found with the 9 kDa CNBR fragment [27]. Although sequence homology between the Fc-binding region of protein A and protein B exists it is unclear as yet whether the same set of amino acids is responsible for binding [28].

Protein B is an extracellular protein and released into the medium during mid to late logarithmic growth phase [11]. In contrast proteins A and G remain cell wall associated, but can be released by treatment with proteases [9,29]. Cell wall attachment of proteins A and G is supposed to be mediated by a hydrophobic C-terminal sequence, which is preceded by an extremely hydrophilic proline-rich region. These 'membrane anchor sequences' of proteins A and G as well as the C-terminal sequence of streptococcal M 6 protein, which is also cell wall associated, exhibit a high degree of homology [17]. As anticipated for protein B which is released into the medium a comparable anchor sequence has not been found.

Thus, the most important structural differences between protein B and proteins A and G are the absence of extended internal repeats and the lack of a membrane anchor sequence. Furthermore, protein B has been characterized as a protein active in synergistic hemolysis [4–6], a function which is not present in either protein A or G.

Acknowledgements: We thank Dr R. Kamp for helpful discussions and Mrs I Housirek for excellent technical assistance. We would also like to express our appreciation of the kind help of Mr D. Kamp and Mr A. Lehmann in carrying out the amino acid analysis and solid-phase sequencing, respectively. This work was supported by the Deutsche Forschungsgemeinschaft, Sfb 9/D2, Technische Universität Berlin.

## REFERENCES

[1] Jürgens, D., Sterzik, B. and Fehrenbach, F.J. (1987) J. Exp. Med. 165, 720–732.
[2] Christie, R., Atkins, N.E. and Munch-Petersen, E. (1944) Aust. J. Exp. Biol. Med. Sci. 22, 197–200.
[3] Jürgens, D., Shalaby, Y.I. and Fehrenbach, F.J. (1985) J. Chromatogr. 348, 363–370.
[4] Doery, H.M., Magnusson, B.J., Cheyne, I.M. and Gulasekharam, J. (1963) Nature 198, 1091–1092.
[5] Bernheimer, A.W., Linder, R. and Avigad, L.S. (1979) Infect. Immun. 23, 838–344.
[6] Fehrenbach, F.J., Schmidt, C.-M., Sterzik, B. and Jürgens, D. (1984) in: Bacterial Protein Toxins (Alouf, J.E. et al. eds) pp. 317–324, Academic Press, London.
[7] Huser, H., Goeke, L., Karst, G. and Fehrenbach, F.J. (1983) J. Gen. Microbiol. 129, 1295–1300.
[8] Boyle, M.D.P. and Reis, K.J. (1987) Biotechnol. 5, 697–703.
[9] Sjöquist, J., Meloun, B. and Hjelm, H. (1972) Eur. J. Biochem. 29, 572–578.
[10] Kronvall, G. (1973) J. Immunol. 111, 1401–1406.
[11] Takaisi, K. and Fehrenbach, F.J. (1988) in: Bacterial Protein Toxins (Fehrenbach, F.J. et al. eds) Zentralbl. Bakteriol. Mikrobiol. Hyg. I. Abt., Suppl. 17, pp. 165–166, Gustav Fischer Stuttgart.

[12] Ashman, K. and Bosserhof, A. (1985) in: Modern Methods in Analytical Protein Chemistry (Tschesche, H. ed.) pp. 155–171. De Gruyter, Berlin.
[13] Chang, J.Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[14] Wittmann-Liebold, B. and Lehmann, A. (1980) in: Modern Methods in Peptide Sequence Analysis (Birr, C. ed.) pp. 49–72, Elsevier, Amsterdam.
[15] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[16] Uhlén, M., Guss, B., Nilsson, G., Gatenbeck, S., Philipson, L. and Lindberg, M. (1984) J. Biol. Chem. 259, 1695–1702 (Authors correction 259, 13628).
[17] Fahnestock, S.R., Alexander, P., Nagle, J. and Filpula, D. (1986) J. Bacteriol. 167, 870–880.
[18] Dayhoff, M.O. (1978) in: Atlas of Protein Sequence and Structure, Vol. 5, Suppl. 3, National Biomedical Research Foundation, Washington, DC.
[19] Jürgens, D., Huser, H. and Fehrenbach, F.J. (1984) in: Bacterial Protein Toxins (Alouf, J.E. et al. eds) pp. 245–246, Academic Press, London.
[20] Offord, R.E. (1966) Nature 211, 591–593.
[21] Vanderkooi, G. and Capaldi, R.A. (1972) Ann. NY Acad. Sci. 195, 135–138.
[22] Sterzik, B., Jürgens, D., and Fehrenbach, F.J. (1985) in: Bacterial Protein Toxins (Falmagne, P. et al. eds) Zentralbl. Bakteriol. Mikrobiol. Hyg. I. Abt., Supp. 15, pp. 101–108, Gustav Fischer Stuttgart.
[23] Gray, G.S. and Kehoe, M. (1984) Infect. Immun. 46, 615–618.
[24] Kehoe, M.A., Miller, L., Walker, J.A. and Boulnois, G.J. (1987) Infect. Immun. 55, 3228–3232.
[25] Moks, T., Abrahmsén, L., Nilsson, B., Hellmann, U., Sjöquist, I. and Uhlén, M. (1986) Eur. J. Biochem. 156, 637–643.
[26] Guss, B., Eliasson, M., Olsson, A., Uhlén, M., Frej, A.-K., Jörnvall, H., Flock, J.I. and Lindberg, M. (1986) EMBO J. 5, 1567–1575.
[27] Fehrenbach, F.J., Jürgens, D., Rühlmann, J., Sterzik, B. and Özel, M. (1988) in: Bacterial Protein Toxins (Fehrenbach, F.J. et al. eds) Zentralbl. Bakteriol. Mikrobiol. Hyg. I. Abt., Suppl. 17, pp. 351–357, Gustav Fischer, Stuttgart.
[28] Deisenhofer, J. (1981) Biochemistry 20, 2361–2370.
[29] Björck, L. and Kronvall, G. (1984) J. Immunol. 133, 969–974.