Amino acid sequence of pilin from *Bacteroides nodosus* (strain 198), the causative organism of ovine footrot

Neil M. McKern, Ian J. O'Donnell, Adam S. Inglis, David J. Stewart[+] and Bruce L. Clark[+]

CSIRO Division of Protein Chemistry, 343 Royal Parade and [+]*CSIRO Division of Animal Health, Park Drive, Parkville, 3052, Victoria, Australia*

Received 7 October 1983

Abstract and keywords not received

1. INTRODUCTION

Bacteroides nodosus is the organism responsible for transmitting footrot in sheep grazing on wet pastures [1–3]. Vaccines prepared against homologous serotypes are highly effective in both preventing and curing the disease in immunised sheep [4,5]. In common with other gram-negative bacteria, *B. nodosus* possesses flexible filaments known as pili which have a diameter of 5–6 nm and are up to 25 μm in length [6–8]. Each pilus is composed of identical subunit pilin proteins of M_r ~18 400 [8]. Although pilus vaccines are highly effective against infection with homologous strains of bacteria [9–12], a vaccine is required which will protect against all 8 pilus serogroups recognised in Australia to date [13]. This work aims at providing structural information that will lead to the development of a cheap and highly effective vaccine against footrot. The current strategy is based on the premise that there exists a region of amino acid sequence in pilin common to the 8 serogroups, which could be used to make a peptide vaccine that is host protective. A similar approach is being tried in the search for an effective vaccine against *Neisseria gonorrhoeae* [14]. We here report the complete amino acid sequence of pilin protein from *B. nodosus* and compare it with those of some other bacteria whose pilin proteins have been completely or partially sequenced.

2. EXPERIMENTAL

Pili from strain 198 (CSIRO Animal Health culture collection) of *B. nodosus* were isolated as in [15] and were further purified by dissolving the pili in 98–100% formic acid and passing the solution through a Biogel P6 column, using 98–100% formic acid as eluent. The pili subunits were shown to be homogeneous by SDS–polyacrylamide gel electrophoresis. The protein was reduced and carboxymethylated with iodo[^{14}C]acetic acid as in [16]. Enzymic digestion with TPCK-trypsin, chymotrypsin and V8 protease from *Staphlococcus aureus* was carried out in 0.05 M ammonium bicarbonate at 37°C overnight at an enzyme:substrate ratio of 1–50 (w/w). Thermolytic digestion was carried out under similar conditions, but with 5 mM $CaCl_2$ included in the buffer. Cleavage with 2% formic acid was performed at 108°C for 2 h at pH 2 [17]. Insoluble material contained in digest solutions after incubation was removed by centrifugation and soluble peptides isolated by HPLC in 0.05% trifluoroacetic acid on a Waters μBondapak C_{18} column, using an acetonitrile gradient. Some peptides (designated *) were purified by high voltage paper electrophoresis and subsequent chromatography in *n*-butanol–acetic acid–water–pyridine (15:3:12:10), by vol.).

Amino terminal sequencing of undigested SCM-pilin was carried out using an automated sequencer

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

as in [18]. Peptides were manually sequenced by a modified Edman procedure [19], using 50% pyridine as the coupling buffer and extracting with *n*-heptane–ethyl acetate (2:1, v/v) instead of benzene [20]. Residues were identified as PTH-amino acid derivatives by HPLC gradient elution on a Zorbax C_{18} column (Du Pont), or as the dansyl derivatives by two-dimensional chromatography [21]. The carboxyl terminal was tritium-labelled as in [22].

3. RESULTS AND DISCUSSION

The complete amino acid sequence of pilin from strain 198 of *B. nodosus* is shown in fig.1. The protein contains 151 amino acid residues, corresponding to an M_r in the unreduced form of 16290. Sequence data necessary for the unequivocal determination of the primary structure are illustrated in fig.1. The N-terminal amino acid of pilin was identified as *N*-methylphenylalanine by co-chromatography of the dansyl derivatives and this identification was confirmed by the finding that PTH-derivatives of the N-terminal amino acid and *N*-methylphenylalanine had the same retention time on reversed-phase HPLC. It was found that about half of the pilin molecules lacked the *N*-methylphenylalanine residue, as has been observed for pilin from *N. gonorrhoeae* [23]. Automatic sequenator analysis of intact protein provided sequence data up to residue 47, residues 20–47 being confirmed by the sequences of peptides C1, C2, T1 and SA2. The overlap peptide SA4 connected peptide T3 to T4 and the sequence was extended to residue 98 by peptide C5 overlapping with peptide T5. Peptide T6 was aligned by the

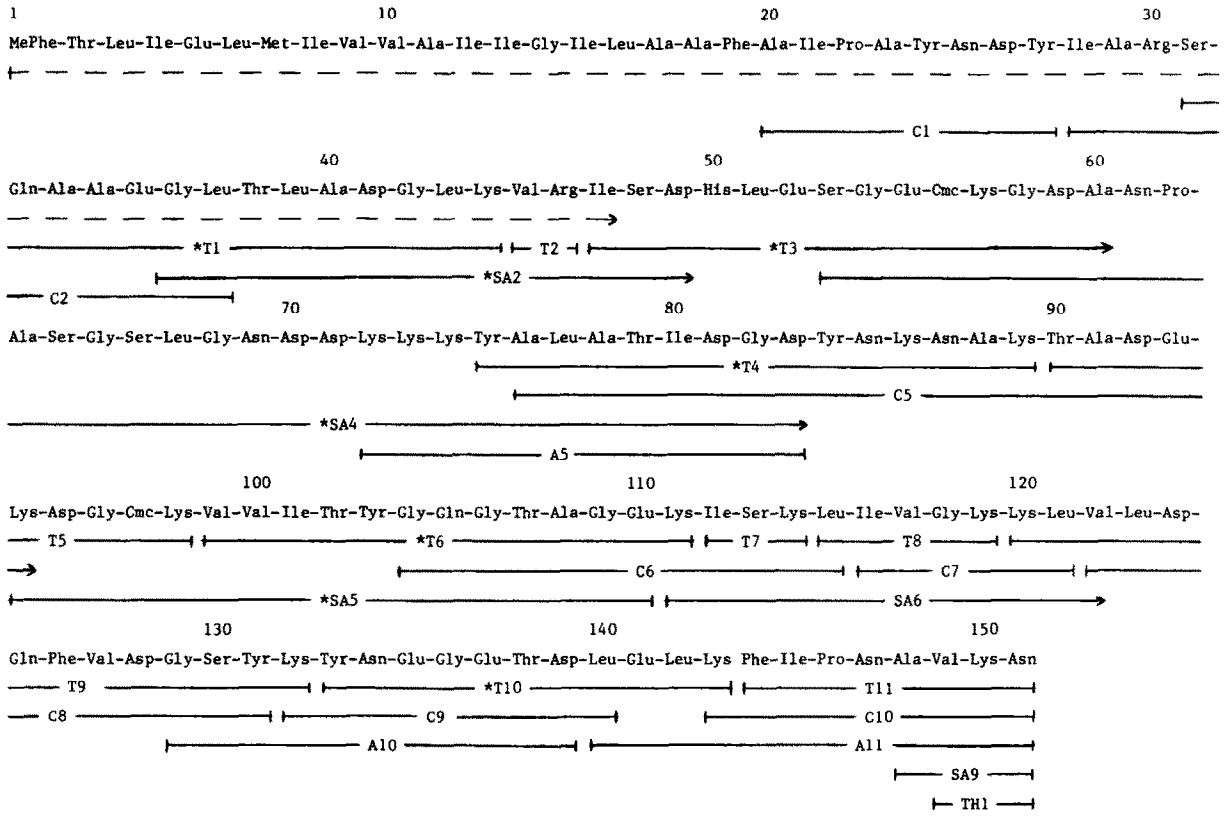

Fig.1. Primary structure of pilin from *B. nodosus* strain 198; automated sequencing indicated by dashed line, manual sequencing by solid lines; T, tryptic peptides; C, chymotryptic peptides; SA, *Staphlococcus aureus* V8 protease peptides; A, peptides from dilute acid cleavage; TH, thermolytic peptide. Only those peptides necessary for the elucidation of the sequence are shown. Peptides not sequenced to completion are shown by →.

sequence of the peptide SA5, peptides C6 and SA6 extending the known structure to residue 122.

Difficulty was encountered in sequencing manually past the Asp–Gly bond (residues 128 and 129) in peptides T9 and C8. Problems with Asp–Gly bonds in other proteins have been reported before [24–26] and were alleviated in an automated sequencing procedure by reducing the time of trifluoroacetic acid cleavage [25]. A similar modification to the manual technique used here enabled the determination of the complete sequences of peptides T9 and C8. None of the other three Asp–Gly bonds in *B. nodosus* pilin presented any sequence difficulties.

The sequence was extended to residue 143 by the overlap of peptide A10 with T9 and T10, the overlap peptides A11 and C10 then continuing the sequence to residue 151. Evidence that the C-terminal residue of the protein is Asn 151 is provided by 3 sets of data. First, 4 different enzyme digests, as well as dilute acid cleavage of the protein, contained peptides ending at Asn 151. Second, the C-terminal amino acid was identified as either asparagine or aspartic acid by selective tritiation [22], and third, upon digestion of the protein with carboxypeptidase Y, asparagine was the first amino acid released.

Table 1

Amino acid composition of pilin from *B. nodosus* strain 198

Amino acid	Number of residues per subunit	
	Analysis[a]	Sequence
CMC	1.7	2
Asx	21.5	21
Thr	7.2	7
Ser	7.3	7
Glx	13.0	12
Pro	4.2	3
Gly	17.3	15
Ala	17.3	17
Val	7.9	9
Met	0.5	1
Ile	9.5	13
Leu	13.3	14
Tyr	7.6	7
Phe	3.7	3
Lys	15.5	16
His	1.4	1
Arg	2.5	2
Total	151.4	151

[a] Samples hydrolysed for 18 h at 110°C under vacuum in constant-boiling HCl containing 0.01% phenol

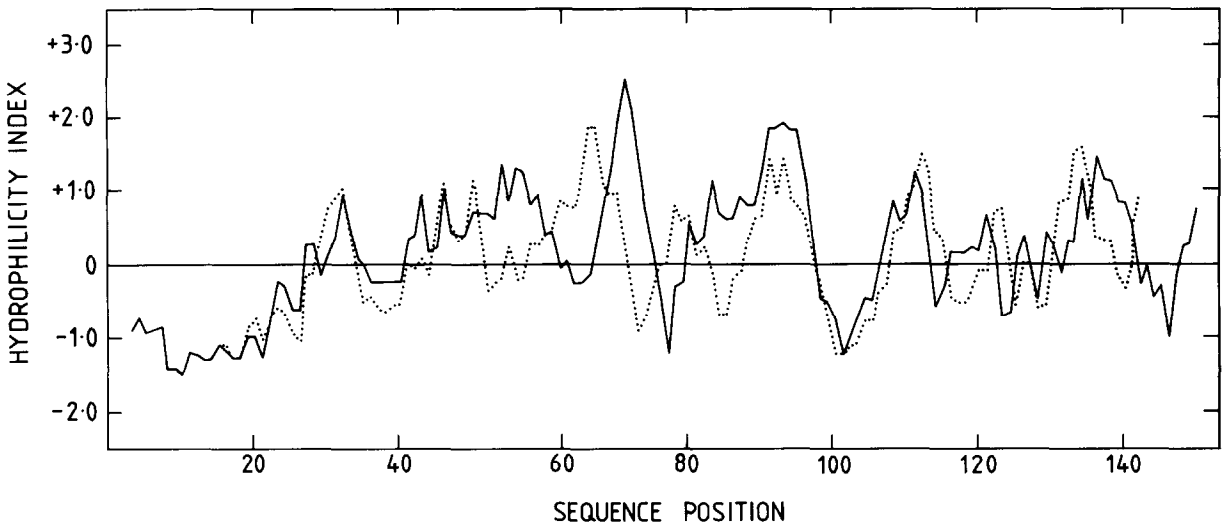

Fig.2. Profile of the hydrophilicity index for *B. nodosus* strain 198 pilin. Hydrophilicity values of each residue were averaged sequentially over spans of 6 residues along the protein [31] and these values plotted at the midpoint of the six residues averaged. *N*-Methylphenylalanine was assigned the hydrophilicity value of phenylalanine. For comparison, the profile of the hydrophilicity index for *P. aeruginosa* pilin is shown by the dashed line [30].

The sequence obtained comprises satisfactorily with the amino acid analysis of the protein hydrolysate (table 1), except for the low values obtained for valine and particularly for isoleucine. The sequence, however, contains several peptide bonds with either valine or isoleucine residues in juxtaposition (residues 8–10, 12–13, 99–101, 116–117) and these bonds are known to be resistant to acid hydrolysis [27,28].

Pilin from strain 198 of *B. nodosus* contains 13 aspartic acid, 9 glutamic acid, 16 lysine, 1 histidine and 2 arginine residues, indicating that the protein is acidic in nature. There are two half-cystine residues per molecule and these do not form an intramolecular disulphide bond, since SCM-pilin and unreduced pilin showed similar mobility when electrophoresed in SDS–polyacrylamide gels (thus indicating the absence of intermolecular bridges in the unreduced protein).

The N-terminal 24-residue sequence is very hydrophobic and is highly homologous with that of pilin from *N. gonorrhoeae, Moraxella nonliquefaciens* and *Pseudomonas aeruginosa* (table 1). This highly conserved hydrophobic region may facilitate transport of the protein through cell membranes and promote interactions between the subunits to form filamentous structures [23,29,30]. Considerable homology between the pilin sequences is evident up to about residue 44, but is infrequent for the remaining portions of sequences depicted in table 1.

When a hydrophilicity index [31] was constructed from the sequence of strain 198 of *B. nodosus* using average values along spans of 6 residues, the profile shown in fig.2 was obtained. As expected, the N-terminal region exhibits a negative hydrophilicity profile, very similar to that obtained for *P. aeruginosa* PAK pilin [30]. The remainder of the *B. nodosus* profile contains a number of hydrophilic regions, at least three of which have peaks in positions close to those in *P. aeruginosa* (residues 94, 112 and 137).

It is of interest that partial [23,32,33] or complete [34,35] amino acid sequences of pilin from several strains of *Escherichia coli* show no significant homology with any regions of the molecules listed in table 2. Furthermore, the two known C-terminal sequences of *E. coli* pilin are quite hydrophobic, whereas in pilin from *B. nodosus* (and the other pilins in table 2) it is the N-terminal sequence which is particularly hydrophobic. Hence it appears that the subunit proteins comprising the

Table 2

Amino-terminal amino acid sequences of pilin from *B. nodosus* (Family Bacteroidaceae), *N. gonorrhoeae* [23] and *Moraxella nonliquefaciens* [29] (Family Neisseriaceae) and *Pseudomonas aeruginosa* [30] (Family Pseudomonadaceae)

	1				5					10					15					20
B. nodosus	MePhe	Thr	Leu	Ile	Glu	Leu	Met	Ile	Val	Val	Ala	Ile	Ile	Gly	Ile	Leu	Ala	Ala	Phe	Ala
N. gonorrhoeae	MePhe	Thr	Leu	Ile	Glu	Leu	Met	Ile	Val	Ile	Ala	Ile	Val	Gly	Ile	Leu	Ala	Ala	Val	Ala
M. nonliquefaciens	MePhe	Thr	Leu	Ile	Glu	Leu	Met	Ile	Val	Ile	Ala	Ile	Ile	Gly	Ile	Leu	Ala	Ala	Ile	Ala
P. aeruginosa	MePhe	Thr	Leu	Ile	Glu	Leu	Met	Ile	Val	Val	Ala	Ile	Ile	Gly	Ile	Leu	Ala	Ala	Ile	Ala

	25					30							35							40
B. nodosus	Ile	Pro	Ala	Tyr	Asn	Asp	Tyr	Ile	Ala	Arg	Ser	Gln	Ala	Ala	Glu	Gly	Leu	Thr	Leu	Ala
N. gonorrhoeae	Leu	Pro	Ala	Tyr	Gln	Asp	Tyr	Thr	Ala	Arg	Ala	Gln	Val	Ser	Glu	Ala	Ile	Leu	Leu	Ala
M. nonliquefaciens	Leu	Pro	Ala	Tyr	Gln	Asp	Tyr	Ile	Ile	Arg	Ala	Gln	Val	Ser	Glu	Ala	Phe	Thr	Leu	Ala
P. aeruginosa	Ile	Pro	Gln	Tyr	Gln	Asn	Tyr	Val	Ala	Arg	Ser	Glu	Gly	Ala	Ser	Ala	Leu	Ala	Ser	Val

	45					50							55							60
B. nodosus	Asp	Gly	Leu	Lys	Val	Arg	Ile	Ser	Asp	His	Leu	Glu	Ser	Gly	Glu	CMC	Lys	Gly	Asp	Ala
N. gonorrhoeae	Glu	Gly	Gln	Lys	Ser	Ala	Val	Thr	Glu	Tyr	Tyr	Leu	Asn	His	Gly	Lys	Trp	Pro	Glu	
M. nonliquefaciens	Asp	Gly	Leu	Lys	Thr	Gly	Ile	Ser	Thr											
P. aeruginosa	Asn	Pro	Leu	Lys	Thr	Thr	Val	Glu	Glu	Ala	Leu	Ser	Arg	Gly	Trp	Ser	Val	Lys	Ser	Gly

Boxed residues are identical to those of pilin from *B. nodosus*

highly specialized bacterial appendage, the pilus, can consist of quite different amino acid sequences for different bacteria, but with one end necessarily being hydrophobic.

ACKNOWLEDGEMENTS

We would like to thank Dr Charles Roxburgh and Mr Neil Cook for HPLC analysis of PTH-amino acids. The expert technical skill of Miss Sue Arendshorst is greatly appreciated. We are grateful to Dr Colin Ward for helpful discussions and advice.

REFERENCES

[1] Beveridge, W.I.B. (1941) Bull. Coun. Sci. Ind. Res. Melb. No. 140.
[2] Egerton, J.R., Roberts, D.S. and Parsonson, I.M. (1969) J. Comp. Pathol. 79, 207–216.
[3] Roberts, D.S. and Egerton, J.R. (1969) J. Comp. Pathol. 79, 217–227.
[4] Egerton, J.R. and Burrell, D.H. (1970) Aust. Vet. J. 46, 517–522.
[5] Egerton, J.R. and Roberts, D.S. (1971) J. Comp. Pathol. 81, 179–185.
[6] Stewart, D.J. (1973) Res. Vet. Sci. 14, 132–134.
[7] Walker, P.D., Short, J., Thomson, R.O. and Roberts, D.S. (1973) J. Gen. Microbiol. 77, 351–361.
[8] Every, D. (1979) J. Gen. Microbiol. 115, 309–316.
[9] Stewart, D.J. (1978) Res. Vet. Sci. 24, 14–19.
[10] Stewart, D.J. (1978) Res. Vet. Sci. 24, 293–299.
[11] Every, D. and Skerman, T.M. (1982) N.Z. Vet. J. 30, 156–158.
[12] Stewart, D.J., Clark, B.L., Peterson, J.E., Griffiths, D.A. and Smith, E.F. (1982) Res. Vet. Sci. 32, 140–147.
[13] Claxton, P.D. (1981) Report of Workshop on Ovine Footrot at University of Sydney, pp. 16–34.
[14] Schoolnik, G.K., Tai, J.Y. and Gotschlich, E.C. (1983) Prog. Allergy 33, 314–333.
[15] O'Donnell, I.J., Stewart, D.J. and Clark, B.L. (1983) Aust. J. Biol. Sci. 36, 15–20.
[16] O'Donnell, I.J. (1973) Aust. J. Biol. Sci. 26, 401–413.
[17] Inglis, A.S. (1983) Methods Enzymol. 91, 324–332.
[18] Inglis, A.S., Strike, P.M., Osborne, W.C., Burley, R.W. (1979) FEBS Lett. 97, 179–182.
[19] Peterson, J.D., Nehrlich, S., Oyer, P.E. and Steiner, D.F. (1972) J. Biol. Chem. 247, 4866–4871.
[20] Tarr, G.E. (1977) Methods Enzymol. 47, 335–357.
[21] Woods, K.R. and Wang, K.-T. (1967) Biochim. Biophys. Acta 133, 369–370.
[22] Matsuo, H. and Narita, K. (1975) in: Protein Sequence Determination (Needleman, S.B. ed) 2nd edn, pp. 104–113, Springer-Verlag, New York.
[23] Hermodson, M.A., Chen, K.C.S. and Buchanan, T.M. (1978) Biochemistry 17, 442–445.
[24] Gough, K.H. and Inglis, A.S. (1980) Aust. J. Biol. Sci. 33, 295–301.
[25] Kolb, E., Harris, J.I. and Brigden, J. (1974) Biochem. J. 137, 185–197.
[26] Kohlmiller, N.A. and Howard, J.B. (1979) J. Biol. Chem. 254, 7309–7315.
[27] Synge, R.L.M. (1945) Biochem. J. 39, 351–355.
[28] Tsugita, A. and Scheffler, J.-J. (1982) Eur. J. Biochem. 124, 585–588.
[29] Frøholm, L.O. and Sletten, K. (1977) FEBS Lett. 73, 29–32.
[30] Sastry, P.A., Pearlstone, J.R., Smillie, L.B. and Paranchych, W. (1983) FEBS Lett. 151, 253–256.
[31] Hopp, T.P. and Woods, K.R. (1981) Proc. Natl. Acad. Sci. USA 78, 3824–3828.
[32] De Graaf, F.K., Klemm, P. and Gaastra, W. (1981) Infect. Immun. 33, 877–883.
[33] De Graaf, F.K. and Roorda, I. (1982) Infect. Immun. 36, 751–758.
[34] Klemm, P. (1981) Eur. J. Biochem. 117, 617–627.
[35] Klemm, P. (1982) Eur. J. Biochem. 124, 339–348.