Amino acid sequence of pilin isolated from *Pseudomonas aeruginosa* PAK

Parimi A. Sastry, Joyce R. Pearlstone*, Lawrence B. Smillie* and William Paranchych+

*Department of Biochemistry and *the Medical Research Council Group in Protein Structure and Function, University of Alberta, Edmonton, T6G 2H7, Canada*

Received 8 November 1982

Amino acid sequence Pseudomonas aeruginosa *pilin* *Hydrophilicity index* *Secondary structure prediction*

1. INTRODUCTION

Pseudomonas aeruginosa strain PAK have polar pili which are flexible filaments of about 5.4 nm diameter and 2.5 μm average length [1–3]. These pili consist of a single subunit, pilin, which was originally reported to have a M_r of 18100 on the basis of SDS–polyacrylamide gel electrophoresis and amino acid compositional studies [4]. The sequence of the first 22 amino acids of the N-terminal region of PAK pilin have been reported [5] and shown to be strikingly homologous to the N-terminus of pili [6] isolated from *Moraxella nonliquefaciens* and *Neisseria gonorrhoeae* [7]. The N-terminal amino acid is the unusual *N*-methylphenylalanine [7,8].

Here we present the complete amino acid sequence of the PAK pilus subunit protein. These sequencing studies show the PAK pilus protein to contain 145 amino acid residues, corresponding to a M_r of 15082. The size of PAK pilin is therefore somewhat smaller than originally estimated.

2. EXPERIMENTAL

P. aeruginosa PAK/2Pfs is a multipiliated mutant of PAK that is resistant to phages PO4 and Pf1 [9]. This strain was kindly provided by Dr D.E. Bradley (Memorial University, St. John's, Newfoundland).

Pili were produced and purified as in [10]. The purified protein was homogeneous as examined by polyacrylamide gel electrophoresis both with and without sodium dodecyl sulfate. Performic acid oxidation was carried out at 0°C for 16 h [11]. Tryptophan was estimated as in [12]. Amino acid compositions were determined with a Durrum D-500 automated amino acid analyzer as in [13]. Cyanogen bromide (CNBr) peptides were prepared by treating the protein with CNBr in 70% formic acid for 30 h at room temperature [14]. The resulting peptides were resolved on a Sepharose G-50 (superfine) column (2.5 × 180 cm) using 25% formic acid.

Prior to cleavage with trypsin, chymotrypsin and pepsin, pili protein was reduced and carboxymethylated [15]. Tryptic and chymotryptic digestions of PAK pilin were carried out in 0.1 M ammonium bicarbonate (pH 8.1) buffer at 37°C with sequential addition of enzyme at 0, 3 and 6 h. The final enzyme concentration was 3:50. Pepsin digestion was at 37°C for 20 h in 5% formic acid at a molar ratio of enzyme to substrate of 1:50. Thermolysin digestion was at a molar ratio of 1:30 in 0.1 M ammonium acetate (pH 8.2) containing 1 mM $CaCl_2$ at 37°C for 20 h. Since all proteolytic

+ To whom correspondence should be addressed

Published by Elsevier Biomedical Press
00145793/83/0000—0000/$3.00 © Federation of European Biochemical Societies

digestion mixtures contained substantial amounts of insoluble peptide material due to incomplete digestion because of poor solubility of pilin, the digests were clarified by centrifugation and only soluble peptides were further purified. Peptide purification usually involved one or more fractionations by gel filtration chromatography on Sephadex G-50, by ion exchange chromatography on Chromobeads type P, or high voltage electrophoresis at pH 6.5, 1.8 and/or 3.5.

Citraconylation of the protein was performed as in [16] using a slight modification. The protein was dissolved in 0.05 M sodium phosphate buffer (pH 8.1). A 40-fold molar excess of citraconic anhydride (Eastman) over the number of lysine residues/monomer was added at 25°C with continuous stirring. The pH of the reaction was maintained at pH 8.1 by the addition of 3 M sodium hydroxide. After completion of the reaction, the citraconylated protein was desalted either by dialysis against 0.1 M ammonium bicarbonate buffer (pH 8.1) or by gel filtration using Sephadex G-25 equilibrated with 0.1 M ammonium bicarbonate (pH 8.1) buffer. Tryptic digestion of the citraconylated protein was carried out at 42°C for 3 h in 0.1 M ammonium bicarbonate buffer (pH 8.1) at an enzyme to substrate molar ratio of 1:50. Following tryptic digestion, the solution was centrifuged and the soluble peptides in the supernatant decitraconylated by incubation in 10% formic acid at 25°C for 6 h.

Amino acid sequencing of peptides was by manual Edman degradations as in [17] and analysis of N-terminal residues by the dansyl method in [18]. Full details of the sequence analysis will be reported subsequently (Sastry, Pearlstone, Smillie and Paranchych, in preparation).

The hydrophilicity values [19] for spans of 6 residues were averaged sequentially along the length of the protein, and these values plotted at the midpoint of the averaged group of amino acids. The secondary structural predictions were obtained using the method in [20].

3. RESULTS AND DISCUSSION

The complete amino acid sequence of *Pseudomonas aeruginosa* PAK pilin is shown in fig.1. The protein has a total of 145 residues, 15 of which are aspartates and glutamates and 15 lysines and arginines, resulting in a net charge of close to zero at neutral pH. Apart from 3 charged residues, the N-terminal region (residues 1–43) is very hydrophobic. The acid and basic residues are relatively evenly distributed throughout the remainder of the molecule. This is illustrated in fig.2, in which the regions of highest local average hydrophilicity (hexapeptide averages) are shown. Fig.2 also shows analysis of the sequence by the method used in [20] for secondary structural predictions. Possible α-helical segments were identified between residues 16–21, 31–40, 46–51, 75–82 and 134–139, totalling 25% α-helix. Four sections of β-sheet structure were predicted at residues 6–15, 23–29, 98–107 and 116–124, totalling 25% of the protein. It is possible that the two areas of β-sheet at the N-terminus are hydrogen-bonded to the two located in the C-terminal portion of the protein, thereby allowing the molecule to fold back onto itself in a more compact form. Seven β-turns were

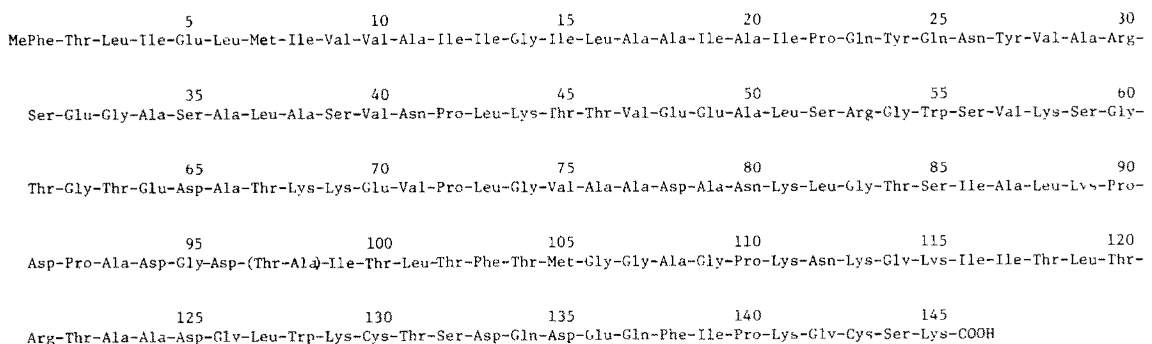

Fig.1. Primary structure of the pilus subunit protein from *Pseudomonas aeruginosa* strain PAK.

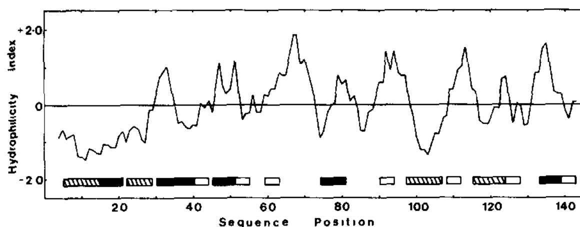

Fig.2. Hydrophilicity index profile and secondary structural predictions for *Pseudomonas aeruginosa* PAK pilin. The hydrophilicity values [19] for spans of 6 residues were averaged sequentially along the length of the protein, and these values plotted at the midpoint of the averaged group of amino acids. The secondary structural predictions were obtained as in [20] for regions of α-helix (▬), β-sheet (▨) and β-turn (▭).

predicted between residues 41–44, 52–55, 60–63, 91–94, 109–112, 125–128 and 140–143.

As indicated in [21], considerable caution must be exercised when using amino acid sequence information to predict the secondary structural characteristics of proteins. The accuracy of the predictions for both β-sheet regions and for turns are generally lower than for the helices, and in a number of instances the agreement between prediction and observation is no better than would be expected for a random selection of residues.

The N-terminal region of PAK pilin (residues 1–30) is particularly interesting because of its extreme hydrophobicity and the fact that it is highly conserved in pili derived from genera as divergent as *Pseudomonas, Neisseria* and *Moraxella* [7]. Although current studies in this laboratory are aimed at elucidating the function of this 30 residue domain, preliminary observations suggest it may be involved in subunit–subunit interactions. This conclusion is based on pH titrations and absorption spectroscopy at 290 nm to monitor the ionized state of tyrosine residues 24 and 27 (Watts, Kay and Paranchych, unpublished). Titration of native pili revealed a pK of 12.3, suggesting the tyrosine residues are inaccessible to titration. Parallel CD measurements show that no denaturation of pili occurs as a function of pH until about pH 12.3. Therefore, native pili are somewhat denatured before the tyrosines are accessible to hydroxyl ions. Alternatively, when pili are dissociated into a mixture of monomers and dimers with octyl glucoside [22], the tyrosine residues titrate with a normal pK of 10.05, suggesting that in this case the tyrosine residues are completely exposed to the aqueous environment. Since CD studies have shown that octyl glucoside treatment of pili does not cause any denaturation or change in secondary structure [23], the foregoing studies suggest that at least the region involving residues 24–27 are involved in subunit–subunit interactions.

In addition to PAK pili, complete amino acid sequence information is available for K88 [24] and CFA/I [25] pili. K88 pilin contains 264 residues, while CFA/I pilin contains 147 residues. We have searched for amino acid homologies between these 3 pilus proteins using a variety of computer programs but in all cases little or no homology was evident. Thus, these 3 types of pili appear to be unrelated.

ACKNOWLEDGEMENTS

We would like to thank Mike Nattriss for excellent technical assistance. This investigation was supported by the Medical Research Council of Canada.

REFERENCES

[1] Bradley, D.E. (1972) Genet. Res. 19, 39–51.
[2] Folkhard, W., Marvin, D.A., Watts, T.H. and Paranchych, W. (1981) J. Mol. Biol. 149, 79–93.
[3] Weiss, R.L. (1971) J. Gen. Microbiol. 67, 135–143.
[4] Frost, L.S. and Paranchych, W. (1977) J. Bacteriol. 131, 259–269.
[5] Paranchych, W., Frost, L.S. and Carpenter, M. (1978) J. Bacteriol. 134, 1179–1180.
[6] Frøholm, L.O. and Sletten, K. (1977) FEBS Lett. 73, 29–32.
[7] Hermodsen, M.A., Chen, K.C.S. and Buchanan, T.M. (1978) Biochem. 17, 442–445.
[8] Frost, L.S., Carpenter, M. and Paranchych, W. (1978) Nature 271, 87–89.
[9] Bradley, D.W. and Pitt, T.L. (1974) J. Gen. Virol. 24, 1–15.
[10] Paranchych, W., Sastry, P.A., Frost, L.S., Carpenter, M., Armstrong, G.D. and Watts, T.H. (1979) Can. J. Microbiol. 25, 1175–1181.
[11] Moore, S. (1963) J. Biol. Chem. 238, 235–237.
[12] Liu, T.-Y. and Chang, Y.H. (1971) J. Biol. Chem. 246, 2842–2848.

[13] Hodges, R.S. and Smillie, L.B. (1972) Can. J. Biochem. 50, 312–329.
[14] Gross, E. and Witkop, B. (1962) J. Biol. Chem. 237, 1856–1860.
[15] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–627.
[16] Gibbons, I. and Perham, R.N. (1970) Biochem. J. 116, 843–849.
[17] Gray, W.R. (1967) Methods Enzymol. 11, 469–475.
[18] Hartley, B.S. (1970) Biochem. J. 119, 805–822.
[19] Hopp, T.P. and Woods, K.R. (1981) Proc. Natl. Acad. Sci. USA 78, 3824–3828.
[20] Chou, P.Y. and Fasman, G.D. (1974) Biochem. 13, 222–245.
[21] Matthews, B.W. (1975) Biochim. Biophys. Acta 405, 442–451.
[22] Watts, T.H., Scraba, D.G. and Paranchych, W. (1982) J. Bacteriol. 151, 1508–1513.
[23] Watts, T.H., Kay, C.M. and Paranchych, W. (1982) Can. J. Biochem. 60, 867–872.
[24] Klemm, P. (1982) Eur. J. Biochem. 117, 617–627.
[25] Klemm, P. (1982) Eur. J. Biochem. 124, 339–348.