JOURNAL OF BACTERIOLOGY, Feb. 1983, p. 950–954
0021-9193/83/020950-05$02.00/0

Vol. 153, No. 2

# N-Terminal Amino Acid Sequencing of EDP208 Conjugative Pili

LAURA S. FROST, GLEN D. ARMSTRONG,† B. BRETT FINLAY, BRIAN F. P. EDWARDS,‡ AND WILLIAM PARANCHYCH*

*Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7*

Received 10 May 1982/Accepted 26 November 1982

EDP208 conjugative pili contain a single polypeptide subunit of 11,500 daltons with a blocked N-terminus. This N-terminal blocking moiety was identified as an *N*-acetyl group by $^1H$ nuclear magnetic resonance analysis of an N-terminal tripeptide isolated from pronase digests of EDP208 pilin. Limited acid hydrolysis of the tripeptide allowed its sequence to be determined as acetyl-NH-Thr-Asp-Leu. Trypsin digestion of EDP208 pilin resulted in the quantitative release of a fragment containing 12 residues from the N-terminus of the protein. The sequence of this dodecapeptide was determined to be acetyl-NH-Thr-Asp-Leu-Leu-Ala-Gly-Gly-Lys-Asp-Val-Asp-Lys.

EDP208 conjugative pili are encoded by a derepressed derivative of a naturally occurring *lac* plasmid, $F_0$*lac*, originally isolated from *Salmonella typhi* (5). EDP208 pili are serologically unrelated to F pili and do not promote infection by F-specific RNA bacteriophage (1, 3). However, they do confer sensitivity to the F-specific filamentous DNA phages, and they promote bacterial conjugation (5). Armstrong et al. (1) have shown that EDP208 pili contain a single polypeptide subunit of 11,500 daltons with a blocked N-terminus. Although it was originally believed that EDP208 pili contain both sugar and phosphate residues (1), recent studies indicate that these moieties are not covalently linked to the protein (2).

In the present communication, we identify the blocking group on EDP208 pilin and the amino acid sequence of 12 residues at the N-terminus of the protein. This dodecapeptide is the primary product released when EDP208 pilin is digested with trypsin.

## MATERIALS AND METHODS

**Bacteria.** The EDP208 plasmid was carried in the host strain JC6256 (*Escherichia coli* K-12 F⁻ *trp lac*). EDP208 is the derepressed form of $F_0$ *lac* (5) and was kindly donated by N. S. Willetts. Department of Molecular Biology, University of Edinburgh, Edinburgh, Scotland.

**Purification of pili.** The purification of EDP208 pili was as described previously by Armstrong et al. (1). Dissociation of intact pili into pilin and gel filtration through sodium dodecyl sulfate (SDS)–Sephadex G-200 columns to remove tightly associated low-molecular-weight contaminants was as described previously by Armstrong et al. (2). The resulting SDS-treated pilin was subjected to acetone precipitation (2) to remove detergent before digestion with proteolytic enzymes.

**Enzymatic digestion of pilin.** Trypsin, tolylsulfonyl phenylalanyl chloromethyl ketone (TPCK), α-chymotrypsin, and pronase were purchased from Worthington Diagnostics (Freehold, N.J.). Digestion of pilin with these enzymes was carried out in 0.1 M ammonium bicarbonate (pH 8.5) for 8 to 18 h at 37°C. The ratio of protein to enzyme for trypsin and chymotrypsin digestion was 50:1 on a molar basis. The pronase concentration was 1% by weight of the protein.

**High-voltage electrophoresis.** Purification of some proteolytic digestion products was achieved by high-voltage preparative paper electrophoresis (4) at pH 6.5 (3,000 V for 45 min). Side strips were cut off and stained with cadmium-ninhydrin to localize peptide bands, which were eluted with water from appropriate regions of the filter paper.

**Mobility and charge of peptides.** The mobility of the peptides was measured at pH 6.5 and is positive and relative to Lys for basic residues and negative and relative to Asp for acidic residues. The charge of a peptide was determined according to Offord's (12) mobility-molecular weight plots.

**Edman degradation.** Manual Edman degradations were carried out by the method of Gray (8). Amino-terminal residues were analyzed by the dansyl method described previously by Hartley (9).

**Amino acid analyses.** Protein samples were hydrolyzed in vacuo at 110°C in 0.2 ml of constant boiling hydrochloride. Peptide samples were routinely hydrolyzed for 16 h, and compositions were expressed as mole ratios, neglecting those amino acids present in amounts less than 0.2 mol. The amino acid analyses

† Present address: Pharmacology Department, University of Calgary, Calgary, Alberta, T2M 1N4, Canada.
‡ Present address: Department of Biochemistry, School of Medicine, Wayne State University, Detroit, MI 48201.

were performed with a Durrum D-500 automated amino acid analyzer.

**$^1$H NMR spectroscopy.** 270 MHz $^1$H nuclear magnetic resonance (NMR) spectra were obtained on a Bruker HXS-270 NMR spectrometer operating in the Fourier transform mode and equipped with quadrature detection. Spectral parameters were: 500 scans; sweep width, ±2,000 Hz; 8,000 data points; 90° pulse (9 μsec); and line broadening, 0.5 Hz. All chemical shifts are referenced to the methyl resonance of sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS), and the ambient temperature was 299 K.

## RESULTS AND DISCUSSION

**Isolation of blocked amino terminal tripeptide.** Five hundred nanomoles of EDP208 pilin was purified by gel filtration chromatography as described above, then digested with pronase (1% by weight of the protein) for 18 h at 37°C in 0.1 M $NH_4HCO_3$ (pH 8.5). The insoluble residue from the digestion was removed by centrifugation at 2,000 × *g* for 10 min, and the soluble portion was lyopohilized and redissolved in water three times to remove $NH_4HCO_3$. The soluble peptides were then dissolved in 1 ml of deionized water and passed through a cation-exchange column (0.8 by 7 cm) containing AG 50W X8 resin (50 to 100 mesh; Bio-Rad Laboratories, Richmond, Calif.). The acidic peptides were eluted with 15 ml of water whereas the remaining peptides were eluted with 15 ml of 1 M ammonium acetate (pH 6.1). The water eluate contained a tripeptide quantitatively removed from the pilin which failed to react with ninhydrin. Amino acid analysis showed the tripeptide to contain Asp, Thr, and Leu in equimolar amounts.

**Identification of the N-terminal blocking group.** Approximately 500 nmol of the N-terminal tripeptide described above was subjected to $^1$H NMR analysis to determine the nature of the blocking group. As shown in the resulting NMR spectra (Fig. 1), a resonance at 2.1 ppm indicated the presence of an *N*-acetyl moiety. This assignment was made on the following basis. (i) The resonance at 2.1 ppm is similar to the *N*-acetyl resonance obtained with *N*-acetyl-Asp or *N*-acetyl-Thr but different from that obtained with free acetate added subsequently to the pH 6.6 sample (1.9 ppm). (ii) Adjusting the pH of the sample from 2.0 (not shown) to 6.6 caused a large chemical shift of the free acetate resonance, whereas the *N*-acetyl resonance shifted only slightly downfield. It was concluded that EDP208 pilin contains an *N*-acetyl blocking group at its amino terminus. The $^1$H NMR

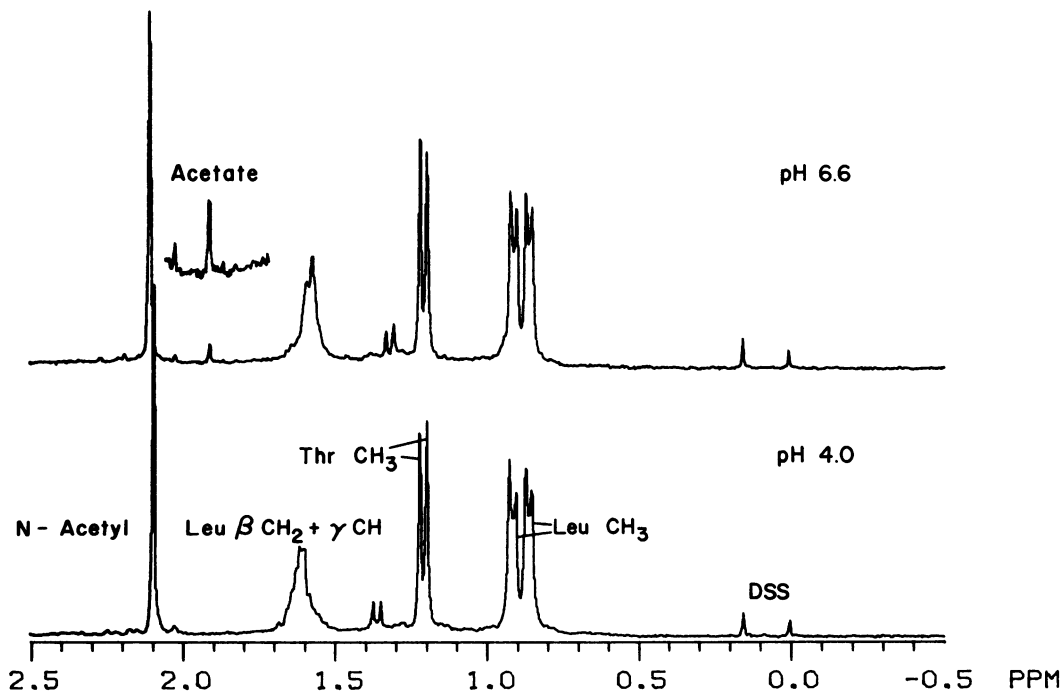


FIG. 1. 270 Mhz $^1$H NMR spectra of the N-terminal tripeptide of EDP208 pilin. Approximately 500 nmol of tripeptide was dissolved in 400 μl of 99.8% $D_2O$ (Bio-Rad) containing 5 mM imidazole as a buffer and 0.05 mM DSS as an internal standard. Spectra were obtained at pH 6.6 and 4.0. Sodium acetate (0.1 mM) was added to the spectrum at pH 6.6.

spectra also indicated that the N-terminal tripeptide (Asx, Thr, Leu) contains Asp rather than Asn (based on the fact that the Asp β protons would not titrate and that the β protons shifted with added $Yb^{3+}$ [see below]).

**Sequence of EDP208 N-terminal tripeptide.** The sequence of the amino terminal tripeptide was deduced by analyzing the products of a limited acid hydrolysis (1 h at 110°C in 0.03 M HCl in an open tube). Separation of the hydrolysis products by paper electrophoresis at pH 6.5 (Fig. 2) produced three spots with mobilities of 0.53 (spot 1), 0.67 (spot 2), and 0 (spot 3). The spots were eluted from the electropherograms as described above and analyzed for amino acid composition. Spot 1 contained Thr, Asp, and Leu (1:1:1); spot 2 contained Thr and Asp (1:1); and spot 3 contained only Leu. Since dansylation of spot 1 showed Thr to be the amino terminal residue, it was deduced that the sequence of the EDP208 N-terminal tripeptide is acetyl-NH-Thr-Asp-Leu.

This sequence was verified, using $^1H$ NMR in the presence of the paramagnetic shift probe $Yb^{3+}$. Ions such as $Yb^{3+}$ markedly affect the magnetic environment of nuclei within their vicinity, producing shifts in line positions and altered relaxation rates. The effect on chemical shifts decreases as a function of the distance between the proton and the $Yb^{3+}$ ion. Since $Yb^{3+}$ binds to carboxyl groups, titration of the tripeptide with this ion allowed the distance between the carboxyl groups and characteristic protons to be determined. This suggested that the sequence of the tripeptide was acetyl-NH-Thr-Asp-Leu, confirming the sequence deduced from analysis of partial acid hydrolysis products.

**Isolation of tryptic peptides from EDP208 pilin.** The purified pilin monomer (2.5 μmol) was digested for 18 h with TPCK-trypsin as described above. The insoluble portion was pelleted by centrifugation (2,000 × *g* for 10 min), whereas the soluble peptides were fractionated on a Bio-Gel P4 column (100 mesh; 1 by 100 cm; exclusion limit, 4,000 daltons), using 0.1 M $NH_4HCO_3$ as a buffer.

Aliquots (2 μl) were removed from even-number fractions (0.5 ml) and spotted on a Polygram Cel 400 nitrocellulose thin-layer sheet (20 by 20 cm) (Brinkmann Instruments, Inc., Westbury, N.Y.) 6 cm from one edge of the sheet. After electrophoresis at 500 V for 45 min with 8% formic acid–2% acetic acid (pH 2.1) in the electrode chambers, the electropherogram was stained with the cadmium-ninhydrin reagent (4).

Three peptides were eluted from the P4 column and labeled ET1, ET2, and ET3. ET2 and ET3 coeluted from the P4 column and were separated by high-voltage preparative paper electrophoresis at pH 1.8, using Whatman 1MM paper. When thin-layer electrophoresis and ninhydrin detection were used to monitor column fractions, minute amounts of free amino acids and possibly dipeptides were detected in fractions eluting later than those containing ET1, ET2, and ET3. These trace contaminants were not detectable by standard amino acid compositional analysis.

A summary of the properties and amino acid

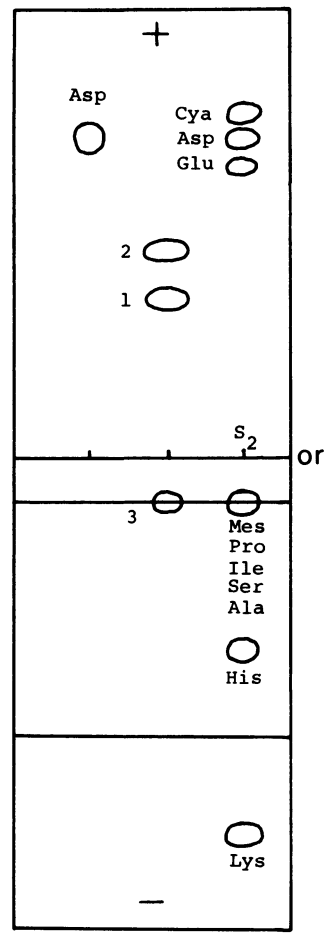


FIG. 2. Scale drawing of paper electropherogram of partial acid hydrolysis products of the EDP208 N-terminal tripeptide. Twenty to thirty nanomoles of peptide was applied to the origin, and electrophoresis was carried out with 10% pyridine–0.3% acetic acid (pH 6.5). The conditions were 3,000 V for 45 min. The spots were visualized with cadmium-ninhydrin spray. The mobilities of spots 1 and 2 were calculated relative to the distance between the center of the neutral spot (spot 3) and the top of the Asp spot. Standard ($S_2$) contained Pro, Ile, Ser, Ala, His, Lys, Glu, Asp, methionine sulfone (Mes), and cysteic acid (Cya). Ori, Origin.

TABLE 1. Summary of properties and amino acid compositions of tryptic, chymotryptic and pronase peptides derived from EDP208 pilin

| Peptide | Yield (%)[a] | Mobility (pH 6.5)[b] | Net charge[c] | No. of residues | Amino acid composition and sequence data[d] |
|---|---|---|---|---|---|
| ET1 | 79.9 | −0.43 | −2 | 12 | Acetyl-Thr-Asp-Leu-Leu-Ala-Gly-Gly-Lys-Asp-Val-Asp-Lys<br>1.1 1.0 1.0 1.0 1.1 1.0 1.0 0.9 1.0 1.0 1.0 0.9 |
| ET2 | 12.0 | −0.30 | −1 | 8 | Acetyl-Thr-Asp-Leu-Leu-Ala-Gly-Gly-Lys<br>1.2 1.0 1.0 1.0 1.0 0.9 0.9 0.9 |
| ET3 | 6.5 | −0.46 | −2 | 4 | Asp-Val-Asp-Lys<br>1.2 0.9 1.2 1.0<br>→ → → → |
| ET1CH1 | 95.9 | ND[e] | −1 | 4 | Acetyl-Thr-Asp-Leu-Leu<br>1.0 1.0 1.0 0.6 |
| ET1CH2 | 76.8 | 0 | 0 | 9 | Leu-Ala-Gly-Gly-Lys-Asp-Val-Asp-Lys<br>0.3 1.1 1.1 1.1 1.0 1.0 0.9 1.0 1.0 |
| EP1[f] | 93.0 | ND | −1 | 3 | Acetyl-Thr-Asp-Leu<br>1.0 1.0 1.0<br>→ → → → → → → → → |

[a] Yields of peptides are based on nanomoles of EDP208 pilin taken for tryptic digestion.
[b] Electrophoretic mobilities are relative to Asp at pH 6.5.
[c] Net charges were calculated from $M_r$ and mobility by the method of Offord (12).
[d] Amino acid compositions are in molar ratios. The notation → is for sequence determination by the dansyl-Edman procedure. The notation ⇀ is for sequence determination by a combination of $^1$H NMR and amino acid analysis of partial acid digestion products.
[e] ND, Not determined.
[f] N-terminal peptide generated by pronase digestion of ET1. See text for details.

compositions of peptides ET1, ET2, and ET3 is shown in Table 1. Both ET1 and ET2 had blocked amino terminals and were considered to represent related peptides derived from the amino terminal of the protein. To confirm this, 100 nmol of ET1 was digested with pronase, and the acidic peptides were isolated by cation-exchange chromatography as described above. As expected, amino acid analysis and partial acid hydrolysis studies showed the water eluate to contain the acetylated amino terminal tripeptide EP1.

**Amino acid sequence determination of tryptic peptide ET1.** Chymotrypsin digestion of peptide ET1 for 8 h as described above generated a single ninhydrin-reactive peptide (ET1CH2) and a blocked N-terminal peptide (ET1CH1). ET1CH2 was purified by high-voltage preparative paper electrophoresis at pH 1.8, whereas the acidic N-terminal peptide was purified by cation-exchange chromatography. Amino acid composition analysis of ET1CH1 indicated that this was the N-terminal tetrapeptide containing a Leu residue in position four. Peptides ET3 and ET1CH2 were sequenced manually, using Edman degradation and dansylation as described above. The results of the sequencing studies are summarized in Table 1. It may be seen that the sequence of ET1 was deduced to be acetyl-Thr-Asp-Leu-Leu-Ala-Gly-Gly-Lys-Asp-Val-Asp-Lys. In a separate study, Worobec et al. (14) confirmed the accuracy of this sequence by showing that synthetically prepared ET1 is strongly antigenic toward antipilus antiserum, whereas antibodies raised against the synthetic peptide react strongly with native pili.

Table 2 shows the amino acid compositions of the soluble and insoluble fractions derived from trypsin digests of EDP208 pilin. A total composition for the intact pilin is also shown, as well as

TABLE 2. Amino acid compositions of the soluble and insoluble fractions derived from trypsin digests of EDP208 pilin

| Amino acid | EDP208 pilin (no.) | Trypsin insoluble (no.) | Trypsin soluble (no.) | N-Terminal dodecapeptide (no.) |
|---|---|---|---|---|
| Lys | 7 | 5 | 2 | 2 |
| His | 0 | 0 | 0 | 0 |
| Arg | 2 | 2 | 0 | 0 |
| Asp | 10 | 7 | 3 | 3 |
| Thr | 12 | 11 | 1 | 1 |
| Ser | 2 | 1 | 0 | 0 |
| Glu | 2 | 1 | 0 | 0 |
| Pro | 0 | 0 | 0 | 0 |
| Gly | 11 | 9 | 2 | 2 |
| Ala | 9 | 8 | 1 | 1 |
| 1/2 Cys | 2 | ND[a] | ND | 0 |
| Val | 10 | 8 | 2 | 1 |
| Met | 5 | 5 | 0 | 0 |
| Ile | 10 | 8 | 0 | 0 |
| Leu | 12 | 10 | 2 | 2 |
| Tyr | 2 | 2 | 0 | 0 |
| Phe | 7 | 7 | 0 | 0 |
| Trp | 0 | 0 | 0 | 0 |

[a] ND, Not determined.

that of the N-terminal dodecapeptide. It may be seen that trypsin treatment of EDP208 pilin resulted in the quantitative removal of the N-terminal dodecapeptide, with little else being removed from the protein. Studies are in progress to characterize the trypsin-insoluble fraction, using other proteolytic enzymes and cyanogen bromide cleavage.

It is of interest to compare the EDP208 N-terminal pilin sequence with that of other types of pilins. Whereas the EDP208 N-terminus contains a high proportion of polar amino acids, pilins derived from *Pseudomonas aeruginosa* (13), *Neisseria gonorrhoea* (10), and *Moraxella nonliquifaciens* (6) have a strikingly homologous sequence of nonpolar amino acids in the N-terminal region and the unusual amino acid *N*-methylphenylalanine (7, 10). Three examples of pilins with unblocked N-termini are type 1, K88, and CFA/I (11). The N-terminus of F pilin, on the other hand, is acetylated (unpublished data).

## ACKNOWLEDGMENTS

We thank M. Nattriss for excellent technical assistance and L. B. Smillie for making available the amino acid analyzer facilities. Special thanks are due to B. D. Sykes for use of the NMR facility and for critical review of the manuscript.

This investigation was supported by the Medical Research Council of Canada.

## LITERATURE CITED

1. **Armstrong, G. D., L. S. Frost, P. A. Sastry, and W. Paranchych.** 1980. Comparative biochemical studies on F and EDP208 conjugative pili. J. Bacteriol. **141**:333–341.
2. **Armstrong, G. D., L. S. Frost, H. J. Vogel, and W. Paranchych.** 1981. Nature of the carbohydrate and phosphate associated with ColB2 and EDP208 pilin. J. Bacteriol. **145**:1167–1176.
3. **Bradley, D. E., and E. Meynell.** 1978. Serological characteristics of pili determined by the plasmids R711b and F *lac*. J. Gen. Microbiol. **108**:141–149.
4. **Dreyer, W. J., and E. Bynum.** 1967. High-voltage paper electrophoresis. Methods Enzymol. **11**:32–39.
5. **Falkow, S., and L. S. Baron.** 1962. Episomic element in a strain of *Salmonella typhosa*. J. Bacteriol. **98**:1598–1601.
6. **Frøholm, L. O., and K. Sletten.** 1977. Purification and N-terminal sequence of a fimbrial protein from *Moraxella nonliquifaciens*. FEBS Lett. **73**:29–32.
7. **Frost, L. S., M. Carpenter, and W. Paranchych.** 1978. N-Methylphenylalanine at the N-terminus of pilin isolated from *Pseudomonas aeruginosa* K. Nature **271**:87–89.
8. **Gray, W. R.** 1967. Sequential degradation plus dansylation. Methods Enzymol. **11**:469–475.
9. **Hartley, B. S.** 1970. Strategy and tactics in protein chemistry. Biochem. J. **119**:805–822.
10. **Hermodson, M. A., K. C. S. Chen, and T. M. Buchanan.** 1978. *Neisseria* pili proteins: amino-terminal amino acid sequences and identification of an unusual amino acid. Biochemistry **17**:442–445.
11. **Klem, P.** 1979. Fimbrial colonization factor CFA/I protein from human enteropathogenic *Escherichia coli* strains. FEBS Lett. **108**:107–110.
12. **Offord, R. E.** 1966. Electrophoretic mobilities of peptides on paper and their use in the determination of amide groups. Nature (London) **211**:591–593.
13. **Paranchych, W., L. S. Frost, and M. Carpenter.** 1978. N-terminal amino acid sequence of pilin isolated from *Pseudomonas aeruginosa*. J. Bacteriol. **134**:1179–1180.
14. **Worobec, E. A., A. K. Taneja, R. S. Hodges, and W. Paranchych.** 1983. Localization of the major antigenic determinant of EDP208 pili at the N-terminus of the pilus protein. J. Bacteriol. **153**:955–961.