Bacterial Adhesion Pili Are Heterologous Assemblies of Similar Subunits

Esther Bullitt* and Lee Makowski#
*Department of Biophysics, Boston University School of Medicine, Boston, Massachusetts 02118-2526, and #Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306 USA

ABSTRACT P-pili on uropathogenic bacteria are 68-Å-diameter rods typically 1 μm in length. These structures project from the outer membrane of *Escherichia coli*, and contain on their distal tip a thin fibrillum, 25 Å in diameter and 150 Å long, displaying an adhesin protein responsible for the binding of the bacterium to the surface of epithelial cells lining the urinary tract. Operationally, it is possible to identify three morphologically distinct states of the 68-Å-diameter P-pili rods, based on the degree of curvature each can adopt. These states are designated "straight," "curved," and "highly curved." The rods can also be unwound to form thin "threads" that are very similar to the tip fibrillae. Electron microscope data are used to distinguish among these four morphological states and to define limits on the shapes of the pilus proteins. The mechanical properties of the PapA polymers are assessed, and implications of rod polymorphism for pilus function are discussed. A wide variety of data are considered in light of the possibility that all pilins are similar in molecular architecture, with specific differences designed to optimize their specialized functions in the pilus assembly.

INTRODUCTION

Adhesion to host tissue receptors is the first step in successful colonization by many Gram-negative pathogens. Adhesion pili are specialized surface structures responsible for the successful recognition and binding of these bacteria to their host receptors. In addition, these pili are responsible for maintaining this contact during the first stages of bacterial colonization. P-pili found on strains of *Escherichia coli* that cause pyelonephritis (upper urinary tract infections that involve the kidneys) bind specifically to globoseries of glycolipids that are found on epithelial cells that line the urinary tract (Källenius et al., 1980; Leffler and Svanborg-Edén, 1980). These pili are composed of six distinct structural proteins and require at least two additional gene products for assembly. The structural organization of a P-pilus is diagrammed in Fig. 1. The major structural protein is PapA (Båga et al., 1984), a 16,500-Da protein that forms a 68-Å-diameter rod typically 1 μm long that projects from the outer membrane of *E. coli*. PapA is anchored in the outer membrane by PapH (Båga et al., 1987). At its distal end is attached a thin "fibrillum" structure made up largely of PapE (Kuehn et al., 1992) and attached to the rod by PapK (Jacob-Dubuisson et al., 1993). The PapG adhesin (Lund et al., 1987) that carries the specific binding site of the digalactoside target is the most distal protein in the pilus and is attached to the fibrillum by PapF (Jacob-Dubuisson et al., 1993).

Examination of the three-dimensional reconstruction of P-pili structure calculated from electron micrographs of negatively stained pili (Bullitt and Makowski, 1995) suggests that the rod is formed by a tight winding of PapA into a coil, much like the winding of a steel wire into a spring. Mechanical shear can result in the unwinding of this coil into a thin, 25-Å-diameter structure reminiscent of the structure of the tip fibrillae. This unwinding occurs without depolymerization of the PapA.

Despite the differences in structure of the rod and the fibrillum, and the differences in the functions of the various pilins that make up P-pili, there are substantial similarities in their amino acid sequences (Normark et al., 1983; Lindberg et al., 1986; Båga et al., 1987). All six have signal sequences for localization of the nascent peptide to the inner cell membrane, and all six are transported across the periplasm by the PapD chaperone protein. With the exception of PapG, the mature forms of these pilins have molecular masses of ~16,000 Da. The sequences of their amino and carboxyl termini are highly conserved, and there are additional homologies throughout their sequences. All pilins except PapG have two cysteines that are disulfide bond linked. PapG has four cysteines, and a molecular weight about double that of the other pilins.

Three distinct morphologies are defined here for the pilus rod according to the degree of bending that each can undergo. These differences are operational, based on different appearances in electron micrographs, but may reflect differences in mechanical properties and structure that imply an adaptability of pili to their microenvironment. Distinct filament morphologies have also been seen in actin, where a conformational change in the actin subunit alters the flexibility of F-actin, which is dependent on bound divalent cations or nucleotide (Orlova and Egelman, 1993). Mechanical properties are critical to the function of adhesion pili, as broken pili cannot maintain attachment of the bacterium to its host cell. Thus an understanding of the impact of environmental conditions on pili is critical for an understanding of the role these adhesion pili play in bacterial pathogenesis. As a first step toward this goal, we estimate here the flexural rigidity and Young's modulus of P-pili that exhibit distinct morphologies, using persistence length calculations from

Received for publication 22 May 1997 and in final form 14 October 1997.
Address reprint requests to Dr. Esther Bullitt, Department of Biophysics, Boston University School of Medicine, 715 Albany Street, Boston, MA 02118-2526. Tel.: 617-638-4242; Fax: 617-638-4041; E-mail: bullitt@bu.edu.

0006-3495/98/01/623/10 $2.00

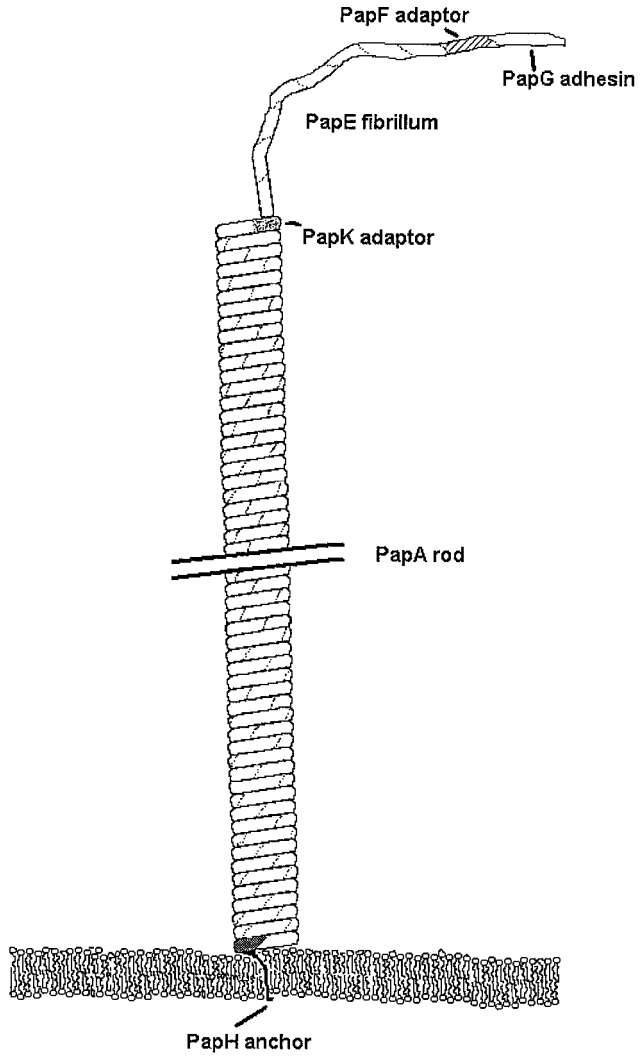

FIGURE 1 Schematic of P-pili. The major structural pilin, PapA, forms an ~1-μm-long helical rod that is anchored in the outer cell membrane by PapH. A thin PapE fibrillum at the tip is connected to the rod by an adaptor protein, PapK. At the distal tip, the host-binding adhesin, PapG, is connected to the fibrillum via an adaptor protein, PapF.

electron micrographs and geometrical information about the pilus structure.

The possibility that all of the proteins making up the P-pili are highly extended, fibrous proteins with similar three-dimensional structures is analyzed in detail. Structural data, sequence data, genetic data, and protein-protein interaction data are all analyzed in the context of this possibility.

METHODS

Isolation of pili

P-pili were purified using procedures described previously (Gong and Makowski, 1992). Hyperpiliated *E. coli* harboring a P-pilus plasmid (pPAP5 for wild-type pili) were grown on LB minimal media plates with 10 μg tetracycline overnight at 37°C. Pili were sheared from the bacteria and purified by cycles of centrifugation of solubilized and precipitated pili, with precipitation initiated by the addition of $MgCl_2$. Specifically, bacteria were grown in soft agar, then scraped into 5 mM Tris, pH 7.5, and blended in a chilled Waring blender (90 s on, 90 s off, four cycles). After centrifugation for 20 min at 8000 $\times$ *g* (Sorvall, SS-34 rotor), pelleted bacteria were discarded. Solid $MgCl_2$ was added to the supernatant to 1 M and incubated at 4°C for at least 10 min. Precipitated pili were pelleted by centrifugation for 40 min at 31,000 $\times$ *g* at 4°C. The pellet was resuspended in 0.5 mM Tris, pH 7.5, and spun for 30 min at 31,000 $\times$ *g* to remove impurities and aggregates. $MgCl_2$ (1 M) was added to the supernatant for a final concentration of 0.1 M $MgCl_2$ to repeat the precipitation, and this purification/precipitation cycle was repeated at least three times. The pure pili were then dialyzed against 0.5 mM Tris, pH 7.5.

Electron microscopy

Pili examined in negative stain were prepared by dilution of the sample to ~0.05 mg/ml. Six microliters were then placed on a carbon-coated 300-mesh copper grid, and the sample was adsorbed to the carbon for 5 min. The grids were washed with 5–20 drops of 10 mM Tris, pH 7.5, stained with 1–2% phosphotungstic acid, pH 7, or 1–2% uranyl acetate. Frozen hydrated pili were prepared by dilution to ~0.1 mg/ml. Six microliters of sample was placed on a holey carbon grid (modified from Fukami and Adachi, 1965) in a humidity-controlled chamber, blotted briefly with filter paper, plunged into liquid nitrogen-cooled liquid ethane (Dubochet et al., 1988), and stored immersed in liquid nitrogen. All grids were examined and photographed with a Philips CM12 electron microscope at 120 kV under low electron dose conditions (~5 $e-/Å^2$). For cryoelectron microscopy, the samples were kept at −175°C in the microscope by use of a Gatan cryoholder and photographed over holes, to reduce the background signal from the carbon film.

Persistence length, flexural rigidity calculations

Persistence lengths, *P*, were calculated for 320 segments of 127 pili from 38 electron micrographs, using the relationship

$$\langle R^2 \rangle = 4PL[1 - (2P/L)(1 - \exp(-L/(2P))]$$

(Rivetti et al., 1996), where *R* is the (straight line) end-to-end distance of the filament segment and *L* is the contour length of the segment. The expectation value of R^2 for each filament morphology was derived from the plot peaks for values of *L*/*R* shown in Fig. 5, using measured values of *L* = 1965 Å+/−143 Å. To expand the data near *L*/*R* = 1, we calculated the frequency of obtaining values for *L*/*R* in bins of width 1 + exp(*x*)/10^4, where *x* = 1, 1.25, 1.5, ... This plot provides a simple method for determining $\langle R^2 \rangle$, and does not require curve-fitting of persistence length to an R^2 to *L* plot, which is known to be unreliable for small values of *L*.

The equation for persistence length (Rivetti et al., 1996) was derived for filaments in two-dimensional (2D) equilibrium, and is appropriate for molecules capable of limited reconfiguration after contact with the substrate. Samples were placed on relatively uncharged surfaces, using grids that were not glow-discharged before negative staining. For negatively stained grids, samples sat for five or more minutes before drying, allowing adequate time for the sample to reach equilibrium. The data for frozen-hydrated pili were analyzed separately. Although the thin ice in which pili were photographed approximates a 2D equilibrium, segment lengths will be underestimated where the filaments curve perpendicular to the plane of the grid. Thus these data were used only to confirm qualitatively the results derived from data from negatively stained specimens.

Segment lengths were chosen to be ~2000 Å (*L* = 1965 Å, SD = 143 Å), with 7–11 data points for each segment, to avoid underestimation of persistence length due to errors in determination of filament centers (Egelman and Stasiak, 1986). Filament center locations were determined by either computer mouse locations on electron micrographs digitized at 5 Å/pixel, or by manual localization on prints of electron micrographs of pili enlarged to 64–67 Å/mm. In all cases, the locations of pili center locations were recorded at ~200–300-Å intervals along the pili.

For thermal bending of flexible polymers, as is the case for pili in 2D equilibrium, the persistence length is given by (Gittes et al., 1993; Reif, 1965)

$$P = EI/kT$$

where E is Young's modulus, I is the geometric moment of inertia, k is Boltzmann's constant, T is temperature, and the flexural rigidity is the product of E and I. Using the known geometry of P-pili, a 68-Å-diameter filament with an elliptical hole in its cross section, the flexural rigidity and Young's modulus of filaments were calculated from the persistence lengths, with the above equation. The geometric moment of inertia, I, of the filaments is equal to $(\pi/4)(\rho^4 - (ab)^2)$, where ρ is the outer radius of the pilus and a and b are the elliptical axes of the hole in the filament. From the three-dimensional reconstruction (Bullitt and Makowski, 1995), ρ, a, and b were taken to be 32 Å, 7.5 Å, and 12.5 Å, respectively. This analysis implicitly assumes that the pili have no intrinsic curvature. This is consistent with our electron microscopic observations of thousands of pili under widely varying conditions. Intrinsic curvature, if it were present, would lead to an underestimation of both the persistence length (Trifonov et al., 1988; Bednar et al., 1995) and Young's modulus.

RESULTS

Polymorphism in PapA assemblies

Electron micrographs of P-pili are included in Figs. 2–4. Figs. 2 and 3 contain electron micrographs of negatively stained specimens. Fig. 4 contains electron micrographs of frozen hydrated pili. In Fig. 2 *a*, most pili appear to be very straight, with occasional sharp bends that suggest punctate failures of the pilus rods. Short segments of isolated pili, apparently broken from longer pili, are also present in this micrograph. In Fig. 2, *b*, *c*, and *e*, most pili appear to be gently curved. Curved pili of this type seldom, if ever, exhibit punctate failures like those seen in Fig. 2 *a*. Fig. 2 *d* shows a threadlike structure over 1500 Å in length, that appears to extend from the end of one pilus rod to the end of another. Fig. 2 *f* shows a curved pilus with a relatively short, thin structure, possibly a fibrillum, extending from one end. Fig. 2 *g* is a field in which a curved rod is immediately adjacent to a straight rod that exhibits a single failure point.

Fig. 2 *h* contains a highly curved pilus. This pilus appears to have about the same diameter as the straighter pili, but has a substructure that is more highly visible on its surface. Fig. 3 contains additional images of highly curved pili that also have enhanced substructure. Fig. 4 contains images of frozen-hydrated pili.

The P-pili in these micrographs have been classified into four structural forms: 1) straight rods, 2) curved rods, 3) highly curved rods, and 4) threads. The "straight" rods are commonly found in samples that have been freeze-dried for scanning transmission electron microscopy (Gong and Makowski, 1992), negatively stained for transmission electron microscopy, or preserved in vitreous ice for transmission electron microscopy. As can be seen in Fig. 2 *a*, these rods are very linear, but exhibit occasional sharp turns in which it appears that a single rod has broken and the bonds between the turns of the coil have given way, leaving only a narrow point of attachment between the two resulting rod fragments. These rods appear to be brittle, with little elastic response to torque or shear until punctate failure. At other places, two straight rods appear to be joined by a thin "thread" that extends between their ends (Fig. 2 *d*). These threads appear to be formed by the unwinding of the straight rods (Bullitt and Makowski, 1995).

Sharp bends and threads are observed mainly in association with straight rods, presumably indicating punctate failure of the pilus. Curved and highly curved pilus structures appear to be capable of elastic bending in response to shear or torque, leading to a reduction in the number of punctate failures. Highly curved pili appear to have a more open architecture than either straight or curved rods, with surface substructure visualized by the penetration of negative stain between the subunits of the rods. All structural forms are seen in specimens

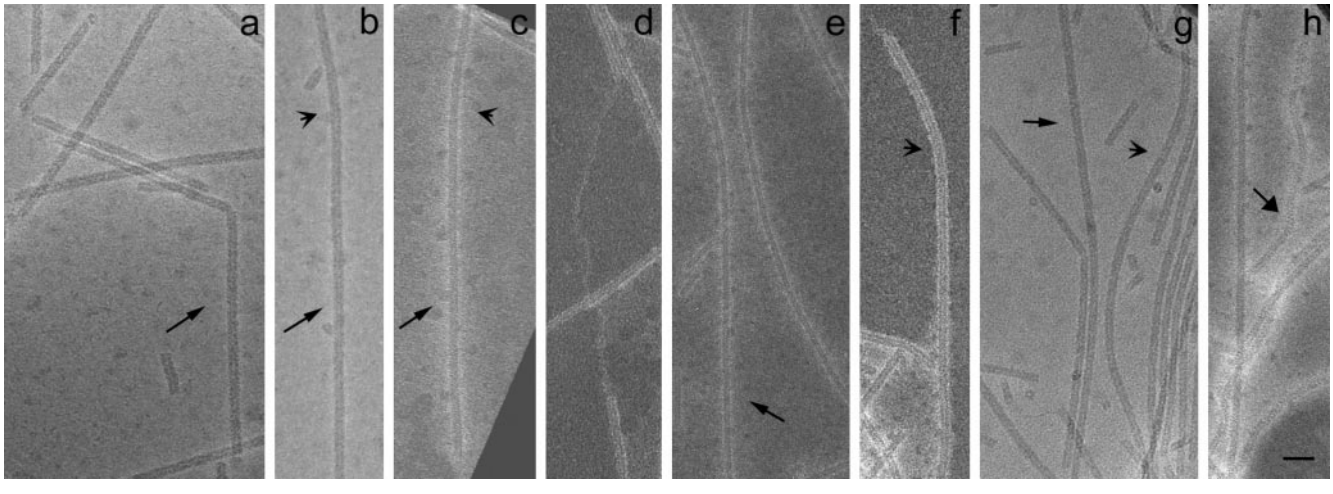

FIGURE 2 Four structural morphologies are visible in these micrographs of negatively stained P-pili, including straight rods (*small arrow*), curved rods (*arrowhead*), highly curved rods (*large arrow*), and thin threads (*thin fiber in d*). Sharp angle turns in P-pili (*a*) and the attachment point of threads to the pilus rod (*d*) indicate punctate failure of the filament. Magnification bar, 250 Å.

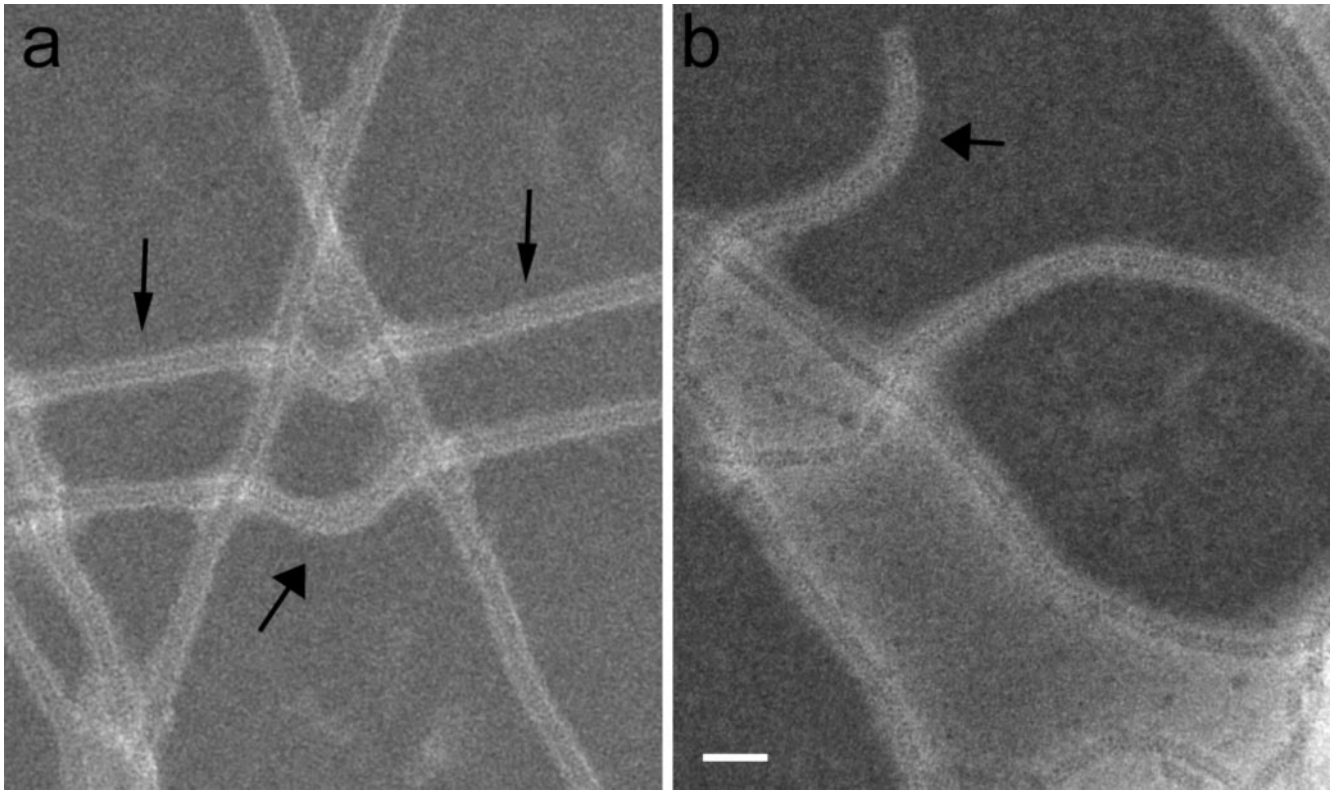

FIGURE 3 Two examples of highly curved pili. (*a*) The straight rod morphology switches to a highly curved rod while crossing other pili. (*b*) Highly curved pili appear to have a more open architecture than either straight or curved rods, with a surface substructure visualized by the increased penetration of negative stain. Magnification bar, 250 Å.

that have been negatively stained, whereas curved or highly curved rods appear to be rare or nonexistent in specimens that have been freeze-dried.

Persistence length

Different degrees of curvature suggest different degrees of rigidity. To quantitate the impression of different flexural rigidity that is given by examination of the electron micrographs, the persistence lengths of P-pili were calculated from measurements made on electron micrographs of negatively stained and frozen hydrated samples. The persistence, or correlation length, measures how far along a filament its direction remains correlated to that of the starting direction. This distance is inversely proportional to flexibility; the more flexible the filament, the shorter its persistence length. As the shortest distance between two points is a straight line, any curve within a filament will increase its contour length, L, as compared to its end-to-end distance, R. Because the filaments are highly concentrated about $L/R = 1$, the data are best visualized quantitatively by a histogram of L/R frequency against a function $(1 + \exp(x)/10^4)$, expanding the region near $L/R = 1$. Examination of this plot, shown in Fig. 5 for data from negatively stained samples, suggests the presence of three distinct populations, corresponding, respectively, to straight, curved, and highly curved pili. The histogram, however, fits equally well to two peaks or three (χ^2 test; data not shown). Consequently, this analysis does not distinguish between the straight and the curved pili. Because our observations indicate that "straight" pili exhibit punctate failure, whereas curved pili do not, we have chosen to make the distinction between straight and curved pili in our analysis of persistence lengths. This separation may not, however, define a population of curved pili that are structurally distinct from straight pili. Highly curved pili are clearly in a class of their own, both qualitatively and quantitatively. The localization of the peaks makes possible an accurate estimate for the average $\langle R^2 \rangle$ for each population, and the subsequent calculation of persistent lengths. The first peak corresponds to $L/R = 1.002$, representing essentially straight pili. The calculated persistence length for these filaments is ~ 8 μm. The putative peak corresponding to $L/R = 1.015$ lies in the region that includes curved pili, and yields a persistence length of ~ 1 μm. The final peak is extremely broad (note the increasing bin width of L/R ratios as x increases), as expected for the wide range of conformations sampled by filaments having short persistence lengths. The peak of these data corresponds to $L/R = 1.160$, yielding a persistence length of ~ 0.1 μm for the highly curved pili.

The designations of "straight" and "curved" pili do not necessarily imply a significant structural difference between them. The helical symmetries of the straight and curved rods were determined by computer analysis of electron micro-

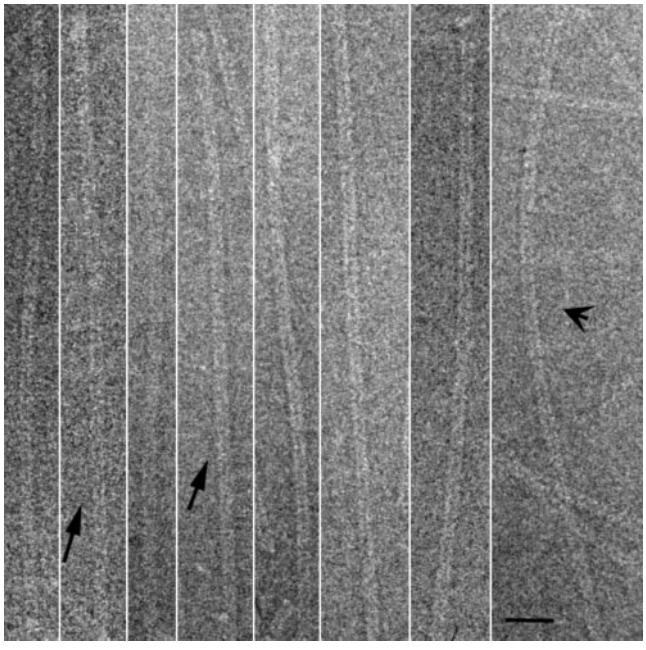

FIGURE 4 Electron micrographs of P-pili preserved in vitreous ice, showing straight-rod and curved-rod morphologies. Magnification bar, 250 Å.

graphs and found to be essentially identical (data not shown). The apparent differences may reflect the mechanical properties of a single kind of structure that has a high likelihood of either bending slightly or remaining straight and breaking into two (straight) rods as the specimen is dried onto the electron microscope grid (see Discussion).

Flexural rigidity and Young's modulus

Flexural rigidity and Young's modulus were calculated from the persistence length and the geometrical moment of inertia of pili, assuming thermal bending of flexible polymers. These results place the flexural rigidity of highly curved pili (4×10^{-28} N m^2) between RecA-DNA complexes (Egelman and Stasiak, 1986) and DNA (Smith et al., 1996); indicate that curved pili (4×10^{-27} N m^2) are slightly more flexible than actin (Oosawa, 1980); and show straight pili (3×10^{-26} N m^2) to be within the range of rigidity determined for actin (Oosawa, 1980). The calculated Young's moduli for the pili morphologies are 5 MPa for highly curved pili, 50 MPa for curved pili, and 400 MPa for straight pili. Highly curved pili have a Young's modulus comparable to that of rubber (E_{rubber} = 1–10 MPa) and will resist bending in a manner similar to that of an equivalently shaped rod made of rubber.

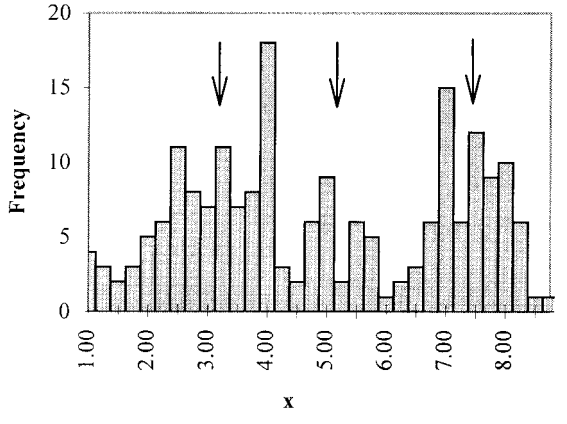

FIGURE 5 Histogram of L/R for P-pili. To expand the data near L/R = 1, we calculated the frequency of obtaining values for L/R in bins of width $1 + \exp(x)/10^4$, where x = 1, 1.25, 1.5, ... The equation is used to calculate L/R at any position x on the graph. Data fall into three peaks, corresponding to L/R ratios of 1.002, 1.015, and 1.160, for straight and curved filaments, and highly curved pili, respectively (*arrows from left to right on the plot*).

Highly curved rods

The highly curved rods appear to have a structure different from that of the straight or curved pili. No attempt has been made to determine their helical symmetry or to calculate a three-dimensional reconstruction of their structure. Nevertheless, the impression obtained from examination of electron micrographs of these structures is that there is much greater penetration of negative stain between the subunits of these rods than for the straight or curved rods. This suggests a significant difference in their structure compared to that of the straight or curved pili rods.

Threads

The thin diameter of threads (25 Å) defines them as a distinct morphological class. These threads appear to be present either in the middle of a pilus, connecting two straight rods, or at the end of a pilus, extending from a rod. Presumably, many or all of those extending from the end of a rod correspond to the fibrillum structure that is present at the distal end of the pilus, and which is composed of PapK, PapE, PapF and PapG. The threads that connect two rods cannot be made up of this material (PapG is always the most distal protein of the pilus), and are most likely made up of PapA that has unwound from the rods in response to stress. In electron micrographs, the fibrillae at the ends of rods cannot be distinguished from the threads connecting two rods. In both cases, these thin structures often appear to take on an open helical structure with a pitch of 100–200 Å. The mass per unit length is indistinguishable for these two structures. Nevertheless, biochemical data indicate that they are made up of different pilins—largely PapE at the ends and largely PapA in the intermediate structures. Our conclusion is that PapA, normally found in rods, is capable of unwinding to form threads, fibrillum-like structures that at low resolution are indistinguishable from the fibrillae composed largely of PapE.

Size and shape of subunits

The results of electron microscopy indicate that the cross-sectional areas of the threads and of the fibrillae are both

that of a ~25-Å-diameter cylinder of protein. Scanning transmission electron microscopy (STEM) was used to determine that their mass per unit length was ~0.4 kDa/Å (Gong and Makowski, 1992). The density of each protein can be calculated from the amino acid composition and the volumes occupied by each amino acid in crystalline, globular proteins (Chothia, 1975). The densities calculated for the pilus proteins from amino acid composition and amino acid volumes are tabulated in Table 1, along with the corresponding lengths for each protein, assuming a mass per unit length equal to that observed for the threads or fibrillae. Given the molecular weights of the mature PapE, PapF, and PapA proteins, this corresponds to lengths of 38 Å, 39 Å, and 41 Å, respectively.

To test the hypothesis that the threads are formed from rods that have unwound, the mass per unit length of PapA in the rods was calculated assuming that it is wound around the 1-start helix of the rod. The dimensions and mass per unit length of the rod itself are accurately determined from the electron microscopic analysis (Bullitt and Makowski, 1995). Given that PapA is an elongated protein wrapped into a rod with known outside diameter, inner cavity size and location, and helical pitch, the PapA protein cannot have a cross-sectional area larger than $\sim \pi P \Delta r/4$, where P is the pitch, equal to 24 Å, and Δr is the difference between the outer and inner radii of the rod (~30 Å). Given the calculated density of PapA, 0.787 Da/Å^3, these dimensions correspond to a mass per unit length of 0.45 kDa/Å, within the expected error of the STEM measurements of threads and fibrillae. Consequently, all of the information we have about the rod structure is consistent with the possibility that threads are formed by unwinding of the rods.

At this resolution, boundaries between individual pilins are unobserved, obscuring the cross-sectional areas of individual proteins. Each subunit could occupy the entire cross section of the threads/fibrillae, and be ~40 Å long. Alternatively, each subunit could occupy half that cross section, have twice that length, and maintain extended contact regions with the subunits immediately distal and proximal to it. Given the importance of the mechanical properties of the fibrillae and threads for survival of the bacterium, the possibility that the proteins are highly elongated (substantially more than 40 Å in length) must be seriously considered. This is certainly the case for the gonococcal pilin, although it is arranged very differently in the pilus rod (Parge et al., 1995).

TABLE 1 Size, shape, and density of P-pilins

Pilin	Molecular mass (Daltons)	Density (Da/Å^3)*	Min. length (Å)
PapG	35,520	0.777	89
PapF	15,560	0.782	39
PapE	15,930	0.780	40
PapK	17,260	0.777	43
PapA	16,520	0.782	41
PapH	19,250	0.777	48

*Volumes of amino acids are from Chothia (1975).

Sequences of pilins

There are substantial similarities between the amino acid sequences of P-pili and some of the other known pili. These similarities are apparent among analogous pilins from different strains (e.g., comparison of PapA sequences; Denich et al., 1991); different pili (e.g., the sequence of PapA in P-pili is similar to that of FimA in type 1 pili); and among structurally distinct pilins within a single pilus (e.g., PapA, PapE, PapF . . .). Although this similarity does not extend to pilins in the type IV class, new pilins are often identified by their similarities to this large family of proteins that includes the P-pilins. The regions of greatest similarity are the C- and N-termini and the regions immediately adjacent to the cysteines that form a single disulfide bond in the intact proteins. Fig. 6 shows one possible sequence alignment for five of the six structural pilins on the P-pilus. Although there are substantial differences in these sequences, they share a number of structural motifs. The similarities among these sequences suggest the possibility that these five proteins are structurally homologous.

Substantial differences are also apparent in the sequences of the pilins, as might be expected for proteins with distinct functional roles. These differences are highlighted in the data in Table 2. Table 2 includes the number of charged residues in each pilin, and one estimate of the percentage of different secondary structures within each protein. Examination of the data in Table 2 suggests that the fibrillae-forming pilins (PapE and PapF) constitute a subclass rather different from that of the other pilins in the P-pili. They have less likelihood of containing α-helical structure and far fewer charged residues than the other pilins. This comparison also suggests that the PapH anchor protein and the PapK adaptor protein are rod-forming pilins with structures more similar to that of PapA than to that of PapE or PapF. This possibility is reflected in the sketch of the P-pilus structure in Fig. 1, in which both PapH and PapK are depicted as integral parts of the pilus rod. The PapH anchor is apparently formed by the highly hydrophobic region at the amino-terminal end of the sequence of the mature protein, the remainder of the protein exhibiting similarity to PapA.

The predicted secondary structure content of the pilins appears to suggest that the structures of the fibrillae-forming pilins are rather different from those of the rod-forming pilins. However, use of sequences to predict secondary structures has proved highly unreliable for proteins that assemble to form macromolecular complexes (e.g., Finer-Moore et al., 1984; Kishchenko and Makowski, 1997). Consequently, these predictions reflect real differences only in the sequences of the different pilins; they may not reflect substantial structural differences.

Rare codons and protein expression

In the biogenesis of P-pili, *E. coli* face the problem of synthesizing the structural proteins of the pili in ratios that correspond to the ratios of their usage within the pili. There are an average of ~1200 PapA per pilus; 50 PapE; and one each of PapG, PapF, PapK, and PapH. Expression of gene

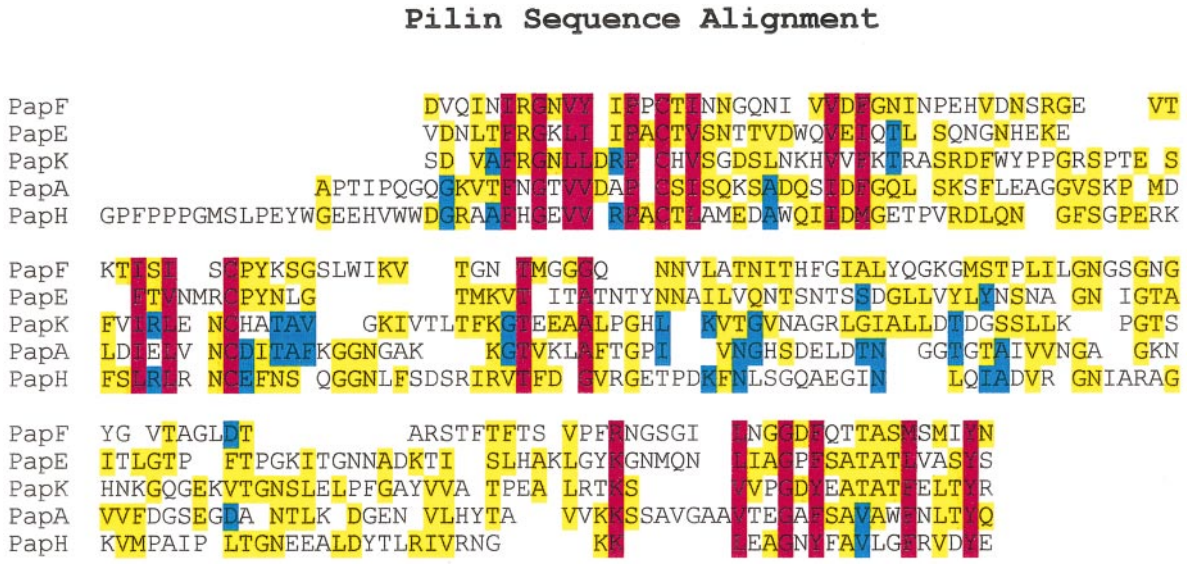

FIGURE 6 One possible sequence alignment of the five structural pilins. Sequences were aligned to maximize the correspondence of highly conserved features, including the disulfide linkage and the amino and carboxy terminal regions. Dark highlighting is shown for residues that are similar for all pilins; light highlighting is for sequences that are similar for some pilins.

products is presumably regulated at the levels of both transcription and translation, and the mechanisms regulating their expression levels are likely to be multiple and complex. Nevertheless, it is interesting to note that the usage of rare codons in the Pap gene cluster correlates well with the number of copies of each gene product that appears in the final pilus structure.

E. coli use several rare codons to control the expression levels of gene products. The levels of tRNAs corresponding to the rare arginine codons AGG and AGA have been proposed as a means for regulating the expression of genes enriched in these codons (Chen and Inouye, 1994). To investigate the possible role that these codons may have in regulating the relative levels of expression of the P-pili genes, the number of AGG and AGA codons in the P-pili genes was tabulated. Table 3 lists the frequencies of the rare arginine codons AGA and AGG in the six pilus structural proteins and in the genes for the nonstructural proteins PapD, PapC, PapB, PapI, and PapJ. PapA, which makes up the bulk of the pili structure, has neither of these codons. PapE, which is expressed at a much lower level than PapA, has two AGA codons and no AGGs. Pili contain only one copy each of the other structural proteins, and the genes coding for these proteins each contain at least one AGG codon. The occurrences of these rare codons correlate well with the level of incorporation of each of these gene products into the pilus structure. The genes for the nonstructural proteins all contain multiple AGG codons, consistent with their role as regulatory or morphogenetic proteins. These numbers suggest that *E. coli* use rare codons to regulate the levels of synthesis of structural proteins to correspond with their level of incorporation into macromolecular assemblies.

TABLE 2 Sequence characteristics

Pilin	Predicted (% alpha)*	Predicted (% beta)*	(+) charges#	(−) charges
PapG	19	25	30	27
PapF	0	46	8	7
PapE	0	38	9	8
PapK	21	32	18	16
PapA	7	24	12	17
PapH	17	14	20	23

*SSP program (Solovyev and Salamov, 1994).
#Assumes His is polar, but uncharged.

DISCUSSION

The results reported here address two fundamental issues involving the P-pili. The first is the relationship between the structure, structural polymorphism, and mechanical functions of the pili. The second is the structure of individual pilins and their structural relationships to one another. Both of these issues depend on the size and shape of the PapA monomer.

Structure of the PapA monomer

There are currently no data that directly address the dimensions or conformation of the PapA monomer. The structure of the *Neisseria gonorrhoeae* pilin has been determined to 2.6-Å resolution by x-ray crystallography (Parge et al., 1995). These data, with constraints from fiber diffraction and electron microscopy data, have been used to model the type IV fiber. The striking differences between type IV and P-pili preclude the construction of a molecular model of P-pili based on the structural model of type IV pili. Currently, the resolution of our electron microscope studies is insufficient to identify boundaries between subunits in the pilus structure. The monomer has not been crystallized, and

TABLE 3 Rare codons in pilins

Protein	Approximate no. of proteins per pilus	AGG codons	AGA codons
PapG	1	2	3
PapF	1	1	0
PapE	50	0	2
PapK	1	3	1
PapA	1200	0	0
PapH	1	4	5
PapB	—	2	1
PapI	—	2	0
PapD	—	4	2
PapC	—	1	5

although it has been visualized in the electron microscope in complex with the PapD chaperone protein (Bullitt et al., 1996), the shape of individual proteins cannot be ascertained from those data.

The cross-sectional area of the PapA threads is virtually identical to the area available to subunits wound around the one-start helix of the rod. The structure of the rods is thereby compatible with the possibility that they are made of a tight winding of a threadlike structure. In that case, the PapA protein in the unwound state will be a minimum of ~40 Å long and, at most, 25 Å in diameter. All available data are consistent with this possibility.

The function of the P-pilus requires it to have mechanical strength adequate to maintain the attachment of the bacterium to the epithelial cells lining the urinary tract during the flow of urine. For the PapA monomers to remain bound to one another during mechanical stress, they need to have a high affinity for one another. A structure in which the PapA molecules are highly extended (more than 40 Å in length), with extensive interactions with one another, may have the potential for being stronger than one constructed from molecules with relatively small interactive surfaces.

Structural switching in PapA

For the PapA protein to make the transition from being coiled in the pilus rod to being extended into a threadlike structure, it must undergo a structural transition. In the rod there are 3.28 subunits per turn, meaning that identical points on adjacent PapA monomers are rotated ~110° (360°/3.28) relative to one another. In the extended threads, identical points on adjacent monomers are likely to have an orientation within 30° of one another. (In a straight structure, identical points will be identically oriented. The threads seldom change direction from this orientation by more than 30° in a 40-Å-length of filament.) Consequently, for the unwinding of the rod to take place, the two ends of the PapA monomer must rotate 80°-140° relative to one another. This rotation could involve the rotation of two domains within a single protein, the rotation of the bonding surfaces between proteins, or a more extended structural transition distributing the rotation throughout the protein. This rotation will occur in concert with the breaking of bonds between the turns of the rods. Fig. 7 is a sketch of an axial view of the end of a pilus, demonstrating the different structural requirements for a subunit in the rod and in a thread or fibrillum. PapA appears to be able to switch from one conformation to the other during the unraveling of the rod into threads.

Structural polymorphism of PapA polymers

We have classified the observable structural forms of polymers of PapA into four categories, probably representing three distinct structural forms of the pili. These categories are straight rods, curved rods, highly curved rods, and threads. The straight rods have a persistence length of ~8 μm. They appear to be rigid but brittle, and subject to punctate failures at which the turns of the PapA coil making up the rod split apart, leaving a single point of attachment between two straight rods, or maintaining a long, thin, threadlike structure between the two straight rods.

The rods classified as "curved" are morphologically indistinguishable from the straight rods, with a helical symmetry identical to that of the straight rods. The difference between the curved and straight rods may be purely operational—a difference in calculated persistence length created by choosing to fit three rather than two peaks to the graph that quantifies contour length to end-to-end length ratios (Fig. 5). We make this distinction because of the distinct failure modes of which the rods are capable: rods that are not subject to a bending moment during specimen formation are imaged as straight rods. Rods that are subject to extreme bending moments break, resulting in two shorter, straight rods connected by a failure point. Rods that are subject to intermediate bending moments are preserved as the curved rods observed in the electron micrographs. These curved rods appear to have a persistence length of ~1 μm. Straight

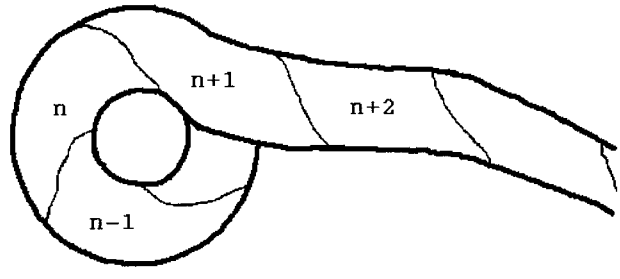

FIGURE 7 Axial view of the top of a pilus rod. Proteins n, $n-1$, and so on are portions of the rod in a curved conformation. Proteins $n+1$, $n+2$, and so on are in an extended conformation in a thread or fibrillum. The binding surfaces of the proteins in the thread are nearly parallel to one another, resulting in an open, extended form for the thread. The binding surfaces of the pilins in the rod are rotated ~110° relative to one another. If this drawing represents the end of a rod and the beginning of a fibrillum, protein n will be the PapK; proteins $n+1$, $n+2$, and so on will be PapE; and protein $n-1$ will be PapA. If this drawing represents the end of a rod of PapA molecules being unwound, all of the proteins represented correspond to PapA. Proteins $n+1$, $n+2$, and so on then represent PapA molecules that have undergone a conformational change. Proteins n, $n-1$, and so on represent PapA molecules still in the rod-forming conformation.

rods with punctate failures are seen in electron micrographs of freeze-dried and negatively stained specimens, but seldom in electron micrographs of frozen-hydrated specimens. It is possible that the interaction of the pili with a surface (which does not occur in the frozen-hydrated specimens) leads to the observed punctate failures.

The "highly curved" pili rods appear to be a distinct state of the pili that has not previously been reported. There is distinct substructure on the surface of these rods, suggestive of substantially more penetration of stain between subunits than in the curved or straight pili. They are also capable of much tighter curvature than other morphological forms of pili. They occur in the same fields of electron micrographs as straight or curved pili, indicating that they form under very similar if not identical conditions. This result differs from the variable flexibility seen in actin filaments; F-actin flexibility is altered by changes in bound divalent cation or nucleotide, which effects the entire filament population (Orlova and Egelman, 1993). The presence of highly curved rods of P-pili adjacent to straight P-pili may represent a second type of mechanical failure of the pili, an alternative to punctate failure as a response to stress. But existing data provide no clue as to why pili should appear to take on this morphologically distinct form under conditions essentially identical to those in which straight or curved rods are found.

The response of a thin rod to a bending moment may result in a local failure, as in the bending of a piece of uncooked spaghetti, or in a distributed curvature, such as the bending of a piece of cooked spaghetti. The difference between these two types of responses involves the intermolecular interactions along the direction parallel to the axis of the rod. If these interactions are rigid, the rod undergoes very little distortion until punctate failure. If these interactions are elastic, the rod bends readily in response to the moment.

The feature that may distinguish the "highly curved" pili from the "straight" pili is the nature of the axial bonds that allow adjacent turns of the helix to adhere to one another. These bonds must break at a single point for the formation of a punctate failure; they must break along the entire length of a thread for the rod to unravel; and they must bend, but not break for the formation of the highly curved pili. Because highly curved pili appear to have an open structure with more space between subunits at high radii, it is reasonable to think that they may also have an altered set of interactions between the proteins in adjacent turns of the helix. It is possible that these interactions are very sensitive to small changes in the ionic strength or pH of the surrounding media. Consequently, the structural and mechanical properties of these pili should be investigated under conditions similar to those existing within the urinary tract.

Subunit-subunit binding among the structural pilins

Electron microscopy and sequence analyses both suggest that PapF, PapE, PapK, PapA, and PapH have similar architectures. Genetic experiments in which the function and morphology of deletion mutants have been analyzed also suggest that these pilins are very similar. In mutants lacking PapE, rods are terminated with ~100-Å-long fibrillar stubs (Kuehn et al., 1992). The stubs are presumably made of PapG and PapF, suggesting that PapK can interact with PapF to attach it to the PapA rod. This implies that the binding surfaces of PapE and PapK are very similar (they can both bind PapF). Mutants lacking PapK also exhibited fibrillae at the ends of the rods (Kuehn et al., 1992), suggesting that PapA can interact with PapE, thus implying that the distal binding surfaces of PapA and PapK are similar (they can both bind PapE). Other genetic experiments of this type (Uhlin et al., 1985; Jacob-Dubuisson et al., 1993) provide further support for this point of view.

The PapG adhesin

The PapG sequence is significantly different from that of the other P-pilins, except for the highly conserved carboxyl terminus. It is composed of an N-terminal lectin-like domain and a C-terminal pilin-like domain (Jones et al., 1995). PapG is not visible as a thickening or thinning of the fibrillum at its distal end. Given that it has twice the molecular weight of the other pilins, arguments analogous to those used above indicate that if PapG occupies the entire 25-Å-diameter cross section of the fibrillae, it must be ~80 Å in length.

Electron micrographs of type 1 pili (Jones et al., 1995) exhibit a fibrillum-like structure on their tip that is only ~160 Å in length. In deletion mutants lacking the fimH gene (the adhesin analogous to PapG in P-pili), a 30-Å-long stub is observed. Because fimH has a molecular weight similar to that of PapG, it is unlikely to be able to fill a 130-Å-length of a fibrillus that is ~25 Å in diameter. This suggests that the type 1 fibrillae are thinner than the P-pili fibrillae, that there are multiple copies of fimH, or that the absence of fimH at the tip of the type 1 pili may prevent the attachment of another structural component that is ~50 Å long.

Quasiequivalence in pili

P-pili are heterologous polymers of six different proteins with common structural motifs and disparate functions. Pilins are able to carry out their various functions because of differences in their structures that represent variations on a single structural theme. If the evolutionary relationship between the pilins is ever traced, we may find that they have all evolved from a single ancestral pilin, and that the division of labor and specialization of function have evolved to produce a more efficient and effective macromolecular assembly for adherence to target tissues.

Adherence of bacteria to the epithelial lining of the urinary tract cannot be effectively promoted or maintained under physiological conditions by a rigidly constructed pilus. The structural plasticity of the pilus is essential for it to carry out its primary function. This plasticity is dependent

on polymorphism of the PapA polymers, which is, in turn, dependent on the structural switching and adaptability of PapA. Satisfaction of functional requirements through the use of structural polymorphism represents a variation on the themes of quasiequivalence first put forward by Donald Caspar and Aaron Klug in their analysis of virus structure in 1963 (Caspar and Klug, 1963). The P-pilus is an adaptive assembly in which multiple conformations of its components contribute significantly to its functional effectiveness.

We thank Dr. M. Schmid for helpful discussions and sharing of unpublished data, and C. England for technical assistance.

We also thank the Charles A. King Trust of the Medical Foundation (EB) and the National Science Foundation (LM) for support for this work.

REFERENCES

Båga, M., M. Norgren, and S. Normark. 1987. Biogenesis of *Escherichia coli* Pap-pili: PapH, a minor pilin subunit involved in cell anchoring and length modulation. *Cell.* 49:241–251.

Båga, M., S. Normark, J. Hardy, P. O'Hanley, D. Lark, O. Olsson, G. Schoolnik, and S. Falkow. 1984. Nucleotide sequence of the gene encoding the Pap pilus subunit of human uropathogenic *E. coli*. *J. Bacteriol.* 157:330–333.

Bednar, J., P. Furrer, V. Katritch, A. Z. Stasiak, J. Dubochet, and A. Stasiak. 1995. Determination of DNA persistence length by cryo-electron microscopy. Separation of the static and dynamic contributions to the apparent persistence length of DNA. *J. Mol. Biol.* 254:579–594.

Bullitt, E., C. H. Jones, R. Striker, G. Soto, F. Jacob-Dubuisson, J. Pinkner, M. J. Wick, L. Makowski, and S. J. Hultgren. 1996. Development of pilus organelle subassemblies in vitro depends on chaperone uncapping of a beta zipper. *Proc. Natl. Acad. Sci. USA.* 93:12890–12895.

Bullitt, E., and L. Makowski. 1995. Structural polymorphism of bacterial adhesion pili. *Nature.* 373:164–167.

Caspar, D. L. D., and A. Klug. 1963. Design and construction of icosohedral viruses. *In* Viruses, Nucleic Acids and Cancer, 17th M. D. Anderson Symposium. Williams and Wilkins, Baltimore, MD. 27.

Chen, G. T., and M. Inouye. 1994. Role of AGA/AGG codons, the rarest codons in global gene regulation in *Escherichia coli*. *Genes Dev.* 8: 2641–2652.

Chothia, C. 1975. Structural invariants in protein folding. *Nature.* 254: 304–308.

Denich, K., L. B. Blyn, A. Craiu, B. A. Braaten, J. Hardy, D. A. Low, and P. D. O'Hanley. 1981. DNA sequences of three PapA genes from uropathogenic *E. coli* strains: evidence of structural and serological conservation. *Infect. Immunol.* 59:3849–3858.

Dodson, K. W., F. Jacob-Dubuisson, R. T. Striker, and S. J. Hultgren. 1993. Outer-membrane PapC molecular usher discriminately recognizes periplasmic chaperone-pilus subunit complexes. *Proc. Natl. Acad. Sci. USA.* 90:3670–3674.

Dubochet, J., M. Adrian, J. Chang, J. C. Homo, J. Lepault, A. W. McDowall, and P. Schultz. 1988. Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.* 21:129–228.

Egelman, E. H., and A. Stasiak. 1986. Structure of helical RecA-DNA complexes. Complexes formed in the presence of ATP-gamma-S or ATP. *J. Mol. Biol.* 191:677–697.

Finer-Moore, J., R. M. Stroud, B. Prescott, and G. J. Thomas, Jr. 1984. Subunit secondary structure in filamentous viruses: predictions and observations. *J. Biomol. Struct. Dyn.* 2:93–100.

Fukami, A., and K. Adachi. 1965. A new method of preparation of a self-perforated micro plastic grid and its application. *J. Electron Microsc. (Tokyo).* 14:112–118.

Gittes, F., B. Mickey, J. Nettleton, and J. Howard. 1993. Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. *J. Cell Biol.* 120:923–934.

Gong, M., and L. Makowski. 1992. Helical structure of P pili from *Escherichia coli*. Evidence from X-ray fiber diffraction and scanning transmission electron microscopy. *J. Mol. Biol.* 228:735–742.

Jacob-Dubuisson, F., J. Heuser, K. Dodson, S. Normark, and S. Hultgren. 1993. Initiation of assembly and association of the structural elements of a bacterial pilus depend on two specialized tip proteins. *EMBO J.* 12:837–847.

Jones, C. H., J. S. Pinkner, R. Roth, J. Heuser, A. V. Nicholes, S. N. Abraham, and S. J. Hultgren. 1995. FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. *Proc. Natl. Acad. Sci. USA.* 92:2081–2085.

Källenius, G., R. Mollby, S. B. Svensson, J. Winberg, A. Lundblad, S. Svensson, and B. Cedergren. 1980. The P^k antigen receptor for haem-agglutinin of pyelonephritic *Escherichia coli*. *FEMS Microbiol. Lett.* 7:297–302.

Kishchenko, G., and L. Makowski. 1997. Shuffling of structural elements in filamentous bacteriophages. *Proteins.* 27:405–409.

Kuehn, M. J., J. Heuser, S. Normark, and S. J. Hultgren. 1992. P pili in uropathogenic *E. coli* are composite fibres with distinct fibrillar adhesive tips. *Nature.* 356:252–255.

Landau, L. D., and E. M. Lifschitz. 1958. Structural Physics: Course of Theoretical Physics. Pergamon, London.

Leffler, H., and C. Svanborg-Edén. 1980. Chemical identification of a glycosphingolipid receptor for *Escherichia coli* attaching to human urinary tract epithelial cells and agglutinating human erythrocytes. *FEMS Microbiol. Lett.* 8:127–134.

Lindberg, F., B. Lund, and S. Normark. 1986. Gene products specifying adhesion of uropathogenic *Escherichia coli* are minor components of pili. *Proc. Natl. Acad. Sci. USA.* 83:1891–1895.

Lund, B., F. Lindberg, B. I. Marklund, and S. Normark. 1987. The PapG protein is the α-D-galactopyranosyl-(1$\rightarrow$4)-β-D-galactopyranose-binding adhesin of uropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. USA.* 84:5898–5902.

Normark, S., D. Lark, R. Hull, M. Norgren, M. Båga, P. O'Hanley, G. Schoolnik, and S. Falkow. 1983. Genetics of digalactoside binding adhesin from a uropathogenic *Escherichia coli* strain. *Infect. Immunol.* 41:942–949.

Oosawa, F. 1980. The flexibility of F-actin. *Biophys. Chem.* 11:443–446.

Orlova, A., and E. H. Egelman. 1993. A conformational change in the actin subunit can change the flexibility of the actin filament. *J. Mol. Biol.* 232:334–341.

Parge, H. E., K. T. Forest, M. J. Hickey, D. A. Christensen, E. D. Getzoff, and J. A. Tainer. 1995. Structure of the fibre-forming protein pilin at 2.6 Å resolution. *Nature.* 378:32–38.

Reif, F. 1965. Fundamentals of Statistical and Theoretical Physics. McGraw-Hill, New York.

Rivetti, C., M. Guthold, and C. Bustamante. 1996. Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis. *J. Mol. Biol.* 264:919–932.

Smith, S. B., Y. Cui, and C. Bustamante. 1996. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science.* 271:795–799.

Solovyev, V. V., and A. A. Salamov. 1994. Predicting alpha-helix and beta-strand segments of globular proteins. *Comput. Appl. Biosci.* 10: 661–669.

Takebayashi, T., Y. Morita, and F. Oosawa. 1977. Electron microscopic investigation of the flexibility of F-actin. *Biochim. Biophys. Acta.* 492: 357–363.

Trifonov, E. N., R. K. Z. Tan, and S. C. Harvey. 1988. Static persistence length of DNA. *In* Structure and Expression. W. K. Olson, M. H. Sarma, R. H. Sarma, and M. S. Sundaralinjam, editors. Adenine Press, Inc., Albany, NY. 243–254.

Uhlin, B. E., M. Norgren, M. Båga, and S. Normark. 1985. Adhesion to human cells by *Escherichia coli* lacking the major subunit of a digalactoside-specific pilus-adhesin. *Proc. Natl. Acad. Sci. USA.* 82: 1800–1804.