J Mol Evol (1987) 25:261–269

Journal of
Molecular Evolution
© Springer-Verlag New York Inc. 1987

An Analysis of the Organization and Evolution of Type 4 Fimbrial (MePhe) Subunit Proteins

Brian Dalrymple and John S. Mattick

CSIRO Division of Molecular Biology, PO Box 184, North Ryde, NSW 2113, Australia

Summary. We have analyzed and compared the amino acid sequences of the type 4 fimbrial subunits from *Pseudomonas aeruginosa, Moraxella bovis, M. nonliquefaciens, Bacteroides nodosus, Neisseria gonorrhoeae,* and *N. meningitidis.* We propose a consensus sequence for the highly conserved amino-terminal regions of these proteins. In the variable regions, a domain corresponding to an epitope common to *N. gonorrhoeae* and *N. meningitidis* fimbriae is conserved, both in sequence and in environment, in fimbrial subunits from *B. nodosus.* The subunits from *M. bovis* and *P. aeruginosa* do not show any homologies to this sequence. In all of the subunits, the carboxy-terminal half of the molecule consists of a series of fairly hydrophobic domains. The last three domains, two of which include the cysteines of the disulfide bridge in *N. gonorrhoeae, P. aeruginosa,* and *M. bovis,* are more or less conserved in sequence in all of the proteins including that of *B. nodosus.* We propose that these conserved hydrophobic regions, which have the potential to form a series of beta-sheets, form a structural framework around which more variable hydrophilic sequences determining immunological profile are arranged. The evolutionary relationships of the contemporary proteins and the distribution of type 4 fimbriae are also discussed.

Key words: Fimbriae — Pili — Protein structure — Microbial phylogeny — Evolution

Introduction

A wide variety of bacteria produce long extracellular protein filaments termed fimbriae (or pili), which consist of several hundred or more predominantly identical subunits. A number of different types of fimbriae have been characterized (see Jones and Isaacson 1983). The members of one group of these fimbriae have the unusual amino acid *N*-methylphenylalanine as the first residue of the mature protein, followed by a highly conserved hydrophobic amino-terminal sequence (see Marrs et al. 1985). The first example of this group of fimbrial proteins was isolated from *Moraxella nonliquefaciens* (Froholm and Sletten 1977). Subsequently, similar proteins have been characterized from a wide range of gram-negative bacteria: *Neisseria gonorrhoeae* and *Neisseria meningitidis* (Hermodson et al. 1978), *Pseudomonas aeruginosa* (Frost et al. 1978; Paranchych et al. 1978), *Bacteroides nodosus* (McKern et al. 1983), and *Moraxella bovis* (Marrs et al. 1985). These fimbriae always have a polar location on the cells and mediate a form of bacterial motility known as twitching motility (Henrichsen 1975; Henrichsen and Blom 1975; Bradley 1980). A number of other bacterial species with these phenotypes have also been identified (Henrichsen 1975; see also Discussion). Fimbriae with these characteristics have been designated as type 4 (Ottow 1975; Mattick et al. 1987) or, in those cases where sequence data are available, the MePhe group (see Marrs et al. 1985).

Many of the species expressing type 4 fimbriae are mammalian pathogens. For example, *N. gonorrhoeae* causes venereal disease in man, *B. nodosus* causes footrot in sheep, and *M. bovis* causes keratoconjunctivitis in cattle. Others, such as *P. aeruginosa,* are opportunistic pathogens in a range of species. The fimbriae are probably involved in the attachment of bacteria to cells of the host during the course of infection (see Hermodson et al. 1978; Woods et al. 1980), and potential binding regions have been identified (Rothbard et al. 1985). The

Offprint requests to: B. Dalrymple

fimbriae are also prominent antigens, and vaccination with purified fimbriae can protect against homologous infection (Pugh et al. 1977; Brinton et al. 1978; Every and Skerman 1982). However, the fimbriae also exhibit considerable antigenic variation, apparently as a mechanism for avoiding host immune responses (Sandhu et al. 1974; Bradley and Pitt 1975; Buchanan 1975; Anderson et al. 1986). This variation is due to amino acid substitutions as well as small insertions or deletions in the carboxy-terminal two-thirds of the molecules (Hagblom et al. 1985; Marrs et al. 1985; McKern et al. 1985; Sastry et al. 1985a). Not surprisingly, perhaps, the variable regions of these proteins from different species appear to be unrelated (Schoolnik et al. 1984; Marrs et al. 1985). On the other hand, it has recently been shown that *P. aeruginosa* can process and assemble *B. nodosus* subunit proteins into fimbriae that are structurally and immunologically indistinguishable from authentic *B. nodosus* fimbriae (Mattick et al. 1987). This observation, taken together with the physical and functional similarities of the fimbriae, suggests that the subunits may have common structural features in their variable regions. To examine this possibility, and to explore in general terms the relationships between the type 4 fimbrial subunits expressed by this diverse range of bacteria, we have analyzed the structural and functional organization of the protein and the phylogenetic distribution of the system.

Methods

Primary protein sequences were taken from *M. bovis* alpha (Marrs et al. 1985) and *B. nodosus* E (McKern et al. 1985). Other protein sequences were translated from the DNA sequences of cloned fimbrial subunit genes as follows: *P. aeruginosa* PAK and PAO (Sastry et al. 1985a), *M. bovis* beta (Marrs et al. 1985), *N. gonorrhoeae* MS11 (Meyer et al. 1984), *N. gonorrhoeae* variants (Hagblom et al. 1985; Haas and Meyer 1986), and *B. nodosus* A (Elleman and Hoyne 1984). The unpublished DNA sequence of the *B. nodosus* F serotype (strain VCS 1017) subunit gene was provided by B. Anderson. The short amino-terminal sequences and amino acid compositions were taken from the following sources: *M. nonliquefaciens* NCTC7784 (Froholm and Sletten 1977), *N. meningitidis* ATC13090 (Hermodson et al. 1978), and *N. meningitidis* SP3428 (Olafson et al. 1985).

Amino acid sequences were aligned using dot matrix plots generated by the MTXDOT program (Reisner and Bucholtz 1986), and the significance of the alignments was calculated using the SEQDP program (Kanehisa et al. 1984). One hundred random sequences were generated for each pair, and a bias of zero and a deletion penalty of four were used. To compute the correlation statistics for amino acid composition of the variable regions of the proteins, the MTXAACOR program of the MTX package was used (Reisner and Bucholtz 1986). For the proteins for which the total amino acid sequence is unknown, but for which amino acid composition and amino-terminal sequence data are known, the amino acid composition of this sequence was from the total composition. Secondary structure plots were calculated using the program of Novotny and Auffray (1984) with the Kyte and Doolittle (1982) hydropathy parameters. To combine the predictions for two or more sequences, the unaligned amino acids were removed from the relevant sequences.

Results and Discussion

The available amino acid sequences of the type 4 fimbrial subunit proteins were aligned using two-dimensional dot matrix plots (Fig. 1). The proteins can be divided into three regions (I–III), which will be described in more detail below. The first is highly conserved, and we have derived a consensus sequence representative of the different type 4 subunits (Fig. 1). In an attempt to reduce bias in the derivation of the consensus sequence for region I, due regard was given to the fact that *Neisseria* and *Moraxella* are both members of the family Neisseriaceae (see Bovre 1984).

Region I itself can be further divided into three domains (regions Ia, Ib, and Ic) on the basis of amino acid composition and level of sequence conservation. Region Ia comprises a short positively charged, partially conserved hydrophilic leader peptide. These amino acids are not present in mature subunits and presumably play a role in the transport of the subunit to the site of assembly. The next amino acid, phenylalanine, is usually methylated in mature subunits (Frost et al. 1978; Hermodson et al. 1978). Some of the conserved amino acids may constitute the recognition site, or sites, for these processes.

The *N*-methylphenylalanine is the first residue of region Ib, a 27-amino acid hydrophobic region that is extremely conserved between all of the fimbrial subunits. Region Ib appears to play an important role in intersubunit contacts and maintenance of the structure of the mature fimbrial strands (see Watts et al. 1983; Schoolnik et al. 1984; Sastry et al. 1985b). Furthermore, regions Ia and Ib together may contain all of the information required for the processing of the subunits and assembly of the fimbriae (see Mattick et al. 1987). Thus, the extreme conservation of this entire region probably stems from the pressure of recognition of the components by the assembly system, and the correct polymerization of the subunits into mature fimbriae.

Region Ic extends from amino acid 28 to the glycine at residue 54 (55 in *N. gonorrhoeae*). Within a species, region Ic is almost entirely conserved, with no more than three changes observed (Fig. 1). The only exceptions are the sequences predicted from a number of the *N. gonorrhoeae* incomplete fimbrial gene copies that occur in the *pilS1* locus (Haas and Meyer 1986). However, these related sequences appear to constitute a reservoir of variable information, and there is no evidence that their Ic regions

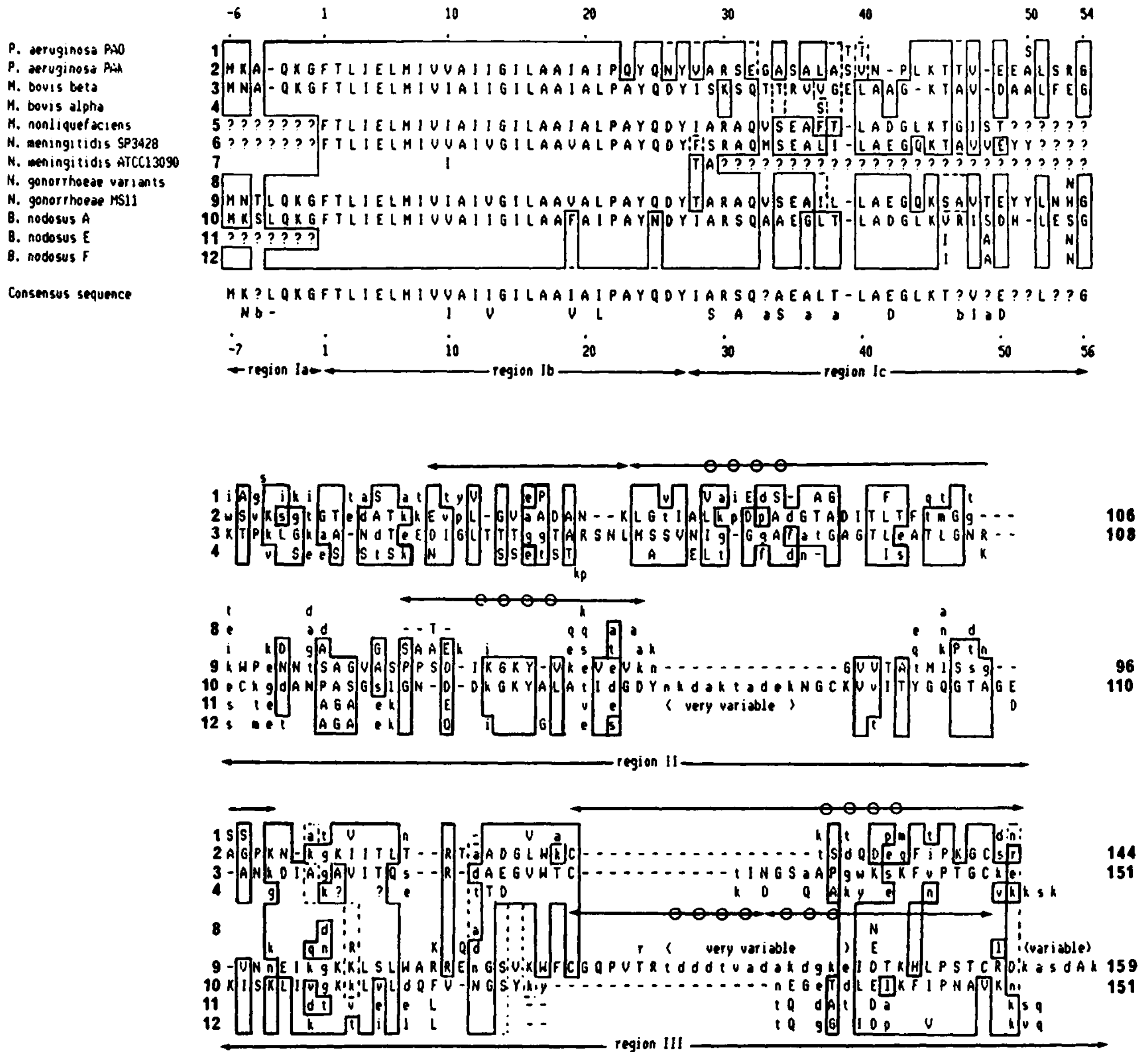

Fig. 1. Alignment of the sequences of the type 4 fimbrial subunit proteins. The sequence of one protein from each species is given in full; for the other proteins only residues that are different from the first sequence are shown. For *N. gonorrhoeae* variant proteins, all of the observed amino acid substitutions observed are shown, but no particular subunit protein is represented. Dashes indicate gaps introduced to maximize the alignments and question marks indicate unknown residues. Insertions are shown just above or below the relevant sequence. Region I: in the consensus sequence the amino acids most likely to be present in the (putative) ancestral protein are shown, and where there are two amino acids listed, the alternatives are considered equally likely. The lowercase letters indicate positions where a number of related amino acids are present: (a) small or hydrophobic amino acids, (b) small amino acids, and (c) charged or hydrophilic amino acids. All residues homologous to the consensus sequence are boxed, and conservative substitutions are boxed with dashed lines. Amino acid substitutions that have a positive score in the 250 PAMs substitution matrix (Dayhoff et al. 1978) were counted as conservative. The coordinates above the sequences are for *P. aeruginosa* and those below the consensus sequence are the standard positions (SP). Regions II and III: conserved residues and conservative substitutions in all of the examples of proteins from one species are in capital letters and nonconservative substitutions are in lowercase letters. For *N. gonorrhoeae*, positions in the MS11 sequence that are rarely substituted with a nonconservative amino acid are also in capital letters. Conserved residues and conservative substitutions between the pairs *P. aeruginosa–M. bovis* and *N. gonorrhoeae–B. nodosus*, or between three or more of the species are boxed. Subsets of conserved amino acids are boxed in dashed lines. The amino acid sequences of the very variable regions are only shown for the protein listed in full. The arrows over the sequences indicate the extent of some of the peptides, or fragments of proteins, used to determine the epitopes. The open circles indicate the probable location of the epitopes (see Marrs et al. 1985; Rothbard et al. 1985; Sastry et al. 1985b).

are ever actually translated into mature proteins (Haas and Meyer 1986; Segal et al. 1986). The sequences of the proteins from the closely related species *N. gonorrhoeae* and *N. meningitidis* are also very similar. There is more variability among the other species. Nevertheless, a number of residues are conserved or semiconserved between all of the proteins, although small gaps have to be introduced

to maximize the alignments (Fig. 1). Aligned positions are referred to as SP (standard position). At some positions in region Ic (such as SP33, 36, 38, and 49), a number of different amino acids are represented, but these amino acids are either small (G, A, T, or S) or hydrophobic (M, V, I, L, or F). Indeed, only SP51, 52, 54, and 55 can apparently accommodate a wide variety of amino acids. Studies with *N. gonorrhoeae* proteins have shown that region Ic contains a number of epitopes and may also be involved in fimbrial binding (Rothbard et al. 1984, 1985), with different fimbrial variants possibly exhibiting different binding properties (Lambden et al. 1980). Fimbriae from different species also bind to different cell types or receptors (Schoolnik et al. 1983). Thus, the somewhat limited variability of the sequence of region Ic may reflect the different organizations required for different binding specificities, while the conserved amino acids (such as a hydrophobic residue at SP40, lysine at SP45, hydrophobic residue at SP48, leucine at SP53, and glycine at SP56) may provide a structural framework or fulfill other functions in this region.

Finlay et al. (1986) have proposed that regions Ib and Ic may form a pair of antiparallel helices during the transport of the *P. aeruginosa* subunit across the cell membranes. These workers have also proposed that salt bridges, which would reduce the free energy required to bury the region in the membrane (Engelman and Steitz 1981), may occur in region Ic of *P. aeruginosa* PAK protein. Although different, the sequences of the fimbrial subunits from other species are consistent with this idea (e.g., the conserved glutamic or aspartic acid at SP42 and lysine at SP45), and some of the conservation observed in charged residues in region Ic may be due to such a requirement.

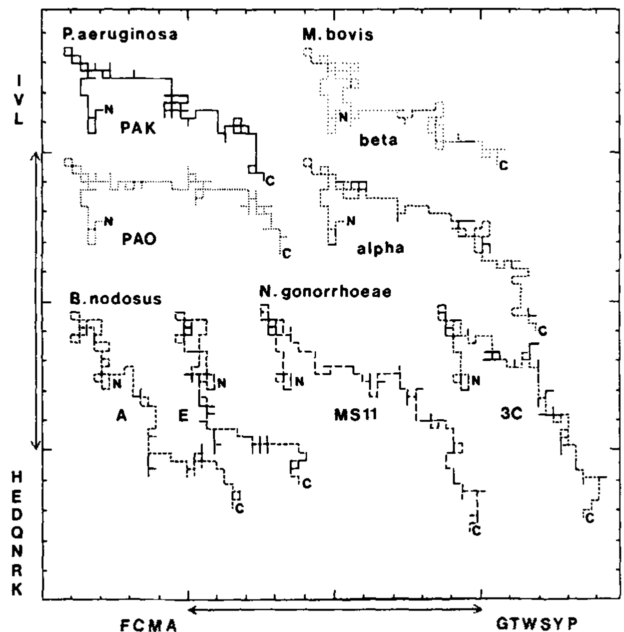

Fig. 2. Two-dimensional hydropathy plots of the complete fimbrial subunit amino acid sequences. The plots were calculated using the GPLINE program (Reisner and Bucholtz, unpublished). This program plots a line of unit length for each amino acid in one of four directions depending on the nature of the amino acid. The four groups of amino acids are shown on the axes of the plot. The amino-terminal amino acids were placed arbitrarily on the surface to ensure separated plots.

Comparison of the Variable Regions of Proteins from the Same Species

The glycine at residue 54 [55 in *N. gonorrhoeae* (SP 56)] appears to represent the border between the conserved amino-terminal region I and regions II and III, which together encompass the variable carboxy-terminal two-thirds of the protein (see Fig. 1). To determine the degree of similarity between the variable regions of the proteins from different serotypes of the same species, the amino acid sequence alignments were analyzed statistically using the SEQDP program (Kanehisa et al. 1984). The *N. gonorrhoeae* proteins are the most closely related group of those analyzed (≈28 SD), while the two representatives of each of the other species were more divergent but still very significantly related (>20 SD). In addition, the two-dimensional hydropathy plots of the amino acid sequences were clearly similar for each of the pairs of serological variants from one species and distinct for each of the different species (Fig. 2). Circular dichroism studies of *P. aeruginosa* PAK and PAO fimbrial proteins suggested that both subunits have very similar secondary structures (Watts et al. 1982). The variants of *N. gonorrhoeae* proteins have also been predicted, from the amino acid sequences, to have similar secondary structures (Schoolnik et al. 1984). Our own analyses of secondary structure predicted from amino acid sequences (data not shown) indicate that this is also the case for the fimbrial subunits of *B. nodosus* and *M. bovis*. Therefore, by the criteria of amino acid sequence alignments, hydropathy, and predicted secondary structure, it seems that, despite the extensive sequence differences, variant proteins from the same species have very similar overall structural organizations.

Comparison of the Variable Regions of Proteins from Different Species

The *P. aeruginosa* and *M. bovis* amino acid sequences can be aligned with each other throughout the variable regions with little need for compensating deletions or insertions (Fig. 1). In contrast, while the *N. gonorrhoeae* and *B. nodosus* proteins align better with each other than with either the *P. aeruginosa* or *M. bovis* proteins, large insertions and

deletions are required for optimal matching (Fig. 2). Perhaps as a consequence of these gaps, only the *P. aeruginosa* and *M. bovis* alignments are significantly better than the mean of random alignments with the same amino acid composition (>7 SD). Amino acid composition correlation coefficients provide a different measure of the potential gross relationship of protein sequences. This approach also allows the inclusion of proteins for which only the amino acid composition is known. Correlation coefficients between all of the proteins were calculated using the amino acid composition of the variable regions (data not shown). Proteins from different serotypes of one species were clearly related ($r > 0.7$), especially the *N. gonorrhoeae* proteins ($r > 0.8$). However, the amino acid compositions of the *N. meningitidis* and *M. nonliquefaciens* proteins did not appear to be significantly related to that of any of the other proteins. The only significant correlations in amino acid composition between proteins produced by different species were between the examples from *P. aeruginosa* and *M. bovis* ($r \approx 0.7$).

Although the *P. aeruginosa* and *M. bovis* proteins have a number of apparently conserved amino acids and conservative substitutions in region II (Fig. 1), no significant homologies were highlighted by the dot matrix plots (data not shown). Indeed, statistical analysis of the significance of the alignments of the sequences in region II showed that they are no more significant than random alignments of sequences with the same amino acid composition. Thus, the similarities between these proteins may be due as much, or more, to conservation of particular types of amino acids rather than to sequence conservation per se. However, there is a demonstrable relationship in this region between the *N. gonorrhoeae* and *B. nodosus* proteins, specifically the sequence (D/E)-KGKY between amino acids 72 and 77 (*N. gonorrhoeae* MS11) and 70 and 75 (*B. nodosus* A). In both the *N. gonorrhoeae* and *B. nodosus* proteins, this sequence lies in an almost identical position relative to the conserved glycine SP56. Additionally, in both cases this region is predicted to be alpha-helical and to be preceded by a region with a high beta-turn probability (Fig. 3). Thus, the secondary structure and environment of the conserved sequences are probably quite similar. The conserved sequence is located within the *N. gonorrhoeae* peptide 69–84 and is the most conserved region of the peptide (Hagblom et al. 1985; Haas and Meyer 1986). This peptide contains an immunorecessive epitope common to all gonococcal fimbrial serotypes and many serotypes of *N. meningitidis* (Schoolnik et al. 1984; Stephens et al. 1985), and is also involved in the binding of fimbriae to human endocervical and endometrial carcinoma cells (Rothbard et al. 1985). It is probable that the conserved sequence is a major part of the common epitope and is involved in the binding of the fimbriae to cells. The equivalent region in *B. nodosus* fimbriae may perhaps have a similar role, although binding of these fimbriae to cells has not been studied.

In all of the proteins, region III contains a number of fairly hydrophobic domains that are more or less conserved in sequence (Fig. 1). Surprisingly, even the *B. nodosus* proteins have significant homology to the other proteins in this region, despite the different location of the cysteine residues. In contrast, there is no amino acid sequence homology between the sequences surrounding the cysteines in *B. nodosus* and the sequences surrounding the cysteines in the other proteins. Thus, the different location of the disulfide bridge in the *B. nodosus* proteins does not appear to be a consequence of a gross rearrangement of the DNA or protein sequences. This observation was recently confirmed (after submission of this manuscript) with the publication of the sequence of a novel *B. nodosus* subunit gene that encodes a protein with cysteines in positions equivalent to the *P. aeruginosa* and *M. bovis* proteins (Elleman et al. 1986).

Structural and Functional Relationships

We have examined the secondary structures of all of the proteins, as predicted from their amino acid sequences, to identify potential secondary structure homologies among proteins from different species. The composite predictions of each species were aligned with the sequences of all of the proteins (data not shown). In several regions, the same secondary structure was predicted for the subunit proteins for more than one species (see Fig. 3). In the carboxy-terminal half of all of the proteins there appear to be a number of beta-sheets separated by turns (Fig. 3). Three of the potential beta-sheet regions show extensive sequence homology among the proteins (Fig. 1). Potential alpha-helices lie in region II and the disulfide loops and equivalent regions. The overall composition of the predicted secondary structure is not inconsistent with the prediction from circular dichroism studies that the *P. aeruginosa* proteins contain 40–55% alpha-helix and about 40% beta-sheet (Folkhard et al. 1981; Watts et al. 1982, 1983), although alpha-helix is clearly underrepresented in the sequence-derived predictions.

Interestingly, while the *N. gonorrhoeae* proteins have only an immunorecessive epitope in region II, the major serotype-specific epitope of *P. aeruginosa* PAK appears to lie between amino acids 82 and 110 (a turn region), with a minor epitope between amino acids 70 and 81 (Sastry et al. 1985b). From our alignments of the sequences (Fig. 1), the *N. gonorrhoeae* proteins are predicted to have a deletion in

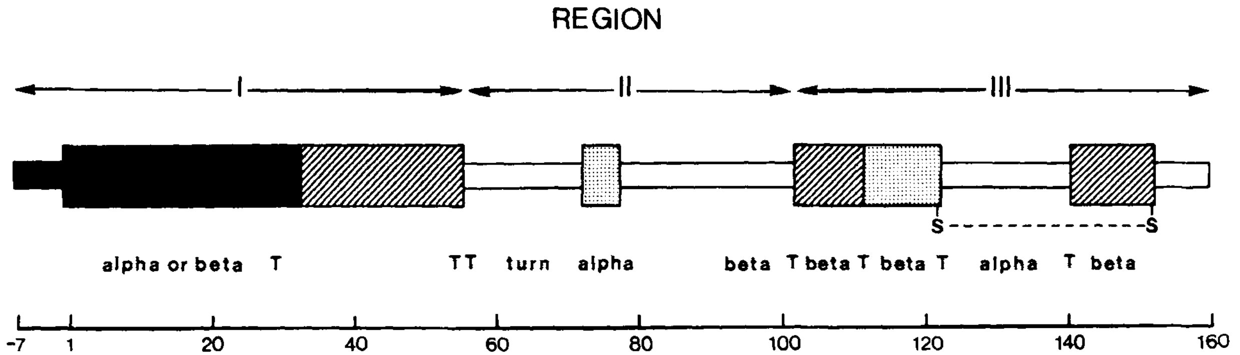

Fig. 3. The general organization of type 4 fimbrial subunit proteins. Within region I, the thinner dark block represents the partially conserved six- to seven-amino acid leader sequence. The thicker dark block represents the highly conserved first 32 amino acids of the mature proteins, including the strongly hydrophobic region Ib. The shaded areas indicate regions that, although they show some divergence, are broadly conserved in all of the subunit proteins. The dotted regions are conserved in some but not all of the proteins. The other areas are not conserved between proteins from different species. The consensus secondary structure prediction for all of the proteins is shown where there is a high probability for that structure compared with any other structure: alpha-predicted alpha-helix, beta-beta sheet, T-turn. Regions without a designation had an ambiguous secondary structure prediction. The location of the disulfide bridge in *N. gonorrhoeae, P. aeruginosa,* and *M. bovis* is marked S---S. The coordinates shown refer to the *N. gonorrhoeae* MS11 sequence.

the region equivalent to the major *P. aeruginosa* PAK epitope. However, the deletion in the *N. gonorrhoeae* proteins is balanced by a large (30 or more amino acids) disulfide loop that contains the major serotype-specific epitopes (Rothbard et al. 1985), while the small *P. aeruginosa* loop contains only a minor epitope (Sastry et al. 1985b). *Moraxella bovis* apparently also has a serotype-specific epitope in the disulfide loop (see Marrs et al. 1985). Little is known about the epitopes of *B. nodosus* fimbriae, but the proteins have a hydrophilic region equivalent to the disulfide loop, and, like *M. bovis* and *P. aeruginosa,* have no deletion in region II. From the sequence alignments and the data from the other proteins, we would predict that the *B. nodosus* proteins carry serotype-specific epitopes both in region II, just preceding the second cysteine residue, and in region III, around amino acid 135.

The subunits of *P. aeruginosa* type 4 fimbriae are stacked into a very regular polymeric quaternary structure. The strands have a hydrophobic central region with a hydrophilic exterior surface that appears to be comprised of alpha-helices arranged approximately parallel to the axis of the fimbriae (Folkhard et al. 1981). The observations presented in this section suggest that the primary epitopes, and probably also the binding sites, are contained within surface alpha-helical domains or turns of variable sequence (primarily region Ic, region II, and the disulfide loops), which are themselves arranged around a conserved core of beta-sheets of less variable sequence (primarily region III). Therefore, despite the differences that exist among type 4 fimbriae from different species, these proteins would appear to have a similar overall organization and structure. This is also consistent with the fact that fimbriae from these species are physically very similar, and that one subunit may be successfully substituted for another as exemplified by the morphogenetic expression of *B. nodosus* fimbriae in *P. aeruginosa* (Mattick et al. 1987).

Evolution of Type 4 Fimbriae

The structural analyses and sequence alignments presented above allow us to consider the evolutionary relationships among type 4 fimbriae, a subject that has not hitherto been addressed. The sequence of region Ic varies little among representatives of the same species (Fig. 1), and also varies among species much less than the carboxy-terminal two-thirds of the proteins (Fig. 1). Thus, organisms producing proteins with closely related sequences in this region would be expected to be related phylogenetically and vice versa. In support of this, *N. gonorrhoeae* and *N. meningitidis,* which are closely related members of the same genus (Kingsbury et al. 1969), encode antigenically related proteins with similar amino-terminal sequences (Virji and Heckels 1983; Olafson et al. 1985; Stephens et al. 1985). Thus, we would also expect the overall organization of the *N. meningitidis* proteins to be similar to that of the *N. gonorrhoeae* proteins. The family Neisseriaceae also includes the genus *Moraxella* (see Bovre 1984), within which *M. nonliquefaciens* and *M. bovis* are closely related (Bovre 1970). However, although in region Ic the *M. nonliquefaciens* protein has good sequence homology with the *Neisseria* spp. proteins and the consensus sequence, the *M. bovis* proteins exhibit a much weaker relationship. Furthermore, except in region III, the variable region of the *M. bovis* proteins is not very similar to the variable region of the *N. gonorrhoeae* proteins. Thus, considerable change in sequence has occurred since

the division of the two lines, although in region Ic this has presumably occurred after the branching of the lines to *M. nonliquefaciens* and *M. bovis.*

The Neisseriaceae (represented by the genus *Acinetobacter*) and *P. aeruginosa* are in the gamma-3 subdivision of the purple bacteria, although *P. aeruginosa* is only distantly related to *Acinetobacter* (Woese et al. 1985). The similarities in region II of *P. aeruginosa* and *M. bovis,* which are not very significant, may reflect the vestiges of a common ancestral sequence or similar but distinct evolutionary pressures. However, the conservation of much of region III between the *M. bovis* and *P. aeruginosa* proteins suggests that this organization may be the most similar to that of the (putative) ancestral protein.

Phylogenetically, *B. nodosus* is thought to be more distantly related to *N. gonorrhoeae* than *N. gonorrhoeae* is to *M. bovis* and *P. aeruginosa,* although *B. nodosus* is probably not a true *Bacteroides* (Shah and Collins 1983) and is therefore of uncertain phylogenetic position. However, the *N. gonorrhoeae* and *B. nodosus* proteins share a domain of apparently conserved sequence and structure in region II, for which there is no equivalent in the *M. bovis* and *P. aeruginosa* proteins. Thus, the conserved features of region II may reflect the ancestral organization of the protein, which has subsequently been independently lost from the lines of descent leading to *M. bovis* and *P. aeruginosa.* Interestingly, of the three most distinct sequences in region Ic (as represented by *M. bovis, P. aeruginosa,* and the *Neisseria* species; Fig. 1), the *B. nodosus* sequence is most similar to the *Neisseria* species, and to a certain extent this is also true in region III. This observation suggests that this sequence in region I may be the closest to the ancestral or that, alternatively, *B. nodosus* may be a misclassified member of the Neisseriaceae possibly related to the *Neisseria* species, in which case the conservation in region II may represent a recently evolved feature. There is also the possibility that exchange of fimbrial subunit genes, or subsections of the information therein, may have occurred by lateral transfer between compatible species that express type 4 fimbriae. However, there is no good evidence for a recent horizontal transfer of the entire type 4 fimbrial system. Species expressing polar fimbriae and the associated twitching motility phenotype are found in all of the genera in the Neisseriaceae and in the majority of species that have been tested. Such fimbriae are also widely distributed in the *Pseudomonas* species that are closely related to *P. aeruginosa* (see Palleroni et al. 1973; Henrichsen 1975; Henrichsen and Blom 1975). The position is not so clear for *B. nodosus,* but several other species also presently classified in the genus *Bacteroides* and related genera (but which are not true *Bacteroides*) probably express type 4 fimbriae (see below).

Indeed, based on the association of the polar location/twitching motility phenotype with the conserved amino terminal amino acid sequence, it appears that type 4 fimbriae are in fact widely distributed in the beta (Woese et al. 1984) and gamma subdivisions of the purple bacteria. In the gamma subdivision these species include *Xanthomonas* (ex *Pseudomonas*) *maltophilia, Alteromonas* (ex *Pseudomonas*) *putrefaciens,* and *Pasteurella multocida* (Henrichsen 1975; Henriksen and Froholm 1975). In the beta subdivision these species include *P. testosteroni* and *P. acidovorans* (Henrichsen 1975; Dalrymple, unpublished). Gram-negative bacterial species of unknown phylogenetic position, which have fimbriae with type 4 characteristics, include *Eikenella corrodens, B. ureolyticus,* and *B. gracilis* (Tanner et al. 1981; Henrichsen 1983). These species also have the characteristic of corroding the surface of agar plates, which is a feature of many type 4 fimbriate bacteria (see Froholm and Bovre 1972; Henrichsen et al. 1972; Henriksen and Froholm 1975). A number of species classified as *Wolinella* that are anaerobic gram-negative mammalian pathogens (Tanner et al. 1981, 1984) and presently unclassified nongonococcal urethritis strains have also been reported to have twitching motility and agar-corroding activity (Fontaine et al. 1984), and on this basis probably also have type 4 fimbriae. A similar phenotype is also exhibited by a number of strains isolated from soft tissue infections of cats and dogs (Love et al. 1984). From the known phylogenetic positions of type 4 fimbriate bacteria, the evidence suggests that these fimbriae were present in the ancestor of the beta and gamma subdivisions of the purple bacteria, and it is possible that the less well characterized species presently classified in the genera *Bacteroides, Eikenella,* and *Wolinella* may also lie within this group. Some isolates of the gram-positive species *Streptococcus sanguis* (Henriksen and Henrichsen 1975) may have type 4 fimbriae, which suggests that this system may be even more ancestral, being kept in those species whose ecology may require colonization of eukaryotic cells.

Acknowledgments. This work was supported in part by the Australian Meat and Livestock Research and Development Corporation. We thank Alex Reisner and Carolyn Bucholtz for their help with the computing, Margaret Dowell and Anne McGill for the typing, and Beau Anderson for providing the sequence of the *B. nodosus* serotype F fimbrial subunit prior to publication.

References

Anderson BJ, Kristo CL, Egerton JR, Mattick JS (1986) Variation in the structural subunits and basal protein antigens of *Bacteroides nodosus* fimbriae. J Bacteriol 166:453–460.

Bovre K (1970) Pulse-RNA-DNA hybridization between rod-shaped and coccal species of the *Moraxella-Neisseria* groups. Acta Pathol Microbiol Scand Sect B Microbiol 78:565–574

Bovre K (1984) Neisseriaceae. In: Kreig NR, Holt JG (eds) Bergey's manual of systematic bacteriology. Williams and Wilkins, Baltimore, p 288

Bradley DE (1980) A function of *Pseudomonas aeruginosa* PAO pili: twitching motility. Can J Microbiol 26:146–154

Bradley DE, Pitt TL (1975) An immunological study of the pili of *Pseudomonas aeruginosa*. J Hyg Camb 74:419–430

Brinton CC, Bryan J, Dillon JA, Guerina N, Jacobson LJ, Labik A, Lee S, Levine A, Lim S, McMichael J, Polen S, Rogers K, To AC-C, To SC-M (1978) Uses of pili in gonorrhea control: role of bacterial pili in disease, purification and properties of gonococcal pili, and progress in the development of a gonococcal pilus vaccine for gonorrhea. In: Brooks GF, Gotschlich EC, Holmes KK, Sawyer WS, Young FE (eds) Immunobiology of *Neisseria gonorrhoeae*. American Society for Microbiology, Washington DC, p 155

Buchanan TM (1975) Antigenic heterogeneity of gonococcal pili. J Exp Med 141:1470–1475

Dayhoff MO, Schwartz RM, Orcutt BC (1978) A model of evolutionary change in proteins. In: Dayhoff MO (ed) Atlas of protein sequence and structure, vol 5. National Biomedical Research Foundation, Washington DC, p 345

Elleman TC, Hoyne PA (1984) Nucleotide sequence of the gene encoding pilin of *Bacteroides nodosus*, the causal organism of ovine footrot. J Bacteriol 160:1184–1187

Elleman TC, Hoyne PA, McKern NM, Stewart DJ (1986) Nucleotide sequence of the gene encoding the two-subunit pilin of *Bacteroides nodosus* 265. J Bacteriol 167:243–250

Engelman DM, Steitz TA (1981) The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis. Cell 23:411–422

Every D, Skerman TM (1982) Protection of sheep against experimental footrot by vaccination with pili purified from *Bacteroides nodosus*. NZ Vet J 30:156–158

Finlay BB, Pasloske BL, Paranchych W (1986) Expression of the *Pseudomonas aeruginosa* PAK pilin gene in *Escherichia coli*. J Bacteriol 165:625–630

Folkhard W, Marvin DA, Watts TH, Paranchych W (1981) Structure of polar pili from *Pseudomonas aeruginosa* strains K and O. J Mol Biol 149:79–93

Fontaine EAR, Borriello SP, Taylor-Robinson D, Davies HA (1984) Differential characteristics of a small gram-negative anaerobe associated with non-gonococcal urethritis which morphologically resembles *Bacteroides ureolyticus*. Scand J Urol Nephrol Suppl 86:157–165

Froholm LO, Bovre K (1972) Fimbriation associated with the spreading-corroding colony type in *Moraxella kingii*. Acta Pathol Microbiol Scand Sect B Microbiol 80:641–648

Froholm LO, Sletten K (1977) Purification and N-terminal sequence of a fimbrial protein from *Moraxella nonliquefaciens*. FEBS Lett 73:29–32

Frost LS, Carpenter M, Paranchych W (1978) N-methylphenylalanine at the N-terminus of pilin isolated from *Pseudomonas aeruginosa* K. Nature 271:87–89

Haas R, Meyer TF (1986) The repertoire of silent pilus genes in *Neisseria gonorrhoeae*: evidence for gene conversion. Cell 44:107–115

Hagblom P, Segal E, Billyard E, So M (1985) Intragenic recombination leads to pilus antigenic variation in *Neisseria gonorrhoeae*. Nature 315:156–158

Henrichsen J (1975) The occurrence of twitching motility among gram-negative bacteria. Acta Pathol Microbiol Scand Sect B Microbiol 83:171–178

Henrichsen J (1983) Twitching motility. Annu Rev Microbiol 37:81–93

Henrichsen J, Blom J (1975) Examination of fimbriation of some gram-negative rods with and without twitching and gliding motility. Acta Pathol Microbiol Scand Sect B Microbiol 83:161–170

Henrichsen J, Froholm LO, Bovre K (1972) Studies on bacterial surface translocation 2. Correlation of twitching motility and fimbriation in colony variants of *Moraxella nonliquefaciens, M. bovis* and *M. kingii*. Acta Pathol Microbiol Scand Sect B Microbiol 80:445–452

Henriksen SD, Froholm LO (1975) A fimbriated strain of *Pasteurella multocida* with spreading and corroding colonies. Acta Pathol Microbiol Scand Sect B Microbiol 83:129–132

Henriksen SD, Henrichsen J (1975) Twitching motility and possession of polar fimbriae in spreading *Streptococcus sanguis* isolates from the human throat. Acta Pathol Microbiol Scand Sect B Microbiol 83:133–140

Hermodson MA, Chen KCS, Buchanan TM (1978) *Neisseria* pili proteins: amino-terminal amino acid sequences and identification of an unusual amino acid. Biochemistry 17:442–445

Jones GW, Isaacson RE (1983) Proteinaceous bacterial adhesins and their receptors. Int Rev Microbiol 10:229–260

Kanehisa M, Klein P, Greif P, DeLisi C (1984) Computer analysis and structure prediction of nucleic acids and proteins. Nucleic Acids Res 12:417–428

Kingsbury DT, Fanning GR, Johnson KE, Brenner DJ (1969) Thermal stability of interspecies *Neisseria* DNA duplexes. J Gen Microbiol 55:201–208

Kyte J, Doolittle RE (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132

Lambden PR, Robertson JN, Watt PJ (1980) Biological properties of two distinct pilus types produced by isogenic variants of *Neisseria gonorrhoeae* P9. J Bacteriol 141:393–396

Love DN, Jones RF, Bailey M, Calverley A (1984) Comparison of strains of gram-negative, anaerobic, agar-corroding rods isolated from soft tissue infections in cats and dogs with type strains of *Bacteroides gracilis, Wolinella recta, Wolinella succinogenes* and *Campylobacter concisus*. J Clin Microbiol 20: 747–750

Marrs CF, Schoolnik G, Koomey JM, Hardy J, Rothbard J, Falkow S (1985) Cloning and sequencing of a *Moraxella bovis* pilin gene. J Bacteriol 163:132–139

Mattick JS, Bills MM, Anderson BJ, Dalrymple B, Mott MR, Egerton JR (1987) Morphogenetic expression of *Bacteroides nodosus* fimbriae in *Pseudomonas aeruginosa*. J Bacteriol 169: 33–41

McKern NM, O'Donnell IJ, Inglis AS, Stewart DJ, Clark BL (1983) Amino acid sequence of pilin from *Bacteroides nodosus* (strain 198), the causative agent of ovine footrot. FEBS Lett 164:149–153

McKern NM, O'Donnell IJ, Stewart DJ, Clark BL (1985) Primary structure of pilin protein from *Bacteroides nodosus* strain 216: comparison with corresponding protein from strain 198. J Gen Microbiol 131:1–6

Meyer TF, Billyard E, Haas R, Storzbach S, So M (1984) Pilus genes of *Neisseria gonorrheae*: chromosomal organization and DNA sequence. Proc Natl Acad Sci USA 81:6110–6114

Novotny J, Auffray C (1984) A program for prediction of protein secondary structure from nucleotide sequence data: application to histocompatibility antigens. Nucleic Acids Res 12:243–255

Olafson RW, McCarthy PJ, Bhatti AR, Dooley JSG, Heckels JE, Trust TJ (1985) Structural and antigenic analysis of meningococcal piliation. Infect Immun 48:336–342

Ottow JCG (1975) Ecology, physiology, and genetics of fimbriae and pili. Annu Rev Microbiol 29:79–108

Palleroni NJ, Kunisawa R, Contopoulou R, Doudoroff M (1973)

Nucleic acid homologies in the genus *Pseudomonas*. Int J Syst Bacteriol 23:333–339

Paranchych W, Frost LS, Carpenter M (1978) N-terminal amino acid sequence of pilin isolated from *Pseudomonas aeruginosa*. J Bacteriol 134:1179–1180

Pugh GW, Hughes DE, Booth GD (1977) Experimentally induced infectious bovine keratonconjunctivitis: effectiveness of a pilus vaccine against exposure to homologous strains of *Moraxella bovis*. Am J Vet Res 38:1519–1522

Reisner AH, Bucholtz CA (1986) The MTX package of computer programmes for the comparison of sequences of nucleotides and amino acid residues. Nucleic Acids Res 14:233–238

Rothbard JB, Fernandez R, Schoolnik GK (1984) Strain-specific and common epitopes of gonococcal pili. J Exp Med 160: 208–221

Rothbard JB, Fernandez R, Wang L, Teng NNH, Schoolnik GK (1985) Antibodies to peptides corresponding to a conserved sequence of gonococcal pilins block bacterial adhesion. Proc Natl Acad Sci USA 82:915–919

Sandhu TS, White FH, Simpson CF (1974) Association of pili with rough colony type of *Moraxella bovis*. Am J Vet Res 35: 437–439

Sastry PA, Finlay BB, Pasloske BL, Paranchych W, Pearlstone JR, Smillie LB (1985a) Comparative studies of the amino acid and nucleotide sequences of pilin derived from *Pseudomonas aeruginosa* PAK and PAO. J Bacteriol 164:571–577

Sastry PA, Pearlstone JR, Smillie LB, Paranchych W (1985b) Studies on the primary structure and antigenic determinants of pilin isolated from *Pseudomonas aeruginosa* K. Can J Biochem Cell Biol 63:284–291

Schoolnik GK, Tai JY, Gotschlich EC (1983) A pilus peptide vaccine for the prevention of gonorrhea. Prog Allergy 33:314–331

Schoolnik GK, Fernandez R, Tai JY, Rothbard J, Gotschlich EC (1984) Gonococcal pili primary structure and receptor binding domain. J Exp Med 159:1351–1370

Segal E, Hagblom P, Seifert HS, So M (1986) Antigenic variation of gonococcal pilus involves assembly of separated silent gene segments. Proc Natl Acad Sci USA 83:2177–2181

Shah HN, Collins MD (1983) Genus *Bacteroides*. A chemotaxonomical perspective. J Appl Bacteriol 55:403–416

Stephens DS, Whitney AM, Rothbard J, Schoolnik GK (1985) Pili of *Neisseria meningitidis*: analysis of structure and investigation of structural and antigenic relationships to gonococcal pili. J Exp Med 161:1539–1553

Tanner ACR, Badger S, Lai CH, Listgarten MA, Visconti RA, Socransky SS (1981) *Wolinella* gen. nov., *Wolinella succinogenes* (*Vibrio succinogenes* Wolin et al.) comb. nov., and description of *Bacteroides gracilis* sp. nov., *Wolinella recta* sp. nov., *Campylobacter concisius* sp. nov., and *Eikenella corrodens* from humans with periodontal disease. Int J Syst Bacteriol 31:432–445

Tanner ACR, Listgarten MA, Ebersole JL (1984) *Wolinella curva* sp. nov.: "*Vibrio succinogenes*" of human origin. Int J Syst Bacteriol 34:275–282

Virji M, Heckels JE (1983) Antigenic cross-reactivity of *Neisseria* pili: investigations with type- and species-specific monoclonal antibodies. J Gen Microbiol 129:2761–2768

Watts TH, Kay CM, Paranchych W (1982) Dissociation and characterization of pilin isolated from *Pseudomonas aeruginosa* strains PAK and PAO. Can J Biochem 60:867–872

Watts TH, Kay CM, Paranchych W (1983) Spectral properties of three quaternary arrangements of *Pseudomonas* pilin. Biochemistry 22:3640–3646

Woese CR, Weisburg WG, Paster BJ, Hahn CM, Tanner RS, Krieg NR, Koops HP, Harms H, Stackebrandt E (1984) The phylogeny of the purple bacteria: the beta subdivision. Syst Appl Microbiol 5:327–336

Woese CR, Weisburg WG, Hahn CM, Paster BJ, Zablen LB, Lewis BJ, Macke TJ, Ludwig W, Stackebrandt E (1985) The phylogeny of the purple bacteria: the gamma subdivision. Syst Appl Microbiol 6:25–33

Woods DE, Straus DC, Johanson WG Jr, Berry VK, Bass JA (1980) Role of pili in adherence of *Pseudomonas aeruginosa* to mammalian buccal epithelial cells. Infect Immun 29:1146–1151

Received August 25, 1986/Revised November 24, 1986