ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/6730578

Evidence for a novel domain of bacterial outer membrane ushers

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · DECEMBER 2006

Impact Factor: 2.63 · DOI: 10.1002/prot.21147 · Source: PubMed

CITATIONS
20

READS
13

3 AUTHORS, INCLUDING:

Oliv Eidam
Roche
17 PUBLICATIONS **408** CITATIONS

SEE PROFILE

Available from: Oliv Eidam
Retrieved on: 11 January 2016

PROTEINS: Structure, Function, and Bioinformatics 65:816–823 (2006)

PREDICTION REPORT

Evidence for a Novel Domain of Bacterial Outer Membrane Ushers

Guido Capitani,* Oliv Eidam, and Markus G. Grütter
Biochemisches Institut der Universität Zürich, Zürich CH-8057, Switzerland

ABSTRACT **Many pathogenic bacteria possess adhesive surface organelles (called pili), anchored to their outer membrane, which mediate the first step of infection by binding to host tissue. Pilus biogenesis occurs via the "chaperone–usher" pathway: the usher, a large outer membrane protein, binds complexes of a periplasmic chaperone with pilus subunits, unloads the subunits from the chaperone, and assembles them into the pilus, which is extruded into the extracellular space. Ushers comprise an N-terminal periplasmic domain, a large transmembrane β-barrel central domain, and a C-terminal periplasmic domain. Since structural data are available only for the N-terminal domain, we performed an in-depth bioinformatic analysis of bacterial ushers. Our analysis led us to the conclusion that the transmembrane β-barrel region of ushers contains a so far unrecognized soluble domain, the "middle domain", which possesses a β-sandwich fold. Two other bacterial β-sandwich domains, the TT0351 protein from *Thermus thermophilus* and the carbohydrate binding module CBM36 from *Paenibacillus polymyxa*, are possible distant relatives of the usher "middle domain". Several mutations reported to abolish *in vivo* pilus formation cluster in this region, underlining its functional importance. Proteins 2006;65:816–823.**
© 2006 Wiley-Liss, Inc.

Key words: bacterial pili; outer membrane assembly platforms (ushers); chaperone–usher pathway; structure prediction; new domain; fold recognition

INTRODUCTION

Pili (or fimbriae) are adhesive surface organelles, found in many pathogenic bacteria, which mediate binding to host tissue. They are highly oligomeric, filamentous protein complexes anchored to the outer bacterial membrane.[1] Type 1 pili from uropathogenic *Escherichia coli* strains are a prototype for such organelles and are the subject of extensive investigation. Type 1 pili enable *E. coli* to attach to mannose units of the glycoprotein receptor uroplakin Ia on the surface of urinary epithelium cells, thus mediating the first step of infection in the urinary tract. They are also responsible for bacterial invasion and persistence in target cells.[2] Type 1 pili consist of a rod and a tip: the pilus rod is a right-handed, helical array of 500–3000 copies of the subunit FimA. The 6.9-nm-wide rod is surmounted by a linear tip fibrillum, composed of three kinds of subunits: the adhesin FimH (one copy) and several copies of FimG and FimF[3] (see Fig. 1).

Biogenesis of type 1 pili takes place through the so-called chaperone–usher pathway,[4] an assembly machinery conserved in many gram-negative bacteria. The chaperone–usher pathway consists of a periplasmic chaperone and an outer membrane assembly platform (FimC and FimD, respectively, in the case of type 1 pili). The assembly platform is often referred to as the usher. The periplasmic chaperone fulfils three tasks: first, it forms stoichiometric complexes with pilus subunits, thereby preventing premature pilus assembly in the periplasm; second, it catalyzes the folding of subunits; third, it transports them to the assembly platform in the outer membrane[5] (see Fig. 1).

Structural studies established that pilus subunits have an incomplete immunoglobulin-like fold (greek-key β-sandwich) lacking the seventh, C-terminal β-strand (referred to hereafter as "pilin fold"). In chaperone–subunit complexes, the chaperone provides the subunit with the missing β-strand: a segment of the chaperone, the "donor strand", is inserted parallel to the sixth β-strand of the subunit. In the assembled pilus, the donor strand function is fulfilled by an N-terminal extension (~15 residues long), which precedes the pilin fold of each subunit. This way, a subunit provides its own donor strand to the preceding subunit (completing its fold) and accepts a donor strand from the following subunit (reviewed by Sauer et al., 2004).

Abbreviations: TM, transmembrane; PMD, predicted middle domain.

The Supplementary Material referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

*Correspondence to: Guido Capitani, Biochemisches Institut der Universität Zürich, Zürich CH-8057, Switzerland.
E-mail: capitani@bioc.unizh.ch

Received 6 May 2006; Accepted 13 June 2006

Published online 25 October 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21147

© 2006 WILEY-LISS, INC.

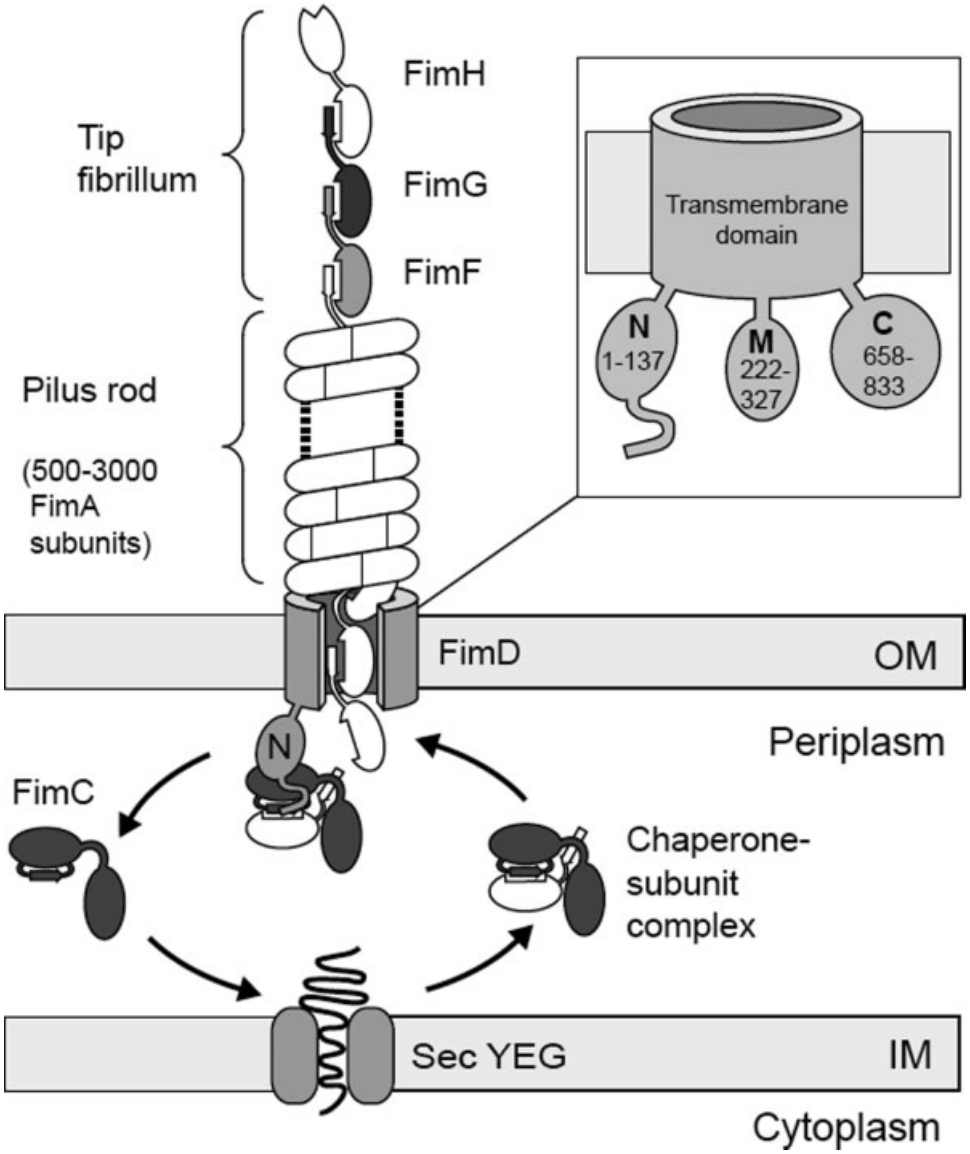

Fig. 1. Schematic drawing of the chaperone–usher pathway of type 1 pilus biogenesis. The inset shows the proposed domain structure of FimD, where M denotes the PMD.

Pilus assembly takes place at the usher ("assembly platform"). In the case of type 1 pili, the assembly platform (FimD) is a 833-residue-long outer membrane protein,[6] performing multiple tasks: it anchors the pilus to the cell membrane, it recognizes FimC–subunit complexes in the periplasm, and translocates folded subunits through the outer membrane. So far three distinct parts have been identified in FimD: (1) an N-terminal, 139-residue-long periplasmic domain, called $FimD_N$,[7] which binds and recognizes chaperone–subunit complexes; (2) a central transmembrane (TM) region, featuring a β-barrel fold; (3) a C-terminal periplasmic domain, which has been studied in the usher PapC (a paralogue of FimD) and has been demonstrated to be necessary for pilus biogenesis *in vivo*.[8] Currently, structural information at the atomic level is available only for $FimD_N$.[9] Low-resolution electron microscopy studies, also in conjunction with biochemical analyses, have been carried out on FimD[10] and PapC.[11] PapC is the paralogue of FimD involved in the biogenesis of P pili[12] and shares 26% identity and 41% similarity with FimD. Based on the results by Li et al., PapC is assumed to be a dimer, with each monomer forming a pore of about 2-nm diameter in the outer membrane. To gain more information on the TM domain of the usher and to guide future biochemical experiments, we undertook an in depth bioinformatic analysis.

MATERIALS AND METHODS

The protein sequence of *E.coli* FimD (P30130) was edited to remove the experimentally determined signal peptide[7] and was used as query for a NCBI BLAST2[13]

search against the UniProt database,[14] with the low-complexity filter turned off. The top 250 hits were analyzed, down to an *E*-value of 0.03, which still corresponded to a protein annotated as a bacterial usher. Forty-one sequences were manually selected from the BLAST output using the NiceBlast tool. The selected sequences were chosen throughout the range of *E*-values of the BLAST output, to ensure a good coverage of the sequence diversity of the usher protein family. Other exclusion criteria were incompleteness of the sequences and analysis of a preliminary multiple sequence alignment with T-COFFEE[15]: the N-terminal part of the alignment, corresponding to the well-characterized domain $FimD_N$[9] was inspected for conservation of residues known to be important for structure and function. Sequences lacking those residues were excluded and replaced by others from the BLAST output. The 41 sequences meeting the selection criteria were then edited to remove signal peptides (referring to the UniProt annotation or using SignalP 3.0[16] with the HMM-method and the gram-negative bacteria option). For FIMD_ECOLI the experimentally determined start of the mature protein, as determined by Nishiyama et al. (2003), was used. In the case of Q8FIX1_ECOL6 (FocD), a very weak and ambiguous signal in the SignalP output (no signal with HMM, weak with the neural network method) was resolved by analogy to FimD. For CSD1_ECOLI there was in practice no signal and the starting point was assigned by analogy with the CSD2_ECOLI sequence. Mature sequences were then realigned with T-COFFEE. The resulting alignment was displayed and edited with GeneDoc.[17] Secondary structure predictions for the FIMD_ECOLI sequence were assigned using PSIPRED,[18] JPRED[19], and SABLE.[20] Predictions of TM segments of FIMD_ECOLI were carried out with the HMM-B2TMR predictor[21] of TRAMPLE[22] and with PROFtmb.[23] PHYRE (http://www.sbg.bio.ic.ac.uk/phyre/) was used for threading, while for *ab initio* structure prediction we resorted to the ROBETTA server.[24] Searches for remote homologues were carried out using the FFAS03 server,[25] HMMER,[26] and PSI-BLAST.

RESULTS AND DISCUSSION

Overall Features and Implications of the Multiple Sequence Alignment of Bacterial Ushers

The multiple sequence alignment of 41 bacterial ushers described in the Methods section (see Supplementary Fig. 1) was analyzed in conjunction with HMM-B2TMR and PROFtmb TM β-strand predictions for the FimD sequence. The central part of the FimD sequence is predicted to form a 22-stranded TM β-barrel and it may be similar, in its spatial architecture and overall dimensions, to the ferrichrome receptor FhuA.[27] The central part of the FimD sequence exhibits two distinct features with respect to similarity to its bacterial homologues/paralogues: most of the region, as expected for a TM β-barrel fold, has a comparatively low degree of conservation and is characterized by an alternation of variable loops and moderately conserved TM β-strands.

On the contrary, residues 222–327 of the mature FimD sequence show a higher degree of sequence identity and similarity throughout all aligned usher sequences (see Fig. 2 and Supplementary Fig. 1, which shows the full-length alignment); the TM region of FimD (residues 138–658 according to HMM-B2TMR) contains 27 residues that are completely or very strongly conserved (according to GeneDoc): residues 222–327 represent only about 20% of that sequence in length but account for 52% of its completely or very strongly conserved residues (14).

The stretches predicted to be loops are in general shorter and better conserved than those outside the 222–327 region. An independent prediction of TM β-strands carried out with two different programs (HMM-B2TMR and PROFtmb, Figs. 2 and 3 and Supplementary Fig. 1) indicates that residues 222–327 should not contain TM strands.

Figure 3 displays the intensity of the TM prediction signal (HMM-B2TMR) along the FimD sequence: all predicted TM strands (in yellow) feature a signal well above the threshold line (only the last TM has a slightly weaker signal) and, importantly, the program does not assign any TM strand in the region 1–137, corresponding to the experimentally characterized N-terminal soluble domain.[9] The PROFtmb prediction[23] yielded a whole protein score of 10.6.

Novel Soluble Domain of Bacterial Outer Membrane Ushers: The PMD

Taken together, the above results indicate that the 222–327 region is very likely to be a so-far unrecognized domain of FimD and of its homologous/paralogous bacterial ushers: the predicted "middle" domain (PMD). The experimentally known periplasmic location of usher N-terminal domains and the presence in both predictions of four TM strands between the N-terminal and the middle domain hint at a periplasmic location of the latter as well (see Fig. 3), even though an extracellular location cannot be ruled out in case of inaccuracy in the number of predicted strands separating the two domains (four separating strands according to both HMM-B2TMR and PROFtmb). The two programs predict slightly different spans for the two TM strands preceding and following, respectively, the PMD. As a result, the PMD of FimD may extend beyond the 222–327 region, which was defined based on the consensus of the HMM-B2TMR and PROFtmb predictions. The PMD may encompass about ten more residues at the N-terminus and/or about three more residues at the C-terminus (Fig. 2 and Supplementary Fig. 1). Both additional stretches are essentially gapless and exhibit a degree of conservation comparable to that of residues 222–327.

The PMD of bacterial ushers is a 105–120-residue-long module, characterized by a high content of secondary structure (with short loops). The consensus of three prediction programs (see Fig. 2) indicates seven β-strands

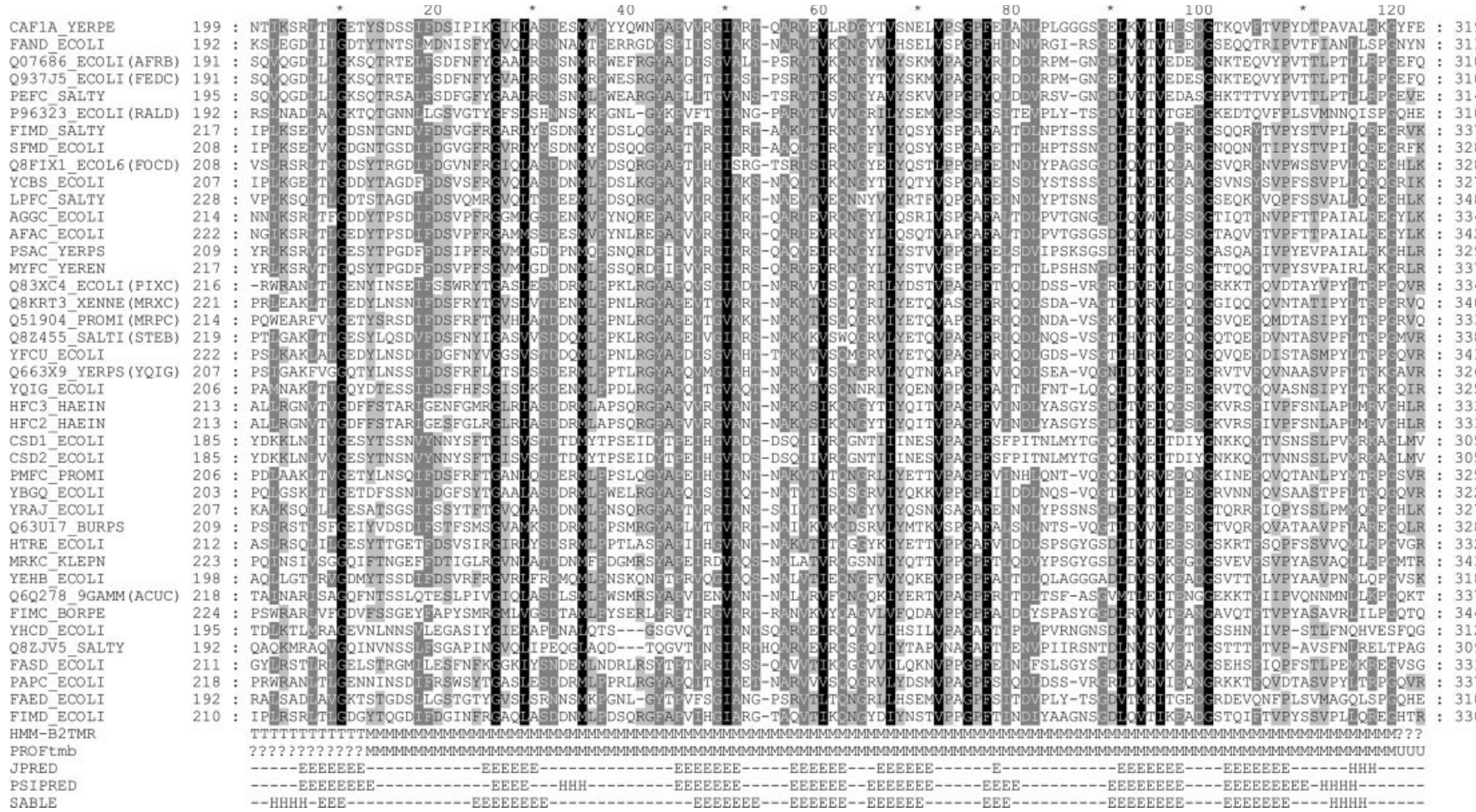

Fig. 2. Multiple sequence alignment of the PMD region of 41 bacterial ushers (indicated by their UniProt ID codes; gene names in parenthesis follow those entry names that are not indicative of gene names). Completely or nearly completely conserved residues are shaded in black, those retaining 80 and 60% similarity in grey and light grey, respectively. TM β-strand predictions (HMM-B2TMR and PROFtmb) are reported: “M” denotes residues assigned to the PMD by both methods, “?” residues assigned by only one method. “T” denotes a predicted TM β-strand, “U” a predicted TM β-strand with periplasm-to-outside direction. The last three lines contain secondary structure predictions (JPRED, PSIPRED, and SABLE) for the PMD.

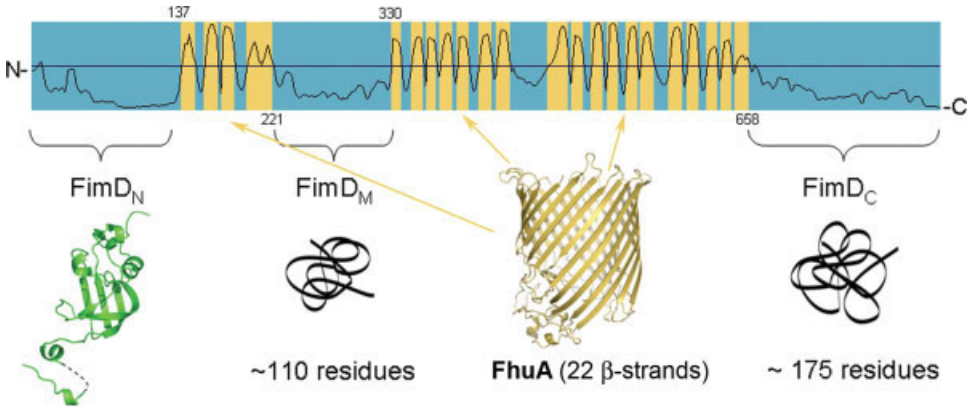

Fig. 3. TM β-strand prediction and FimD domain architecture. The diagram with blue background shows the HMM-B2TMR signal for TM β-strands along the FimD sequence: yellow bars represent predicted TM β-strands. The FhuA barrel is shown as a reference for the fold of the TM domain of FimD. The horizontal line crossing the diagram at mid-height indicates the TM-prediction threshold.

and at least one helix for the region 222–327, while an additional β-strand is predicted in the boundary stretch 210–221 (see Fig. 2).

The presence of a highly conserved region in usher sequences had already been reported in an early study of *E. coli* FasD,[28] but was not interpreted as a possible domain. In spite of the low number of sequences analyzed (10) and of a sequencing error that at the time affected the FimD sequence (residues 352–402 of the full-length, see http://ca.expasy.org/uniprot/FIMD_ECOLI), the region of highest conservation identified by Schifferli and Alrutz substantially coincides with the PMD of the usher.

As mentioned above, the PMD accounts for ~20% of the usher sequence, but contains 52% of its completely or very strongly conserved residues: it is conceivable that this high percentage of conserved residues reflects both an important role for the biological function of the usher and stringent structural constraints.

Fold Recognition Analysis of the PMD and Search for Remote Homologues

Threading analysis with PHYRE was carried out using both the "long" and the "short" FimD stretches (residues 210–330 and 222–327, respectively). In the case of the "short" stretch, PHYRE identified the carbohydrate binding module CBM36 from *Paenibacillus polymyxa* (PDB entry 1W0N) as top hit, albeit with a very low degree of confidence (*E*-value 3.6). Interestingly, when we submitted both the "long" and the "short" stretch to the ROBETTA server,[24] the GINZU domain recognition routine identified exactly the same CBM36 module (this time from PDB entry 1UX7) as the most likely fold for both queries. CBM36 possesses a jelly roll β-sandwich fold with nine strands. It seems reasonable to speculate that also the PMD of FimD could assume a β-sandwich fold, with seven or eight β-strands (depending on its effective length). Interestingly, several β-sandwich structures appear among the ROBETTA models for both the "long" and the "short" query.

A remote homologue search was carried out using the FFAS03 server, scanning each sequence of Figure 2 against the PDB[29] and the PFAM[30] databases. In 21 cases out of 41, the search against the PDB yielded a hit to be considered significant according to the authors' criteria (i.e., FFAS03 score lower than −9.5): in three cases, the significant hit was the carbohydrate binding module CBM36 from *Paenibacillus polymyxa* (PDB entry 1W0N), already found by PHYRE; in 18 cases, the significant hit was another bacterial protein, the TT0351 protein from *Thermus thermophilus* (PDB entry 1V8H, Kunishima et al., to be published). The TT0351 protein possesses a β-sandwich (Greek-key) fold, like pilus subunits, and is annotated by similarity as putative sulphur oxidation protein SoxZ. SoxZ is part the *sox* gene cluster, which confers sulfur-oxidizing (Sox) ability and was first described from the alphaproteobacterium *P. pantotrophus*, which is a facultative chemolithoautotroph and grows with thiosulfate.[31]

It appears, then, that modules structurally similar to the PMD (see Fig. 4 for a multiple sequence alignment), and most likely performing a different function, are found in prokaryotes. The above mentioned FFAS03 search versus the PFAM database did not yield any significant hit (apart from fimbrial ushers).

An iterative HMMER search versus the UniProt database (v76) was also carried out, using as initial seed the multiple sequence alignment of 41 PMDs (see Fig. 2) and an *E*-value cutoff of 0.05, but it converged after having detected only usher sequences.

A PSI-BLAST search of the PMD against the NCBI nr database was carried out to identify possible distant relatives of the PMD in eukaryotes. Two iterations with *E*-value threshold 0.005 and number of descriptions = 1000 were run against the full nr database, while the third iteration was performed against eukaryotes only and yielded no significant hit (*E*-value for the first hit was 8.9). This finding is not surprising if one considers

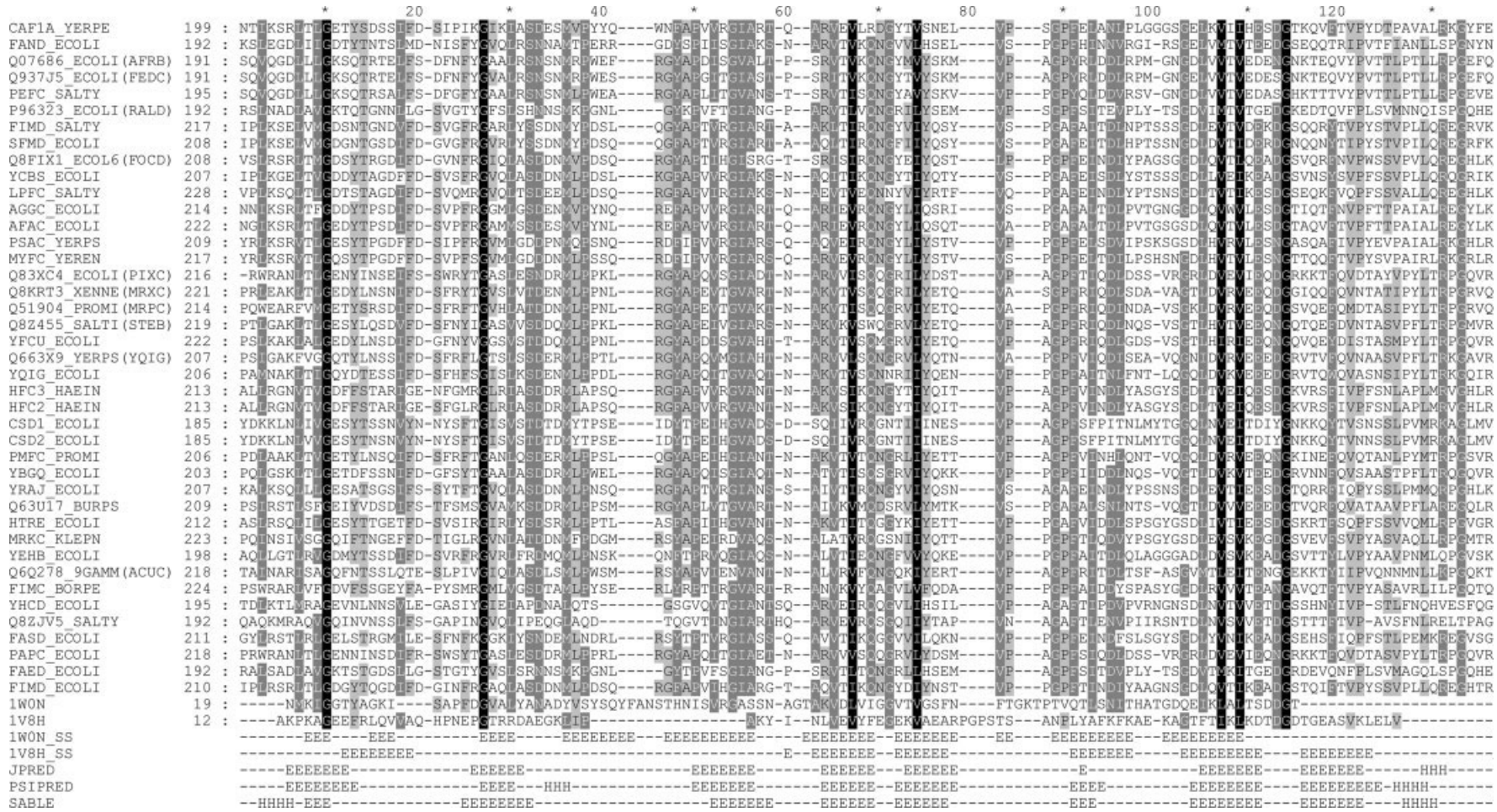

Fig. 4. Alignment of the PMD region of 41 bacterial ushers with the CBM36 domain (PDB code 1W0N) and the TT0351 protein (PDB code 1V8H) as detected by FFAS03. The alignment was produced by manually combining and editing in GeneDoc two FFAS03 pairwise alignments (FimD with 1W0N and PapC with 1V8H) with the multiple sequence alignment of Figure 2. Only the segments of the 1W0N and 1V8H sequences that FFAS03 aligns to the PMD sequences are shown. The fourth and the fifth line from the bottom show the experimentally determined secondary structures of 1W0N and 1V8H, respectively. The last three lines contain secondary structure predictions (JPRED, PSIPRED, and SABLE) for the PMD.

that the structurally well-characterized $FimD_N$ domain[7,9] leads to very similar results when used as a query in the above-described BLAST search protocol.

Topology Studies and Predicted Location of the PMD

An experimental topology study on FaeD,[32] based on site-directed insertion of c-myc epitopes, is consistent with our prediction on the periplasmic location of the PMD. FaeD shares 22% identity and 37% similarity (according to the GeneDoc definition of juxtaposition with score >0) with FimD and gives rise to a nearly identical topology prediction in HMM-B2TMR. Residues 174 and 321 of the mature FaeD sequence (corresponding to 187 and 341 of FimD, respectively) were both experimentally identified as belonging to external loops. Both HMM-B2TMR and PROFtmb predict for the FaeD sequence that Asp174 and Asp321 belong to external loops and are separated from the PMD by one TM β-strand each, indicating a periplasmic location for the middle domain.

Further confirmation comes from an earlier topology study of FaeD[33] carried out using C-terminally truncated variants of the usher fused to PhoA: a construct encompassing the first 215 residues of FaeD and C-terminally fused to PhoA turned out to have its C-terminus, which should belong to the PMD, in the periplasm, consistently with the predictions of HMM-B2TMR and PROFtmb.

Functional Importance of the Usher PMD and Conclusions

A recent experimental study on the topology of PapC,[34] based on a new site-directed fluorescence labelling technique, is not in accord both with our predictions and with crystallographic and biochemical studies of $FimD_N$ (that shares 22% identity and 39% similarity to the N-terminal region of PapC) as an autonomously folding, periplasmic entity. The authors, in fact, proposed that PapC has no periplasmic N-terminal domain and the structure consists of a 26-stranded β-barrel encompassing the first 586 residues of the PapC sequence, with no middle domain.

These contrasting results notwithstanding, the work of Henderson et al. contains findings that hint at an important functional role for the PMD: the authors characterized in vivo, using an agglutination test,[8] the pilus biogenesis activity of many Cys mutants that they had created for fluorescent labelling. The mutations (of two types, insertion of a Cys or mutation to Cys) were scattered throughout the PapC sequence and most of them, in the TM region predicted by HMM-B2TMR (residues 144–635), did not impair the usher function, with one remarkable exception: a set of six mutations in the Asp250-Gly259 stretch (belonging to the N-terminal half of the PMD), which completely abolished agglutination activity. The six mutations are a replacement, Met 252Cys, and five insertions of a Cys residue, after Asp250, Met252, Pro255, Leu257, and Gly259.

Of the above residues, only Met252 is very highly conserved, while the other four exhibit lower conservation (see Fig. 2). In two sequences, Q8ZJV5_SALTY and YHCD_ECOLI, a three-residue deletion is observed in the region corresponding to PapC 256–258. A plausible interpretation, also in light of the secondary structure predictions (see Fig. 2) is that the stretch Asp250-Gly259 (PapC) or Asp242-Gly251 (FimD) contains a loop that is important for the correct folding or pilus assembly function of the domain. The FFAS03-derived alignments with CBM36 and TT3051 (see Fig. 4) show that the Asp250-Gly259 region is specific to bacterial ushers: in CBM36 the corresponding stretch exhibits a four-residue insertion and in TT0351 a long deletion.

Another study on PapC,[11] carried out using cryo-electron microscopy, shed light on the quaternary structure of the usher (dimeric) and on the role of the C-terminal periplasmic domain. That domain had been demonstrated to be essential for pilus biogenesis in vivo but not for in vitro chaperone–subunit interaction.[8] The cryoelectron microscopy results showed that PapC640Δ is still a dimeric protein but has a nonnative arrangement. Thus, the usher C-terminal domain is also necessary for formation of the native dimer interface of the usher.

At present, there is no indication of an involvement of the PMD in the quaternary assembly of the usher. The experiments by Henderson et al. demonstrate that, whatever the exact role of the PMD is, it possesses a stretch of residues that are fundamental for pilus biogenesis. Our bioinformatics results on the PMD will guide future biochemical experiments aimed at characterizing this important region of the bacterial assembly platform. Those experiments will include attempts of producing the PMD in a soluble form and also the deletion of the entire PMD from the usher sequence.

ACKNOWLEDGMENTS

This project was supported by the NCCR Structural Biology program of the Swiss National Science Foundation. OE is affiliated with the Molecular Life Sciences Ph.D. Program of the University/ETH Zürich.

REFERENCES

1. Jones CH, Pinkner JS, Roth R, Heuser J, Nicholes AV, Abraham SN, Hultgren SJ. FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. Proc Natl Acad Sci USA 1995;92:2081–2085.
2. Vetsch M, Glockshuber R. Formation of adhesive pili by the chaperone/usher pathway. In:Buchner J,Kiefhaber T, editors. Protein folding. Weinheim: Wiley-VCH; 2005. pp 965–986.
3. Hahn E, Wild P, Hermanns U, Sebbel P, Glockshuber R, Haner M, Taschner N, Burkhard P, Aebi U, Muller SA. Exploring the 3D molecular architecture of *Escherichia coli* type 1 pili. J Mol Biol 2002;323:845–857.
4. Sauer FG, Remaut H, Hultgren SJ, Waksman G. Fiber assembly by the chaperone–usher pathway. Biochim Biophys Acta 2004; 1694:259–267.

5. Vetsch M, Puorger C, Spirig T, Grauschopf U, Weber-Ban EU, Glockshuber R. Pilus chaperones represent a new type of protein-folding catalyst. Nature 2004;431:329–333.
6. Klemm P, Christiansen G. The fimD gene required for cell surface localization of *Escherichia coli* type 1 fimbriae. Mol Gen Genet 1990;220:334–338.
7. Nishiyama M, Vetsch M, Puorger C, Jelesarov I, Glockshuber R. Identification and characterization of the chaperone-subunit complex-binding domain from the type 1 pilus assembly platform FimD. J Mol Biol 2003;330:513–525.
8. Thanassi DG, Stathopoulos C, Dodson K, Geiger D, Hultgren SJ. Bacterial outer membrane ushers contain distinct targeting and assembly domains for pilus biogenesis. J Bacteriol 2002;184: 6260–6269.
9. Nishiyama M, Horst R, Eidam O, Herrmann T, Ignatov O, Vetsch M, Bettendorff P, Jelesarov I, Grutter MG, Wuthrich K, Glockshuber R, Capitani G. Structural basis of chaperone-subunit complex recognition by the type 1 pilus assembly platform FimD. Embo J 2005;24:2075–2086.
10. Saulino ET, Bullitt E, Hultgren SJ. Snapshots of usher-mediated protein secretion and ordered pilus assembly. Proc Natl Acad Sci USA 2000;97:9240–9245.
11. Li H, Qian L, Chen Z, Thibault D, Liu G, Liu T, Thanassi DG. The outer membrane usher forms a twin-pore secretion complex. J Mol Biol 2004;344:1397–1407.
12. Mulvey MA. Adhesion and entry of uropathogenic *Escherichia coli*. Cell Microbiol 2002;4:257–271.
13. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25: 3389–3402.
14. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS. UniProt: the Universal Protein knowledgebase. Nucleic Acids Res 2004;32: D115–D119.
15. Notredame C, Higgins DG, Heringa J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 2000;302: 205–217.
16. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004;340: 783–795.
17. Nicholas KB, Nicholas HBJ, Deerfield DWI. GeneDoc: analysis and visualization of genetic variation. embnetnews 1997; 4:2.
18. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics 2000;16:404, 405.
19. Cuff JA, Clamp ME, Siddiqui AS, Finlay M, Barton GJ. JPred: a consensus secondary structure prediction server. Bioinformatics 1998;14:892, 893.
20. Adamczak R, Porollo A, Meller J. Combining prediction of secondary structure and solvent accessibility in proteins. Proteins 2005;59: 467–475.
21. Martelli PL, Fariselli P, Krogh A, Casadio R. A sequence-profile-based HMM for predicting and discriminating β barrel membrane proteins. Bioinformatics 2002;18 (Suppl. 1):S46–S53.
22. Fariselli P, Finelli M, Rossi I, Amico M, Zauli A, Martelli PL, Casadio R. TRAMPLE: the transmembrane protein labeling environment. Nucleic Acids Res 2005;33:W198–W201.
23. Bigelow HR, Petrey DS, Liu J, Przybylski D, Rost B. Predicting transmembrane β-barrels in proteomes. Nucleic Acids Res 2004; 32:2566–2577.
24. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 2004;32: W526–W531.
25. Jaroszewski L, Rychlewski L, Li Z, Li W, Godzik A. FFAS03: a server for profile–profile sequence alignments. Nucleic Acids Res 2005;33:W284–W288.
26. Eddy SR. Profile hidden Markov models. Bioinformatics 1998; 14:755–763.
27. Locher KP, Rees B, Koebnik R, Mitschler A, Moulinier L, Rosenbusch JP, Moras D. Transmembrane signaling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. Cell 1998;95: 771–778.
28. Schifferli DM, Alrutz MA. Permissive linker insertion sites in the outer membrane protein of 987P fimbriae of *Escherichia coli*. J Bacteriol 1994;176:1099–1110.
29. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
30. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, Eddy SR. The Pfam protein families database. Nucleic Acids Res 2004;32:D138–D141.
31. Quentmeier A, Hellwig P, Bardischewsky F, Grelle G, Kraft R, Friedrich CG. Sulfur oxidation in *Paracoccus pantotrophus*: interaction of the sulfur-binding protein SoxYZ with the dimanganese SoxB protein. Biochem Biophys Res Commun 2003;312: 1011–1018.
32. Harms N, Oudhuis WC, Eppens EA, Valent QA, Koster M, Luirink J, Oudega B. Epitope tagging analysis of the outer membrane folding of the molecular usher FaeD involved in K88 fimbriae biosynthesis in *Escherichia coli*. J Mol Microbiol Biotechnol 1999;1:319–325.
33. Valent QA, Zaal J, de Graaf FK, Oudega B. Subcellular localization and topology of the K88 usher FaeD in *Escherichia coli*. Mol Microbiol 1995;16:1243–1257.
34. Henderson NS, So SS, Martin C, Kulkarni R, Thanassi DG. Topology of the outer membrane usher PapC determined by site-directed fluorescence labeling. J Biol Chem 2004;279:53747–53754.