Protein Structures

International Edition: DOI: 10.1002/anie.201505065
German Edition: DOI: 10.1002/ange.201505065

Hybrid Structure of the Type 1 Pilus of Uropathogenic *Escherichia coli*

Birgit Habenstein, Antoine Loquet, Songhwan Hwang, Karin Giller, Suresh Kumar Vasa, Stefan Becker, Michael Habeck, and Adam Lange**

Abstract: *Type 1 pili are filamentous protein assemblies on the surface of Gram-negative bacteria that mediate adhesion to host cells during the infection process. The molecular structure of type 1 pili remains elusive on the atomic scale owing to their insolubility and noncrystallinity. Herein we describe an approach for hybrid-structure determination that is based on data from solution-state NMR spectroscopy on the soluble subunit and solid-state NMR spectroscopy and STEM data on the assembled pilus. Our approach is based on iterative modeling driven by structural information extracted from different sources and provides a general tool to access pseudo atomic structures of protein assemblies with complex subunit folds. By using this methodology, we determined the local conformation of the FimA pilus subunit in the context of the assembled type 1 pilus, determined the exact helical pilus architecture, and elucidated the intermolecular interfaces contributing to pilus assembly and stability with atomic detail.*

Uropathogenic *Escherichia coli* adhesion to host cell tissue is mediated by the biogenesis of type 1 pili.[1–3] Such pili contain approximately 3000 copies of the major subunit FimA, which form the pilus rod that serves as a bacterial fishing line, and an adhesive tip (Figure 1 A).[4,5] The subunit FimA adopts an immunoglobulin-like all-β fold,[6] although it contains an N-terminal extension and lacks the C-terminal G1 strand. Prior to expulsion to the extracellular space, the N-terminal extension of the succeeding pilus subunit is inserted into the position of the missing G1 strand, thus causing polymerization into the pilus.[2,7] The subunits are arranged into a right-handed helical filament, whose structure is largely determined by noncovalent N-terminal donor-strand exchange,[7a,8–10] involving several highly conserved residues (see Figure S1 in the Supporting Information). Insolubility and the noncrystalline nature of the pili have so far prevented the determination of the atomic structure of the pilus rod. Puorger et al.[11] solved the structure of a monomeric soluble FimA construct called FimAa[12,13] by solution NMR spectroscopy. The missing G1 β-strand of the immunoglobulin-like fold was replaced in this monomeric structure by the FimA N-terminal extension artificially inserted at the C terminus.

We herein report the atomic structure of the assembled uropathogenic *Escherichia coli* pilus rod, as guided by solid-state NMR (ssNMR) restraints collected on intact pilus filaments. ssNMR spectroscopy is an emerging technique for structural studies of filamentous protein assemblies.[14] It can be directly applied to intact pili and provides atomic-resolution structural data. We determined the structure of the intact pilus containing the 159 residue FimA construct by using probabilistic structural modeling[15] to combine complementary techniques in a hybrid approach: Restraints from solution NMR spectroscopy defined the FimA monomer structure and the interactions with the artificial G1 strand, which represents the donor strand of the adjacent subunit;[11] scanning transmission electron microscopy (STEM) data revealed the axial rise per subunit and thus the symmetry of the pilus; and ssNMR spectroscopic data confirmed the monomeric structure in the pilus and provided the intermolecular interfaces between the FimA subunits.

To confirm that pili present on the *E. coli* cell surface (Figure 1 A) shared a common fold with pili assembled in vitro, we compared the symmetry parameters encoded in STEM mass-per-length (MPL) measurements. A typical dark-field STEM image of FimA pili assembled in vitro is shown in Figure 1 B. The MPL histogram of the FimA pili assembled in vitro is centered at 2.23 kDa $Å^{-1}$ (Figure 1 C). Accounting for the mass of FimA (15 958.5 Da), the axial rise per subunit was calculated to be 7.16 Å. The axial rise per subunit of the pili polymerized in vitro is therefore extremely close to the value observed in vivo (7.2 Å).[5] Thus, we could safely assume that the recombinant pili have assembled correctly and utilize them for ssNMR spectroscopic analysis.

We successfully combined a tailored three-dimensional assignment strategy to assign the backbone (^{15}N, $^{13}C'$, $^{13}C\alpha$) and $^{13}C\beta$ side-chain atoms with classical two-dimensional

[*] Dr. B. Habenstein, Dr. A. Loquet, S. Hwang, K. Giller, Dr. S. K. Vasa, Dr. S. Becker, Prof. Dr. A. Lange
Department of NMR-Based Structural Biology
Max Planck Institute for Biophysical Chemistry
Am Fassberg 11, 37077 Göttingen (Germany)
Dr. B. Habenstein, Dr. A. Loquet
Université de Bordeaux/CBMN UMR5248
Institut Européen de Chimie et Biologie (IECB)
2 rue Robert Escarpit, 33600 Pessac (France)
Dr. M. Habeck
Felix Bernstein Institute for Mathematical Statistics in the Biosciences, University of Göttingen
Goldschmidtstrasse 7, 37077 Göttingen (Germany)
and
Statistical Inverse Problems in Biophysics
Max Planck Institute for Biophysical Chemistry
Am Fassberg 11, 37077 Göttingen (Germany)
E-mail: mhabeck@gwdg.de
S. Hwang, Prof. Dr. A. Lange
Institut für Biologie, Humboldt-Universität zu Berlin
Invalidenstrasse 110, 10115 Berlin (Germany)
and
Department of Molecular Biophysics
Leibniz-Institut für Molekulare Pharmakologie (FMP)
Robert-Rössle-Strasse 10, 13125 Berlin (Germany)
E-mail: alange@fmp-berlin.de

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201505065.

 © 2015 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

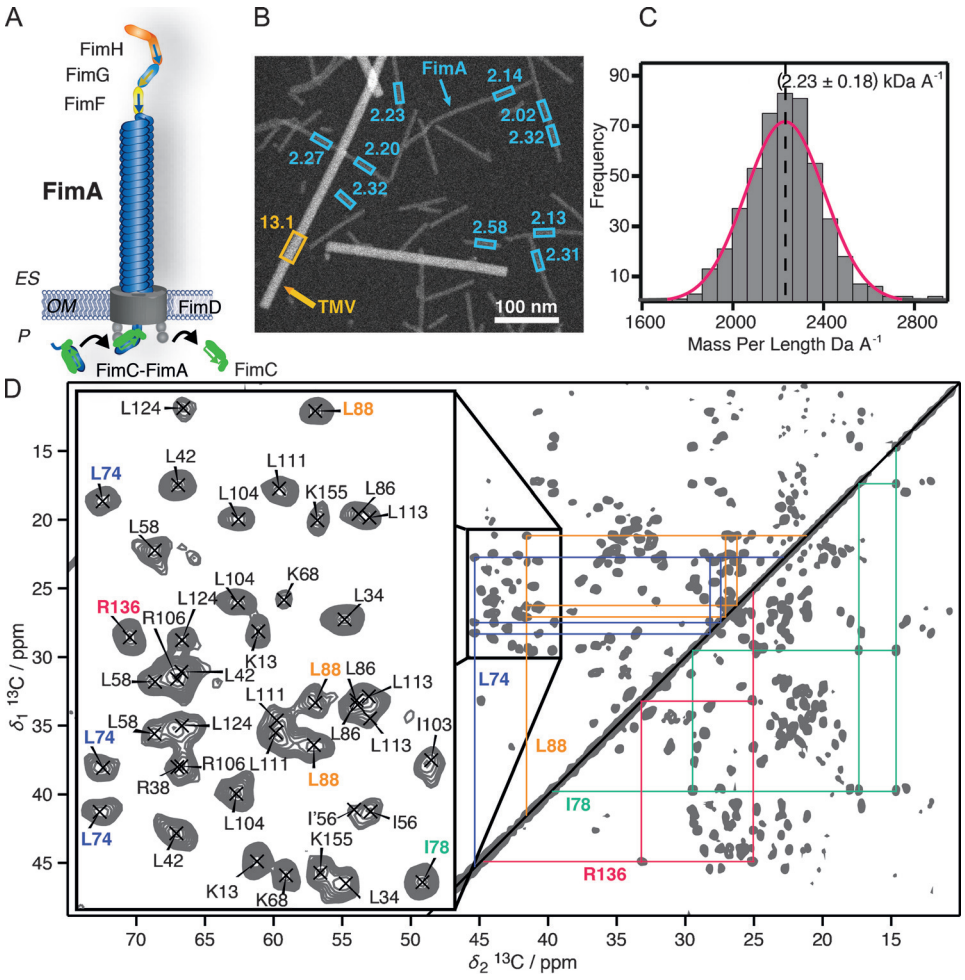

Figure 1. STEM and solid-state NMR spectroscopic analysis of the type 1 pilus. A) Structural diagram of the type 1 pilus, including the pilus tip (FimF, FimG, FimH), the rod (FimA), and the usher (FimD). ES = extracellular space, OM = outer membrane, P = periplasm. B) Dark-field STEM images of FimA pili and tobacco mosaic virus (TMV) reference particles. C) Mass-per-length histogram fitted with a Gaussian distribution function centered at (2.23 ± 0.18) kDa Å^{-1}. D) Determination of the ssNMR resonances on the basis of a 3D sequential-assignment strategy (see Figure S4) and complementary high-resolution 2D spectra, as exemplified with an excerpt of a ^{13}C–^{13}C correlation spectrum (50 ms mixing time proton-driven spin diffusion, PDSD) of uniformly [^{13}C,^{15}N] labeled FimA pili. Highlighted in the excerpt are side-chain assignments of residues L74, I78, L88, and R136.

techniques for the assignment of further side-chain resonances (Figure 1D; see also Figures S2–S4). The narrow line widths (full width at half-height, FWHH (^{13}C) ≈ 100 Hz for a uniformly [^{13}C,^{15}N] labeled sample) and the quality of the ssNMR spectra reflect a high structural order and homogeneity in the pilus architecture. We found that the FimA subunits adopt a unique conformation in the pilus assembly, as encoded in a single set of ssNMR resonance frequencies, with minimal doubling of Ile56 Cβ, Cγ, and Cδ resonances. Structural polymorphism, which has been observed for class 5 pili,[16] can therefore be excluded on the atomic scale in type 1 pili of uropathogenic *E. coli*. We assigned the ssNMR resonances of all backbone atoms of the 160-residue FimA subunit (excluding the flexible N-terminal Met residue). Altogether, we assigned 100% of ^{15}N, ^{13}C′, ^{13}Cα, and ^{13}Cβ atoms, 97% of Cγ atoms, and 96% of all carbon atoms (see Table S1 in the Supporting Information; BMRB entry 25334).

The pilus protein FimA ranges among the largest assembled-state proteins for which the NMR resonance frequencies have been assigned completely.

The secondary structure of FimA in the pilus (Figure 2A) can be derived from the secondary chemical shifts (see Figure S5) as well as database predictions with the Talos+ software[17] (see Figure S6). Although the chemical shifts (CSs) differ between solution NMR spectroscopy of the artificial FimAa construct[11] and FimA in the pili by up to about 3 ppm for Cα (Figure 2B) and 4 ppm for C′ atoms (see Figure S7), the FimA secondary structure is globally conserved (see Figures S5 and S6). In particular, the N-terminal extension of FimA in the intact pili adopts a highly similar secondary structure to that of the artificial G1 strand in the FimAa construct (Figure 2C,D). The conserved secondary structure justifies the assumption that the overall 3D fold of FimA is conserved between the artificial monomeric FimAa and the filamentous wild-type (wt) FimA state. Therefore, the solution NMR restraints obtained for FimAa can be used with caution for the structure calculation of the FimA pilus. Nevertheless, studies of the soluble monomers cannot fully reflect the structure of the monomer in the context of the supramolecular assembly owing to a lack of subunit–subunit contacts. Interestingly in this regard, the backbone root-mean-square deviation (r.m.s.d.; dashed pink line in Figure 2B) for the solution NMR structure shows significantly higher values around residues 90–95, thus revealing flexibility.[11] In ssNMR spectroscopy, the peak volume in dipolar-coupling-based experiments primarily depends on the protein rigidity.[18] The volume of the peaks in a 3D ^{15}N–^{13}C′–^{13}Cα spectrum (CP, BSH-CP[19]) of FimA pili reveals an almost entirely rigid protein backbone (see Figure S8), without any fluctuations that would indicate flexibility around residues 90–95. The differences in backbone mobility of FimA in the context of the pilus assembly and in the soluble monomeric form, as well as the striking ^{13}Cα and ^{13}C′ CS differences in this five-residue stretch (Figure 2B; see also Figure S7), indicate a crucial role of these residues in the FimA assembly.

© 2015 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

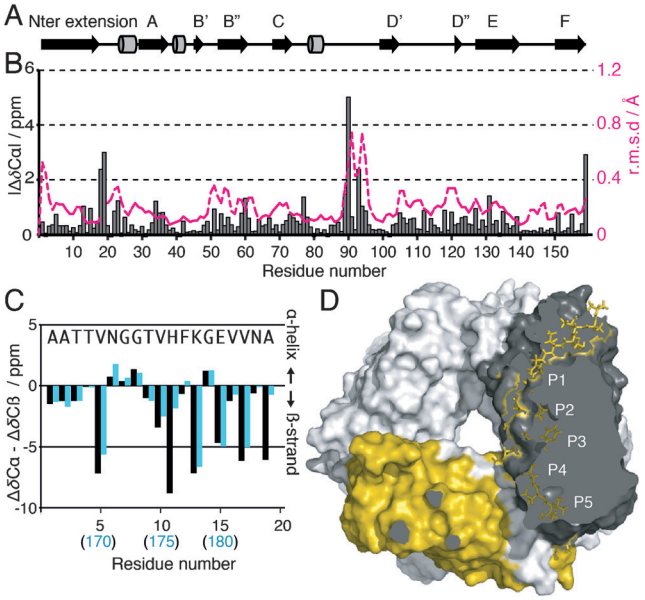

Figure 2. A) Secondary structure of the type 1 FimA pilus. B) ^{13}C chemical-shift differences between the solution NMR data of self-complemented FimAa monomers[11] and ssNMR chemical shifts of wt FimA pili (absolute values of Δ(solution NMR−ssNMR)$\delta C\alpha$). The dashed pink line indicates the backbone r.m.s.d. of the solution NMR structure of the self-complemented FimAa monomer (PDB entry 2JTY). Residues A1–N18 in 2JTY have been replaced with residues A166–N183 representing the donor strand in 2JTY. The r.m.s.d. for the connecting residue A19 has been discarded because it is found in different environments in the monomeric construct FimAa and the FimA pili. C) Secondary structure of the FimA donor strand, as determined by secondary chemical shifts ($\Delta\delta C\alpha - \Delta\delta C\beta$), compared for the assembled pilus (black) and the self-complemented FimAa construct (cyan). Negative values (>3 in a row) indicate β-strand conformation. D) Surface view of a FimA pilus cut at the level of subunits i-1 (yellow) and i (gray). The donor strand of subunit i-1 is detailed in a stick representation. For clarity, subunits i-2 to i-5 are depicted in white, and the donor strand of subunit i is omitted.

Sparse ^{13}C labeling of FimA pili by the use of [2-^{13}C]glycerol as the carbon source during recombinant protein expression facilitates the collection of long-range ssNMR restraints by increasing the spectral resolution and enhancing the polarization transfer between distant carbon atoms.[20] ^{13}C–^{13}C spectra with long mixing times were recorded at different magnetic fields (600 and 850 MHz proton frequency) to obtain such meaningful restraints (Figure 4A; see also Figures S9 and S10).

To determine the type 1 pilus structure at high resolution, we developed a protocol based on the Inferential Structure Determination software (ISD),[15,21] which allowed us to combine the ssNMR restraints with other sources of structural information. The flowchart shown in Figure 3 illustrates our modeling approach to optimally integrate the available data from MPL with solid-state and solution NMR measurements (see the Supporting Information for more details).

In the first step, a structure ensemble of monomeric FimA was calculated on the basis of the solution NMR data (in the form of hydrogen-bonding restraints derived from the reference 2JTY[11]) by starting from an extended conformation and further refining it with intraresidue ssNMR restraints. We only used the backbone hydrogen bonds derived from 2JTY and not the full set of structural restraints, since local discrepancies in the chemical shifts indicated local structural changes between the solution FimAa structure and the structure of FimA in the pili. In a second step, we estimated the helix symmetry of the pilus by rigid-body modeling. The docking procedure was guided by the MPL data and the ambiguously assigned ssNMR distance restraints. We used a single FimA subunit to build up the pilus rod, which reflects the presence of a single protomer conformation, as seen by ssNMR chemical-shift analysis. Finally, the assembly of the FimA pilus structure was computed by incorporating all ssNMR restraints and the helix symmetry. On the basis of the docking results, the structure calculation took into account interactions between the subunit representing the protomer and six successive symmetry-related subunits. Two iterative rounds of cross-peak assignment followed by structure refinement were applied to produce the final restraint list (see Table S2). In the assignment step, systematically violated contributions to the ambiguous ssNMR distance restraints

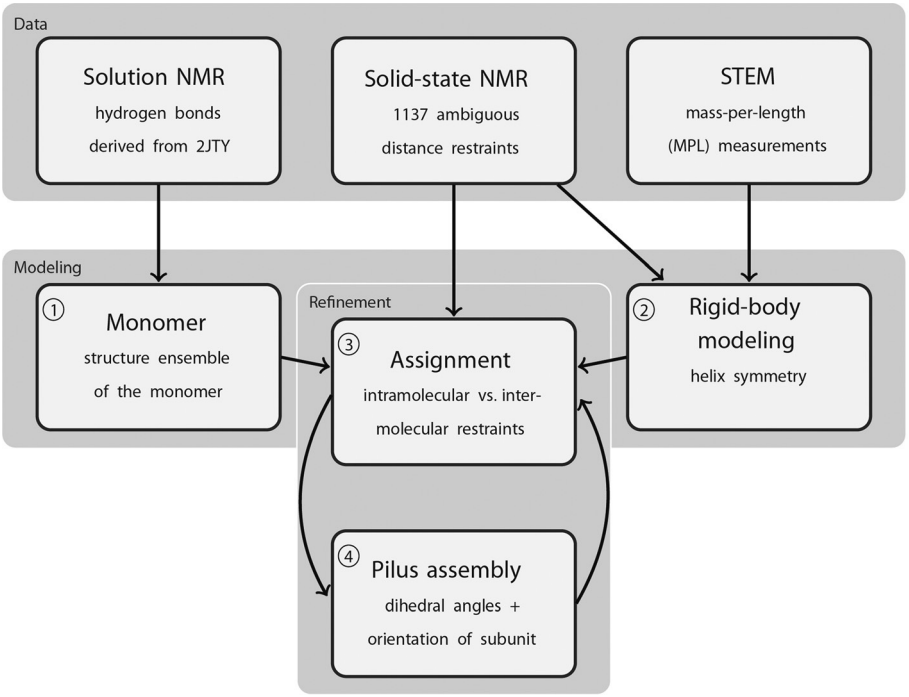

Figure 3. Flowchart of the structure-determination procedure based on a combination of solution NMR, ssNMR, and STEM data.

© 2015 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

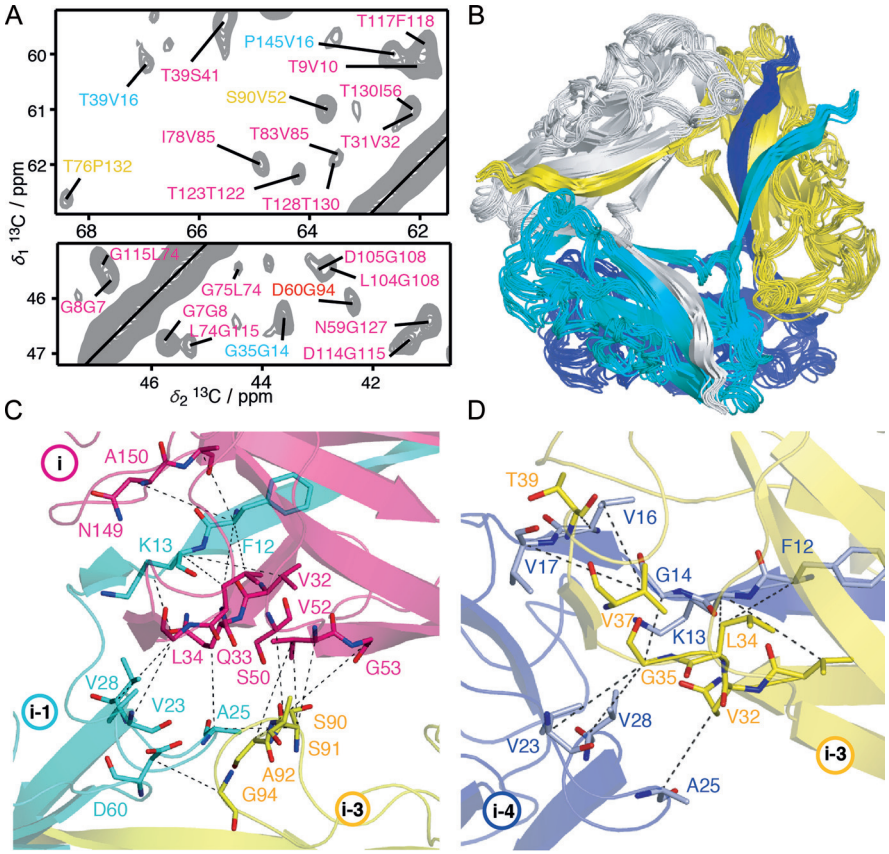

Figure 4. Structure determination of the *E. coli* type 1 pilus. A) Long-range contacts detected in a ^{13}C–^{13}C correlation spectrum of [2-^{13}C]glycerol-labeled FimA pili. Contacts are colored in pink (i–i), cyan (i–i-1), red (i–i-2), and yellow (i–i-3). See Figure S9 for the full spectrum. B) Structural ensemble of six successive subunits. C) Dashed lines mark atoms interacting at the interface between subunits i, i-1, and i-3. D) Donor-strand entry site of subunit i-4 (blue) into subunit i-3 (yellow).

were discarded (see the Supporting Information for further details of the procedure).

In total, 1137 ^{13}C–^{13}C restraints were assigned by a fully automated procedure (see Table S2 for details), including 278 nonredundant long-range contacts that were unambiguous in the first structure ensemble (illustrated by examples in Figure 4A). A total of 946 and 74 restraints result from intra- and intermolecular proximities, respectively; 117 contacts have been modeled as arising from intra- or inter-subunit contacts by the use of ambiguous distance restraints. On the basis of the final restraint list, we computed a high-resolution structural ensemble of the type 1 pilus (Figure 4B). The ensemble shows an r.m.s.d. of (0.9 ± 0.2) Å for backbone atoms and (1.2 ± 0.3) Å for heavy atoms of a single monomeric FimA subunit, and (1.1 ± 0.3) Å and (1.4 ± 0.3) Å, respectively, across the assembly of six subunits (PDB ID: 2MX3).

The type 1 pilus of uropathogenic *E. coli* contains a one-start helix formed by all-β subunits that are connected by donor-strand complementation (DSC). The donor-strand insertion register was found to be conserved between the solution NMR FimAa structure and the assembled pilus structure. Alternating hydrophobic amino acids provide an extremely stable zipper with five binding pockets P1–P5 (Figure 2D).[22,11] The donor strands are located at the interior of the pilus, as hypothesized previously.[10] Conserved residues (see Figure S1) are located in the hydrophobic core of FimA, on the donor strand and at the subunit interfaces (see Figure S11). We characterized the intermolecular interfaces between subunits i and i-1, i-2, and i-3 (Figure 4C,D). The extremely high stability of the pilus against denaturants has been reported to not only arise from DSC.[23] Our structure reveals interactions at the i–i-2 and i–i-3 interfaces that complement DSC. Overall, a comparison between side-chain solution and solid-state NMR resonances reveals that the vast majority of side-chain chemical-shift changes are located at the intermolecular interfaces (see Figure S12 and Table S3).

Our method allowed us to solve a hybrid structure of the type 1 pilus from *E. coli* (Figure 5A,B), with a subunit of roughly twice the size of the previously determined structure of type 3 secretion system needles.[14a,b] As compared to that system, the FimA pilus is characterized by a much more complex subunit fold, which makes the availability of a subunit structure necessary (either from solution NMR spectroscopy or X-ray crystallography), to reduce the ambiguities of the ssNMR restraints. Comparison of the solution NMR (if available) and ssNMR chemical shifts of the monomeric subunit and the assembly, respectively, also enables the extraction of valuable information on the conformational changes involved in the formation of macromolecular assemblies.

Acknowledgements

We thank Joe Wall for collecting the STEM data. This research was supported by the Max Planck Society, the Leibniz-Institut für Molekulare Pharmakologie (FMP), the European Research Council (ERC Starting Grant to A. Lange), the German Research Foundation (DFG; Emmy Noether Fellowships to M.H. and A. Lange, SFB 860 project B09 to M.H.), and the Agence Nationale de la Recherche (ANR 13-PDOC-0017-01 to B.H.).

Keywords: filamentous protein assemblies · iterative modeling · pilus structure · solid-state NMR spectroscopy · structure elucidation

 © 2015 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

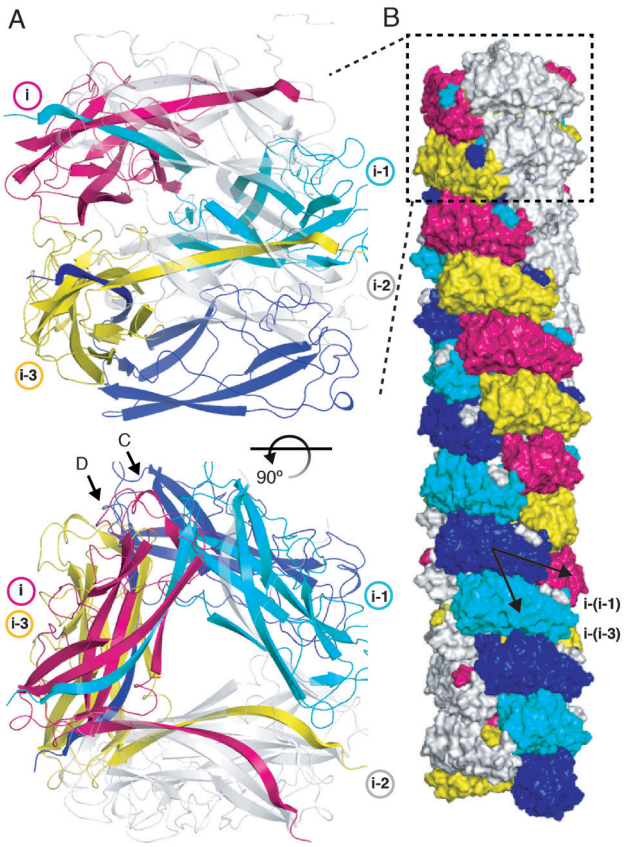

Figure 5. Hybrid structure of the *E. coli* type 1 pilus (PDB ID: 2MX3). A) Pilus structure represented by six FimA subunits. The two front FimA subunits have been colored white and rendered transparent to facilitate the view of the pilus architecture. Arrows C and D indicate the viewing directions in Figure 4 C,D. B) Representation of the pilus surface.

How to cite: *Angew. Chem. Int. Ed.* **2015**, *54*, 11691–11695
Angew. Chem. **2015**, *127*, 11857–11861

[1] G. Waksman, S. J. Hultgren, *Nat. Rev. Microbiol.* **2009**, *7*, 765–774.
[2] M. Nishiyama, T. Ishikawa, H. Rechsteiner, R. Glockshuber, *Science* **2008**, *320*, 376–379.
[3] S. Geibel, E. Procko, S. J. Hultgren, D. Baker, G. Waksman, *Nature* **2013**, *496*, 243–246.
[4] C. H. Jones, J. S. Pinkner, R. Roth, J. Heuser, A. V. Nicholes, S. N. Abraham, S. J. Hultgren, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2081–2085.
[5] E. Hahn, P. Wild, U. Hermanns, P. Sebbel, R. Glockshuber, M. Häner, N. Taschner, P. Burkhard, U. Aebi, S. A. Müller, *J. Mol. Biol.* **2002**, *323*, 845–857.
[6] P. Bork, L. Holm, C. Sander, *J. Mol. Biol.* **1994**, *242*, 309–320.
[7] a) F. G. Sauer, K. Fütterer, J. S. Pinkner, K. W. Dodson, S. J. Hultgren, G. Waksman, *Science* **1999**, *285*, 1058–1061; b) M. Vetsch, C. Puorger, T. Spirig, U. Grauschopf, E. U. Weber-Ban, R. Glockshuber, *Nature* **2004**, *431*, 329–333; c) H. Remaut, R. J. Rose, T. J. Hannan, S. J. Hultgren, S. E. Radford, A. E. Ashcroft, G. Waksman, *Mol. Cell* **2006**, *22*, 831–842.
[8] Y. Eshdat, F. J. Silverblatt, N. Sharon, *J. Bacteriol.* **1981**, *148*, 308–314.
[9] E. Bullitt, L. Makowski, *Nature* **1995**, *373*, 164–167.
[10] D. Choudhury, A. Thompson, V. Stojanoff, S. Langermann, J. Pinkner, S. J. Hultgren, S. D. Knight, *Science* **1999**, *285*, 1061–1066.
[11] C. Puorger, M. Vetsch, G. Wider, R. Glockshuber, *J. Mol. Biol.* **2011**, *412*, 520–535.
[12] M. M. Barnhart, J. S. Pinkner, G. E. Soto, F. G. Sauer, S. Langermann, G. Waksman, C. Frieden, S. J. Hultgren, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7709–7714.
[13] M. Vetsch, D. Erilov, N. Molière, M. Nishiyama, O. Ignatov, R. Glockshuber, *EMBO Rep.* **2006**, *7*, 734–738.
[14] a) A. Loquet, N. G. Sgourakis, R. Gupta, K. Giller, D. Riedel, C. Goosmann, C. Griesinger, M. Kolbe, D. Baker, S. Becker, A. Lange, *Nature* **2012**, *486*, 276–279; b) J. P. Demers, B. Habenstein, A. Loquet, S. Kumar Vasa, K. Giller, S. Becker, D. Baker, A. Lange, N. G. Sgourakis, *Nat. Commun.* **2014**, *5*, 4976; c) S. Vasa, L. Lin, C. Shi, B. Habenstein, D. Riedel, J. Kühn, M. Thanbichler, A. Lange, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E127–E136; d) O. Morag, N. G. Sgourakis, D. Baker, A. Goldbourt, *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 971–976.
[15] W. Rieping, M. Habeck, M. Nilges, *Science* **2005**, *309*, 303–306.
[16] V. E. Galkin, S. Kolappan, D. Ng, Z. Zong, J. Li, X. Yu, E. H. Egelman, L. Craig, *J. Bacteriol.* **2013**, *195*, 1360–1370.
[17] Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, *J. Biomol. NMR* **2009**, *44*, 213–223.
[18] A. Lange, Z. Gattin, H. Van Melckebeke, C. Wasmer, A. Soragni, W. F. van Gunsteren, B. H. Meier, *ChemBioChem* **2009**, *10*, 1657–1665.
[19] a) V. Chevelkov, C. Shi, H. K. Fasshuber, S. Becker, A. Lange, *J. Biomol. NMR* **2013**, *56*, 303–311; b) C. Shi, H. K. Fasshuber, V. Chevelkov, S. Xiang, B. Habenstein, S. K. Vasa, S. Becker, A. Lange, *J. Biomol. NMR* **2014**, *59*, 15–22.
[20] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, H. Oschkinat, *Nature* **2002**, *420*, 98–102.
[21] W. Rieping, M. Nilges, M. Habeck, *Bioinformatics* **2008**, *24*, 1104–1105.
[22] F. G. Sauer, J. S. Pinkner, G. Waksman, S. J. Hultgren, *Cell* **2002**, *111*, 543–551.
[23] C. Puorger, O. Eidam, G. Capitani, D. Erilov, M. G. Grütter, R. Glockshuber, *Structure* **2008**, *16*, 631–642.

Received: June 3, 2015
Revised: July 16, 2015
Published online: August 12, 2015

 © 2015 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim