A Comparative Structure/Function Analysis of Two Type IV Pilin DNA Receptors Defines a Novel Mode of DNA Binding

Graphical Abstract

Highlights

- ComP orthologs are type IV pilins binding DNA, key for natural transformation
- Two high-resolution structures reveal a unique pilin fold
- DNA binding involves a novel motif

Authors

Jamie-Lee Berry, Yingqi Xu, Philip N. Ward, Susan M. Lea, Stephen J. Matthews, Vladimir Pelicic

Correspondence

s.j.matthews@imperial.ac.uk (S.J.M.), v.pelicic@imperial.ac.uk (V.P.)

In Brief

ComP is the first type IV pilin with DNA-binding ability, key for natural transformation in Neisseriaceae. By reporting high-resolution structures for two ComP orthologs, Berry et al. shed light on a novel mode of DNA binding.

Accession Numbers

5HZ7
2NBA
A Comparative Structure/Function Analysis of Two Type IV Pilin DNA Receptors Defines a Novel Mode of DNA Binding

Jamie-Lee Berry, Yingqi Xu, Philip N. Ward, Susan M. Lea, Stephen J. Matthews, and Vladimir Pelicic

1MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London SW7 2AZ, UK
2Centre for Structural Biology, Imperial College London, London SW7 2AZ, UK
3Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK
*Correspondence: s.j.matthews@imperial.ac.uk (S.J.M.), v.pelicic@imperial.ac.uk (V.P.)
http://dx.doi.org/10.1016/j.str.2016.04.001

SUMMARY

DNA transformation is a widespread process allowing bacteria to capture free DNA by using filamentous nano-machines composed of type IV pilins. These proteins can act as DNA receptors as demonstrated by the finding that Neisseria meningitidis ComP minor pilin has intrinsic DNA-binding ability. ComP binds DNA better when it contains the DNA-uptake sequence (DUS) motif abundant in this species genome, playing a role in its trademark ability to selectively take up its own DNA. Here, we report high-resolution structures for meningococcal ComP and Neisseria subflava ComPsub, which recognize different DUS motifs. We show that they are structurally identical type IV pilins that pack readily into filament models and display a unique DD region delimited by two disulfide bonds. Functional analysis of ComPsub defines a new mode of DNA binding involving the DD region, adapted for exported DNA receptors.

INTRODUCTION

Numerous bacterial species, defined as naturally competent, are able to capture free DNA from the environment and import it into their cytoplasm across formidable permeability barriers (Chen and Dubnau, 2004). Although imported DNA can be used as a source of food or as a template for the repair of DNA damage (Chen and Dubnau, 2004), when the DNA is new and it is stably acquired, this process is called transformation, since the bacteria are “transformed” by exhibiting novel phenotypic traits. Transformation allows competent bacteria to evolve rapidly by promoting transfer of DNA between different species, an important evolutionary process known as horizontal gene transfer (HGT) (Thomas and Nielsen, 2005).

Most naturally competent species import extracellular DNA using the same two-step process (Chen and Dubnau, 2004). The first step involves DNA uptake mediated by type IV filamentous (Tff) nano-machines composed of type IV pilins (Berry and Pelicic, 2015), either long filaments known as type IV pil (Tfp) or elusive competence (pseudo)pil. It is widely accepted (Chen and Dubnau, 2004; Maier et al., 2004) that filaments bind free DNA and pull it across the outer membrane (in Gram-negative bacteria) and/or peptidoglycan (in Gram-positive species). This scenario is supported by the finding that in species where bona fide retractable Tfp are involved, as in the human pathogens Neisseria meningitidis and Neisseria gonorrhoeae, there is no DNA uptake when pilus retraction is abolished (Brown et al., 2010; Wolfgang et al., 1998). In the second step, once DNA is in the periplasm/pseudo-periplasm, it is bound by the ComE/ComEA receptor (Chen and Gotschlich, 2001; Provvedi and Dubnau, 1999; Seitz et al., 2014) and translocated across the cytoplasmic membrane by the Com machinery (Chen and Dubnau, 2004). Eventually, imported DNA is integrated in the chromosome in a RecA-dependent manner.

Until recently, no filament-localized DNA receptor had been identified and it was unknown how Tff nano-machines bound extracellular DNA. However, recent findings clearly showed that some type IV pilins can act as DNA receptors. We discovered that the minor Tfp component ComP, which is identical in meningococci and gonococci and is crucial for their competence (Brown et al., 2010; Wolfgang et al., 1999), is the only one of four N. meningitidis type IV pilins with intrinsic DNA-binding ability (Cehovin et al., 2013). Furthermore, quantitative DNA-binding experiments (Berry et al., 2013; Cehovin et al., 2013) showed that although purified ComP can interact with any DNA sequence, it displays a preference for the 12-bp DNA-uptake sequence (DUS) motif (Ambur et al., 2007). This motif is highly repeated in meningococcal and gonococcal genomes and enhances uptake by these species of DNA containing it (Goodman and Scocca, 1988). Selective uptake of their own DNA is a characteristic feature of Neisseriaceae (Frye et al., 2013) and Pasteurellaceae in which a different motif (termed USS) is found (Danner et al., 1980). This property is thought to protect these competent species from indiscriminate transformation by foreign DNA. Identification of ComP as the DUS receptor shed light on the long-standing mystery of how pathogenic Neisseria species manage to recognize and import their own DNA during transformation. The presence of ComP homologs in all other Neisseria species and most Neisseriaceae, some of which were shown to exhibit preferential uptake of their own DNA-containing DUS variants (Frye et al., 2013), suggests that co-evolution of DUS variants and cognate ComP receptors is an elegant mechanism for modulating HGT between competent species.
sharing the same environmental niche (Berry et al., 2013). This scenario is further supported by the finding that DUS variants present in other Neisseria species, which differ from meningococcal DUS by as little as 1 bp, show suboptimal transformation of N. meningitidis (Berry et al., 2013).

High-resolution structural information is necessary to advance our understanding of how the ComP family of type IV pilins recognizes DNA. Although a nuclear magnetic resonance (NMR) analysis generated a low-resolution global fold for meningococcal ComP and highlighted an electropositive surface important for DNA binding and transformation (Cehovin et al., 2013), the finer details were lacking. We therefore embarked upon a comparative structural analysis of two ComP orthologs. In the present study, we shine new light on this poorly understood and fascinating phenomenon by reporting high-resolution 3D structures for ComP from N. meningitidis and ComP_{sub} from the non-pathogenic species Neisseria subflava, which is a common inhabitant of the human upper respiratory tract, and by performing an in-depth functional analysis of ComP_{sub} DNA-binding ability, which shows specificity for DUS_{var1} differing from meningococcal DUS by 1 bp.

RESULTS

DNA Binding with a Preference for Their Cognate DUS Is a Conserved Property in ComP Homologs

Although purified ComP_{sub} was previously shown to bind DNA (Berry et al., 2013), quantitative DNA-binding data were needed to strengthen the notion that all ComP homologs bind their cognate DUS specifically. To overcome previous protein stability problems, we fused the 118-amino-acid (aa) long soluble portion of ComP_{sub} (Figure 1) to non-cleavable N-terminal maltose-binding protein (MBP) or hexahistidine tag (His_{6}). This allowed us to purify well-folded proteins, which we used to test whether ComP_{sub} has a higher affinity for its cognate DUS_{var1}. First, we used acrylamide electrophoretic mobility shift assays (EMSA) to perform competition reactions. We assessed the effect of an excess of unlabeled double-stranded (ds) primer on a pre-formed complex between purified MBP-ComP_{sub} and a biotinylated DUS_{var1} ds primer, which produces a characteristic shift on gel (Figure 2A). While unlabeled DUS_{var1} efficiently out-competed bound biotinylated DUS_{var1} in a dose-dependent manner, as demonstrated by the gradual and eventually complete disappearance of the biotinylated complex, a scrambled SUD primer (in which every second base is altered) had no effect (Figure 2A). Next, the affinity of His_{6}-ComP_{sub} for DUS_{var1} and two different scrambled primers was quantified in real time using surface plasmon resonance (SPR), as previously done for meningococcal ComP (Cehovin et al., 2013). In brief, equivalent amounts of ds biotinylated DUS_{var1}, SUD, and SDU primers were coupled to adjacent channels on a neutravidin-coated sensor chip. Increasing amounts of purified His_{6}-ComP_{sub} (10, 25, 50, 100, or 200 μM) were then injected and the responses at equilibrium (R_{eq}) were measured for each protein concentration (Figure 2B). This clearly showed that ComP_{sub} binds DNA in a dose-dependent fashion, and that the affinity for DUS_{var1} is much higher than the affinity for SUD or SDU, as indicated by the higher R_{eq} at each protein concentration. Unlike binding to the scrambled primer, binding to DUS_{var1} was approaching saturation, allowing us to estimate (using a non-linear least-squares fit) a dissociation constant (K_{D}) of 52.7 ± 2.2 μM. Taken together, these findings confirm that DNA binding, which is tighter in the presence of their cognate DUS, is a conserved property in ComP orthologs.

ComP Orthologs Share Similar 3D Structures with a Highly Distinctive DD Region Stabilized by Two Disulfide Bonds

Since there is no high-resolution structural data for this class of DNA-binding pilins, we first endeavored to solve the 3D structure of its defining member, N. meningitidis ComP. To facilitate purification and crystallization, we used a synthetic comP gene, codon-optimized for expression in Escherichia coli, and fused the 115-aa-long soluble portion of ComP to an MBP modified
to promote crystallization by surface entropy reduction (Moon et al., 2010). This soluble portion of ComP excludes the 6-aa-long leader peptide (Figure 1A), which is processed by the dedicated prepilin peptidase PilD (Brown et al., 2010), and the first 28 residues of the mature protein, most of which correspond to the hydrophobic residues that form the protruding part (α1N) of N-terminal α1 helix in type IV pilins (Berry and Pellic, 2015). The MBP-ComP protein crystallized readily in multiple conditions. After optimizing the best diffracting crystals, we collected a complete dataset on crystals formed in the spacegroup P2₁2₁2₁, which diffracted to a resolution of 1.43 Å (Table S1), and solved the structure of the fusion protein (Figure S1). As can be seen in Figure 3, the ComP moiety adopts the classical type IV pilin fold (Giltner et al., 2012) with the C-terminal part of the long α1 helix (α1C) packed against a β sheet of four antiparallel β strands. This conserved core is highly similar to that of the major pilin PilE (Parge et al., 1995) and the minor pilin PilX (Helaine et al., 2007) (Figure 3B). Subtle differences in ComP’s conserved core include the lack of curvature in α1C (as previously observed for PilX) and the rather long loop (14 residues) connecting β1 and β2, the first two β strands of the β sheet. In contrast, major differences exist in the structurally variable “edges” that usually distinguish type IV pilins (Giltner et al., 2012), the α1β1 loop connecting α1 and β1 and the C-terminal D region delimited by a disulfide bond (hence its name). These regions in ComP are unique (Figure 3). While ComP’s α1β1 loop is an extended unstructured region, its D region is particularly striking. Unlike in (most) other type IV pilins, where one disulfide bond stabilizes the C terminus of the protein by stapling it to β4, in ComP the C terminus forms a long unstructured loop held in place across the face of the β sheet by two disulfide bonds between the four Cys residues in the protein (C76-C125 and C118-C139). The C118-C139 bond delimits the loop by stapling its C terminus back to β4 and is therefore equivalent to the single disulfide bond found in other pilins (Figure 3B). In contrast, C76-C125 is unique and pins the middle of the loop back to β1. To highlight the unique nature of this region, we termed it the DD region. Finally, three of the last five residues of the protein form a short β strand (β5), which is incorporated into the β sheet (Figure 3A).

To determine the structural relationship between ComP orthologs that recognize different DUS, we also solved the structure of ComPsub from N. subflava. The 118-aa-long soluble portion of this protein, which has a binding preference for DUSvar1 (see Figure 2), displays ~52% sequence identity to ComP (Figure 1). Since we could not obtain crystals for the MBP-ComPsub fusion because the protein was too soluble, we decided to explore whether a solution structure determination by NMR would be possible for His6-ComPsub. We isotopically labeled our His6-ComPsub with 13C and 15N for NMR assignment and obtained a high-resolution nuclear Overhauser effect (NOE)-derived structure in solution (Table S1). The ComPsub structures within the NMR ensemble superpose well onto each other, with a root-mean-square deviation (rmsd) of 0.28 Å for all backbone atoms, which suggests that there is no significant flexibility in the structure (Figure 4A). As can be seen in Figure 4B, ComPsub adopts a 3D structure highly similar to that of ComP. The two structures align over their whole length with an rmsd of 2.41 Å for all backbone atoms (Figure 4C). Importantly, the distinctive features highlighted in ComP are also present in ComPsub, namely, the unstructured α1β1 loop, the long β1-β2 loop, and the DD region that is held in place across the face of the β sheet by two disulfide bonds between the four Cys residues (C76-C127 and C118-C141) in ComPsub.

Taken together, these structural findings show that ComP orthologs adopt very similar 3D structures and present a unique structural feature previously unreported in type IV pilins, i.e. a DD region that sits on top of the β sheet and is delimited by two disulfide bonds between four conserved Cys residues.

**ComP Orthologs Bind DNA Mainly via Their DD Region and β1-β2 Loop, which Are Predicted to Be Exposed on the Surface of Tfp**

As already mentioned, ComP’s DNA-binding activity has been characterized to some extent by NMR (Cehovin et al., 2013). However, this analysis was limited by relatively poor NMR data with many key resonances broadened through conformational exchange, which rendered them unassignable and precluded the calculation of a high-resolution solution structure for ComP. In contrast, the NMR data for ComPsub is of much higher quality, leading to complete assignments of 1H, 15N, and 13C nuclei and eventually to the calculation of a high-resolution solution structure (Figure 4). We therefore characterized by NMR the atomic resolution features of DNA binding by His6-ComPsub. We titrated increasing amounts of an unlabeled DUSvar1 ds primer...
Figure 3. High-Resolution 3D Structure of the DUS Receptor ComP from *N. meningitidis*

(A) Crystal structure of the soluble portion of ComP at 1.43-Å resolution. Two different views are shown as a cartoon. The conserved core in type IV pilins (the N-terminal α helix and four-stranded antiparallel β sheet) is depicted in gray. Distinctive/key structural features such as the α1β1 loop (red), the large β1-β2 loop (blue), and the DD region delimited by two disulfide bonds that sits on top of the β sheet (green) are also highlighted.

(B) Structural similarity/differences between the soluble portions of ComP, major pilin PilE from *N. gonorrhoeae* (Parge et al., 1995), and minor pilin PilX from *N. meningitidis* (Helaine et al., 2007). The distinctive/key structural features shown in (A) are also highlighted using the same coloring scheme.

The interaction between ComP<sub>sub</sub> and DUS<sub>var1</sub> led to the broadening of many signals even under saturating DNA concentrations, suggesting that even in the bound state there are significant conformational dynamics. Since this precluded the direct measurement of intermolecular NOEs, we used the HADDOCK software suite (de Vries et al., 2007; Dominguez et al., 2003) to generate a structural model for the ComP<sub>sub</sub>-DUS<sub>var1</sub> complex. All the bases of DUS<sub>var1</sub> were considered as important, because the mutagenesis of even a single base significantly impairs DNA binding and/or transformation (Berry et al., 2013; Frye et al., 2013). In contrast, active residues in ComP<sub>sub</sub> were defined as those experiencing CSP at DNA concentrations below 20 μM (1/5 ratio of DNA to protein), to avoid giving weight to residues showing CSP at later titration points. The lowest overall energy cluster contained ten structures with an rmsd of 2.5 Å and an overall HADDOCK score of 70.7 ± 14.5 (Figure 6A). Complex formation in this cluster resulted in an average buried surface area of 1,558.1 ± 137.3 Å². Closer inspection of this complex reveals that the DNA docks onto the vertical stack of residues on ComP<sub>sub</sub> just above the ledge feature introduced by the DD region. The DD region, β1-β2 loop, and α1-β1 loop intercalate with successive grooves of the DNA, establishing contacts with multiple bases of DUS<sub>var1</sub> (Figure 6A). Finally, using Modeller (Webb and Sali, 2014), we produced a full-length ComP<sub>sub</sub> model using the gonococcal PilE structure (Craig et al., 2006) as a template for the missing N-terminal α1N helix and were able to model its packing within a Tfp structural model (Craig et al., 2006). This revealed that ComP<sub>sub</sub> not only fits readily into the filaments, but that the regions of the protein involved in DNA binding are clearly exposed on the surface of Tfp (Figure 6B).

Taken together, these findings suggest that ComP<sub>sub</sub> binds DNA using unique structural features conserved in this class of type IV pilins and exposed on the surface of the Tfp. These features, i.e. the large β1-β2 loop and DD region, define a new DNA-binding motif differing dramatically from well-known motifs (Luscombe et al., 2000).

**DISCUSSION**

Natural transformation is a widespread biological property playing a key role in bacterial physiology. Although it has been the subject of intense study for almost a century (Griffith, 1928) its very first step, during which Tff nano-machines bind free DNA to subsequently promote its uptake, remains one of the least understood. The discovery that ComP is the only purified meningococcal pilin capable of binding DNA showed that type IV pilins (the Tff subunits) can act as DNA receptors (Cehovin et al., 2013), highlighting a new property for these versatile molecular modules (Giltner et al., 2012). Furthermore, the finding that ComP shows better binding to the 12-bp DUS sequence motif (Ambur et al., 2007), which is highly abundant in pathogenic *Neisseria* species genomes and enhances uptake of their own DNA (Goodman and Scocca, 1988), solved a long-standing mystery. It revealed an elegant mechanism for limiting transformation by foreign DNA that is widespread in competent *Neisseria* species (Frye et al., 2013). A significant barrier to our understanding of how and why the ComP family can bind DNA is the absence of high-resolution structural information for this class of type IV pilins. In this report, we have addressed this limitation, which led to significant findings.
We first provide evidence suggesting that all ComPs are DNA-binding pilins displaying binding preference for their cognate DUS. Our previous finding that meningococcal ComP has a higher affinity for its cognate DUS (Cehovin et al., 2013) could readily be extended to other Neisseria species that contain the same DUS and encode ComPs with ~80%–100% aa identity (such as N. gonorrhoeae, Neisseria lactamica, Neisseria polysaccharea, and Neisseria cinerea). However, to generalize our findings to this entire class of proteins, it was still to be demonstrated that a ComP from more distant Neisseriaceae, which usually display ~30%–50% aa identity to meningococcal ComP (Cehovin et al., 2013), would show a similar binding preference to its cognate DUS that differs from meningococcal DUS by one to four bases (Frye et al., 2013). A previous competition EMSA with purified ComPsub fell short of this goal, although it showed a slightly better competition by its cognate DUSvar1 (Berry et al., 2013). No quantitative DNA-binding data could be obtained because of protein stability/folding issues. These problems were solved here by using novel expression/purification strategies. Using SPR, the affinity of ComPsub for its cognate DUSvar1 was found to be much higher than its affinity for scrambled primers. The KD of ComPsub for DUSvar1 (53 μM) was found to be comparable with the KD of meningococcal ComP (Cehovin et al., 2013), which shares only 27% aa identity with ComP and is expected to recognize a DUS differing from meningococcal DUS by three bases (Figure S2A). ComPsub structural model was found to be virtually identical to meningococcal ComP structure since they align over their whole length with an rmsd of 0.26 Å for all backbone atoms (Figure S2B). Once other DNA-binding pilins are identified (that are not homologous to ComP), it will be interesting to explore whether they share a similar structure to ComP and whether the DD region is a conserved DNA-binding motif. This will be particularly interesting once the USS receptor in Pasteurellaceae is identified, because no ComP homolog is present in these species and their USS bears no resemblance to DUS motifs.

We have also derived high-resolution structures for ComP and ComPsub orthologs. Consistent with the notion that they are minor components of Tfp (Aas et al., 2002; Brown et al., 2010), these structures revealed that ComP and ComPsub are bona fide type IV pilins and can pack efficiently within the available model for Tfp (Craig et al., 2006). Interestingly, these proteins also exhibit a distinctive structural feature exposed on the surface of the pilus, the DD region delimited by two disulfide bonds, which sits across the β sheet. The fact that this structural motif is not found in the dozens of type IV pilin structures available in the databases (Giltnner et al., 2012) is consistent with ComP orthologs being (so far) the only known type IV pilins that bind DNA. Considering that (1) the 3D structures of ComP and ComPsub are more distant ComPKor ortholog from Kingella oralis, which shares only 27% aa identity with ComP and is expected to recognize a DUS differing from meningococcal DUS by three bases (Figure S2A). ComPsub structural model was found to be virtually identical to meningococcal ComP structure since they align over their whole length with an rmsd of 0.26 Å for all backbone atoms (Figure S2B). Once other DNA-binding pilins are identified (that are not homologous to ComP), it will be interesting to explore whether they share a similar structure to ComP and whether the DD region is a conserved DNA-binding motif. This will be particularly interesting once the USS receptor in Pasteurellaceae is identified, because no ComP homolog is present in these species and their USS bears no resemblance to DUS motifs.

The last, and perhaps most important, finding in this study is that the mode whereby ComP orthologs interact with DNA has not been seen before. Rather than well-known and widespread DNA-binding motifs (helix-turn-helix, zinc finger, leucine zipper, and so forth) (Luscombe et al., 2000), ComP orthologs exploit a series of residues in two distinct domains, i.e. the DD region.
and the tip of the β1-β2 loop (although a small part of the α1-β1 loop is also involved), to establish contacts with multiple bases of DUS in successive grooves of the ds DNA. These residues form an almost vertical stack on a face of the protein predicted to be exposed on the surface of the filaments, which was previously implicated in meningococcal ComP binding to DUS (Cehovin et al., 2013). As previously noted, this surface is highly positively charged (Figure S3) and is likely to be involved initially in electrostatic attraction of the negatively charged DNA. Interestingly, the DD region and the β1-β2 loop, which are flanked by almost invariable residues, are among those regions that differ the most between ComP orthologs (see Figures 1 and S2). This is likely to be the reason why these proteins recognize different DUS motifs. It is now possible to propose a model for the mode of action of these DNA receptors in natural transformation (Figure 7). When Neisseriaceae encounter free DNA they might attract it, electrostatically at first, and use ComP subunits in their Tfp to “scan” it for the presence of their cognate DUS. Once a cognate DUS has been recognized (it is unknown at this stage which specific aa differences in different ComPs are responsible for their different DUS specificities), ComP “docks” onto it using the vertical stack residues exposed on the filament surface, which is probably accompanied by conformational changes further increasing the strength and specificity of the interaction. Upon rotation of the pilus during PilT-powered retraction (Nivaskumar et al., 2014), the DNA docked at ComP anchor point(s) might wrap around the Tfp following the newly noted grooves on their corrugated surface (Craig et al., 2006). Eventually, upon translocation across the secretin pore, this would result in DNA uptake.

In conclusion, by providing high-resolution structural information for the ComP class of proteins, this study has shed light on the atomic basis for their DNA-binding ability, which is yet an additional property for the highly versatile type IV pilins. This has led to the discovery of a novel DNA-binding domain, which is consistent with the fact that members of the ComP family are, to the best of our knowledge, the only known surface-exposed DNA receptors that bind specific DNA sequences.

EXPERIMENTAL PROCEDURES

Protein Production and Purification of N. meningitidis ComP
A synthetic gene, codon-optimized for E. coli expression, encoding ComP from N. meningitidis 8013 (Rusniok et al., 2009), was synthesized by GeneArt. The portion of optimized comP encoding residues 29–143 from the mature protein was PCR amplified using optcomP-F and optcomP-R primers (Table S2), cut with EcoRI and HindIII and cloned into the pMALX(E) vector cut with the same enzymes. This resulted in the fusion of the soluble portion of ComP with a modified MBP carrier containing several mutations designed to promote crystallization (Moon et al., 2010). The resulting plasmid was verified by sequencing and transformed into chemically competent E. coli SHuffle T7 express cells (New England Biolabs), which enables formation of disulfide bonds in the cytoplasm. A single colony was transferred to 5 ml of Luria-Bertani (LB) medium (Difco) supplemented with ampicillin (100 μg ml⁻¹) and allowed to grow to saturation at 30 °C overnight in an orbital shaker. The
Protein Production and Purification of ComP

The affinity-purified MBP-ComP was further washed extensively with lysis buffer A to remove unbound material before equilibrated in lysis buffer A, to bind MBP-ComP. The column was then FPLC through a 5-ml MBPTrap HP column (both from GE Healthcare), pre-washed in lysis buffer B, was then added to the clarified lysate and exchanged into 50 mM Na2HPO4/NaH2PO4 (pH 6) and 50 mM NaCl. This lysate was then centrifuged at 8,000 x g for 20 min and subjected to one freeze/thaw cycle in lysis buffer A (50 mM Tris-Cl [pH 8], 100 mM NaCl, 1 x SIGMAFAST EDTA-free protease inhibitor cocktail [Sigma]). This lysate was further disrupted by repeated cycles of sonication, i.e. pulses of 5 s on and 5 s off during 3–5 min, until the cell suspension was visibly less viscous. The cell lysate was then centrifuged at 30 min for 18,000 x g to remove cell debris. The clarified lysate was then passed using an Akta Purifier FPLC through a 5-ml MBPTrap HP column (both from GE Healthcare), pre-equilibrated in lysis buffer A, to bind MBP-ComP. The column was then washed extensively with lysis buffer A to remove unbound material before the fusion protein was eluted using TSM buffer (50 mM Tris-Cl [pH 8], 500 mM NaCl, 20 mM imidazole, 1 x SIGMAFAST EDTA-free protease inhibitor cocktail). 2 ml of Ni-NTA (nitrilotriacetic acid) agarose resin (Qiagen), pre-washed in lysis buffer B, was then added to the clarified lysate and incubated for 2 hr at 4°C with gentle agitation. This chromatography mixture was then filtered through a Poly-Prep gravity-flow column (Bio-Rad) and washed several times with lysis buffer B, before eluting the protein with lysis buffer B containing 500 mM imidazole. The affinity-purified MBP-ComP was further purified by gel-filtration chromatography on a HiLoad 16/60 Superdex 200 column using TSM buffer for elution.

**Figure 6.** Modeling of the ComPsub-DUSvar1 Complex and ComPsub Packing within Tfp

(A) HADDOCK model of the interaction between ComPsub (in which the α1β1 loop, β1-β2 loop, and DD region are highlighted) and DUSvar1 (in cyan) in cartoon representation. Two different views are shown. (B) Packing of full-length ComPsub into Tfp. A full-length ComPsub model was generated using Modeller. One PilE subunit in the Tfp model was then replaced by this full-length ComPsub. The structural features of ComPsub involved in DNA binding are highlighted on the surface representation.

**Protein Production and Purification of N. subflava ComP**

For assessment of the DNA-binding activity of ComPsub by EMSA, the portion of N. subflava comP encoding residues 29–146 from the mature ComPsub protein was PCR amplified from N. subflava NJ9703 genomic DNA (Marri et al., 2010) using comPsub-pMalF and comPsub-pMalR primers (Table S2) and cloned as above into the pMALX(E) vector. This resulted in the fusion of the soluble portion of ComPsub with a non-cleavable MBP carrier. The fusion MBP-ComPsub protein was purified similarly to MBP-ComP.

**Figure 7.** Model for the Role of ComP DNA-Binding Pilins in Natural Transformation in Neisseriaceae

For determination of the 3D structure of ComPsub, the above portion of N. subflava comP was amplified using hiscomPsub-pETF and hiscomPsub-pETR primers (Table S2) and cloned into the pET-28b vector (Novagen). The forward primer was designed to fuse a non-cleavable N-terminal His6 tag to ComPsub. The resulting plasmid was back-diluted 1:50 into 10 ml of M9 minimal medium supplemented with a mixture of vitamins and trace elements. This was grown to saturation overnight at 30°C and an orbital shaker, then back-diluted 1:500 into 1 l of the same medium containing 13C D-glucose and 15N NaCl for isotopic labeling. Cells were grown in an orbital shaker at 30°C until the OD600 reached 0.8, before adding 0.4 mM IPTG to induce protein production overnight at 30°C. Cells were then harvested and disrupted as above in lysis buffer B (50 mM NaHPO4/NaH2PO4 [pH 7.4], 500 mM NaCl, 20 mM imidazole, 1 x SIGMAFAST EDTA-free protease inhibitor cocktail). 2 ml of Ni-NTA (nitrilotriacetic acid) agarose resin (Qiagen), pre-washed in lysis buffer B, was then added to the clarified lysate and incubated for 2 hr at 4°C with gentle agitation. This chromatography mixture was then filtered through a Poly-Prep gravity-flow column (Bio-Rad) and washed several times with lysis buffer B, before eluting the protein with lysis buffer B containing 500 mM imidazole. The affinity-purified His6-ComPsub was further purified by gel-filtration chromatography on an Akta Purifier using a Superdex 75 10/300 Gl column (GE Healthcare), and simultaneously buffer-exchanged into 50 mM NaHPO4/NaH2PO4 (pH 6) and 50 mM NaCl.

**Crystallization and Structure Determination of Meningococcal ComP**

Purified MBP-ComP was concentrated to 7, 14, and 28 mg ml–1 and placed through sitting-drop vapor diffusion crystallization trials in MRC two-well crystallization plates (Hampton Research) using a wide range of commercially available kits. Trials were initially conducted using 100 nl of protein and 100 nl of mother liquor, over a reservoir of 80 µl of mother liquor. The trials produced a large number of initial hits, which were varied and optimized to yield larger and better diffracting crystals by altering the pH, precipitant concentration, and drop size. The crystals used for structure determination were obtained when the purified protein at 14 mg ml–1 was mixed 1:1 with crystallization liquor.
containing 0.1 M sodium cacodylate (pH 6.5) and 25% polyethylene glycol 4000, and left to equilibrate with the reservoir solution at 20°C. Crystals were cryoprotected using crystalization liquor containing 20% ethylene glycol and flash-cooled in liquid nitrogen. Data were collected on beamline I02 at the Diamond Light Source and processed with xia2 (Winter, 2010). Molecular replacement was carried out using Phaser MR (McCoy et al., 2007) with MBP structure (PDB: 1HSJ) as a search model. ComP was then built manually into the remaining density using Buccaneer (Cowtan, 2006) and Coot (Emsley et al., 2010). Multiple iterative rounds of refinement and model building were carried out using Phenix (Adams et al., 2010) and Coot, resulting in a final R_work/R_free of 0.16/0.19.

**NMR Structure Determination of ComPsub**

Isotopically labeled purified His6-ComPsub was concentrated to ~750 μM in NMR buffer (50 mM Na2HPO4/NaH2PO4 [pH 6], 50 mM NaCl, 10% D2O). A full set of triple-resonance NMR spectra were recorded on a Bruker Avance III 800-MHz spectrometer equipped with triple-resonance cryoprobes at 295 K, and processed with NMRPipe (Delaglio et al., 1995). Backbone assignments were completed using a combination of HBHA, HN(CA)CO, HNCACB, and CBCA(CO)NH experiments using NMRView (One Moon Scientific) as previously described (Johnson and Blevins, 1994). Side-chain resonance assignments were obtained from a combination of C(CO)NH, HC(C)H-total correlation spectroscopy (TOCSCY), and (HC-H-H)-correlation spectroscopy experiments using an in-house software developed within NMRView (Marchant et al., 2008). Distance restraints were obtained from 2D 1H1H15N-NOE and 1H1H13C-NOE spectra and used for structure calculations in ARIA 2.3 (Rieping et al., 2007), along with dihedral angle restraints obtained from chemical-shift values calculated using the TALOS+ server (Shen et al., 2009). During later rounds of calculations, disulfide restraints for C76-C127 and C118-C141 were introduced, after these had already formed in the calculated structures. For each round of calculations, 20 structures were calculated over eight iterations. In the final iteration, the 20 lowest-energy structures were submitted to a water-refinement stage to form the final structural ensemble.

**NMR Chemical-Shift Perturbation Analysis in ComPsub upon DNA Binding**

The residues in ComPsub important for DNA binding were identified by NMR. Complementary primers corresponding to DUSvar1 were first dissolved in NMR buffer to a final concentration of 6 mM, combined in equimolar amounts, heated for 5 min at 95°C and left to cool overnight to produce 3 mM ds target DNA. The ds primer was washed with NMR buffer using G-25 spin columns (GE Healthcare) to remove residual buffer components from the oligonucleotide synthesis. The ds target DNA was titrated into a 100 μM sample of His6-ComPsub in NMR buffer, at increasing concentrations (4, 5, 10, 20, 40, 80, and 120 μM), and 2D 1H1H15N-HSQC spectra were recorded at each titration point revealing changes to backbone chemical shifts.

**DNA-Binding Assays**

DNA binding by ComPsub was assessed using EMSA or SPR. EMSA was performed as follows (Berry et al., 2013; Cehovin et al., 2013). Biotin-labeled and non-labeled ds primers corresponding to DUSvar1 or scrambled sequences (Table S2) were prepared by mixing equimolar amounts of two complementary oligonucleotides in 50 mM Tris (pH 8.0) and 100 mM NaCl, heating for 5 min at 95°C, and leaving to cool overnight. Purified MBP-ComPsub was prepared in the same buffer. An MBP-ComPsub/MBP-ComPsub complex was generated by mixing 5 fmol of biotinylated DUSvar1 ds primer and 0.8 μM MBP-ComPsub (except for a DNA-only control) in 20 μl of 20 mM Tris-Cl (pH 8), 50 mM NaCl, and 2.5 mM MgCl2, and incubating for 20 min at ambient temperature. Increasing concentrations of unlabeled competitor DNA (0.7, 2.8, 11.25, 45, and 180 pmol) were then added to the DNA-binding reactions, which were further incubated for 20 min at ambient temperature before being analyzed by native gel electrophoresis as described previously (Cehovin et al., 2013).

SPR was performed as follows (Berry et al., 2013; Cehovin et al., 2013). Equivalent amounts of biotin-labeled ds primers, prepared as above in 20 mM Tris (pH 8) and 150 mM NaCl, were coupled to neutravidin on the surface of a ProteOn NLC sensor chip resulting in ~265 RU as assessed on a ProteOn XPR36 protein interaction array system instrument (Bio-Rad). SPR was then performed by passing 10, 25, 50, 100, and 200 μM His6-ComPsub (in 20 mM Tris [pH 8], 150 mM NaCl, 0.05% Tween 20) across the six available analyte channels of the chip at 60 μl min⁻¹, and the responses at equilibrium (Rₑq) were recorded. A control trace was also collected using an empty ligand channel and used to normalize for non-specific binding effects. Four independent analyte injections were performed, with a regeneration step performed between each using 0.5 M NaCl. All experiments were carried out at 25°C.

**ACCESSION NUMBERS**

The X-ray crystal structure of ComP and NMR structure of ComPsub have been deposited in the PDB under accession codes PDB: 5HZ7 and 2NBA, respectively.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes three figures and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.str.2016.04.001.

**AUTHOR CONTRIBUTIONS**

S.J.M. and V.P. designed and directed the research. All the experiments were done by J.L.B. Y.X. helped with NMR, P.N.W. and S.M.L. helped with SPR. J.L.B., S.J.M., and V.P. wrote the paper.

**ACKNOWLEDGMENTS**

This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) and supported by the Wellcome Trust (via a Senior Investigator Award and a multi-user equipment grant to S.J.M.). We are grateful to Vivianne Goosens and Sophie Helaine (both from Imperial College London) for critical reading of the manuscript. We are grateful to Nathan J. Weyand (University of Arizona) for the gift of N. subflava genomic DNA.

Received: February 3, 2016

Revised: March 19, 2016

Accepted: April 4, 2016

Published: May 5, 2016

**REFERENCES**


