THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

pubs.acs.org/JPCA

# Molecular and Electronic Structure of the Peptide Subunit of *Geobacter sulfurreducens* Conductive Pili from First Principles

Gustavo T. Feliciano,[†,‡] Antonio J. R. da Silva,[†] Gemma Reguera,*[,§,∥] and Emilio Artacho[‡,⊥,#,¶,∥]

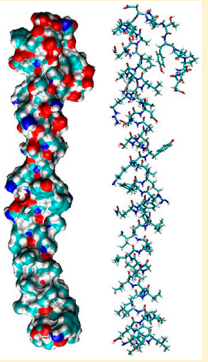
[†]Departamento de Física dos Materiais e Mecânica, Instituto de Física, Universidade de São Paulo, SP, Brazil

[‡]Department of Earth Sciences, University of Cambridge, Cambridge CB2 3EQ, United Kingdom

[§]Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan 48824, United States

Supporting Information

**ABSTRACT:** The respiration of metal oxides by the bacterium *Geobacter sulfurreducens* requires the assembly of a small peptide (the GS pilin) into conductive filaments termed pili. We gained insights into the contribution of the GS pilin to the pilus conductivity by developing a homology model and performing molecular dynamics simulations of the pilin peptide in vacuo and in solution. The results were consistent with a predominantly helical peptide containing the conserved $\alpha$-helix region required for pilin assembly but carrying a short carboxy-terminal random-coiled segment rather than the large globular head of other bacterial pilins. The electronic structure of the pilin was also explored from first principles and revealed a biphasic charge distribution along the pilin and a low electronic HOMO−LUMO gap, even in a wet environment. The low electronic band gap was the result of strong electrostatic fields generated by the alignment of the peptide bond dipoles in the pilin's $\alpha$-helix and by charges from ions in solution and amino acids in the protein. The electronic structure also revealed some level of orbital delocalization in regions of the pilin containing aromatic amino acids and in spatial regions of high resonance where the HOMO and LUMO states are, which could provide an optimal environment for the hopping of electrons under thermal fluctuations. Hence, the structural and electronic features of the pilin revealed in these studies support the notion of a pilin peptide environment optimized for electron conduction.



## INTRODUCTION

Extracellular electron transfer to Fe(III) oxide minerals is the hallmark of the physiology of bacteria in the family *Geobacteraceae*.[1] The insoluble nature of the electron acceptor has selected for a complex respiratory network across the cell envelope that electronically connects the cell with its environment.[2] Much of what we know about this respiratory assembly comes from studies in *Geobacter sulfurreducens*,[3] which serves as a genetically tractable model representative of the family.[4] Electron transfer across the multilayered cell envelope of *G. sulfurreducens* involves metalloproteins, mostly *c*-cytochromes,[2] and also requires the assembly of thin protein filaments known as pili.[5] The pili of *G. sulfurreducens* (herein called GS pili) are conductive and required for growth with Fe(III) oxides, suggesting that they are the main electrical connection between the cell and the extracellular Fe(III) oxides.[5] The GS pili maximize the cellular surface available for electron transfer reactions and have also been proposed to promote efficient electrical contact with the Fe(III) oxides, which are often heterogeneously dispersed as coatings on small clay particles and other particulate matter in soils.[5] Although bacterial pili are the most widespread cell structure for cell attachment and surface motility,[6] evidence today suggests that the GS pili do not mediate these functions.[5,7] Rather, they carry out a specialized role as electronic conduits between the cell and extracellular electron acceptors (such as insoluble Fe(III) oxide minerals[5] and toxic, soluble contaminants such as uranium[8]) or across electroactive biofilms.[7,9]

Scanning tunneling microscopy (STM) spectroscopy along the GS pili revealed electronic states near the Fermi level, consistent with a conducting material, but did not detect any electronic signatures of heme groups, which could have reflected the contribution of redox-active proteins such as cytochromes.[10] Furthermore, topographic and electronic features intrinsic to the GS pilus fiber were readily apparent in STM images, consistent with the contribution of the pilus protein to its conductivity.[10] As the GS pili are homopolymers of a single peptide subunit, the GS pilin,[8] the electronic structure of the pilus fiber resolved by STM reflects the contribution of the density of states of the pilin subunit. This suggests that the GS pilin has evolved to promote both intramolecular and intermolecular charge transport.

Much of what is currently known about the GS pilin comes from amino acid sequence analyses of the encoding gene or *pilA*.[5] The GS pilin is synthesized as a precursor or prepilin carrying a signal peptide with the distinctive features that enable recognition and processing by a type IV pilin peptidase or PilD,[11] which also *N*-methylates the newly exposed N-terminal amino acid residue.[6,12] The amino(N)-terminal amino acid of the mature GS pilin is a phenyalanine, as in pilins of the type IVa subgroup.[12] However, unlike other type IVa pilins, which have an average length of ca. 150 amino acids,[12] the mature GS

**Received:** March 7, 2012
**Revised:** July 4, 2012
**Published:** July 10, 2012



   dx.doi.org/10.1021/jp302232p | *J. Phys. Chem. A* 2012, 116, 8023−8030

pilin is only 61 amino acids long. Furthermore, despite sharing homology with the N-terminal region of other type IVa pilins, phylogenetic analyses placed the GS pilin in an independent line of descent along with pilin subunits from other members of the family *Geobacteraceae*.[5] This is consistent with a divergent amino acid sequence that could be related to the peptide's specialized function in electron transfer.

Studies in well-characterized biological systems (such as ruthenium-modified *c*-cytochrome or azurin, photosynthetic reaction centers, cytochrome *c* oxidase, and DNA photolyase enzyme) or synthetic peptides have provided evidence for the involvement of structural motifs and individual amino acids in intramolecular and intermolecular electron transfer reactions.[13−17] Seminal work by Beratan, Onuchic, and collaborators,[18,19] for example, successfully predicted the relative rates of electron transfer in rhuthenium-modified metalloproteins by regarding the protein as a complex tridimensional tunneling barrier for electrons, with specific residues mediating electronic coupling. However, single-step tunneling through proteins is limited by the distance (~20 Å) that electrons can tunnel at biologically relevant rates.[20] For larger distances, relay amino acids are often involved that promote the hopping of electrons[14] and accelerate the rates of electron transfer.[21] Furthermore, the protein's structural dynamics and electrostatic properties, which also depend on the protein's environment, also influence the electron transfer rates greatly.[22,23] Hence, the protein's structure and amino acid composition can tune the mechanism and rate of electron transfer.

Interestingly, the GS pilin contains several aromatic amino acids in positions conserved in the pilin subunits of other bacteria in the family *Geobacteraceae*,[5] which could provide a relay pathway for intramolecular and intermolecular electron transfer. Furthermore, although structural information of the GS pilin is lacking, its assembly into the pilus fiber suggests structural conservation of the pilin's $\alpha$-helix domain that is necessary for pilin polymerization,[24] which could provide an optimal peptide environment for electronic coupling.[25] We, therefore, investigated the structure of the GS pilin by developing a homology model and a molecular mechanics method based on empirical potentials. We also took a first-principles, quantum-mechanical approach to enable the theoretical exploration of the main electronic structural features of the GS pilin around the electronic band gap. This is necessarily connected to structural and electrostatic characteristics of the molecule in close to natural conditions. The electrostatics affects the electronic signatures substantially due to the strong fields generated by ions, both in solution and within the protein, the latter including the end amino and carboxylic groups (the zwitterion form), and similar ions related to particular residues specific to the GS pilin. In addition, the dipoles associated with peptide bonds, which align in the pilin's $\alpha$-helix, gave rise to additional electrostatic effects and contributed to the low band gaps. Hence, the results are consistent with a divergent pilin structure and amino acid composition evolved to favor electron transfer reactions.

## METHODS

**A. Homology Modeling of the GS Pilin.** A homology model of the GS pilin was built using the Swiss-PdbViewer (DeepView) software[26] available at the ExPASy Bioinformatics Web Portal (http://www.expasy.org/spdbv). The high-resolution X-ray crystal structure of the full-length pilin of *Pseudomonas aeruginosa* strain K or PAK pilin[27] was retrieved from the Protein Data Bank Europe (PDBe; PDB entry 1oqw) and used as a template. The PyMOL software was then used to generate the 3D images of the PAK and GS pilin structures shown in Figure 1. The atomic coordinates of the GS pilin can be found in the Supporting Information.

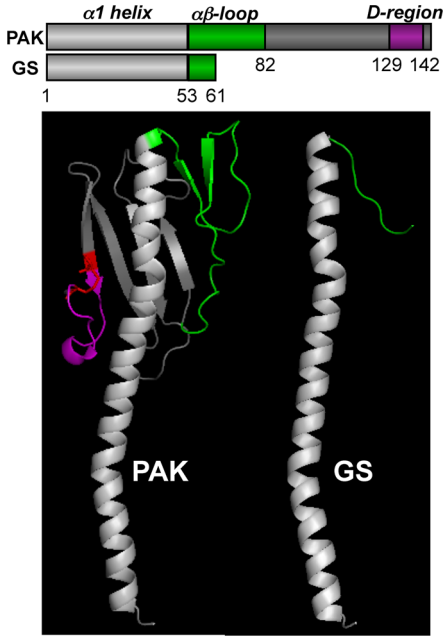



**Figure 1.** Structural alignment of the PAK and GS pilins showing the relative lengths (in amino acids) of the $\alpha$1-helix (light gray), the $\alpha\beta$-loop (green) and the D-region (magenta) (top). Structure of the GS pilin modeled against the structure of the PAK pilin and color-coded to match the location of the $\alpha$-helix, $\alpha\beta$-loop, and D-region (bottom).

**B. Molecular Mechanics (MM) Simulations.** The homology model of the GS pilin provided the initial atomic coordinates for structural and dynamical studies (molecular dynamics) using the standard all-atom version of the optimized potentials for liquid simulations (OPLS) empirical force field.[28] OPLS consists of empirical potential terms for binary interactions (including Coulombic among charges and Lennard-Jones and bond-harmonic-stretching contributions), bond-bending and dihedral angle terms. All simulations were done in periodic boundary conditions, with an Ewald-summation treatment of the long-range electrostatics. All molecular mechanics (MM) simulations were done using the GROMACS code.[29]

For in vacuo studies, the homology model, which comprised 944 atoms in total, was confined to a cubic box of 104.04 Å per side, such that the smallest distance between neighboring molecules was 10 Å. The system was relaxed by molecular dynamics (MD) using the empirical force field at 1 ns, with 1 fs time steps. A thermostat was used to control the temperature around 300 K. For conformational analysis, we took statistically independent snapshots (sufficiently far apart in time) of the MD run. The conformers were then relaxed by minimizing the potential to the nearest minimum. The atomic coordinates obtained by the MM and quantum-mechanical (QM) relaxations in vacuo (see below) can be found in the Supporting Information. As a control, we also modeled an 18 amino acid-long polyalanine $\alpha$-helix and a 10 amino acid long polyproline helix in dry conditions, to access the specific

contributions of the amino acid composition and the helical structure.

The molecular-dynamics simulations including liquid water used the empirical TIP3P force field for pure liquid water[30] which considers rigid water molecules, a Lennard-Jones potential for the intermolecular interactions, as well as three effective point charges accounting for the charge distribution in each water molecule. The protonation states of the residues were set to match those corresponding to a pH of 7. This resulted in an overall charge of −2|e|, which was compensated by countercations in solution by dissolving Na and Cl ions into the water at a 0.2 M concentration. The same cell in periodic boundary conditions used for the in vacuo calculations was used as well, except that the whole box was filled with water, which surrounded the pilin molecule so as to avoid water surfaces and bubbles. The number of water molecules was chosen so that the overall pressure remained close to room pressure in average for the original box (with the shortest distance among periodic images of the pilin being ∼10 Å). The total number of atoms in the cell was 4580. A first ns of MD was run with the empirical MM potential for the protein with a fixed structure, so as to allow the water and solute ions to relax around it, at 300 K. The linear constraint solver algorithm (LINCS18) was used for the purpose. A second 1 ns run allowed for a full relaxation, but using a *NPT* ensemble, i.e., with a thermostat at 300 K and a barostat for $P = 1$ bar to allow for the pressure relaxation of the wet system. The zwitterion form of the polyalanine peptide was embedded in a NaCl solution following the same procedures used for the GS pilin, the whole system consisting in 1041 atoms.

**C. First Principles.** The quantum-mechanics (QM) first-principles calculations used in this study are based on density-functional theory (DFT), using the generalized gradient approximation for the exchange and correlation potential, in its PBE form.[31] Core electrons were replaced by norm-conserving pseudopotentials,[32] fully factorized[33] and using nonlinear partial-core corrections for pseudowave function smoothness close the nuclei. The Kohn−Sham wave functions for the valence electrons were expanded in a double-polarized (DZP) basis set of soft-confined, finite-support numerical atomic orbitals[34] and specific basis set and pseudopotentials.[49] The calculations were done using the SIESTA method,[35] using the Γ-point approximation for *k*-point sampling, and a finite real-space grid corresponding to a 150 Ry plane-wave cutoff for the integrals in real space.[35] Energy minimizations for structure relaxations were performed using conjugate gradients, to within a tolerance of 30 meV/Å in the maximum force component for the whole system. For the studies of the protein in vacuo, a classical conformation analysis was made to identify the relevant structures and a subsequent QM optimization was carried out in each representative structure. The GS pilin was also studied in solution to mimic the pilin's natural environment. For these studies, statistically independent snapshots of the MM simulations were extracted for QM single-point calculations and used to calculate statistical averages of the electronic structure and electrostatics from first principles. Ten snapshots, spaced by 50 ps, were used, and the last 0.5 ns of the last simulation were sampled.

## RESULTS AND DISCUSSION

**A. Homology Model of the GS Pilin in Reference to the PAK Pilin.** The structure of the GS pilin was investigated by constructing a homology model using the X-ray crystal structure of the full-length PAK pilin[27] as a template (Figure 1). The PAK pilin serves as a structural model for type IVa pilins,[27,36−38] as it shares with them a conserved modular architecture: an α-helix spanning approximately the 53 N-terminal amino acids (α1 domain) and a globular domain containing an antiparallel β-sheet region and two conserved cysteines, which form a disulfide bond between the β-sheet and the C-terminal segment of the helix.[24] The globular head of bacterial pilins also contains an αβ-loop region between the α1 helix and the β-sheet region and a C-terminal D-region that is flanked by the two conserved cysteines (Figure 1). The structure of the GS pilin is unique among all known type IVa pilins because, although it contains the hydrophobic N-terminal α-helix (α1 domain), it lacks the globular head, which is replaced by a short random-coiled segment at the C-terminus (Figure 1). The α-helix is a flexible region that facilitates subunit assembly into the hydrophobic core of the pilus filament and confers on the fiber both strength and flexibility.[12] Thus, conservation of the α1 domain in the GS pilin is consistent with a conserved mechanism of pilin assembly and the formation of a flexible and mechanically strong pilus fiber. However, the GS pilin lacks the C-terminal globular head with the αβ-loop and D-region that determine the shape and chemistry of the exposed globular head and confer on the pilus fiber specific biological functions such as adhesion, surface (twitching) motility and cell aggregation.[24,27,39] Not surprisingly, pili are not required for adhesion and surface motility in *G. sulfurreducens*.[5,7,9] As with other bacterial pili, the GS pili are required for cell−cell aggregation and biofilm formation, yet in *Geobacter* these functions also require the pili to function as electronic conduits between the cells.[5,7,9]

The lack of a C-terminal globular head also reduces the size of the GS pilin peptide considerably (61 amino acids, compared to 144 amino acids in the PAK pilin) and locates most of the pilin's amino acids in the α-helix. Interestingly, peptide length and helical structure contribute greatly to peptide conductivity.[40] Electron transfer is faster, for example, in α-helices than in extended β-strands because electronic coupling is higher in helical structures.[25] Furthermore, the permanent dipole of helical peptides generates an electrostatic field along the helix axis dominated by a positive charge at the N-terminus and a negative charge at the C-terminus that promotes intramolecular electron transfer.[41] Such electrostatic effect is minimized in PAK-like pilins because the globular head embeds much of the C-terminal half of the α-helix (the so-called α1-C domain) and promotes electrostatic interactions between the α1-C residues and the β-sheet region.[24] This effectively neutralizes the negative charge at the α1-C end and the electrostatic field along the helix axis. Reduced electronic coupling is also expected in PAK-like pilins due to the presence of β-strands in the globular head.[25] Not surprisingly, the PAK pili are not conductive by conductive probe atomic force microscopy (CP-AFM) whereas the GS pili are.[5] Hence, the divergent molecular structure of the GS pilin compared to other bacterial pilins is consistent with a peptide environment optimized for electron transfer.

**B. Molecular Dynamics of the GS Pilin in Vacuo.** As the conformational dynamics of a protein are critical to their function,[42] we used the atomic coordinates provided in the homology model of the GS pilin (Figure 1) for in vacuo calculations after neutralizing the side chains and the N- and C-termini of the peptide. As shown in Figure 2A, the MD simulations show that the angle between the C-terminal

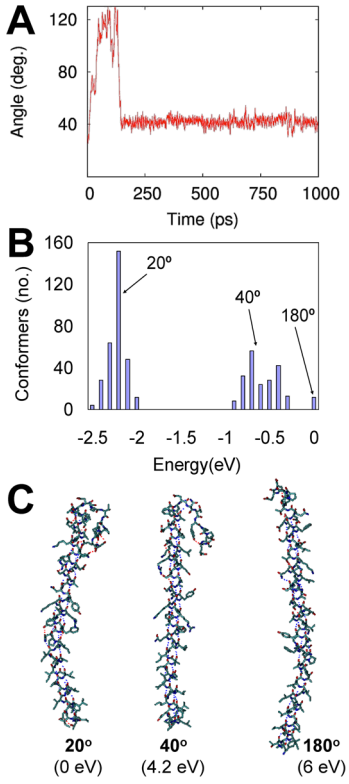



**Figure 2.** (A) Distribution of the C-terminal segment overtime in 1 ns MD. (B) Histogram of conformer distribution from the conformational analysis study. (C) QM-optimized conformers derived from representative conformers generated in (B), each having the C-terminal segment at a different angle (20, 40, or 180°) with respect to the pilin's $\alpha$-helix. The energy values associated with each conformer are shown in brackets (eV).

segment and the $\alpha$-helix opens up with temperature but stabilizes at 40°. We identified three types of conformers, each having the C-terminal, random-coiled segment at a different angle (20 °, 40 °, and 180°) with respect to the $\alpha$-helix, with the 20° conformer displaying the highest stability and the 180 ° being the most unstable (Figure 2B).

One conformer from each representative type was chosen for further QM optimization. The energies obtained for each conformer after QM optimization revealed a similar pattern of stability among the conformers (20° > 40° > 180°) (Figure 2C). However, higher energy differences were observed between the conformers. The conformer with the lowest-energy (20°), which was used as a reference, was 4.2 eV more stable than the 40° conformer and 6 eV more stable than the conformer with the highest energy (180°). Although lower energy structures cannot be ruled out by this method, the low-energy structure (20°) obtained represents a plausible approximation to a stable conformer in these conditions. Not only does the C-terminal segment contain amino acid residues known to destabilize helical conformations, but the 20° conformer also matches the position where the number of hydrogen bonds are maximized in the neutral protonation state. These results can be interpreted as there being at least three clear basins in the energy landscape for the molecule, each capturing the relative disposition of the tail with respect to the helix. Although the deeper basin is the one with 20° angle, the 40° one allows for more entropy and, therefore, stabilizes with temperature. These calculations thus validated the homology model, which showed the C-terminal segment at an angle with respect to the $\alpha$-helix. They also confirmed the predominantly helical structure of the GS pilin and conservation of the pilin's $\alpha$-helix region, which displays robust stability in vacuo at 300 K as expected of a conserved domain involved in pilin assembly.

**C. Electronic Structure of the GS Pilin in Vacuo.** The charge distribution along the GS pilin modeled in vacuo revealed a substantial net electric field along the peptide (Figure 3), which is likely to be larger once corrected for periodic

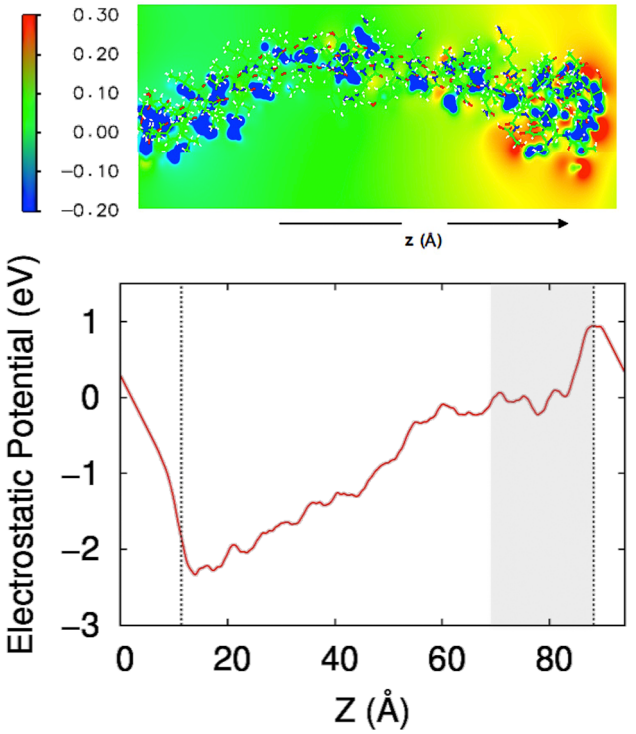



**Figure 3.** Color-coded map (top) and plot (bottom) of the electrostatic potential for the lowest-energy conformer of the GS pilin in vacuo obtained from first principles. The map shows a plane generated as a longitudinal cross section of the pilin molecule, and the plot shows the electrostatic potential versus $z$ averaged in the perpendicular $x$, $y$ plane (pilin length is highlighted by vertical dashed lines and the C-terminal segment, by the gray boxed area).

boundary effects. The imposed periodicity on the potential generates a potential drop across the vacuum gap, to compensate for the ramp within the molecule. It is analogous to having an isolated molecule with an applied electric field countering the intrinsic one. Although methods exist to correct the electric field, e.g., by introducing a dipole correction in vacuo (a narrow capacitor), such corrections are appropriate for arrays of two-dimensional slabs, but not for the system at hand. It is, however, not essential for extracting qualitative information. Such an electric field tends to push up electronic states from one end, and bring down the ones at the opposite end. Consistent with this, the electronic structure of the GS pilin in vacuo gives an electronic band gap of 0.2 eV. This is much smaller than expected for any protein, even when considering the tendency of generalized gradient approximations density functional theory (GGA-DFT) to underestimate the gap. One can thus see what happens with the simplified picture of a typical band gap in such a protein (a few electronvolts) being locally shifted along the helix axes, thus giving a reduced global gap.


dx.doi.org/10.1021/jp302232p | J. Phys. Chem. A 2012, 116, 8023−8030

A short (18 residues) polyalanine $\alpha$-helix with intact carboxylic and amine ends was simulated as a control. The electrostatic field (Figure S1, Supporting Information) and band gap (0.4 eV) were similar to the GS pilin. In contrast, a calculation of an infinitely periodic $\alpha$-helix of polyalanine gave a much larger gap (3.6 eV). The electric field is completely absent in this infinite polyalanine as the dipole is neutralized by the imposed periodicity of the infinite continuous protein. Physically, it is the same as the potential and electronic structure of a long polyalanine connected to shorted metal leads at both ends. In contrast, a finite polyproline peptide modeled in vacuo had a homogeneous distribution of charges (Figure S1, Supporting Information) and a larger band gap (3.9 eV). This is because the steric hindrance between the cyclic side chains in the proline peptide induces distortions in the direction of the helix axis and prevents the alignment of the peptide bond dipole moments (Figure S2, Supporting Information). The data therefore suggest that effects intrinsic to the $\alpha$-helical structure of the GS pilin such as one-dimensional polarization effects by the aligned dipoles associated to the peptide bonds in helices and related electrostatics (Figure S2, Supporting Information) contributed to the low band gap.

Despite similarities in the electrostatic potential and energy band gap noted for the GS pilin and the polyalanine peptide modeled in vacuo, the spatial distribution of the HOMO and the LUMO was different. The polyalanine peptide had the HOMO and LUMO orbitals located at the N-terminal and C-terminal ends, respectively, due to the ramp in potential induced by the helix field. A similar distribution of frontier orbitals was modeled in the polyproline peptide due to the contribution of the positively charged N-terminus and negatively charged C-terminus (Figure S3, Supporting Information). However, the frontier orbitals displayed weight at both ends in the GS pilin (Figure 4), consistent with the contribution of specific residues of the pilin to the HOMO and LUMO states. These residues included the N-terminal phenylalanine (F1) in HOMO and the aspartates from the C-terminal segment (D53, D54) in LUMO, which are all residues containing delocalized electrons (phenylalanine has a benzene ring and aspartate has a carboxylate group).

**D. Molecular Dynamics of the GS Pilin in Solution.** The dry, in vacuo structure of a protein can display very different properties from that displayed in solution at the relevant temperature. The macrodipole of the helical peptide, for example, is expected to be screened in solution, and specific contributions from amino acid residues and protonation states are expected to arise. This is due to the effect that water and ions have on the protein's electrostatics, as described in more detail below, as well as to thermal fluctuations of water and ions in addition to those of the protein itself. Calculations and simulations in solution are also more biologically relevant, as the GS pili function as electronic conduits in aqueous environments. Hence, the GS pilin was simulated in water using MD. Once equilibrated, the structure of the pilin remained essentially unchanged: an extended $\alpha$-helix and a C-terminal, random-coiled segment (Figure 5). The favored conformer in solution was that in which the C-terminal segment was at a 40° angle with the $\alpha$-helix (Figure 5A). However, the angle of the C-terminal segment fluctuated more in water (Figure 5A) than in vacuo (Figure 2A). This is not surprising because water can compensate for the hydrogen bonds lost when the helix-tail angle opens. A snapshot from

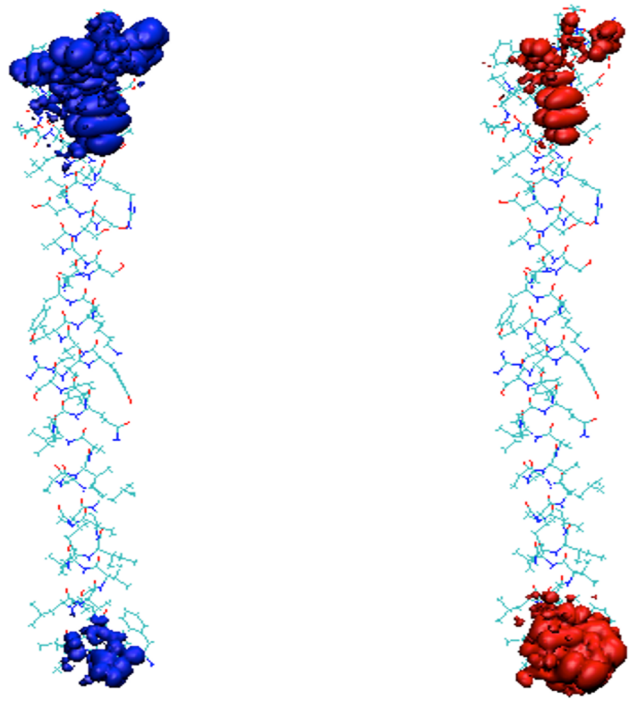


**Figure 4.** Density of states around the HOMO (left) and LUMO (right) of the lowest energy conformer of the GS pilin generated by QM optimization in vacuo.

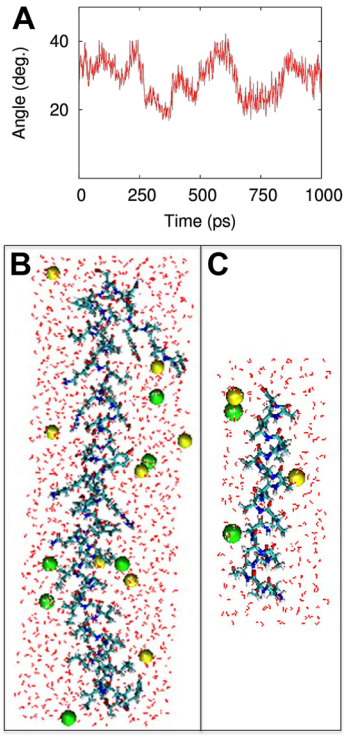



**Figure 5.** (A) Angle distribution of the C-terminal segment of the GS pilin in solution with respect to the $\alpha$-helix in 1 ns MD. (B, C) Classical MD snapshots of the GS pilin (B) and polyalanine (C) peptides in solution.

classical MD simulations of the GS pilin and a control 18 amino acid-long polyalanine peptide in solution is shown in Figure 5B,C, respectively. The MD simulations of the GS pilin in solution matched well with the homology model and the in


dx.doi.org/10.1021/jp302232p | J. Phys. Chem. A 2012, 116, 8023−8030

vacuo structural predictions, again revealing the short C-terminal segment at approximately a 40° angle with respect to the $\alpha$-helix (Figure 1). This conformation exposes the C-terminal segment on the pilus exterior during pilin assembly and contributes to the shape and charge of the pilus fiber and, ultimately, to its specific functions.

**E. Electronic Structure of the GS Pilin in Solution.** As with the in vacuo simulations, we used the statistically independent snapshots extracted for the GS pilin and polyalanine peptide (Figure 5B,C) for QM single-point calculations and to generate statistical averages of the electronic structure and electrostatics from first principles. The average electrostatic potential along the GS pilin modeled in solution revealed a biphasic charge distribution (Figure 6A), rather than

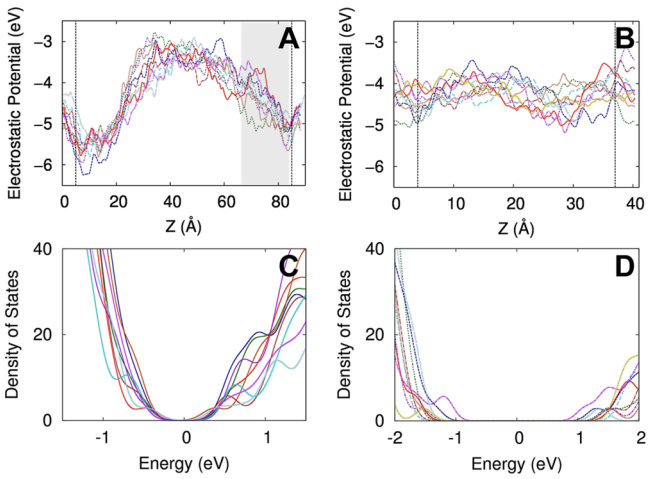



**Figure 6.** $z$-averaged electrostatic potential along the peptide length (vertical dashed line) from QM calculations (A and B) and density of states of the MD 50-ps snapshots (C and D) of the GS pilin (A and C) and polyalanine (B and D). The C-terminal segment of the GS pilin is highlighted in the shaded area in (A).

the steady increase in potentials from the N-terminal to the C-terminal end of the peptide observed under dry conditions (Figure 3). Although ions from the solution are close to the pilin's charges, the potential was more positive in the mid region and more negative in the C-terminal segment (Figure 6A). This profile contrasted with the electrostatic potential profile of the polyalanine helix in the same time scales (Figure 6B). Due to the screening effect by water, there were no steep increases or decreases in potentials in the polyalanine helix. Furthermore, due to the absence of polar or charged side chains in the polyalanine helix, the counterions are not bound to the protein and they can move more freely during the MD simulations, which lead to more variability. The differences in electrostatic potentials correlated well with differences in electronic structure as well. The electronic band gap in the GS pilin ranged from 0.9 to 1.1 eV (Figure 6C), which is higher than in the in vacuo calculations but still relatively low for a common peptide. In contrast, the electronic band gap of the polyalanine helix was much higher (between 2.7 and 3.2 eV) (Figure 6D). Hence, it was possible to differentiate between the energy band gaps of the GS pilin and the polyalanine control, despite the general tendency of DFT to underestimate the band gap. These differences in potential and band gap between the GS pilin and the polyalanine peptides, which share the helical structure but differ in amino acid composition, therefore reflect the contribution of specific residues in the GS pilin to its characteristic charge distribution.

The HOMO and LUMO states also matched the distribution of charges in the GS pilin modeled in solution and allowed to better appreciate the specific contribution of specific amino acids to the projected density of states over the HOMO−LUMO states (Figure 7). The LUMO was located in the

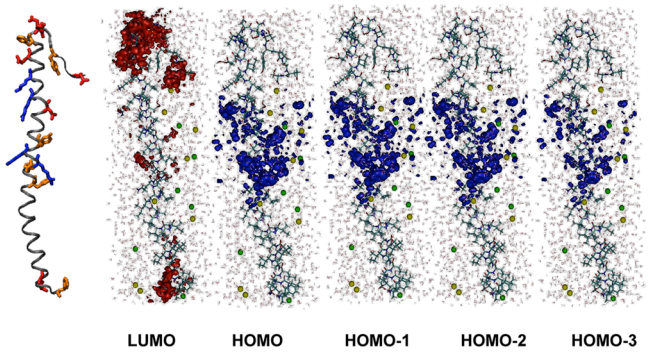



**Figure 7.** Density of states of the GS pilin in solution, from LUMO to HOMO−3. The distribution of positively (blue) or negatively (red) charged and aromatic (orange) amino acids and tyrosines is also shown using the homology model of the GS pilin (left).

regions of low electrostatic potential, at the N- and C-termini, whereas the HOMO was located in the mid region of high potential. This contrasts with the localization of the HOMO of the polyalanine to the N-terminus and the LUMO dispersion throughout the mid region and C-terminal end (Figure S4, Supporting Information). Thus, the unique HOMO−LUMO distribution in the GS pilin is consistent with the contribution of specific amino acids.

The mid region of the GS pilin where the HOMO localized corresponds to the region of highest electrostatic potential (Figure 6A), containing positively charged residues (R28, K30, R41) and aromatic residues (F24, Y27, and Y32) (Figure 7). The low potential regions at the N- and C-terminal ends (Figure 6A) where the LUMO localized are the regions where negatively charged residues (E5, E48, D53, D54, E60, and the C-terminal S61) and aromatic residues (Y27, Y32, Y57) localize. Thus, all the residues that contributed to the frontier states come from side chains where electrons are usually delocalized. The HOMO−LUMO distribution in the GS pilin is determined by the aqueous environment, which influences the protonation states of the residues and therefore, their charge. In fact, a calculation of the GS pilin in solution in which the residues were neutralized did not change the HOMO−LUMO distribution but increased the electronic gap to 2.7 eV, supporting our previous conclusion that the distribution of charged amino acids in the GS pilin contributed to the low band gap by establishing a new electrostatic potential profile. It is also important to note that most of the amino acids contributing to the density of states are located in regions of the GS pilin (the $\alpha$1-C domain and the C-terminal segment) that are most divergent at the amino acid sequence and structural level, respectively. This suggests that the amino acid sequence and structure of the GS pilin is adaptive in nature and evolved to promote electron transfer reactions.

Interestingly, the electronic structure of the GS pilin in solution displayed some weight, in both the HOMO and LUMO states, in the same spatial region in the middle of the $\alpha$-helix (Figure 7). This is the region where three of the five

aromatic residues of the GS pilin are located (Figure 7). Resonance in this region is predicted to be high, thus favoring the hopping of electrons between the aromatic residues under thermal fluctuations and facilitating intramolecular electron transfer. Electronic coupling in also expected once the GS pilins assemble as amino acids in the regions of high and low potential mediate electrostatic interactions between the $\alpha$-helices of neighboring pilins in the assembled pilus.[39] For example, the electrostatic interactions between the positively charged N-terminal phenylalanine F1 of the last assembled pilin and the negatively charged E5 residue of the incoming pilin, which is stored in the hydrophobic lipid bilayer, promote the assembly of new pilins into the growing fiber[24] and determine the 10.5 Å axial rise measured for neighboring pilins in the pilus fiber.[39] This rise also positions the HOMO of one subunit closer to the LUMO of the adjacent subunit, which stabilizes pilin−pilin interactions and promotes charge transfer between the contact residues. Furthermore, there is also overlap between the LUMOs of adjacent pilins in this contact region. The LUMO−LUMO contacts are located between the tyrosine rings of neighboring pilins, thus contributing to charge hopping from one pilin−pilin contact point to another in the assembled pilus fiber.

## CONCLUSIONS

In this study, the molecular and electronic structure of the pilin peptide that forms the conductive pili of *G. sulfurreducens* was analyzed under different conditions. The results are consistent with a pilin structure and amino acid sequence that has evolved to optimize intramolecular electron transfer while preserving the conserved features required for pilin assembly into the pilus fiber. The GS pilin's reduced size, the absence of a globular head with $\beta$-strands, and the predominantly helical peptide all are structural features that promote electron transfer in peptides.[25,40,41] Furthermore, specific amino acids in the GS pilin resulted in a distribution of charges and density of states along the peptide that could also favor charge transport. Aromatic residues, including tyrosines, for example, were located in regions of HOMO−LUMO and LUMO−LUMO overlap between adjacent assembled pilins. The oxidation potential for tyrosines is the lowest among the amino acid side chains, and this feature, along with aromatic stabilization, allows them to function as relay amino acids for long-range electron transfer.[43] The results also suggest that charge transfer is possible between assembled pilin subunits through the LUMO of one chain and the LUMO of the next one. The resonance generated by these contacts, where the frontier states of the chains overlap, is expected to optimize the hopping of electrons between the relay amino acids under thermal fluctuations. The rates and direction of electron transfer through peptides are also modulated by the electrostatic potential generated by charged amino acids located in close proximity to relay amino acids.[44] Our studies predict similar electrostatic effects in the GS pilin, which could modulate intramolecular charge transport as well as its directionality, an important consideration in a bacterial electronic conduit evolved to transfer electrons from the cell to the extracellular electron acceptor.

As the intersubunit distance is very short in the pilus fiber,[39] the assembly of the GS pilin is also likely to favor proper stacking of the pilin's delocalized residues such as aromatic and charged amino acids, which may promote intermolecular electron transfer via electron hopping and long-range delocalization through the pilin chains in the pilus fiber. Furthermore, posttranslational modifications have been reported for bacterial pilins[45−48] and could influence the molecular structure and charge distribution of the GS pili to further optimize the peptide environment for electron transfer. Yet charge transport through the GS pili ultimately depends on the ability of the GS pilin subunit to transfer electrons. The structural and electronic features of the GS pilin revealed in these studies support the notion of a peptide environment optimized for electron conduction at both the structural and amino acid sequence level. This information is critical to understand the interplay between the protein's structure and its environment and the protein's biological role, which can provide fundamental knowledge for the engineering of protein nanowires with custom properties.

## ASSOCIATED CONTENT

### Supporting Information

Additional figures (electrostatic potential, dipole and antiparallel effects, density of states models) and the atomic coordinates of the GS pilin homology model (pdb format) and MM and QM conformers (xyz format). This information is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*Tel.: (+1) 517 884-5401. Fax: (+1) 517 353-8957. E-mail: reguera@msu.edu.

### Present Addresses

[⊥]CIC Nanogune and DIPC, Tolosa Hiribidea 76, 20018 San Sebastián, Spain.

[#]Basque Foundation for Science Ikerbasque, 48011 Bilbao, Spain.

[¶]Cavendish Laboratory, University of Cambridge, J. J. Thomson Avenue, Cambridge CB3 0HE, U.K. Tel: (+34) 943 574039. E-mail: e.artacho@nanogune.eu.

### Author Contributions

[‖]These authors contributed equally to this work

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by grant MCB-1021948 from the National Science Foundation and a Strategic Partnership Grant from the Michigan State University Foundation to GR. Calculations were performed on the HPC and CamGRID facilities at Cambridge. GTF acknowledges the hospitality of the Department of Earth Sciences at the University of Cambridge and support from a fellowship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) of Brazil.

## REFERENCES

(1) Lovley, D. R.; Holmes, D. E.; Nevin, K. P. In *Advances in Microbial Physiology*; Poole, R. K., Ed.; Elsevier Academic Press: London, 2004; pp 219−286.

(2) Shi, L.; Squier, T. C.; Zachara, J. M.; Fredrickson, J. K. Respiration of metal (hydr)oxides by *Shewanella* and *Geobacter*: a key role for multihaem *c*-type cytochromes. *Mol. Microbiol.* **2007**, *65*, 12−20.

(3) Caccavo, F., Jr.; et al. *Geobacter sulfurreducens* sp. nov., a hydrogen- and acetate-oxidizing dissimilatory metal-reducing microorganism. *Appl. Environ. Microbiol.* **1994**, *60*, 3752−9.

(4) Coppi, M. V.; Leang, C.; Sandler, S. J.; Lovley, D. R. Development of a genetic system for *Geobacter sulfurreducens*. *Appl. Environ. Microbiol.* **2001**, *67*, 3180−7.
(5) Reguera, G.; et al. Extracellular electron transfer via microbial nanowires. *Nature* **2005**, *435*, 1098−101.
(6) Pelicic, V. Type IV pili: *e pluribus unum*? *Mol. Microbiol.* **2008**, *68*, 827−37.
(7) Reguera, G.; et al. Biofilm and nanowire production lead to increased current in microbial fuel cells. *Appl. Environ. Microbiol.* **2006**, *72*, 7345−7348.
(8) Cologgi, D. L.; Lampa-Pastirk, S.; Speers, A. M.; Kelly, S. D.; Reguera, G. Extracellular reduction of uranium via *Geobacter* conductive pili as a protective cellular mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 15248−52.
(9) Reguera, G.; Pollina, R. B.; Nicoll, J. S.; Lovley, D. R. Possible nonconductive role of *Geobacter sulfurreducens* pilus nanowires in biofilm formation. *J. Bacteriol.* **2007**, *189*, 2125−7.
(10) Veazey, J. P.; Reguera, G.; Tessmer, S. H. Electronic properties of conductive pili of the metal-reducing bacterium *Geobacter sulfurreducens* probed by scanning tunneling microscopy. *Phys. Rev. E* **2011**, *84*, 060901.
(11) Richter, L. V.; Sandler, S. J.; Weis, R. M. Two isoforms of *Geobacter sulfurreducens* PilA have distinct roles in pilus biogenesis, cytochrome localization, extracellular electron transfer, and biofilm formation. *J. Bacteriol.* **2012**, *194*, 2551−63.
(12) Craig, L.; Pique, M. E.; Tainer, J. A. Type IV pilus structure and bacterial pathogenicity. *Nat. Rev. Microbiol.* **2004**, *2*, 363−78.
(13) Brittain, T. Intra-molecular electron transfer in proteins. *Protein Pept. Lett.* **2008**, *15*, 556−61.
(14) Cordes, M.; et al. Influence of amino acid side chains on long-distance electron transfer in peptides: electron hopping via "stepping stones". *Angew. Chem., Int. Ed. Engl.* **2008**, *47*, 3461−3.
(15) Gray, H. B.; Winkler, J. R. Electron tunneling through proteins. *Q. Rev. Biophys.* **2003**, *36*, 341−72.
(16) Lukacs, A.; Eker, A. P.; Byrdin, M.; Brettel, K.; Vos, M. H. Electron hopping through the 15 Å triple tryptophan molecular wire in DNA photolyase occurs within 30 ps. *J. Am. Chem. Soc.* **2008**, *130*, 14394−5.
(17) Winkler, J. R. Electron tunneling pathways in proteins. *Curr. Opin. Chem. Biol.* **2000**, *4*, 192−8.
(18) Beratan, D. N.; Betts, J. N.; Onuchic, J. N. Protein electron transfer rates set by the bridging secondary and tertiary structure. *Science* **1991**, *252*, 1285−8.
(19) Beratan, D. N.; Onuchic, J. N.; Betts, J. N.; Bowler, B. E.; Gray, H. B. Electron-tunneling pathways in ruthenated proteins. *J. Am. Chem. Soc.* **1990**, *112*, 7915−7921.
(20) Gray, H. B.; Winkler, J. R. Long-range electron transfer. *Proc. Natl. Acad. Sci .U. S. A.* **2005**, *102*, 3534−9.
(21) Shih, C.; et al. Tryptophan-accelerated electron flow through proteins. *Science* **2008**, *320*, 1760−2.
(22) Kimura, S. Molecular dipole engineering: new aspects of molecular dipoles in molecular architecture and their functions. *Org. Biomol. Chem. 6* (**2008**).
(23) Sek, S.; Swiatek, K.; Misicka, A. Electrical behavior of molecular junctions incorporating alpha-helical peptide. *J. Phys. Chem. B* **2005**, *109*, 23121−4.
(24) Craig, L.; Li, J. Type IV pili: paradoxes in form and function. *Curr. Opin. Struct. Biol.* **2008**, *18*, 267−77.
(25) Shin, Y.-g. K.; Newton, M. D.; Isied, S. S. Distance dependence of electron transfer across peptides with different secondary structures: The role of peptide energetics and electronic Coupling. *J. Am. Chem. Soc.* **2003**, *125*, 3722−3732.
(26) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **1997**, *18*, 2714−23.
(27) Craig, L.; et al. Type IV pilin structure and assembly: X-ray and EM analyses of *Vibrio cholerae* toxin-coregulated pilus and *Pseudomonas aeruginosa* PAK pilin. *Mol. Cell* **2003**, *11*, 1139−50.
(28) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(29) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926.
(31) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(32) Troullier, N.; Martins, J. L. Efficient pseudopotentials for plane-wave calculations. II. Operators for fast iterative diagonalization. *Phys. Rev. B* **1991**, *43*, 8861−8869.
(33) Kleinman, L.; Bylander, D. M. Efficacious form for model pseudopotentials. *Phys. Rev. Lett.* **1982**, *48*, 1425−1428.
(34) Junquera, J.; Paz, Ó.; Sánchez-Portal, D.; Artacho, E. Numerical atomic orbitals for linear-scaling calculations. *Phys. Rev. B* **2001**, *64*, 235111.
(35) Soler, J. M.; et al. The SIESTA method for ab initio order-N materials simulation. *J. Phys.: Condens. Matter* **2002**, *14*, 2745.
(36) Audette, G. F.; Irvin, R. T.; Hazes, B. Crystallographic analysis of the *Pseudomonas aeruginosa* strain K122-4 monomeric pilin reveals a conserved receptor-binding architecture. *Biochemistry* **2004**, *43*, 11427−35.
(37) Hazes, B.; Sastry, P. A.; Hayakawa, K.; Read, R. J.; Irvin, R. T. Crystal structure of *Pseudomonas aeruginosa* PAK pilin suggests a main-chain-dominated mode of receptor binding. *J. Mol. Biol.* **2000**, *299*, 1005−17.
(38) Winther-Larsen, H. C.; et al. *Pseudomonas aeruginosa* Type IV pilus expression in *Neisseria gonorrhoeae*: effects of pilin subunit composition on function and organelle dynamics. *J. Bacteriol.* **2007**, *189*, 6676−85.
(39) Craig, L.; et al. Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Mol. Cell* **2006**, *23*, 651−62.
(40) Long, Y. T.; Abu-Irhayem, E.; Kraatz, H. B. Peptide electron transfer: more questions than answers. *Chem.—Eur. J.* **2005**, *11*, 5186−94.
(41) Galoppini, E.; Fox, M. A. Effect of the electric field generated by the helix dipole on photoinduced intramolecular electron transfer in dichromophoric α-helical peptides. *J. Am. Chem. Soc.* **1996**, *118*, 2299−2300.
(42) Karplus, M.; Kuriyan, J. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6679−6685.
(43) Giese, B.; et al. Electron relay race in peptides. *J. Org. Chem.* **2009**, *74*, 3621−5.
(44) Gao, J.; et al. Electron transfer in peptides: the influence of charged amino acids. *Angew. Chem., Int. Ed. Engl.* **2011**, *50*, 1926−30.
(45) Marceau, M.; Forest, K.; Beretti, J. L.; Tainer, J.; Nassif, X. Consequences of the loss of O-linked glycosylation of meningococcal type IV pilin on piliation and pilus-mediated adhesion. *Mol. Microbiol.* **1998**, *27*, 705−15.
(46) Power, P. M.; Jennings, M. P. The genetics of glycosylation in Gram-negative bacteria. *FEMS Microbiol. Lett.* **2003**, *218*, 211−22.
(47) Virji, M. Post-translational modifications of meningococcal pili. Identification of common substituents: glycans and alpha-glycero-phosphate–a review. *Gene* **1997**, *192*, 141−7.
(48) Warren, M. J.; Jennings, M. P. Identification and characterization of *pptA*: a gene involved in the phase-variable expression of phosphorylcholine on pili of *Neisseria meningitidis*. *Infect. Immunol.* **2003**, *71*, 6892−8.