Photoemission and Near-Edge X-Ray Absorption Fine Structure Studies of the Bacterial Surface Protein Layer of *Bacillus sphaericus* NCTC 9602

Denis V. Vyalikh,[†] Alexander Kirchner,[‡] Steffen Danzenbächer,[†] Yuriy S. Dedkov,[†] Andreas Kade,[†] Michael Mertig,*[,‡] and Serguei L. Molodtsov*[,†]

Institute of Solid State Physics, Dresden University of Technology, D-01062 Dresden, Germany, and BioNanotechnology and Structure Formation Group, Max Bergmann Center of Biomaterials, Dresden University of Technology, D-01062 Dresden, Germany

Received: February 11, 2005; In Final Form: July 14, 2005

The electronic structure of the regular, two-dimensional bacterial surface protein layer of *Bacillus sphaericus* NCTC 9602 has been examined by photoemission (PE) and near-edge X-ray absorption fine structure (NEXAFS) spectroscopy. Both the O 1s and the N 1s core-level PE spectra show a single structure, whereas the C 1s core-level spectrum appears manifold, suggesting similar chemical states for each oxygen atom and also for each nitrogen atom, while carbon atoms exhibit a range of chemical environments in the different functional groups of the amino acids. This result is supported by the element-specific NEXAFS spectra of the unoccupied valence electronic states, which exhibit a series of characteristic NEXAFS peaks that can be assigned to particular molecular orbitals of the amino acids by applying a phenomenological building-block model. The relative contributions of the C−O, C−N, and C−C bond originating signals into the C 1s PE spectrum are in good agreement with the number ratios of the corresponding bonds calculated from the known primary structure of the bacterial surface protein. First interpretation of the PE spectrum of the occupied valence states is achieved on the basis of electronic density-of-states calculations performed for small peptides. It was found that mainly the π clouds of the aromatic rings contribute to both the lowest unoccupied and the highest occupied molecular orbitals.

1. Introduction

The discovery that proteins and DNA can be assembled *in vitro* to form regular two-dimensional (2D) layers with tailored morphologies and spatially defined chemical and physical surface properties[1−7] is an important achievement of modern biomimetic materials chemistry. The capability to create defect-free 2D crystals by taking advantage of the unique recognition and self-assembly properties of biomolecules can widely be used for the development of bioanalytical sensors, molecular electronics, supramolecular engineering, and nanotechnology.[8] In particular, the crystalline assembly products can be exploited as templates to grow ordered, low-dimensional hybrid structures.[6−17] The morphology of these hybrid structures is highly controlled from the nanometer to the macroscopic level, resulting in complex architectures that provide multifunctional properties. It is readily expected that these novel materials can show new, still not discovered size-dependent quantum phenomena. In analogy to other organic, inorganic, and molecular systems,[18−23] advanced biomimetic hybrid structures might reveal, for example, low-dimensional magnetic or superconducting behavior.

One basic property of the used biomolecular templates, which is important to make further progress in this modern field of materials chemistry, is their electronic structure, because it defines, for example, the association of molecular species at the protein surface according to the local charge-transfer capacities of the template. The electronic transport properties of templates are also determined by their electronic structures. With a few exceptions, the electronic structure of protein and DNA templates is not investigated. Theoretical studies are difficult to perform, because of the large size of the unit cells.[24−26] The application of such spectroscopic techniques like photoemission[27] (PE) and near-edge X-ray absorption fine structure[28] (NEXAFS), which are usually used to measure the structure of the occupied and unoccupied electronic states of solids directly, is not straightforward in the case of biological structures, since those samples are highly sensitive to photon exposure. Thus, they become easily destroyed by intensive photon beams. Moreover, particularly surface-sensitive PE experiments require samples with atomically clean surfaces that might be difficult to ensure by *ex situ* sample preparation. The possibilities to clean the sample surfaces *in situ* are rather limited too, because delicate biological objects can immediately be destroyed by even subtle annealing or ion sputtering.[29] Because of all these reasons, only a few direct measurements of the electronic structures of biomolecular samples were carried out by PE[30−33] and NEXAFS[34−39] investigations so far. In addition, a few rather indirect measurements of the electronic structure by NMR, electron paramagnetic resonance, fluorescence-excitation spectroscopy, and electronic circular dichroism are reported.[40−43]

Here we report PE and NEXAFS measurements of the electronic structure of a regular bacterial surface layer (S layer). S layers are natural 2D crystalline protein layers found in the cell envelope of bacteria and archaea.[1,2,44] They usually form the outermost cell wall layer. S layers are composed of a single protein or glycoprotein species with molecular weights ranging

* Authors to whom correspondence should be addressed. E-mail: mertig@tmfs.mpgfk.tu-dresden.de; molodtso@physik.phy.tu-dresden.de.
[†] Institute of Solid State Physics, Dresden University of Technology.
[‡] Max Bergmann Center of Biomaterials, Dresden University of Technology.

10.1021/jp050748+ CCC: $30.25
Published on Web 09/14/2005

from 40 to 200 kDa. S layer proteins are arranged in *p*2, *p*3, *p*4, and *p*6 symmetries and exhibit lattice constants of 5−30 nm. The protein crystals usually show several pores and gaps per unit cell that have significance for macromolecules to penetrate.[45] S layers exhibit variable structures in detail[46−48] and some of them remarkable stability even under extreme conditions,[49] features that make them attractive templates for the *in vitro* formation of metal arrays in two dimensions.[9−17] Different methods of metal application have been used, for example, metal deposition *in vacuo* by metal decoration or shadowing,[50] deposition of preformed metal particles, for example, gold, in aqueous solution,[14] and more recently, growth of metal clusters generated by chemical redox mineralization processes *in vitro* starting from template-bound metal complexes.[10,13,15] Despite the large progress in the field of template-directed synthesis of advanced hybrid nanostructures, the mechanisms of chemical interactions between the adsorbates and S layer templates, which could, for example, be derived from the electronic-structure data of the constituents, were not known up to now.[17]

By applying both PE and total-electron-yield NEXAFS, we have characterized the core-level PE spectra and both the occupied and the unoccupied electronic states of the S layer of *Bacillus sphaericus* NCTC 9602. The S layer is a 2D protein crystal. Each unit cell of the S layer lattice is composed of four identical protein subunits. Each subunit consists of 1050 amino acids and has a molecular weight of 111.5 kDa.[51] The acquired spectroscopic data of this rather complex protein structure are discussed in comparison to measurements done for individual amino acids and in analogy to calculations of the electron density-of-states (DOSs) of small protein species. In part, this allows the assignment of individual features of the S layer spectra to molecular orbitals of particular functional groups of amino acids. The investigations also reveal information about the chemical states of O, N, and C atoms in the protein. Both oxygen and nitrogen show similar chemical states for each atom, respectively, while carbon atoms exhibit a range of chemical environments in different functional groups of the amino acids. The highest occupied molecular orbital (HOMO) of the studied S layer is located ∼3 eV below the Fermi energy (E_F).

2. Experimental Section

2.1. Sample Preparation. The S layer was isolated from the bacterium *B. sphaericus* NCTC 9602. The conditions for cell cultivation and purification of S layer sheets were described previously.[16] For the spectroscopic measurements, naturally oxidized Si wafers (SiO_x/Si(100), n-type, commercial grade, 25 × 6 mm^2 in size) were used as substrates. Prior to the protein deposition, the substrates were thoroughly cleaned and plasma-treated (radio frequency (RF) power 20 W, 5 min, pressure of ambient gases of 1×10^{-2} Torr). The regular S layer sheets were deposited *ex situ* by placing the protein suspension (1 mg/mL in 25 mM Tris buffer, 10 mM of $MgCl_2$, pH 7.4) onto the substrate for 30 min. Thereafter, the protein suspension was removed, and the sample was rinsed several times with deionized (DI) water to dispose of remaining salts, blown dry, and immediately injected into the spectrometer chamber, which is operated at ultrahigh vacuum (low 10^{-10}-Torr range). Resembling spectroscopic results were obtained for approximately 20 repeats of sample preparations, giving strong evidence of only negligible — if any — surface contaminations. The homogeneity and the degree of substrate coverage were determined after the spectroscopic measurements by scanning force microscopy (SFM) and scanning electron microscopy (SEM). In addition to that, the morphology of the used S layer protein was inspected by transmission electron microscopy (TEM) of negatively stained S layer sheets prepared on carbon-coated copper grids.

2.2. Electron Spectroscopy. The PE and NEXAFS experiments were performed at the Berliner Elektronenspeicherring für Synchrotronstrahlung (BESSY) using radiation from the Russian−German beamline. This ultrahigh-energy-resolution dipole beamline[52,53] was proven to be eligible to investigate "fragile" biological objects with spectroscopic techniques.[33] Particularly, high-order radiation is suppressed in the difficult region of the C 1s excitations.[52,53] In contrast to most currently employed undulator beamlines, which deliver an extremely high photon flux in the form of discrete spectra, it provides rather low-intensity radiation distributed continuously in a wide range of photon energies (30−1500 eV). Valence-band and core-level (O 1s, N 1s, and C 1s) PE spectra were acquired with a VG-CLAM4 electron-energy analyzer. The overall system energy resolution accounting for the thermal broadening was set to 150 and 200 meV (full width at half-maximum (FWHM)) for the valence-band and core-level PE experiments, respectively. The soft X-ray absorption spectra were recorded in a total-electron-yield mode and normalized to the incident photon flux. The resolution for the NEXAFS measurements is determined solely by the performance of the beamline and was varied from 80 to 100 meV FWHM when going from the C 1s (∼283 eV) to the O 1s threshold (∼530 eV), respectively. The acquired spectra were energy-calibrated using the C 1s → π^* and O 1s → π^* photoionization spectra of CO_2 gas.[54] In both (PE and NEXAFS) experiments, the photon incidence angle was selected to be 35° relative to the normal to the sample surface.

The occupied valence electronic states of the S layer were studied by angle-integrated photoemission with a photon energy of 40.8 eV (equal to the energy of the He IIα radiation). This photon energy was selected to make a compromise between the high cross-section of the photoexcitations of primarily 2p states of the second-period elements, forming the valence bands of proteins, and the relatively high surface sensitivity required to increase the contribution from the valence band of the S layer relative to the background signal from the SiO_x/Si substrate.

2.3. Analysis of Spectra. The C 1s PE spectra were least-squares-fitted to discriminate contributions originating from C atoms in different chemical environments. The fit was performed with Lorentzian line shapes convoluted by a Gaussian to account for the finite experimental resolution. The signal of the inelastically scattered electrons was simulated by an integral (Shirley) background.

2.4. Core-Level Binding Energies and Fermi Level Position. The individual binding energies (BEs) of oxygen, nitrogen, and carbon core levels were derived as kinetic-energy differences between the Fermi level and the O 1s, N 1s, and C 1s signals, respectively.

The position of the Fermi level in the PE spectrum was determined by measuring the Fermi energy offset for a reference metallic gold sample. In addition, the binding energy shifts of the S layer PE features induced by charging of the bare protein sample were accounted for by a separate PE experiment with the S layers covered with a thin gold film. The latter system was found to be metallic, and the gold did not noticeably react with the protein template, since the measured PE line shapes were similar to those of the pristine S layer sample.

3. Results and Discussion

3.1. Characterization of the S Layer Samples. *3.1.1. Protein Layer Morphology.* The morphology of negatively stained S

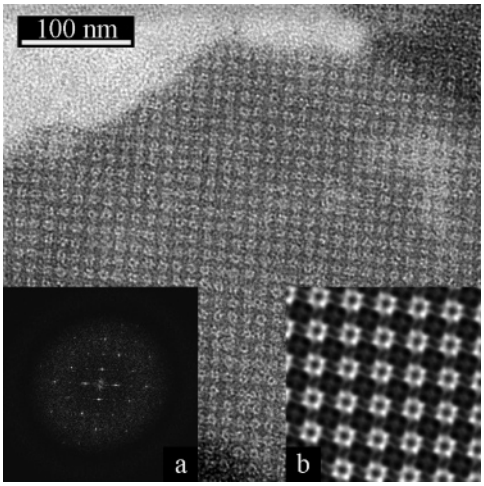

Figure 1. Transmission electron micrograph of a negatively stained S layer of *B. sphaericus* NCTC 9602. The stained S layer covers the whole substrate area imaged except the light area seen in the upper left of the image. (a) Power spectrum of the TEM image, demonstrating the *p*4 symmetry of the bacterial surface protein layer and the periodicity of the regularly arranged protein units. (b) Fourier-filtered image reconstruction of the S layer lattice. The protein appears light; the dark areas mark the heavily stain-filled, regularly arranged pores of the S layer sheet. The periodicity of the protein lattice is 12.5 nm.

layers of *B. sphaericus* NCTC 9602 was inspected by TEM (Figure 1). The S layer shows a *p*4 symmetry, a lattice constant of 12.5 nm, and a complex pattern of pores and gaps that are 2−3 nm wide. To determine the substrate coverage and the homogeneity of the S layer films deposited onto the Si substrates, the prepared samples were examined by SEM and SFM (not shown). The substrate surfaces are homogeneously covered with protein. Typically, a surface coverage with protein of 80−90% was obtained. SFM revealed a thickness of the dried, "as-deposited" S layer sheets of 5.0 nm. No change in sample thickness was detected before of after the spectroscopic measurements.

3.1.2. Core-Level PE Spectra. Overview photoemission spectra in a wide range of photoelectron kinetic energies were acquired to characterize the prepared samples. A typical spectrum taken with 600-eV photons is shown in the upper panel of Figure 2. It reveals mainly sharp features due to photoelectron emission from the 1s core levels of the second-period atoms oxygen, nitrogen, and carbon, which are — in addition to hydrogen — the main atomic constituents of the investigated S layer. Furthermore, we observe PE contributions from the 2p states of Si from the parts of the SiO_x/Si substrate surface, which is not covered with protein. According to the SEM investigations, the uncovered surface area is less than 20% of the total substrate surface.

The individual O 1s, N 1s, and C 1s spectra measured with a higher density of data points are depicted in the lower panels of the figure. The O 1s or N 1s spectra represent single structures pointing to rather similar chemical states of all oxygen and of all nitrogen atoms in the S layer. In contrast to the oxygen and nitrogen signals, the C 1s core-level spectrum reveals a set of structures, which can be attributed to different chemical environments of carbon atoms in different functional groups of the amino acids.[30,31] Figure 3 displays details of the C 1s spectrum taken at a lower photon energy (400 eV) with enhanced energy resolution. The complete analysis of the spectrum is not straightforward, because in general all carbon-containing functional chemical groups in nonequivalent chemical states may contribute specifically to the line shape of the PE

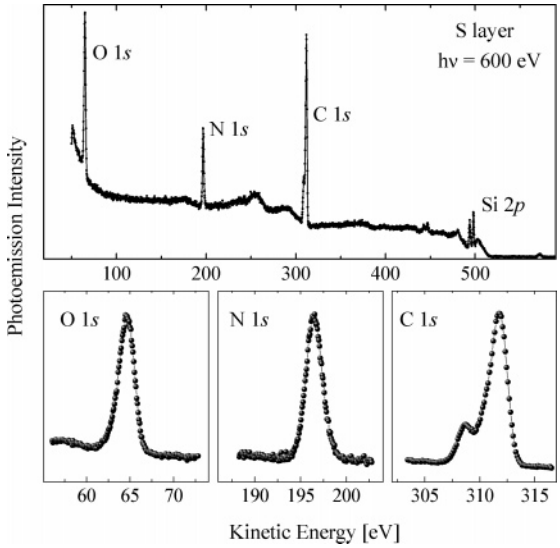

Figure 2. Overview (top panel) and O 1s, N 1s, and C 1s (bottom panels) core-level PE spectra of the S layer adsorbed on a SiO_x/Si(100) substrate ($h\nu$ = 600 eV). A work-function correction is not applied to the kinetic-energy scales.

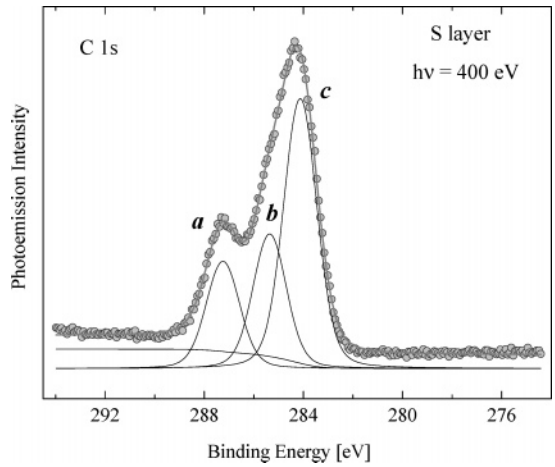

Figure 3. C 1s core-level PE spectrum of the S layer ($h\nu$ = 400 eV). The analysis of the spectrum by a least-squares fit reveals three components, which can be related to three groups of chemically nonequivalent carbon atom sites (C bound to O (***a***), N (***b***), and C (***c***)].

signal. Nevertheless, it is reasonable to assume that — depending on the particular atom, which is bound to carbon (C, N, or O) — basically three distinct chemical states exist for C atoms. In this case, the C 1s spectra can be fitted by three single components. The corresponding fit is depicted in Figure 3. Taking the electronegativity into account, which increases from carbon to oxygen, we expect a larger charge transfer from C to O than that from C to N atoms and thus larger BE shifts for the C−O than those for the C−N component in the C 1*s* spectrum. Therefore, we assign the high-BE component ***a*** to mainly carbon−oxygen (C−O) bonds. The low-BE component ***c*** can be attributed to carbon−carbon (C−C) bonds, whereas the middle component ***b*** can be associated with carbon−nitrogen (C−N) bonds. A similar interpretation was given for the C 1s PE spectrum of the S layer of *B. sphaericus* CCM 2177.[31] Generally, all three signal components contain contributions arising from both single and double bonds located in the peptide chain and in the amino acid residues. According to the known primary protein structure,[51] each S layer protein monomer consists of 1050 amino acid residues connected by 1049 peptide bonds. C−O, C−N, and C−C bonds exist in every peptide unit.

TABLE 1: Comparison between the Relative Intensities of the C−O, C−N, and C−C Contributions into the C 1s PE Spectrum and the Calculated Numbers of the Corresponding Protein Bonds

type of bond[a]	C−O	C−N	C−C
analysis of the C 1s core-level PE spectrum (Figure 3)	21%	27%	52%
calculated number of contributing bonds[b]	1594	2438	3902
calculated relative contribution	20%	31%	49%

[a] All three signal components contain contributions arising from both single and double bonds located in the peptide chain and in the amino acid residues. [b] Numbers calculated according to the known primary protein structure, taking into account that each protein monomer consists of 1050 amino acid residues connected by 1049 peptide bonds. These numbers are the total numbers of existing single and double bonds per monomer. They are derived by considering the number of amino acids, which contain the corresponding C−O, C−N, and C−C bonds, taking further into account how many C atoms of each type of environment exist in these amino acids.

The corresponding amount of bonds in the amino acid residues is less. C−O bonds are present in 452 of the 1050 residues. C−N bonds exist in 252 residues of the monomer. C−C bonds are available in CH chains, aromatic rings, and other environments without strong electronegative substituents. The latter bonds are found in all amino acid residues except glycine, resulting in 975 of the 1050 residues of the monomer.

The performed fit analysis of the line shape of the C 1s spectrum also provides quantitative information about the relative contributions of different carbon sites into the three groups of chemical states. Since the photoemission cross-sections of the C 1s core level do not depend on the chemical environment of the carbon atoms, the relative amount of carbon primarily bound to oxygen, nitrogen, and carbon can be derived from the ratios of the areas underneath the ***a***, ***b***, and ***c*** subspectra. The calculated values are 21%, 27%, and 52%, respectively. In addition, the upper limit for the contribution arising from residual C atoms at the noncovered substrate areas was determined by measuring spectra of bare, freshly plasma-treated Si wafers. In none of the cases did this limit exceed 5% of the intensity acquired for the S layer samples. In Table 1, the relative intensities of the C−O, C−N, and C−C contributions into the C 1s PE spectrum are compared to the calculated numbers of the corresponding protein bonds. Good agreement between the PE data and the number ratios derived from the primary structure of the complex surface layer protein is found.

3.2. NEXAFS Spectra. *3.2.1. General Considerations.* NEXAFS spectroscopy measures the photoabsorption cross-section for the excitation of primarily bound core electrons. In the case of noncorrelated systems containing solely light elements with occupied electron shells of s character, NEXAFS experiments provide according to the dipole selection rules information about p-like partial density of unoccupied electronic states. The corresponding electronic states are localized on the atoms, where the excited electrons were initially bound. The NEXAFS spectra are element-specific because each atom has characteristic binding energies of core levels. The unoccupied, element-specific p-character electronic states of the S layer were probed taken the 1s absorption spectra of oxygen, nitrogen, and carbon atoms.

The measured spectral features correspond to transitions from the ground state to a core excited state. In general this has to be treated as a multielectron process. However, it is common to describe X-ray absorption in closed-shell molecules within an orbital approximation as one-electron transitions into π^* and σ^* unoccupied molecular orbitals. These one-electron C 1s $\rightarrow$ π^* and C 1s $\rightarrow$ σ^* transitions are perturbed by the creation of a core hole.

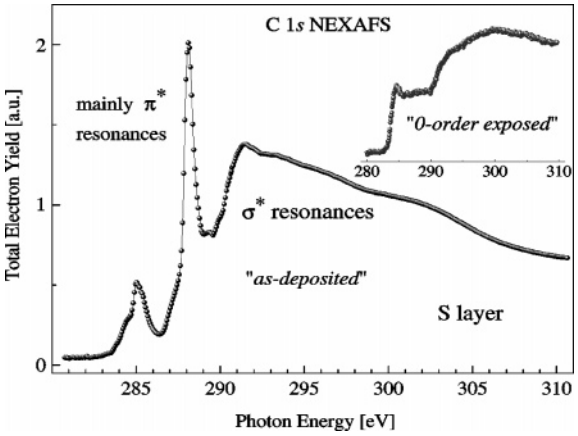

Figure 4. C 1*s* NEXAFS spectrum of the S layer of *B. sphaericus* NCTC 9602, consisting of two regions of photon energy—a low-energy region of π^* transitions and a high-energy region of mainly σ^* resonances. The inset shows the spectrum after the exposure of the S layer to "zero-order" synchrotron radiation.

For solid-state samples with molecular orbitals, which are oriented along particular directions, the relative NEXAFS intensities originating from different bonds can considerably vary upon the change of the projection of the light polarization onto the sample surface. This particular behavior can have a strong impact on the analysis of the contributions of different bonds into the chemical composition of investigated systems. For this reason, first we investigated the dependence of the NEXAFS signals at all three excitation edges (C 1s, N 1s, and O 1s) on the projection of the light polarization. The light provided by the dipole beamline is linearly polarized within the horizontal plane of the storage ring. We have varied the photon incidence angle (angle between the beam direction and the normal to the vertically oriented sample surface) between 0° and 70°. In all cases, the signal intensity was found to be independent of the projection of the light polarization. This can be explained by the fact that the crystallographic axes of the deposited S layer sheets are statistically oriented. Thus, we measure an average over many specific but different orientations of chemical bonds in the regular S layer sheets. In the present study, all NEXAFS and PE spectra were taken at the incidence angle of 35° relative to the normal to the sample surface.

3.2.2. C 1s NEXAFS Spectrum. The C 1s NEXAFS spectrum of the S layer (Figure 4) can be divided into two characteristic parts of photon energy. Rather sharp features are observed in the region between 283 and 289 eV, whereas much broader structures are found in the photon-energy region above 290 eV. The observed spectral features are an intrinsic property of the measured protein, because most of them disappear almost completely after exposure to intense radiation. The inset in Figure 4 shows a C 1s absorption spectrum recorded after the treatment of the S layer with non-monochromatized ("zero-order") synchrotron radiation. Due to the intense radiation, the protein sample decomposes. The huge changes in the NEXAFS signal are caused by extensive bond breaking, reformation, and mass loss.[55]

Previous studies of the C 1s NEXAFS in a series of individual amino acids,[34,35,56] which reveal spectra with similar fingerprints, attributed the sharp low-energy features mainly to 1s $\rightarrow$ π^* transitions into the unoccupied orbitals originating from C atoms in different chemical environments. The high-energy structures

TABLE 2: Energies of the Main NEXAFS Features in the Region of the π^* Resonances

C 1s	O 1s	N 1s
285.0 eV; 287.0 eV; 288.1 eV	531.8 eV	401.3 eV

were related to 1s $\rightarrow \sigma^*$ transitions. These structures are very broad because of the short lifetime of the corresponding exciting states. Contrary to the π^* transitions, there is a large scatter in energy positions and line shape of the σ^* structures. The reason is that the σ^* orbitals are oriented along the bond axes, and the corresponding NEXAFS resonances depend very much on hybridization and bond length. This behavior can be useful for studies of simpler systems than biological ones, where, for example, unknown bonds have to be characterized. However, for complex protein systems, the interpretation of the 1s $\rightarrow \sigma^*$ transitions is difficult so far. Here we concentrate, therefore, on the discussion of the 1s $\rightarrow \pi^*$ resonances.

As shown in refs 34, 35, and 56, a preliminary analysis of complex protein structures containing large amount of amino acids can be obtained within a building-block model, by which the monomer structure is seen as an assembly of smaller pieces. Those building blocks contribute independently to the total spectrum. In this way, specific spectral features can be correlated with contributions arising from individual functional groups of amino acids or subunits of those. For some small peptides, this model allows the treatment of the total NEXAFS spectrum as a superposition of signals solely originating from individual amino acids, suggesting that the peptide bonds, by which the amino acids are linked together, have only a weak effect on the spectra. When the amide bond is formed, the π^* C═O carboxyl-derived NEXAFS peak shifts in the amide carbonyl to a lower energy by only ~0.3 eV.[34,37−39] On one hand, one should be fully aware that this model based on an incoherent sum of contributions from separated building blocks is only a rough approach to understand the observed phenomena. In a more elaborated model, correlations between individual subunits and units have to be considered, which may then serve as basis for the application of a more bandlike description of the electronic structure of proteins. On the other hand, already now the simple building-block model allows the explanation of characteristic features of the π^* region of the C 1s NEXAFS spectrum.

On the basis of the measured spectra of a series of amino acids[34,35] and their interpretation using *ab initio* static-exchange method calculations,[56] we can assign the first feature in the S layer spectrum at ~285 eV (Table 2) predominantly to transitions of electrons localized on the ***c*** component of the C 1s core level (cf. Figure 3) into the lowest unoccupied molecular orbital (LUMO) of the C═C double bonds in aromatic rings of particular amino acids. Obviously, this peak exhibits a multicomponent structure with shoulders on its low- and high-energy sides, respectively. The intensity of the low-energy shoulder at ~284 eV varies slightly from one S layer preparation to another. Note that Si wafers coated with a carbidic layer, which is formed by the cracking of carbon-containing gases during the plasma treatment, may contribute to the 284-eV signal. Apart from that, the multicomponent line shape of the 285-eV feature can be understood by the fact that the energies of the molecular π^* C═C orbitals depend on the specific environment of the C atoms participating in the double bonds.[56] Chemical interaction with radicals leads usually to BE shifts toward higher energies of the C 1s core level in the ground state. Correspondingly, the resonances in NEXAFS studies are found at higher photon energies. The ground-state shifts vary between 0 and ~2 eV for the set of amino acids considered in refs 34, 35, and 56.

Another peak, found in the C 1s NEXAFS spectrum at 288.1 eV photon energy, can be attributed to the excitation of the electrons, related to the ***a*** component of the C 1s core level (cf. Figure 3), mainly in π^* C═O character orbitals, which exist in 228 of the 1050 amino acid residues of the monomer and as amide carbonyls of each peptide bond. Since the chemical environment of this bond does not considerably change across the protein, there is no initial-state splitting of this resonance. Therefore, this component represents less structure than the 285-eV one. From interpretations given in ref 35, the shoulder on the left side of the 288.1-eV feature located at ~287 eV may be assigned to the transitions into the σ^* C−H states. However, it relates in energy with the expected electron transition from the states of the ***b*** component of the C 1s core level into the π^* C═N orbitals,[34] which by the same reason as that discussed for the π^* C═O originating transition is not expected to reveal a pronounced manifold splitting. In the investigated protein, only the amino acid residues of arginine and histidine do contain C═N bonds, which will contribute into the C 1s $\rightarrow \pi^*$ signal around 287 eV.

An additional argument for the application of the building-block model is the observation that the C 1s NEXAFS spectrum taken for the S layer of *B. sphaericus* NCTC 9602 is very similar to those measured for human serum albumin and fibrinogen.[36] This finding also hints at the fact that in NEXAFS experiments the contributions of the building blocks are basically integrated over the whole protein structure, because under this assumption one would expect similar spectral signals for systems with statistically similar counts of amino acids of similar types.

3.2.3. Intensity Analysis of the C 1s NEXAFS. The intensity of the 288.1-eV peak is much higher than that of the 285-eV signal. Because the intensity ratios between the three main features of the π^*-transition region reflect basically the relative numbers of bonds contributing to a particular bond type (C═C-, C═O-, or C═N-derived), we can generally conclude that the relative number of amino acids with aromatic rings in the studied S layer is rather small. This is indeed in agreement with investigations of the primary protein structure. The content of amino acids with aromatic rings of the S layer of *B. sphaericus* NCTC 9602 is <8%.[51] However, a direct quantification of this effect solely from the spectroscopic data is not possible now. An accurate quantitative analysis would require detailed knowledge about cross-sections of all of the different transitions involved.[56] At present, data on oscillator-strength spectra is only partially available; for example, taking benzene and formamide as models, the C 1s *C═C(aromatic) has an intensity of 0.076 per C═C bond, while a C 1s *C═O(amide) has an intensity of 0.042.[57] Cross-section data of transitions involving a series of other environments are still missing.

The latter shows that NEXAFS data, although providing more specific information about molecular bonding, are more challenging to analyze than core-level photoemission spectra. In the C 1s PE spectrum, all bonds between carbon atoms contribute to basically one feature, ***c*** (cf. Figure 3), whereas their contributions are widely spread in energy between the π^* and σ^* regions in the C 1s NEXAFS data (Figure 4). The manifold line shape of the 285-eV feature in Figure 4 provides further insight into details of the electronic structure of the π^* states accounting for certain environments of the C atoms participating in the C═C double bonds.[56]

3.2.4. N 1s and O 1s NEXAFS. Both the N 1s (Figure 5) and the O 1s NEXAFS (Figure 6, top panel) spectra reveal basically one peak each (A_N at 401.3 eV and A_O at 531.8 eV) in the region of the π^* resonances. Together with the single line shape

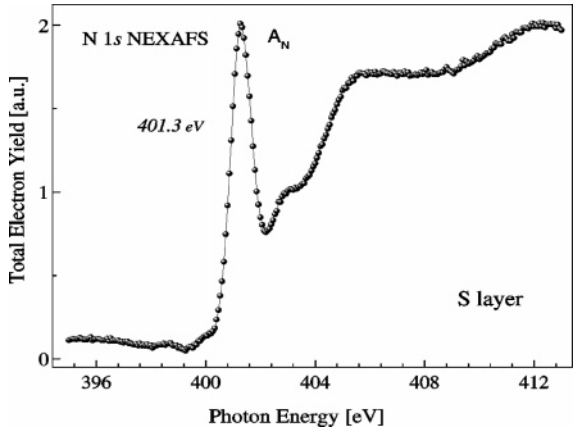

Figure 5. N 1s NEXAFS spectrum of the S layer.

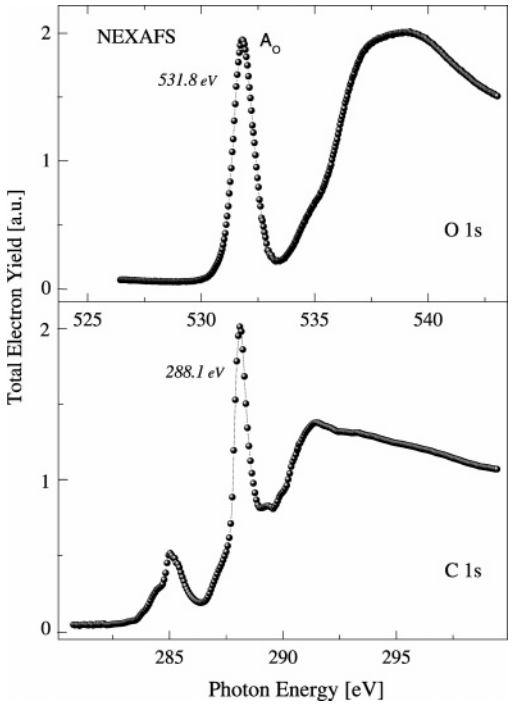

Figure 6. Oxygen (top panel) and carbon (bottom panel) 1s absorption spectra of the S layer. The procedure of energy-scale alignment is explained in the text. Note that the SiO_x substrate contributes to the O 1s signal particularly in the region of 535−542 eV photon energies.

of the N 1s and O 1s PE core levels (Figure 2), this again indicates the existence of rather similar chemical states of the oxygen and nitrogen atoms in the S layer protein, respectively. Contrary to the N 1s and O 1s signals, the C 1s NEXAFS spectrum, presented in Figure 4, is a sum of three subspectra, because at least three types of initial states, shifted in binding energy relative to each other, contribute to it (cf. Figure 3, contributions ***a***, ***b***, and ***c*** to the core-level C 1s PE spectrum). When the initial state is multifold, analyzing NEXAFS data in terms of exclusively unoccupied states is not straightforward. In this case, the NEXAFS signal has to be related to a weighted convolution of occupied and unoccupied states.

Partial contribution of element-specific unoccupied orbitals to the structure of the C 1s NEXAFS spectrum can be distinguished on the basis of pair wise alignment the C 1s signal either to the O 1s or to the N 1s spectrum. The latter two spectra reflect partial densities of the O- and N-derived unoccupied states, respectively, because in both cases only one type of the initial state is available (Figure 2, bottom panels). Usually, the

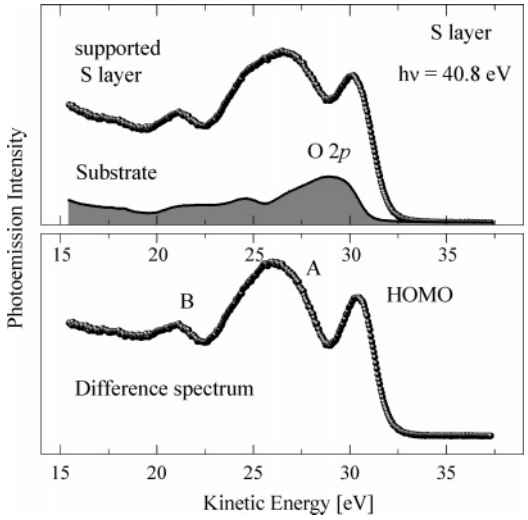

Figure 7. Valence-band PE spectra of the S layer of *B. sphaericus* NCTC 9602 adsorbed at the surface of the SiO_x/Si(100) substrate (top panel) and from a cleaned and plasma-treated bare substrate (shadowed). Both spectra are recorded with 40.8-eV photon energy and normalized to the photon flux. The difference spectrum, which represents the pure PE signal of the S layer, is displayed in the bottom.

alignment is feasible, if the electron−core hole interactions are similar for the NEXAFS transitions from the participating element-specific core levels into the same upper orbital such as for excitations from C 1s and O 1s states into π^* C═O character orbitals.[58] Currently, the situation for excitations from C 1s and N 1s states into unoccupied C−N orbitals is less clear. Therefore, we discuss solely the alignment of the C 1s and O 1s NEXAFS data in the following.

The alignment was done using the differences of the BEs of the corresponding core levels (section 3.1.2). In particular, to align the C 1s and O 1s NEXAFS spectra, the BE difference between the C−O-derived ***a*** component of the C 1s core level, which is supposed to reveal the highest cross-section for the transition into the C−O originating unoccupied orbitals, and the O 1s core level was used. After alignment, the A_O and 288.1-eV signals coincide in position (Figure 6), strongly supporting the conclusion about the π^* C═O origin of the 288.1-eV structure in the C 1s NEXAFS spectrum.

3.3. Valence-Band PE Spectra. A selected valence-band PE spectrum, which contains contributions from the S layer and the substrate, is shown in Figure 7 (top panel). To extract the virginal signal of the S layer, the SiO_x/Si substrate was measured immediately after plasma treatment with the same photon energy (shadowed area in Figure 7). The spectrum taken from the substrate reveals a triplet line shape, which is a combination of a typical SiO_x valence-band spectrum, characterized by a strong O-2p-derived feature, and less intense structures characteristic for admixtures of different carbides resulting from the plasma treatment of Si wafers.[59] The pure S layer PE signal is obtained by subtraction of the shadowed spectrum from the top one after both spectra were normalized to the photon flux. Additionally, the substrate spectrum was scaled accounting for both the calculated attenuation of photoelectrons in the S layer[60] and the degree of the S layer coverage determined from the SEM experiments.

Despite the complex chemical composition of the S layer, its valence-band PE spectrum shown in the bottom of Figure 7 reveals a series of well-resolved structures originating from different molecular orbitals. The relatively sharp maximum at the high-kinetic-energy offset of the spectrum represents the signal from the highest occupied molecular orbital of the S layer.

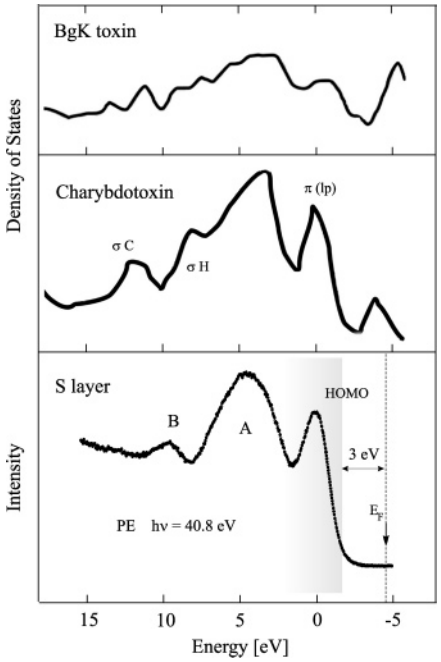

Figure 8. Occupied electronic structure of the S layer of *B. sphaericus* NCTC 9602 as derived from the PE experiment (bottom panel) compared with the calculated DOS for two rather small proteins: peptidic BgK toxin[26] (top panel) and charybdotoxin[25] (middle panel).

A second, intense manifold structure (A) is placed about 4.5 eV below the HOMO. In addition to these features, a less pronounced peak (B) is found at ~21 eV kinetic energy. Since valence-band PE is not element-specific, it is impossible to determine the origin of the occupied molecular orbitals solely from the PE data. Therefore, the measured valence-band PE data are compared to results of the electronic density-of-states for small peptides, calculated using local density approximation (LDA).[25,26] Thus, a first interpretation of the origin of the measured valence-band PE spectra is obtained.

To compare our experimental data with the DOS calculated for charybdotoxin and peptidic BgK toxin,[25,26] which both are small peptides that block potassium channels, the corresponding energy scales are adjusted relative to the positions of the HOMOs. The theoretical and experimental results presented in Figure 8 exhibit a similar pattern of features. As discussed for the NEXAFS spectra in section 3.2, the DOS data are expected to be qualitatively similar for proteins containing statistically similar counts of amino acids of the same type. On the basis of the interpretation given in ref 25, we can assign the occupied valence-band PE features as follows: The main contributions to the HOMO arise from π clouds of the aromatic rings and lone pairs (lp's) of oxygen, nitrogen, and sulfur. The A signal reflects σ bonds, in which one of the atoms is hydrogen, whereas feature B originates from σ bonds between two carbon atoms.

To determine the position of the Fermi level in the PE spectrum, the Fermi energy offset was measured for a reference metallic gold sample. To account additionally for the BE shifts of the PE features, which may be caused by possible charging of the bare protein sample, separate PE experiments were carried out with gold-covered S layers as described in section 2.4. Our analysis shows that the HOMO of the studied S layer is located ~3 eV below the Fermi energy. As a striking feature of these experiments, only a moderate dynamic charging of the S layer at a level similar to that of regular wide-gap semiconductors was observed. This may indicate relatively large charge-carrier mobility in the protein and possibly even suggests a bandlike description of the electronic structure of the S layer.

4. Conclusions

Photoemission and near-edge X-ray absorption fine structure spectroscopies were applied to characterize the electronic structure of the surface protein layer of the bacteria *Bacillus sphaericus* NCTC 9602. This regular protein layer is formed of a single protein monomer, which is composed of 1050 amino acids. Despite the complexity of the investigated protein, the measurements reveal in part quite detailed information about the chemical character of the spectra. In particular, information about the chemical environment of the second-period atoms oxygen, nitrogen, and carbon are derived. Similar chemical states are found for all oxygen atoms in the S layer as well as for all nitrogen atoms. This result is obtained from both O 1s and N 1s core-level PE and NEXAFS spectra. In all cases, sharp, singlet-type spectral features are found. In contrast to the oxygen and nitrogen signals, the C 1s core-level PE and NEXAFS spectra revealed a set of structures, which can be attributed to different chemical environments of carbon atoms in different functional groups of the amino acids. The measured C 1s core-level PE spectrum could be described by a simple superposition of all C−O, C−N, and C−C bond originating contributions derived from the known primary structure of the S layer protein. This result supports the hypothesis that the measured spectra can — in first approximation — be explained by applying a building-block model, where the individual amino acids contribute independently to the total spectrum.

By application of such a building-block model, the main features of the π^* region of the C 1s NEXAFS spectrum could be assigned to transitions into molecular orbitals, which are related to carbon double bonds with other C (LUMO) or O atoms. This result is supported by the comparison of the C 1s NEXAFS spectrum with the photoabsorption spectrum of the S layer, taken at the O 1s edge, demonstrating that corresponding NEXAFS peaks coincide in their energy positions when the NEXAFS spectra are aligned according to the measured binding energies of the respective core levels. In addition, the comparison of the intensities of the individual lines of the C 1s NEXAFS spectrum allows evaluation of the relative numbers of bonds contributing to a particular bond type (C═C or C═O character). Because the oxygen-related peak has a much higher intensity than that of the C═C-related one, we can conclude that the relative number of amino acids with aromatic rings in the studied S layer has to be rather small. This is indeed in agreement with investigations of the primary protein structure. These experiments clearly show the ability of NEXAFS spectroscopy to explore the chemical character of particular functional groups or even individual bonds of amino acids, thus demonstrating the great potential of NEXAFS techniques for the study of even complex protein structures.

For the interpretation of the valence-band PE results, electronic density-of-states calculations performed for small peptides were used, because the calculated DOS data exhibit similar patterns as the measured PE spectra. Thus, the main feature of these spectra could also be assigned to transitions into particular molecular orbitals. The HOMO is mainly formed by the π orbitals of the aromatic rings and located ~3 eV below the Fermi level. In the future, the obtained information about the electronic structure of the S layer can be used to make reliable predictions for possible chemical reactions and other chemical or physical phenomena in S-layer-templated hybrid structures.

Acknowledgment. The study was supported by the DFG (Grant Nos. PO392/18 and SFB463), the BMBF (Grant No. 13N8145), the SMWK, the EU (Grant No. FP5, Contract No.

G5RD-CT-2002-0750), and the bilateral "Russian−German Laboratory at BESSY" program. We acknowledge stimulating discussions with Wolfgang Pompe and Manuel Richter.

References and Notes

(1) Sleytr, U. B.; Messner, P. *Annu. Rev. Microbiol.* **1983**, *37*, 311−339.
(2) Baumeister, W.; Wildhaber, I.; Engelhardt, H. *Biophys. Chem.* **1988**, *29*, 39−49.
(3) Györvary, E.; O'Riordan, A.; Quinn, A.; Redmond, G.; Pum, D.; Sleytr, U. B. *Nano Lett.* **2003**, *3*, 315−319.
(4) Ringler, P.; Schulz, G. E. *Science* **2003**, *302*, 106−109.
(5) Seeman, N. C. *Curr. Opin. Struct. Biol.* **1996**, *6*, 519−526.
(6) Yan, H.; Park, S. H.; Finkelstein, G.; Reif, J. H.; LaBean, T. H. *Science* **2003**, *301*, 1882−1884.
(7) Le, J. D.; Pinto, Y.; Seeman, N. C.; Musier-Forsyth, K.; Taton, T. A.; Kiehl, R. A. *Nano Lett.* **2004**, *4*, 2343−2347.
(8) Pum, D.; Sleyter, U. *Trends Biotechnol.* **1999**, *17*, 8−12.
(9) Douglas, K.; Clark, N. O.; Rothschild, K. J. *Appl. Phys. Lett.* **1990**, *56*, 692−694.
(10) Shenton, W.; Pum, D.; Sleytr, U. B.; Mann, S. *Nature* **1997**, *389*, 585−587.
(11) Moore, J. T.; Beale, P. D.; Winningram, T. A.; Douglas, K. *Appl. Phys. Lett.* **1998**, *72*, 1840−1842.
(12) Dieluweit, S.; Pum, D.; Sleytr, U. B. *Supramol. Sci.* **1998**, *5*, 15−19.
(13) Mertig, M.; Kirsch, R.; Pompe, W.; Engelhardt, E. *Eur. Phys. J. D* **1999**, *9*, 45−48.
(14) Hall, S. R.; Shenton, W.; Engelhardt, H.; Mann, S. *ChemPhysChem* **2001**, *3*, 184−186.
(15) Mertig, M.; Wahl, R.; Lehmann, M.; Simon, P.; Pompe, W. *Eur. Phys. J. D* **2001**, *16*, 317−320.
(16) Wahl, R.; Mertig, M.; Raff, J.; Selenska-Pobell, S.; Pompe, W. *Adv. Mater.* **2001**, *13*, 736−740.
(17) Wahl, R.; Engelhardt, H.; Pompe, W.; Mertig, M. *Chem. Mater.* **2005**, *17*, 1887−1894.
(18) *Organic Superconductivity*; Kresin, V. Z., Little, W. A., Eds.; Plenum Publishers: New York, 1991.
(19) Hebard, A. F.; Rosseinsky, M. J.; Haddon, R. C.; Murphy, D. W.; Glarum, S. H.; Palstra, T. T. M.; Ramirez, A. P.; Kortan, A. R. *Nature* **1991**, *350*, 600−601.
(20) Kelty, S. P.; Chen, C. C.; Lieber, C. M. *Nature* **1991**, *352*, 223−225.
(21) Uji, S.; Shinagawa, H.; Terashima, T.; Yakabe, T.; Terai, Y.; Tokumoto, M.; Kobayashi, A.; Tanaka, H.; Kobayashi, H. *Nature* **2001**, *410*, 908−910.
(22) Huger, E.; Osuch, K. *Europhys. Lett.* **2003**, *63*, 90−96.
(23) Schindler, A.; König, R.; Herrmannsdorfer, T.; Braun, H.; Eska, G.; Gunther, D.; Meissner, M.; Mertig, M.; Wahl, R.; Pompe, W. *Physica B* **2003**, *329*, 1427−1428.
(24) Carloni, P.; Andreoni, W.; Parrinello, M. *Phys. Rev. Lett.* **1997**, *79*, 761−764.
(25) Ireta, J.; Galván, M.; Cho, K.; Joannopoulos, J. D. *J. Am. Chem. Soc.* **1998**, *120*, 9771−9778.
(26) Aparicio, F.; Ireta, J.; Rojo, A.; Escobar, L.; Cedillo, A.; Galván, M. *J. Phys. Chem. B* **2003**, *107*, 1692−1697.
(27) Hüfner, S. *Photoemission Spectroscopy*; Springer-Verlag: Berlin, 1995.
(28) Stöhr, J. *NEXAFS Spectroscopy*; Springer-Verlag: Berlin, 1992.
(29) Panhorst, M.; Brückl, H.; Kiefer, B.; Reiss, G.; Santarius, U.; Guckenberger, R. *J. Vac. Sci. Technol., B* **2001**, *19*, 722−724.
(30) Subirade, M.; Lebugle, A. *Thin Solid Films* **1994**, *243*, 442−445.
(31) Handrea, M.; Sahre, M.; Neubauer, A.; Sleytr, U. B.; Kautek, W. *Bioelectrochemistry* **2003**, *61*, 1−8.
(32) Kato, H. S.; Furukawa, M.; Kawai, M.; Taniguchi, M.; Kawai, T.; Hatsui, T.; Kosugi, N. *Phys. Rev. Lett.* **2004**, *93*, 086403.
(33) Vyalikh, D. V.; Danzenbächer, S.; Mertig, M.; Kirchner, A.; Pompe, W.; Dedkov, Yu. S.; Molodtsov, S. L. *Phys. Rev. Lett.* **2004**, *93*, 238103.
(34) Boese, J.; Osanna, A.; Jacobsen, C.; Kirz, J. *J. Electron Spectrosc. Relat. Phenom.* **1997**, *85*, 9−15.
(35) Kaznacheyev, K.; Osanna, A.; Jacobsen, C.; Plashkevych, O.; Vahtras, O.; Argen, H.; Carravetta, V.; Hitchcock, A. P. *J. Phys. Chem. A* **2002**, *106*, 3153−3168.
(36) Hitchcock, A. P.; Morin, C.; Heng, Y. M.; Cornelius, R. M.; Brash, J. L. *J. Biomater. Sci., Polym. Ed.* **2002**, *13*, 919−937.
(37) Gordon, M. L.; Cooper, G.; Morin, C.; Araki, T.; Turci, C. C.; Kaznatcheev, K.; Hitchcock, A. P. *J. Phys. Chem. A* **2003**, *107*, 6144−6159.
(38) Morin, C.; Hitchcock, A. P.; Cornelius, R. M.; Brash, J. L.; Urquhart, S. G.; Scholl, A.; Doran, A. *J. Electron Spectrosc. Relat. Phenom.* **2004**, *137*−*140*, 785−794.
(39) Cooper, G.; Gordon, M.; Tulumello, D.; Turci, C.; Kaznatcheev, K.; Hitchcock, A. P. *J. Electron Spectrosc. Relat. Phenom.* **2004**, *137*−*140*, 795−799.
(40) Alia; Matysik, J.; Soede-Huijbregts, C.; Baldus, M.; Raap, J.; Lugtenburg, J.; Gast, P.; van Gorkom, H. J.; Hoff, A. J.; de Groot, H. J. M. *J. Am. Chem. Soc.* **2001**, *123*, 4803−4809.
(41) Alia; Hulsebosch, B.; van Gorkom, H. J.; Raap, J.; Lugtenburg, J.; Matysik, J.; de Groot, H. J. M.; Gast, P. *Chem. Phys.* **2003**, *294*, 459−469.
(42) van Oijen, A. M.; Ketelaars, M.; Köhler, J.; Aartsma, T. J.; Schmidt, J. *Science* **1999**, *285*, 400−402.
(43) Hirst, J. D.; Colella, K.; Gilbert, A. T. B. *Phys. Chem. B* **2003**, *107*, 11813−11819.
(44) Beveridge, T. J.; Graham, L. L. *Microbiol. Rev.* **1991**, *55*, 684−705.
(45) Sára, M.; Sleytr, U. B. *J. Bacteriol.* **2000**, *182*, 859−868.
(46) Sleytr, U. B.; Messner, P. *J. Bacteriol.* **1988**, *170*, 2891−2897.
(47) Baumeister, W.; Wildhaber, I.; Phipps, B. M. *Can. J. Microbiol.* **1989**, *35*, 215−227.
(48) Rachel, R.; Pum, D.; Smarda, J.; Smajs, D.; Komrska, J.; Krzyzánek, V.; Rieger, G.; Stetter, K. O. *FEMS Microbiol. Rev.* **1997**, *20*, 13−23.
(49) Engelhardt, H.; Peters, J. *J. Struct. Biol.* **1998**, *124*, 276−302.
(50) Pompe, W.; Mertig, M.; Kirsch, R.; Engelhardt, H.; Kronbach, T. Functionalized biomolecular membranes for microreactors. In *Microreaction Technology*−*Proceedings of the First International Conference on Microreaction Technology*; Ehrfeld, W., Ed.; Springer-Verlag: New York, 1998; p 104.
(51) The structure of the S layer of *B. sphaericus* NCTC 9602 is available from the Nucleotide Sequence Database of the European Molecular Biology Laboratory (http://www.ebi.ac.uk).
(52) Gorovikov, S. A.; Molodtsov, S. L.; Follath, R. *Nucl. Instrum. Methods Phys. Res., Sect. A* **1998**, *441*, 506−512.
(53) Fedoseenko, S. I.; Vyalikh, D. V.; Iossifov, I. F.; Follath, R.; Gorovikov, S. A.; Püttner, R.; Schmidt, J.-S.; Molodtsov, S. L.; Adamchuk, V. K.; Gudat, W.; Kaindl, G. *Nucl. Instrum. Methods Phys. Res., Sect. A* **2003**, *505*, 718−728.
(54) Sham, T. K.; Yang, B. X.; Kirz, J.; Tse, J. S. *Phys. Rev. A* **1989**, *40*, 652−669.
(55) Zubavichus, Y.; Zharnikov, M.; Shaporenko, A.; Fuchs, O.; Weinhardt, L.; Heske, C.; Umbach, E.; Denlinger, J. D.; Grunze, M. *J. Phys. Chem. A* **2004**, *108*, 4557−4565.
(56) Carravetta, V.; Plashkevych, A.; Ågren, H. *J. Chem. Phys.* **1998**, *109*, 1456−1464.
(57) The data on oscillator-strength spectra are available from the database at http://unicorn.mcmaster.ca/corex.html.
(58) Ishii, I.; Hitchcock, A. P. *J. Chem. Phys.* **1987**, *87*, 830−839.
(59) Fang, R.-C.; Ley, L. *Phys. Rev. B* **1989**, *40*, 3818−3829.
(60) Seah, M. P.; Dench, W. A. *Surf. Interface Anal.* **1979**, *1*, 2−11.