Spectroscopic studies of the electronic properties of regularly arrayed two-dimensional protein layers

2006 J. Phys.: Condens. Matter 18 S131

(http://iopscience.iop.org/0953-8984/18/13/S09)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **18** (2006) S131–S144 doi:10.1088/0953-8984/18/13/S09

Spectroscopic studies of the electronic properties of regularly arrayed two-dimensional protein layers

D V Vyalikh[1], A Kirchner[2], A Kade[1], S Danzenbächer[1], Yu S Dedkov[1], M Mertig[2] and S L Molodtsov[1]

[1] Institute of Solid State Physics, Dresden University of Technology, D-01062 Dresden, Germany
[2] BioNanotechnology and Structure Formation Group, Max Bergmann Centre of Biomaterials, Dresden University of Technology, D-01062 Dresden, Germany

Received 11 November 2005
Published 13 March 2006
Online at stacks.iop.org/JPhysCM/18/S131

Abstract

Photoemission (PE) and near-edge x-ray absorption fine structure (NEXAFS) spectroscopy were applied to characterize electronic properties of the regular two-dimensional bacterial surface protein layer (S layer) of *Bacillus sphaericus* NCTC 9602, which is widely used as a protein template for the bottom-up fabrication of advanced metallic and hybrid nanostructures. PE and NEXAFS at the C 1s, O 1s, and N 1s core levels show similar chemical states for each oxygen atom and also for each nitrogen atom, while carbon atoms exhibit a range of chemical environments in different functional groups of the amino acids. A series of characteristic NEXAFS peaks were assigned to particular molecular orbitals of the amino acids by applying a phenomenological building-block model. It was found that the π clouds of aromatic rings make the main contribution to both the lowest unoccupied and highest occupied molecular orbitals. The two-dimensional protein crystal shows a semiconductor-like behaviour with a gap value of $\sim$3.0 eV and the Fermi energy close to the bottom of the LUMO.

1. Introduction

Among the biomolecular structures, regular bacterial surface layers (S layers) can be considered as almost ideal templates for the organization of inorganic material at the nanometre scale [1–13]. S layers are regular 2D protein crystals, which form the outermost cell envelope component of many prokaryotes in almost all phylogenetic branches of bacteria and archaea [1, 2, 14, 15] and exhibit different kinds of lattice symmetry (p1, p2, p3, p4, and p6) with centre-to-centre spacings between neighbouring morphological units in the range of 3–30 nm. They are typically 5–15 nm thick and possess pores of individual but identical size and morphology with diameters in the range between 2 and 6 nm. Most of the currently known S layers are composed of identical proteins or glycoproteins of molecular mass between 30 and 220 kDa. The S layers exhibit structures variable in detail [16–18] and some of them

remarkable stability even under extreme conditions [19], features that make them attractive templates for the *in vitro* formation of metal arrays in two dimensions [4–12].

However, both the understanding and the tailoring of the physical and chemical properties of biotemplated hybrid systems not only require knowledge about the electronic properties of the metallic clusters but also require thorough knowledge of the electronic structure of the protein template. With a few exceptions, these data are lacking. Theoretical investigations are difficult to perform because of the large size of the unit cells [20–22]. Spectroscopic methods like photoemission (PE) [23] and near-edge x-ray absorption fine structure (NEXAFS) [24] spectroscopy provide the most direct experimental information about the structure of the occupied and unoccupied electronic states, respectively. However, so far PE [25–29] and NEXAFS [28–35] have been rarely applied to biological substances, because intense photon irradiation easily destroys the specimen. Another methodical obstacle is that PE experiments in particular require samples with atomically clean surfaces, which might be difficult to ensure by *ex situ* sample preparation, as is necessary for proteins. The possibilities for cleaning the sample surfaces *in situ* are rather limited too, since delicate biological objects can be immediately disintegrated by even subtle annealing or ion sputtering [36].

Here we review recent core-level and valence-band PE as well as NEXAFS studies [28, 29] that were successfully performed to investigate electronic structure of the regular bacterial surface layer of *Bacillus sphaericus* NCTC 9602. Each unit cell of this 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.[3] Analysing light polarization dependent NEXAFS spectra taken at the C 1s, O 1s, and N 1s thresholds of excitation we have found that the crystallographic axes of the deposited S-layer sheets are statistically oriented.

2. Experimental details

The experiments were performed at the Berliner Elektronenspeicherring für Synchrotronstrahlung (BESSY) using radiation from the Russian–German ultrahigh energy resolution dipole beamline. This dipole beamline [37, 38] was proven to be suitable for investigating 'fragile' biological objects with spectroscopic techniques [28, 29]. In particular, high order radiation is suppressed in the difficult region of the C 1s excitations. In contrast to the undulator beamlines mostly employed nowadays, which deliver an extremely high photon flux in the form of a discrete spectrum, this beamline provides moderate intensity radiation continuously distributed over a wide range of photon energies from 30 to 1500 eV. The light spot size at the sample position is $20 \times 80\ \mu m^2$ providing the possibility of studying homogeneous parts of the S layers.

The NEXAFS spectra were acquired with a channeltron in a total electron yield mode. The resolution for the NEXAFS measurements that is determined by the performance of the beamline was varied from 80 to 100 meV (full width at half-maximum, FWHM) going from the C 1s via the N 1s to the O 1s threshold. The spectra acquired were energy calibrated using reference photoionization spectra of CO_2 gas [39] and normalized to the incident photon flux measured with a clean Si wafer.

An experimental station equipped with a VG-CLAM4 electron energy analyser was exploited to measure core-level (O 1s, N 1s, and C 1s) and valence-band PE spectra. The overall system energy resolution accounting for the thermal broadening was set to 150 and 200 meV FWHM for the valence-band and core-level PE experiments, respectively. The occupied

[3] 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).

valence electronic states of the S layer were studied using angle-integrated photoemission with a photon energy of 40.8 eV (He IIα) that was selected as a compromise between the high photoexcitation cross-sections of the 2p states of the second-period elements forming the valence bands of proteins and the high surface sensitivity required to increase the contribution from the valence band of the S layer relative to the background signal from the substrate. It is important to note that the PE experiments revealed evidence for partially extended electronic states in the 2D protein crystals. Namely, if charge carriers were completely localized to individual bonds within the amino acids, it would be impossible to measure PE spectra due to the tremendous positive charging of the S layer in the process of photoelectron emission. Instead, only a moderate S-layer dynamic charging at a level similar to that of regular wide-gap semiconductors was observed with PE. This fact reveals relatively large charge carrier mobility suggesting a band-like description of the electronic structure of the S layer.

3. Data evaluation

Our Win-Fit program [40] was used to evaluate the core-level PE spectra. The spectra were least-squares fitted in order to discriminate contributions originating from atoms in different chemical environments. The fit was performed with Lorentzian lineshapes convoluted with a Gaussian to account for the finite experimental resolution. The contribution from the inelastically scattered electrons was accounted for by integral background. The Fermi energy offset of the PE spectra was used each time in order to obtain the binding energies (BEs) of the oxygen, nitrogen, and carbon core levels. The absolute position of the Fermi level was calibrated in the experiments with a reference metallic gold sample. The photon-induced charging of the wide-gap semiconductor-like S layer, which results in binding energy shifts of the PE structures, was estimated in an experiment with the surface protein layers covered with a thin gold film. This reference system was proven to be metallic. Note that the gold deposition does not destroy the S-layer since the measured PE lineshapes were similar to those of the pristine S-layer sample.

4. Sample preparation and characterization

The S layer was isolated from the bacterium *B. sphaericus* NCTC 9602 as was described previously [11] and deposited onto naturally oxidized Si wafers (SiO_x/Si(100), n-type, commercial grade, 25×6 mm^2 in size), which were thoroughly cleaned and plasma treated (RF power 20 W, 5 min, pressure of ambient gases of 1×10^{-2} Torr) prior to the protein deposition. The protein suspension (1 mg ml^{-1} in 25 mM Tris buffer, 10 mM of $MgCl_2$, pH 7.4) was placed *ex situ* onto the substrate for about 30 min. Thereafter the rest of the protein suspension was removed, the sample was rinsed with deionized water and inserted into the experimental station operating at ultrahigh vacuum conditions. Similar spectroscopic data were obtained for $\sim$20 different sample preparations.

After the spectroscopic measurements, the morphology of the deposited S layers was characterized by transmission electron microscopy (TEM) of negatively stained S-layer sheets prepared on carbon coated copper grids (figure 1). As seen in the figure the S layer shows a p4 symmetry. Both insets have a size of 38×38 nm^2 and depict image-reconstructed parts of the S-layer sheet with different orientations to the surface demonstrating the morphological anisotropy of the 'inner' and 'outer' surface of the S layer. As follows from the Fourier-filtered image reconstruction, the S layer reveals a lattice constant of 12.5 nm, and a complex pattern of pores and gaps that are 2–3 nm wide.

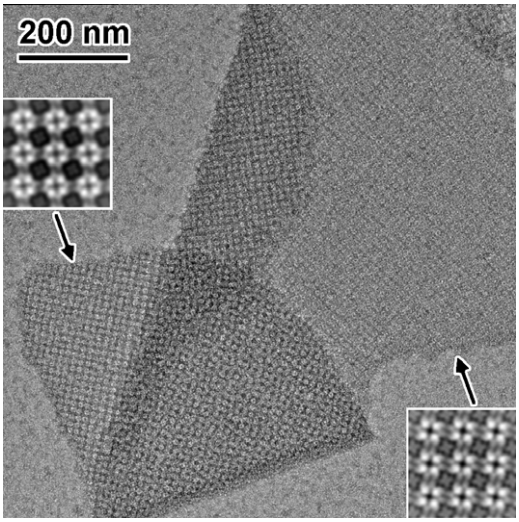

Figure 1. Transmission electron micrograph of a folded, negatively stained S-layer sheet of *B. sphaericus* NCTC 9602. The insets show image-reconstructed parts of the S layer (see the text).

To examine the degree of substrate coverage and the homogeneity of the S-layer films on the SiO_x/Si(100) substrates, the samples grown were examined by scanning electron microscopy (SEM) and scanning force microscopy (SFM). It was found that the wafers are homogeneously covered with protein. Typically, surface coverages with protein between 80 and 90% were obtained. SFM revealed a thickness of the dried, 'as-deposited' S-layer sheets of 5.0 nm. No thickness changes were observed after the spectroscopic measurements.

5. Results and discussion

5.1. Core-level photoemission

To characterize the quality of the samples, first, overview photoemission spectra were taken for a wide range of photoelectron kinetic energies covering both the regions of core-level and valence-band excitations. Figure 2 shows a selected spectrum taken at photon energy of 600 eV. The spectrum contains mainly three sharp features, which can be assigned to O 1s, N 1s, and C 1s core-level signals. In addition to hydrogen, these second-period atoms are the main atomic constituents of the S layer investigated. A weaker peak, close to 500 eV kinetic energy, is caused by 2p states of Si from the parts of the SiO_x/Si substrate surface which are not covered with protein. The contribution from the valence band is negligible owing to very low cross-sections of photoexcitation of the valence states at this photon energy [41]. No additional signals, which could be assigned to sample contaminations, were monitored in the course of our experiments.

Chemical state of atoms in the protein layer can be identified measuring individual O 1s, N 1s, and C 1s spectra. These spectra taken with higher density of data points are shown separately in figure 3. The C 1s core-level spectrum reveals a multiplet structure pointing to different chemical environments of carbon atoms in the chain and different functional groups of the amino acids [25, 26]. Contrasting with that, the O 1s or N 1s spectra are characterized by singlet lineshapes that can be understood in terms of rather similar chemical environment of all oxygen and of all nitrogen atoms in the S layer of *B. sphaericus* NCTC 9602.

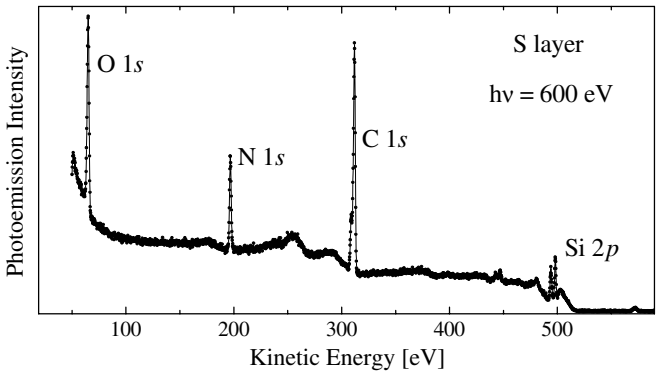

Figure 2. Overview PE spectrum of the S layer of *B. sphaericus* NCTC 9602 adsorbed on a SiO_x/Si(100) substrate. A work-function correction is not applied to the kinetic energy scales.

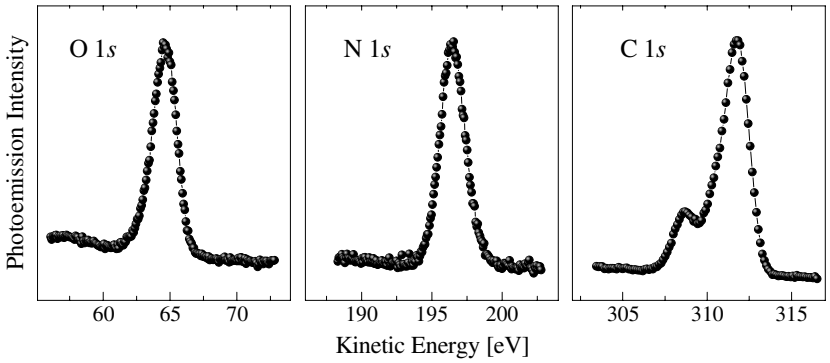

Figure 3. O 1s, N 1s, and C 1s core-level PE spectra of the S layer of *B. sphaericus* NCTC 9602 adsorbed on a SiO_x/Si(100) substrate.

More insight into non-equivalent chemical states of carbon atoms can be obtained from least-squares fit analysis [40] performed for the high resolution C 1s core-level spectrum taken with a lower photon energy (400 eV). As shown in figure 4 three singlet subspectra were necessary to simulate the experimental data. 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 may exist for C atoms. The electronegativity increases when going from carbon to oxygen. We expect, therefore, a larger charge transfer from C to O than from C to N atoms, and thus a larger BE shift for the C–O than for the C–N component in the C 1s spectrum. For this reason the high BE component ***a*** can be assigned to mainly carbon–oxygen (C–O) bonds. The low BE component ***c*** is attributed to carbon–carbon (C–C) bonds, whereas the middle component ***b*** can be associated with carbon–nitrogen (C–N) bonds. Our interpretation of the C 1s core-level spectrum for the surface-layer proteins of *B. sphaericus* NCTC 9602 is consistent with the conclusions previously derived for the S layer of *B. sphaericus* CCM 2177 [26].

The fit analysis of the lineshape of the C 1s spectrum performed may also provide quantitative information about the relative contributions of different carbon sites to the three groups of chemical states. Since, due to strong on-site localization of the core-level wavefunctions, 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

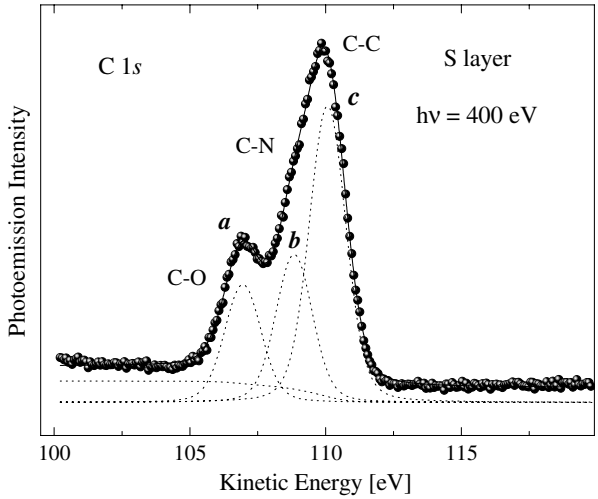

Figure 4. C 1s core-level PE spectrum of the S layer of *B. sphaericus* NCTC 9602 together with results of the least-squares fit analysis (see text).

Table 1. The relative intensities of the C–O, C–N, and C–C contributions to the C 1s core-level PE spectrum in comparison to the calculated numbers of corresponding protein bonds.

Type of bond	C–O	C–N	C–C
Analysis of the C 1s core-level PE spectrum (figure 4)	21%	27%	52%
Calculated number of contributing bonds	1594	2438	3902
Calculated relative contribution	20%	31%	49%

oxygen, nitrogen, and carbon can be derived from the ratios of the areas under the ***a***, ***b***, and ***c*** subspectra. The relative intensities of the O-, N-, and C-derived contributions obtained from the fit analysis amount to 21%, 27%, and 52%, respectively. All three 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 (see footnote 3), 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. C–O bonds are present in 452 of the 1050 residues, while 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 relative intensities of the C–O, C–N, and C–C contributions to the C 1s PE spectrum are compared to the calculated numbers of the corresponding protein bonds in table 1. These calculated numbers are the total numbers of existing single and double bonds per monomer. Considering the number of amino acids, which contain the corresponding C–O, C–N, and C–C bonds, allows one to derive them. The amounts of C atoms of each type of environment existing in these amino acids were taken into account. Good agreement between the PE data and the number ratios derived from the primary structure of the complex surface-layer protein is found.

5.2. NEXAFS spectra

NEXAFS spectroscopy allows probing of the element and angular-momentum-character-specific densities of unoccupied electronic states and, in its light polarization dependent mode,

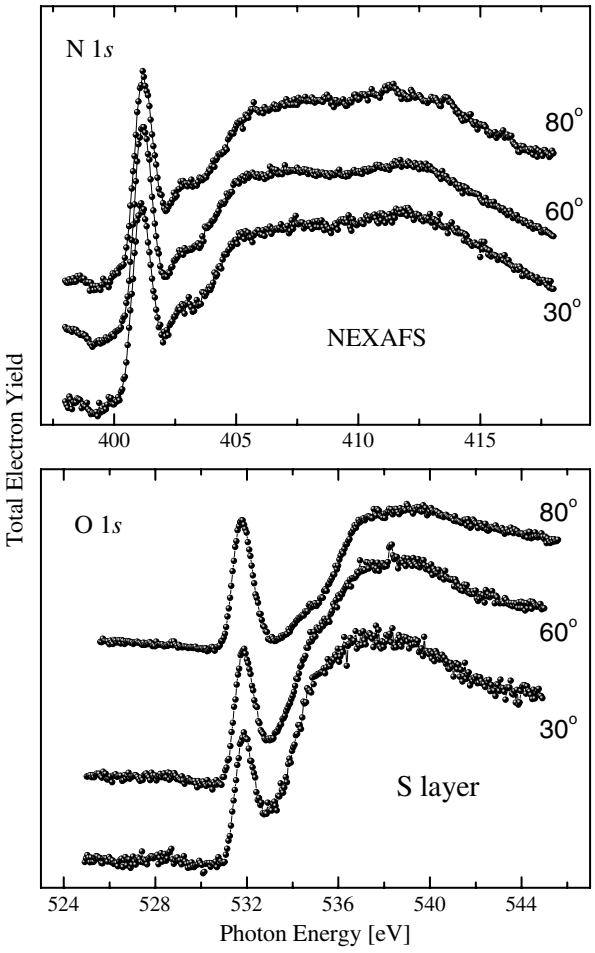

Figure 5. Light polarization dependent N 1s (upper panel) and O 1s (lower panel) NEXAFS spectra of the S layer of *B. sphaericus* NCTC 9602.

may give information about the orientation of chemical bonds in investigated samples [24]. 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 transitions are perturbed, however, by the creation of a core hole.

The unoccupied, element-specific p character electronic states of the S layer were probed, having taken the 1s absorption spectra of oxygen, nitrogen, and carbon atoms. First we investigated the dependence of the lineshape of the NEXAFS signals at all three excitation edges on the projection of the light polarization. The relative intensities of the NEXAFS structures originating from different bonds can considerably vary in light polarization dependent experiments if molecular orbitals are oriented along selected directions in the samples. This particular behaviour can strongly influence the analysis of the contributions of different bonds to the chemical composition of the systems investigated.

The light provided by the dipole beamline is linearly polarized within the horizontal plane of the storage ring. In our light polarization experiments we have changed over a wide range of values the angle between the beam direction and the normal to the vertically oriented sample surface (photon incidence angle). Series of light polarization dependent N 1s and O 1s NEXAFS spectra taken at photon incidence angles of 30°, 60°, and 80° are depicted

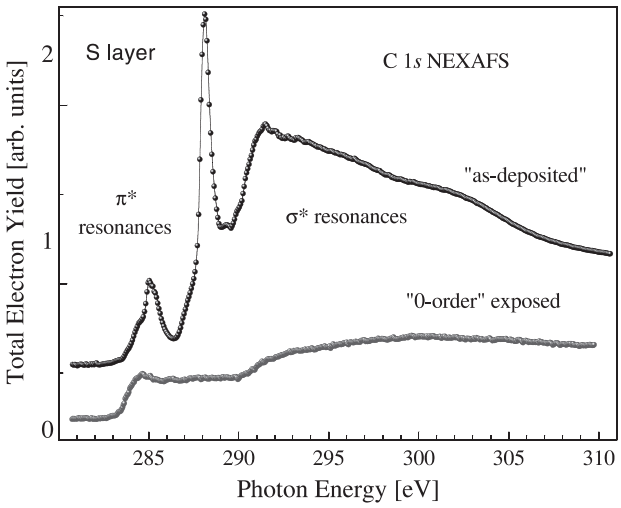

Figure 6. Top: C 1s NEXAFS spectrum of the S layer of *B. sphaericus* NCTC 9602 consisting of two regions of photon energy: 1s $\rightarrow \pi^*$ and 1s $\rightarrow \sigma^*$ resonances. Bottom: the spectrum after exposing the S layer to '0-order' synchrotron radiation.

in figure 5. In both cases as well as in the case of the C 1s NEXAFS spectra the lineshapes of the signals do not noticeably change upon the photon incidence angle variation. Thus, it can be concluded that the crystallographic axes of the deposited S-layer sheets are statistically oriented.

Since S layers contain mainly C atoms involved in different forms of bonding, the most informative NEXAFS spectra to be studied are those taken upon excitation of the C 1s core level. The C 1s spectrum reflecting the partial density of the C 2p unoccupied electronic states of the 'as-deposited' S layer shown in figure 6 (top) is found to be surprisingly simple. It reveals two rather sharp transitions involving π^* symmetry molecular orbitals between 283 and 289 eV and broad smooth structures in the σ^* transition region above 290 eV photon energies. The σ^* transition structures observed for the 'as-deposited' surface-layer proteins are very broad because of the short lifetime of the corresponding exciting states. In the following we will concentrate on the evaluation of the 1s $\rightarrow \pi^*$ resonances. In contrast to the C 1s spectrum, the N 1s and O 1s NEXAFS data (figure 5) reveal basically one peak each in the region of the π^* resonances. Together with the single lineshape of the N 1s and O 1s PE core levels (figure 3), this indicates the existence of rather similar chemical states of the oxygen and nitrogen atoms in the S-layer protein, respectively.

The observed spectral π^* features are an intrinsic property of the measured protein, since most of them disappear almost completely upon exposure to intense radiation. The C 1s NEXAFS spectrum taken after exposure of the S layer to the polychromatic ('0-order') synchrotron radiation is shown at the bottom of figure 6. The huge changes of the NEXAFS signal are caused by extensive bond breaking, reformation, and mass loss [42] of the protein sample. The bottom spectrum is very similar to the NEXAFS spectra measured for amorphous carbon samples [43] giving evidence for the presence of this very thermally stable amorphous phase in the decomposition fragments of the surface-layer proteins. Amorphous carbon remains stuck at the surface even after high temperature treatment. Hence, patterns drawn by the synchrotron radiation beam on the S layer can be fixed at the substrate as a result of the

temperature development. The latter demonstrates the possibility of potential application of the S layers as photoresists for 'dry' x-ray lithography.

The analysis of complex protein NEXAFS structures containing large amounts of amino acids can be performed within a building-block model, where the monomer structure is seen as an assembly of smaller pieces, e.g. individual functional groups of amino acids or subunits of these [30, 31, 44], 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 for amide carbonyl shifts to lower energy by only ~0.3 eV [30, 33–35]. The application of the building-block approach is further justified by 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 [32] as one would expect within this model for systems with statistically similar counts of amino acids of similar types. Although the simple building-block model allows one to assign characteristic features of the π^* region of the C 1s NEXAFS spectrum, one should be fully aware that this model represents only a rough approach for understanding the observed phenomena. In a more elaborate model, correlations between individual subunits have to be considered, which may then lead to a more band-like description of the electronic structure of proteins.

In comparison to experimental data obtained for a set of different amino acids [30, 31] and their interpretation on the basis of the *ab initio* static exchange method calculations [44], we can assign the lowest energy π^* feature of the S-layer spectrum at 285.0 eV to transitions of electrons localized on the $\boldsymbol{c}$ component of the C 1s core level (cf figure 4) mainly into unoccupied orbitals of C=C double bonds located in the aromatic rings of individual amino acids. This lowest energy feature is manifold since the energies of the particular C=C orbitals depend on the specific molecular environment of the C atoms which participate in the double bonds [44]. 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 carbidic and/or amorphous-like carbon layers, which are formed by cracking of carbon-containing gases during the plasma treatment, may contribute to the 284 eV signal. As seen in figure 6, the leading low photon energy edge of the bottom spectrum, acquired for the decomposed sample, overlaps in energy the 284 eV shoulder.

The second peak at 288.1 eV photon energy is of much greater intensity and can predominantly be attributed to the excitation of the electrons, related to the $\boldsymbol{a}$ component of the C 1s core level (cf figure 4) into π^* C=O character orbitals which are found in 228 of the 1050 amino acid residues of the monomer and as the amide carbonyl of each peptide unit. There is no initial state splitting of this resonance since the chemical environment of this bond does not change from one amino acid to another. Correspondingly, this feature appears less structured than the 285 eV one. The shoulder on the low energy side of the C=O peak at ~287 eV coincides in energy with the expected transition from the states of the $\boldsymbol{b}$ component of the C 1s core level into the π^* C=N orbitals [30], which, for the same reason as was discussed for the π^* C=O originating transition, is not expected to reveal a pronounced manifold splitting. On the other hand, this shoulder may be assigned to the transitions into the σ^* C–H states as was concluded in [31].

The relative intensities of the NEXAFS features in the π^* transition region provide information about the relative numbers of bonds contributing to a particular bond type. Since the 288.1 eV peak appears to be much higher than that of the 285 eV signal, we can qualitatively conclude that the relative number of amino acids with aromatic rings in the S layer studied is rather small. This result is in agreement with the primary protein structure revealing that the content of amino acids with aromatic rings of the S layer of *B. sphaericus* NCTC 9602 is lower than 8% (see footnote 3). An accurate quantitative analysis of the spectroscopic data

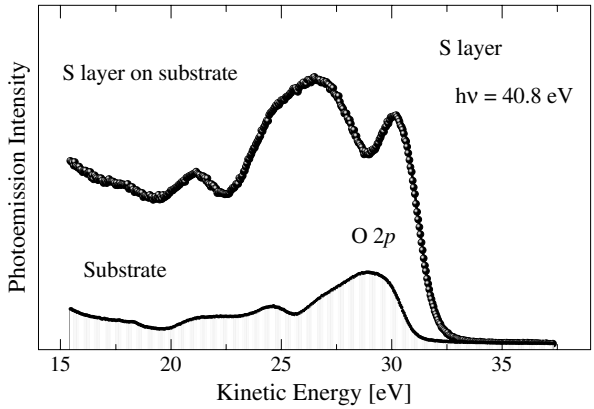

Figure 7. Valence-band PE spectra from the supported S layer of *B. sphaericus* NCTC 9602 (high intensity) and from the plasma-treated SiO_x/Si(100) substrate (low intensity). Both spectra are normalized to the photon flux.

is not possible now, since it requires detailed knowledge about cross-sections of all different transitions involved. At present, the data on oscillator strength spectra are only partially available[4].

5.3. Valence-band photoemission

The occupied valence electronic states were studied by means of angle-integrated valence-band PE. The corresponding spectrum (figure 7, top) contains contributions from the S layer and the substrate.

In order to extract the virginal signal of the S layer, we carried out a prior measurement on the bare plasma-treated substrate at the same photon energy (figure 7, bottom). The substrate spectrum reveals a triplet lineshape, which is a combination of a typical SiO_x valence-band spectrum with a pronounced O 2p-derived feature and structures characteristic for admixtures of different Si carbides [45]. Those are formed at the surface during the initial plasma treatment when carbon-containing gases are cracked and released C is pinned at the SiO_x surface.

The pure S-layer spectrum shown in figure 8 was obtained by subtraction of the bottom spectrum from the top one (figure 7). To this end, the two last spectra were normalized to the photon flux, and the substrate spectrum was additionally scaled, accounting for both the calculated attenuation of photoelectrons in the S layer [46] and the degree of S-layer coverage estimated from SEM data. Similar to results obtained for other organic materials, the S-layer PE spectrum consists of a series of individual peaks that originate from different molecular orbitals. They are separated from each other by rather deep minima. 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. A second, intense manifold structure (**A**) is located ∼4.5 eV below the HOMO, and a less pronounced peak (**B**) is found at 21 eV kinetic energy.

Because valence-band photoemission is not element specific, it is not possible to determine the origin of the occupied molecular orbitals solely from the PE data, particularly since electronic structure calculations do not exist for S layers. Nonetheless, a preliminary interpretation can be derived from electronic density of states (DOS) obtained within the local density approximation (LDA) calculations for rather small proteins [20–22], assuming that

[4] The data on oscillator strength spectra are available from the database (http://unicorn.mcmaster.ca/corex.html).

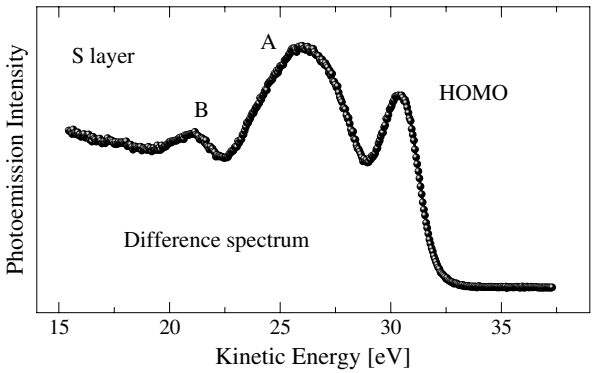

Figure 8. The difference spectrum for the two spectra shown in figure 7, which represents the pure PE signal of the S layer of *B. sphaericus* NCTC 9602.

integral DOS data can be expected to be qualitatively similar for different proteins. Theoretical evidence for this assumption is reported for the comparison of the DOS for peptidic BgK toxin and charybdotoxin [21, 22] as well as for self-assembled peptide nanotubes [20] calculated using the LDA. A first interpretation of the origin of the measured valence-band PE structures is given in the section below.

5.4. Electronic structure of the S layer

To derive the electronic structure of the S-layer protein (bottom panel in figure 9), the PE and C 1s NEXAFS spectra were referred to a common Fermi level. 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, caused by charging of the bare protein sample, separate PE experiments were carried out with gold covered S layers as described in section 3. 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 indicates relatively large charge carrier mobility in the protein and possibly even suggests a band-like description of the electronic structure of the S layer.

The C 1s NEXAFS data were referred to E_F by using the BE of the c component of the C 1s PE spectrum (figure 4) assuming a similar screening of the remaining core holes in x-ray absorption and photoemission processes. Although the effect of the core holes was found to be rather strong in 1D ladder-like structures (e.g. naphthacene, $C_{18}H_{12}$), its magnitude decreases with increasing size and dimensionality of π-conjugated system [47]. The 285 eV C 1s NEXAFS structure, which reflects contributions of the π^* C=C orbitals of the aromatic rings, can be assigned to the LUMO of the S layer, since the Fermi level is found at the bottom of this feature.

The S layer studied is found to reveal semiconductor-like behaviour with a gap value $\Delta \sim 3.0$ eV. To compare our experimental data with the DOS calculated for charybdotoxin and peptidic BgK toxin [21, 22] the corresponding energy scales are adjusted relative to the positions of the HOMOs. The theoretical and experimental results presented in figure 9 exhibit similar lineshapes. The calculated energy gaps between the LUMO and HOMO are slightly smaller than the measured one, but this effect is known from calculations of semiconductor band structures by means of density functional theory [48, 49]. On the basis of the interpretation given in [21], we can conclude that the main contributions to the HOMO arise from π clouds

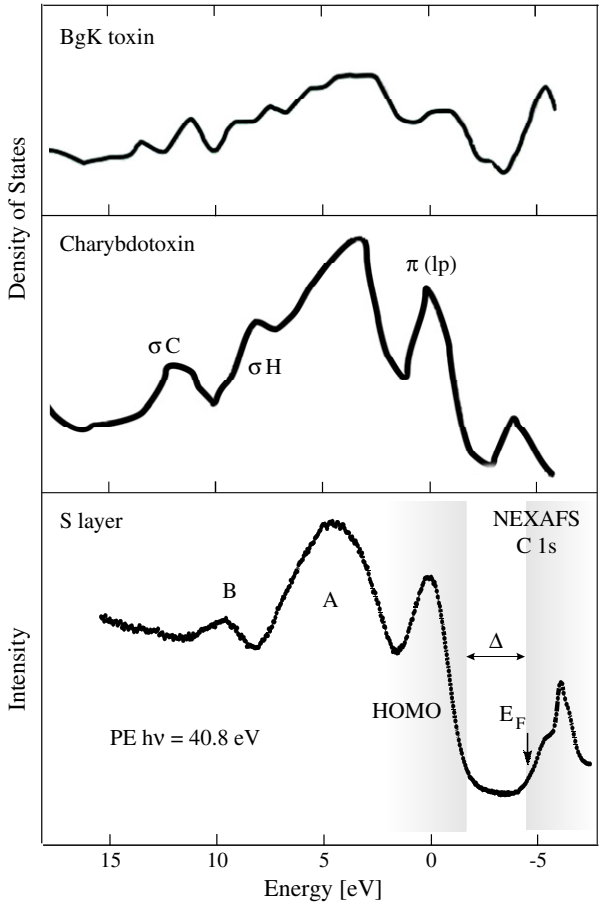

Figure 9. The electronic structure of the S layer of *B. sphaericus* NCTC 9602 as derived from the PE and NEXAFS experiments (bottom) compared with the calculated DOS for two rather small proteins: peptidic BgK toxin [22] (upper) and charybdotoxin [21] (middle).

of the aromatic rings and lone pairs (lp) 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.

The data depicted at the bottom of figure 9 represent the first spectroscopic approach to estimate the valence electronic structure of a protein. The procedure, which was used to refer the C 1s NEXAFS data to the Fermi energy position, assumes similar degrees of electron–hole interaction ('pull down' effect) in absorption and core-level photoemissions. This assumption is supposed to be particularly justified for the Fermi energy referring of the LUMO, which is formed in the S layer by rather delocalized conjugated aromatic-derived orbitals. As to the occupied states, the only moderate dynamic charging of the S layer, which was monitored in the present study, suggests that the 'pull down' effect in the valence-band experiments is largely screened by the charge carriers. The agreement with theoretical calculations of the DOS of small peptides obtained justifies furthermore our approach to constructing the electronic structure of the bacterial surface-layer protein.

6. Conclusions

Photoemission and near-edge x-ray absorption fine structure spectroscopy were applied to characterize the electronic structure of the surface protein layer of the bacteria *B. sphaericus* NCTC 9602. As obtained from both the O 1s and N 1s core-level PE and NEXAFS singlet-type spectra, all oxygen atoms as well as all nitrogen atoms reveal similar chemical states in the S layer. Contrasting to the oxygen and nitrogen signals, the C 1s core-level PE and NEXAFS manifold spectra evidence different chemical environments of carbon atoms in different functional groups of the amino acids.

Within the building-block model the experimental C 1s core-level PE spectrum could be simulated 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. Applying 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. For the interpretation of the valence-band PE results, electronic density-of-states calculations performed for small peptides were used. The HOMO is mainly formed by the π orbitals of the aromatic rings and located $\sim$3 eV below the Fermi level.

The electronic structure of the 2D protein layer studied resembles to that of a moderately wide-gap semiconductor with a gap value of $\sim$3.0 eV. The Fermi level is located close to the bottom of the LUMO. Only a moderate dynamic charging of the S layer at a level similar to that for regular wide-gap semiconductors indicates a relatively large charge carrier mobility in the protein and possibly even suggests a band-like description of the electronic structure of the S layer.

Acknowledgments

The experiments were supported by the DFG (grants PO392/18 and SFB463), the BMBF (grant 13N8145), the SMWK, the EU (FP5, contract: G5RD-CT-2002-0750) and the bilateral programme ‘Russian–German Laboratory at BESSY’.

References

[1] Sleytr U B and Messner P 1983 *Annu. Rev. Microbiol.* **37** 311
[2] Baumeister W, Wildhaber I and Engelhardt H 1988 *Biophys. Chem.* **29** 39
[3] Györvary E, O’Riordan A, Quinn A, Redmond G, Pum D and Sleytr U B 2003 *Nano Lett.* **3** 315
[4] Douglas K, Clark N O and Rothschild K J 1990 *Appl. Phys. Lett.* **56** 692
[5] Shenton W, Pum D, Sleytr U B and Mann S 1997 *Nature* **389** 585
[6] Moore J T, Beale P D, Winningram T A and Douglas K 1998 *Appl. Phys. Lett.* **72** 1840
[7] Dieluweit S, Pum D and Sleytr U B 1998 *Supramol. Sci.* **5** 15
[8] Mertig M, Kirsch R, Pompe W and Engelhardt E 1999 *Eur. Phys. J.* D **9** 45
[9] Hall S R, Shenton W, Engelhardt H and Mann S 2001 *Chem. Phys. Chem.* **3** 184
[10] Mertig M, Wahl R, Lehmann M, Simon P and Pompe W 2001 *Eur. Phys. J.* D **16** 317
[11] Wahl R, Mertig M, Raff J, Selenska-Pobell S and Pompe W 2001 *Adv. Mater.* **13** 736
[12] Wahl R, Engelhardt H, Pompe W and Mertig M 2005 *Chem. Mater.* **17** 1887
[13] Hall S R, Shenton W, Engelhardt H and Mann S 2001 *Chem. Phys. Chem.* **3** 184
[14] Beveridge T J and Graham L L 1991 *Microbiol. Rev.* **55** 684
[15] Sleytr U B, Messner P, Pum D and Sara M 1996 *Crystalline Bacterial Cell Surface Proteins* (San Diego, CA: Academic)
[16] Sleytr U B and Messner P 1998 *J. Bacteriol.* **170** 2891
[17] Baumeister W, Wildhaber I and Phipps B M 1989 *Can. J. Microbiol.* **35** 215
[18] Rachel R, Pum D, Smarda J, Smajs D, Komrska J, Krzyzánek V, Rieger G and Stetter K O 1997 *FEMS Microbiol. Rev.* **20** 13

[19] Engelhardt H and Peters J 1998 *J. Struct. Biol.* **124** 276
[20] Carloni P, Andreoni W and Parrinello M 1997 *Phys. Rev. Lett.* **79** 761
[21] Ireta J, Galván M, Cho K and Joannopoulos J D 1998 *J. Am. Chem. Soc.* **120** 9771
[22] Aparicio F, Ireta J, Rojo A, Escobar L, Cedillo A and Galván M 2003 *J. Phys. Chem.* B **107** 1692
[23] Hüfner S 1995 *Photoemission Spectroscopy* (Berlin: Springer)
[24] Stöhr J 1992 *NEXAFS Spectroscopy* (Berlin: Springer)
[25] Subirade M and Lebugle A 1994 *Thin Solid Films* **243** 442
[26] Handrea M, Sahre M, Neubauer A, Sleytr U B and Kautek W 2003 *Bioelectrochemistry* **61** 1
[27] Kato H S, Furukawa M, Kawai M, Taniguchi M, Kawai T, Hatsui T and Kosugi N 2004 *Phys. Rev. Lett.* **93** 086403
[28] Vyalikh D V, Danzenbächer S, Mertig M, Kirchner A, Pompe W, Dedkov S Yu and Molodtsov S L 2004 *Phys. Rev. Lett.* **93** 238103
[29] Vyalikh D V, Kirchner A, Danzenbächer S, Dedkov Yu S, Kade A, Mertig M and Molodtsov S L 2005 *J. Phys. Chem.* B **109** 18620
[30] Boese J, Osanna A, Jacobsen C and Kirz J 1997 *J. Electron Spectrosc. Relat. Phenom.* **85** 9
[31] Kaznacheyev K, Osanna A, Jacobsen C, Plashkevych O, Vahtras O, Argen H, Carravetta V and Hitchcock A P 2002 *J. Phys. Chem.* A **106** 3153
[32] Hitchcock A P, Morin C, Heng Y M, Cornelius R M and Brash J L 2002 *J. Biomater. Sci. Polym. Edn* **13** 919
[33] Gordon M L, Cooper G, Morin C, Araki T, Turci C C, Kaznatcheev K and Hitchcock A P 2003 *J. Phys. Chem.* A **107** 6144
[34] Morin C, Hitchcock A P, Cornelius R M, Brash J L, Urquhart S G, Scholl A and Doran A 2004 *J. Electron Spectrosc. Relat. Phenom.* **137–140** 785
[35] Cooper G, Gordon M, Tulumello D, Turci C, Kaznatcheev K and Hitchcock A P 2004 *J. Electron Spectrosc. Relat. Phenom.* **137–140** 795
[36] Panhorst M, Brückl H, Kiefer B, Reiss G, Santarius U and Guckenberger R 2001 *J. Vac. Sci. Technol.* B **19** 722
[37] Gorovikov S A, Molodtsov S L and Follath R 1998 *Nucl. Instrum. Methods Phys. Res.* A **441** 506
[38] 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 and Kaindl G 2003 *Nucl. Instrum. Methods Phys. Res.* A **505** 718
[39] Sham T K, Yang B X, Kirz J and Tse J S 1989 *Phys. Rev.* A **40** 652
[40] Win-Fit Program: http://www.physik.tu-dresden.de/ et/winfit/winfit-install.exe
[41] Yeh J J and Lindau I 1985 *At. Data Nucl. Data Tables* **32** 1
[42] Zubavichus Y, Zharnikov M, Shaporenko A, Fuchs O, Weinhardt L, Heske C, Umbach E, Denlinger J D and Grunze M 2004 *J. Phys. Chem.* A **108** 4557
[43] Lynn G J F, Eller L P K, Eser M F, Irick S W and Acobsen C J 2003 *Geochim. Cosmochim. Acta* **67** 4791
[44] Carravetta V, Plashkevych A and Ågren H 1998 *J. Chem. Phys.* **109** 1456
[45] Fang R-C and Ley L 1989 *Phys. Rev.* B **40** 3818
[46] Seah M P and Dench W A 1979 *Surf. Interface Anal.* **1** 2
[47] Oji H, Mitsumoto R, Ito E, Ishii H, Ouchiet Y, Seki K, Yokoyama T, Ohta T and Kosugi N 1998 *J. Chem. Phys.* **109** 10409
[48] Eschrig H, Kikoin K A and Kohn V G 1985 *Solid State Commun.* **56** 773
[49] Stampfl C and Van de Walle C G 1998 *Phys. Rev.* B **59** 5521