Published in final edited form as:
Langmuir. 2010 November 2; 26(21): 16434–16441. doi:10.1021/la1007389.

Probing the Orientation of Surface Immobilized Protein G B1 using ToF SIMS, Sum Frequency Generation, and NEXAFS Spectroscopy

Loren Baugh[¶], **Tobias Weidner**[¶], **J.E. Baio**[§], **Phuong Cac Nguyen**[§], **Lara J. Gamble**[¶], **Patrick S. Stayton**[¶], and **David G. Castner**[¶,§,*]

[¶]National ESCA and Surface Analysis Center for Biomedical Problems, Departments of Bioengineering, University of Washington, Seattle, WA 98195

[§]National ESCA and Surface Analysis Center for Biomedical Problems, Chemical Engineering, University of Washington, Seattle, WA 98195

Abstract

The ability to orient active proteins on surfaces is a critical aspect of many medical technologies. An important related challenge is characterizing protein orientation in these surface films. This study uses a combination of time-of-flight secondary ion mass spectrometry (ToF-SIMS), sum frequency generation (SFG) vibrational spectroscopy, and near edge x-ray absorption fine structure (NEXAFS) spectroscopy to characterize the orientation of surface-immobilized Protein G B1, a rigid 6 kDa domain that binds the Fc fragment of IgG. Two Protein G B1 variants with a single cysteine introduced at either end were immobilized via the cysteine thiol onto maleimide-oligo(ethylene glycol)-functionalized gold and bare gold substrates. X-ray photoelectron spectroscopy was used to measure the amount of immobilized protein and ToF-SIMS was used to measure the amino acid composition of the exposed surface of the protein films and to confirm covalent attachment of protein thiol to the substrate maleimide groups. SFG and NEXAFS were used to characterize the ordering and orientation of peptide or side chain bonds. On both substrates and for both cysteine positions, ToF-SIMS data showed enrichment of mass peaks from amino acids located at the end of the protein opposite the cysteine surface position compared with nonspecifically immobilized protein, indicating end-on protein orientations. Orientation on the maleimide substrate was enhanced by increasing pH (7.0 to 9.5) and salt concentration (0 to 1.5 M NaCl). SFG spectral peaks characteristic of ordered α-helix and β-sheet elements were observed for both variants but not for cysteine-free wild type protein on the maleimide surface. The phase of the α-helix and β-sheet peaks indicated a predominantly upright orientation for both variants, consistent with an end-on protein binding configuration. Polarization dependence of the NEXAFS signal from the N 1s toπ^* transition of β-sheet peptide bonds also indicated protein ordering, with an estimated tilt angle of inner β-strands of 40–50° for both variants – one variant more tilted than the other – consistent with SFG results. The combined results demonstrate the power of using complementary techniques to probe protein orientation on surfaces.

INTRODUCTION

The ability to orient proteins on surfaces with their active sites exposed to the biological environment will enhance a wide range of applications in biotechnology, including protein

[*]Corresponding Author. castner@nb.engr.washington.edu.

Supporting Information Available XPS data for the second set of samples, corresponding to the ToF-SIMS data in Figure 4b, are included in the Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

microarrays, antibody-based diagnostics, affinity chromatography, and biomaterials that present protein ligands to bind cell receptors. Approaches to achieving orientation include using immobilized ligands,[1, 2] covalent tethers,[3–6] hydrophobic patches on the protein,[7] and charge-driven orientation.[8–10] As these methods are developed, techniques are needed to evaluate their effectiveness by accurately measuring protein conformation and orientation on surfaces. Since no single technique provides a complete high-resolution structure of surface species, a combination of surface analytical and spectroscopic techniques is required.

Static time-of-flight secondary ion mass spectrometry (ToF-SIMS) is well suited for characterizing the composition and structure of protein films due to its high chemical specificity and surface sensitivity. ToF-SIMS provides high mass resolution ($m/\Delta m \sim 5000$) and high sensitivity (10^7–10^{11} atoms/cm^2), with a sampling depth of 10–20 Å.[11] The technique uses a pulsed ion beam to bombard a surface and sputter molecular fragments from the surface. Ejected ion fragments are extracted into a time-of-flight mass spectrometer at constant energy for separation by velocity (smaller fragments travel faster) and determination of their exact mass values. Amino acids in proteins have characteristic fragmentation patterns, and spectral fingerprints for all twenty protein amino acids have been recorded for different ion sources.[12, 13] The sensitivity of the technique to orientation depends on the size of the protein relative to the sampling depth and requires asymmetry in the distribution of amino acids in the three-dimensional structure of the protein.[9, 14] ToF-SIMS, combined with multivariate analysis techniques such principal component analysis, has been applied to characterize the composition of protein films,[14–17] detect protein denaturing and conformational changes,[18–21] and probe protein orientation.[9, 22]

Sum frequency generation (SFG) vibrational spectroscopy is a surface specific-spectroscopy that detects vibrational modes from specific chemical bonds in ordered interfacial molecules.[23] Only molecules with ordered bonds in an asymmetric environment such as an interface produce an SFG signal. Vibrational modes of interest for studying proteins include N-H and C-H stretches from peptide bonds and amino acid side chains as well as O-H modes from bound water. Ordered secondary structure elements such as α-helices, β-sheets and β-turns are identified by characteristic peaks in the amide I modes; the orientation of these structural elements – orthogonal versus parallel to the interface – is indicated by the phase of the amide I modes. SFG has been applied to measuring ordering and orientation of side chain bonds and secondary structure elements in peptides and larger proteins.[24–27]

Near-edge X-ray absorption fine structure (NEXAFS) spectroscopy is an element- and bond-specific electron spectroscopy that provides insight into the electronic structure of surface species by sampling unoccupied molecular orbitals.[28] Information about molecular orientation is derived from the dependence of photoexcitation on the orientation of polarized synchrotron light with respect to the transition dipole moment of the probed molecular orbital. The orientation or tilt angles of ordered molecular bonds are calculated by the measuring the change in absorbance with variation of the incidence angle of synchrotron light. NEXAFS has been applied to characterize the surface coverage and orientation of organic molecules and polymers,[29] and more recently, DNA oligomers,[30–32] peptides,[33] and proteins.[34]

In this study, we combine ToF-SIMS, SFG and NEXAFS to probe the orientation of two variants of the barrel-shaped Protein G B1 domain, each with a single cysteine substituted into an exposed loop at one end of the protein or the other. These Protein G B1 mutants were immobilized via the cysteine thiol onto maleimide-oligo(ethylene glycol)-functionalized (MEG) or bare gold substrates (Figure 1). The MEG monolayer presents surface maleimide groups for binding the cysteine thiol in a protein-resistant ethylene glycol background. Bare

gold was used as an alternate thiol-reactive substrate. The combined results indicate opposite end-on orientations of the Protein G B1 variants on both substrates, with the degree of orientation dependent on buffer pH and ionic strength, and different tilt angles for the two variants. The complementary results from differing techniques provide a detailed picture of surface-bound Protein G B1.

MATERIALS AND METHODS

Preparation of Substrates

1×1 cm^2 silicon substrates (Microelectronics Inc., San Jose, CA) were cleaned by sequential sonication in deionized water, dichloromethylene, acetone, and methanol. The substrates were coated with a thin layer of titanium (10 nm) followed by high purity gold (99.99%, 80 nm) in a high vacuum thermal evaporator (pressure < 1×10^{-6} Torr). MEG surfaces were prepared by submerging gold-coated substrates into 1 mM maleimide-ethylene glycol disulfide [$HO(CH_2CH_2O)_4(CH_2)_{11}S\text{-}S(CH_2)_{11}(OCH_2CH_2)_6OCH_2CONH(CH_2)_2C_4H_2NO_2$] (Prochimia, Sopot, Poland) in 200 proof ethanol (Pharmco-Aaper, Shelbyville, KY) for 1 h, and rinsing with ethanol.31

Protein Mutagenesis, Expression and Purification

V21C (Val21 to Cys) and T11C (Thr11 to Cys) mutations were created in a synthetic Protein G B1 gene35 using the QuikChange protocol (Stratagene, La Jolla, CA), and confirmed by sequencing. Protein was expressed using the T7 expression system in BL21(DE3) *E. coli* (Invitrogen, Carlsbad, CA), and isolated by thermal lysis (80 °C for 10 min) and precipitation of bacterial proteins and cell debris (−20 °C for 30 min, centrifuge at 8000 g for 15 min). Expressed protein was purified by two steps of anion exchange chromatography. The first step used diethylaminoethyl media (Bio-RAD, Hercules, CA) with 50 mM sodium phosphate (pH 7.0) as the loading buffer and a step gradient of NaCl from 0 to 1 M for elution. The second step used Poros HQ media and a BioCAD Sprint chromatography system (Applied Biosystems, Foster City, CA) with 50 mM Tris (pH 8.0) as the loading buffer and a continuous NaCl gradient from 0 to 1 M for elution. Protein molecular weight and amino acid composition were confirmed using electrospray mass spectrometry and amino acid analysis (AAA Laboratory, Mercer Island, WA).

Protein Immobilization

The proteins used in this study were Protein G B1 wild type and the cysteine variants, V21C and T11C. Protein solutions, 0.5 mL of 2 mg/mL protein, were added to substrates in 24-well plates. For MEG substrates, three different buffer compositions were used: (1) 50 mM sodium phosphate, pH 7.0; (2) 1.5 M NaCl added to the buffer in (1); (3) 50 mM sodium carbonate, pH 9.5. Binding on MEG proceeded for 3 h at room temperature. For bare gold substrates, the buffer used was phosphate buffered saline (1.54 mM KH_2PO_4, 2.71 mM Na_2HPO_4, 155 mM NaCl) (pH 7.4), and binding proceeded for 16 h at 4 °C.3 To inhibit disulfide bond formation between proteins, Tris(2-carboxyethyl) phosphine (TCEP, Sigma) at 0.2 mM was added to all protein solutions. Following protein immobilization, samples were washed by serial dilution in buffer and aspiration at the air-water interface, followed by submerging samples in a series of water solutions for 1 min each with mixing. Samples were air-dried and stored under nitrogen. Additional protein samples on bare gold were coated with trehalose by soaking in 0.05% trehalose solution in water for 30 min and spinning dry using a spin coater at 4000 rpm for 20 sec. Equimolar mixtures of V21C and T11C were used for comparison with V21C- and T11C-only samples. Control samples with nonspecifically bound protein were prepared using several approaches: (1) using cysteine-free Protein G B1 wild type, (2) evaporating 2 mg/mL V21C and T11C onto silicon without rinsing, and (3) blocking the cysteine thiol in V21C and T11C using iodoacetamide (Sigma-

Aldrich, St Louis, MO). Two sets of samples were prepared and analyzed by XPS, ToF-SIMS and SFG on two separate sets of days; the first set of samples is listed in Table 2, and the second set in Table S1 in the Supporting Information. Additional samples of the two cysteine mutants on the MEG substrate in pH 9.5 buffer (which gave the best results based on ToF-SIMS) were prepared subsequently for *in situ* SFG and *ex situ* NEXAFS analysis.

X-ray Photoelectron Spectroscopy (XPS)

XPS data were acquired on a Kratos AXIS Ultra DLD instrument (Kratos, Manchester, England) in the hybrid mode using a 0° take-off angle (angle between the surface normal and axis of the analyzer lens) and a monochromatic Al $K\alpha_{1,2}$ X-ray source ($h\nu$ = 1486.6 eV). The Kratos Vision software was used to determine atomic compositions from peak areas measured in survey scans (0–1100 eV, C_{1s} and Au_{4f} peaks) and detailed spectra (524–544 eV, O_{1s}; 390–410 eV, N_{1s}; 155–173 eV, S_{2p}) acquired with an analyzer pass energy of 80 eV. All elemental compositions (atomic percentages) are reported as averages and standard deviations from three spots per sample. The molecular composition of MEG substrates was examined using high-resolution spectra from the N_{1s}, C_{1s} and S_{2p} regions (analyzer pass energy = 20 eV). Energy scales were calibrated by normalizing the large CH_x peak in the C_{1s} region to 285.0 eV.

Time-of-Flight Secondary Ion Mass Spectrometry (ToF-SIMS)

Negative and positive secondary ion spectra were acquired on a TOF.SIMS 5–100 instrument (ION TOF, Münster, Germany) using a pulsed 25 keV Bi_3^+ primary ion beam under static conditions (primary ion dose < 10^{12} ions/cm^2). Spectra were collected from five 100 × 100 μm^2 regions per sample. Secondary ions were collected over a range of 0–400 *m/z* at a mass resolution ($m/\Delta m$) between 4000–8000; mass calibration corrections were typically below 20 ppm. Positive spectra *m/z* values were mass calibrated using CH_2^+, $C_2H_2^+$ and $AuSCH_2^+$ peaks, and negative spectra using CH^-, OH^-, and AuS^- peaks. Intensities were normalized to the total ion intensity of each spectrum to correct for differences in total ion yield between spectra.

Principal Component Analysis (PCA)

PCA has been described in detail previously.36–38 This multivariate analysis technique was applied to ToF-SIMS data to identify principle sources of variation between sample spectra using a series of scripts written by NESAC/BIO for MATLAB (MathWorks, Inc., Natick, MA). A list of all significant positive ion peaks derived only from the protein was compiled based on a database of ToF-SIMS ion fragmentation patterns for all twenty amino acids12 and on ToF-SIMS spectra collected from substrate-only and protein samples (Table 1, top). Any peaks that could originate from both protein and MEG layers were not included in the analysis so that PCA would highlight differences between bound proteins. Based on this list, PCA yielded a matrix of peaks and samples that represented all spectra collected. Data sets were mean-centered to insure that variance within the data set was due to differences in sample variances rather than in sample means.

Sum Frequency Generation (SFG) Spectroscopy

SFG vibrational spectra were obtained by overlapping visible and tunable IR laser pulses (25 ps) in time and space. A visible beam at 532 nm was delivered by an EKSPLA Nd:YAG laser operating at 50 Hz, which was also used to pump an EKSPLA optical parametric generation/amplification and difference frequency unit based on barium borate and $AgGaS_2$ crystals to generate tunable IR laser radiation from 1000–4000 cm^{-1}. The bandwidth was 4 cm^{-1} for both visible and IR radiation. Both beams were mildly focused with a diameter of ~2 mm at the sample, and had an energy of 160 μJ per pulse. The SFG signal generated at

the sample was filtered and dispersed using a monochromator, detected with a gated photomultiplier tube and recorded. Spectra were collected using 400 shots per data point in 4 cm^{-1} increments. All spectra were recorded in *ppp* (sum, visible, and infrared) polarization combination. Spectra were normalized by a reference signal generated in a ZnS crystal, and by a reference spectrum of dodecanethiol SAM on gold to eliminate artifacts from water absorption of the IR pump beam. Spectra were recorded *in situ* with buffer-soaked samples pressed against one side of an equilateral calcium fluoride prism in a near-total internal reflection geometry. The fitting routine used for SFG data analysis has been described previously.[39, 40]

Near-Edge X-ray Absorption Fine Structure Spectroscopy (NEXAFS)

NEXAFS spectra were taken at the National Synchrotron Light Source (NSLS) U7A beamline at Brookhaven National Laboratory using an elliptically polarized beam with ~ 85% *p*-polarization. This beam line uses a monochromator and 600 l/mm grating that provide a full-width at half-maximum resolution of ~ 0.15 eV at the carbon K-edge (285 eV). The monochromator energy scale was calibrated using the 285.35 eV C 1s → π* transition on a graphite transmission grid.[41] Nitrogen K-edge NEXAFS spectra were normalized by the signal from a dodecanethiol SAM on Au, and C K-edge spectra were normalized by the spectrum of a clean gold surface. Both reference and signal were divided by the beam flux during data acquisition.[28] Partial electron yield was monitored by a detector with bias voltage at −150 V and −360 V for the C K-edge and the N K-edge, respectively. Samples were mounted to allow rotation about the vertical axis to alter the angle between the incident X-ray beam and the sample surface. Beam damage was tested for by repeatedly scanning and analyzing the 388 eV to 410 eV energy range. In addition, the partial electron yield at 398 eV – the typical photon energy for beam damage-related peaks – was monitored over time to rule out damage-related changes. No changes in the spectra were observed within 15 minutes, which exceeds the exposure time in the longest nitrogen scans performed in this work.

RESULTS AND DISCUSSION

Protein G B1 cysteine mutants, V21C and T11C, each with a single cysteine substituted into an exposed loop at opposite ends of the protein (Figure 1), were created using site-directed mutagenesis. MEG SAMs were used to present maleimide groups for binding the cysteine thiol in an otherwise protein-resistant ethylene glycol background; bare gold was used as an alternate thiol-reactive substrate. Individual samples were analyzed sequentially by XPS, ToF-SIMS and SFG, which required drying samples and subjecting them to ultra-high vacuum conditions. Some samples were dried with a final trehalose solution rinse prior to drying to enhance retention of protein structure.[21] Two sets of samples were prepared and analyzed by XPS, ToF-SIMS and SFG on two separate sets of days; the first set of samples is listed in Table 2, and the second set in Table S1 in the Supporting Information. Additional samples of the two cysteine mutants on the MEG substrate in pH 9.5 buffer (which gave the best results based on ToF-SIMS) were prepared subsequently for *in situ* SFG and *ex situ* NEXAFS analysis.

XPS Analysis

XPS-determined elemental compositions for MEG, bare gold and protein samples are listed in Table 2. The elemental composition of the MEG monolayer, re-normalized without the gold signal, was close to the expected composition based on the molecular stoichiometry of MEG (measured composition, atomic percentages: 73 % C, 22 % O, 4 % N, 1 % S; expected composition: 75 % C, 19 % O, 3 % N, 3 % S). The measured sulfur signal was < 3 % because it is attenuated by the MEG overlayer.[31] No significant change in the composition

of the MEG sample occurred after it was exposed to buffer. Upon protein binding, the measured nitrogen concentration increased due to the higher nitrogen concentration of the protein compared to the MEG SAM, while the gold concentration decreased due to attenuation of the Au_{4f} signal by the added protein overlayer. For protein-covered samples, the nitrogen concentration was lower and the gold concentration higher in cysteine-free wild type and in cysteine mutant samples with thiols chemically blocked using iodoacetamide compared to unblocked cysteine mutant samples (Table 2). (On MEG surfaces at pH 7.0, the atomic % N was 7.5 for WT versus 9.1–9.7 for the cysteine mutants. On MEG surfaces at pH 9.5, % N was 6.2–6.3 for blocked cysteine mutants samples versus 8.7 for the unblocked cysteine mutant samples. On the bare gold surfaces, % N was 13.1 for WT versus 14.3–14.7 for the cysteine mutants.) These results indicate that surface coverage was greater for protein variants with an available reactive thiol. For all proteins, the nitrogen concentration, and thus the protein coverage, was greater on bare Au compared to MEG. While this is consistent with greater nonspecific binding on the gold surface compared to the protein-resistant ethylene glycol background of the MEG surface, no direct comparison can be made since different reaction times and buffer conditions were used for the two different sulfhydryl substrate reactions.

ToF-SIMS Analysis

ToF-SIMS data were analyzed first for surface coverage and successful formation of the MEG maleimide-protein thiol bond. To assess surface coverage, a list of all positive ion peaks derived only from the protein was compiled (Table 1), and the sum of intensities of these peaks (normalized to the total counts) was used to assess the relative amount of protein bound for each sample (Figure 2a). While small differences in this sum were observed between the V21C, T11C and equimolar mixture samples for a given set of buffer conditions, a roughly five-fold decrease in this sum of peak intensities was observed for cysteine-free wild type protein on both substrates. Direct evidence of chemical bonding between the MEG maleimide and protein thiol was found in the negative ion spectra, by comparing peaks from unreacted and reacted maleimide rings (unreacted: $C_4H_2NO_2^-$, m/z = 96.01; reacted, thiolate form: $C_4H_2NO_2S^-$, m/z = 127.98) (Figure 2b).31 The dramatic increase in the ratio of reacted to unreacted MEG peaks upon immobilization indicates specific binding of the mutants to MEG via the cysteine thiol.

ToF-SIMS Principle Components Analysis

PCA (multivariate analysis) was used to reduce the dimensionality of the ToF-SIMS data set into principle components (PCs), which capture the main sources of variation in spectral data between samples. This analysis identified spectral differences between the V21C, T11C, and equimolar mixture samples. On MEG, PC1 (principal component one, PC1, Figure 3a) captured 73% of the variance among the spectra from these samples, using only the protein-related peaks listed in Table 1 in the analysis. Positively and negatively loaded ToF-SIMS peaks are shown in Figure 3b. (A “loading” indicates the contribution of a peak to the score.) Among the largest positively loaded peaks were peaks from amino acids located at the C-terminal end of the protein: peaks 86 from leucine/isoleucine and 98 from asparagine (rounded mass values). Among the largest negatively loaded peaks were peaks from amino acids located the N-terminal end: peaks 107 and 136 from tyrosine. The one exception to the trend on MEG was C-terminal peak 87 from asparagine, which loaded negatively. On bare gold, PC1 captured 70% of the variance among V21C, T11C and equimolar mixture spectra (Figure 3c). C-terminal peaks 86 from leucine/isoleucine and 87 from asparagine were among the largest positively loaded peaks (98 from asparagine also showed a small, positive loading), and N-terminals peak 107 from tyrosine loaded negatively (Figure 3d). Peak 136 from tyrosine showed a small positive loading, an exception to the trend. The overall trend of positively loaded peaks from C-terminal amino

acids and negatively loaded peaks from N-terminal amino acids was stronger on MEG than on bare gold, suggesting a greater degree of end-on protein orientation on the MEG substrate, as discussed below.

These asymmetrically distributed amino acids and their characteristic positive ion peaks are shown in Figure 4 and Table 1 (bottom), respectively. In V21C, the cysteine is located near the N-terminus, while in T11C, the cysteine is near the C-terminus. Amino acids located at the end of the protein opposite the cysteine are expected to be closer to the exposed surface of the protein film (the vacuum interface in ToF-SIMS), and fragments from these amino acids are expected to be preferentially emitted during ToF-SIMS experiments. It is interesting to note that preferential emission of the amino acid fragments is observed even though the height of the protein (~ 3 nm) is similar to the ToF-SIMS sampling depth. This demonstrates that even a few nanometers of protein is sufficient to allow ToF-SIMS to monitor protein orientation. The separation of these peaks from asymmetrically distributed amino acids into positively and negatively loaded components of PC1 (Figures 3b and 3d) is consistent with the adoption of opposite end-on orientations for V21C and T11C on both substrates. Equimolar mixtures of V21C and T11C showed PC1 scores that fell between those of V21C and T11C samples on both substrates (Figures 3a and 3c), consistent with a mixture of “up” and “down” orientations. Similar ToF-SIMS and PCA methods have been applied previously to characterize protein conformation[18, 20, 21] and orientation[9] on surfaces.

ToF-SIMS Peak Ratio Analysis

In addition to PCA analysis, a ratio of secondary ions originating from opposite ends of the protein was created to investigate a wider range of protein immobilization conditions (Figure 5). Peak ratios were calculated as the sum of intensities of C-terminal peaks − 70, 87 and 98 from asparagine and 86 from leucine/isoleucine – divided by the sum of intensities of N-terminal peaks – 107 and 136 from tyrosine (nominal mass values). A greater value for this peak ratio indicates an increased yield of C-terminal fragments relative to N-terminal fragments. In V21C, the C-terminus is opposite the substrate-bound cysteine and thus expected to be closer to the exposed vacuum interface, while in T11C, the N-terminus is opposite the substrate-bound cysteine. Thus, a peak ratio trend of V21C > equimolar mixture > T11C is expected for end-on orientations of the two mutants. As shown in Figure 5a, peak ratios followed this trend on MEG surfaces in two of the three reaction conditions (pH 7.0 with 1.5 M NaCl and at pH 9.5), and on bare gold. Adding 1.5 M NaCl at pH 7 may have improved packing and orientation by inhibiting charge repulsion of adjacent proteins, while increasing pH from 7.0 to 9.5 may have improved orientation by increasing the stability of TCEP, a reducing agent used to prevent protein disulfide formation. Varying temperature from room temperature to 4 °C or 37 °C and adding trehalose to preserve orientation during drying had little effect for samples on MEG at pH 9.5 (Figure 5b). Nonspecifically immobilized samples – V21C and T11C solutions evaporated onto separate silicon wafer substrates – exhibited little difference in peak ratios, as expected. XPS data for all samples used to acquire ToF-SIMS data in Figure 5b are included in Table S1 in the Supporting Information.

SFG Analysis

Amide region SFG spectra of V21C, T11C and Protein G B1 wild type immobilized onto MEG SAMs at pH 9.5 are shown in Figure 6. T11C and, to a lesser extent, V21C spectra exhibit a pronounced amide I feature near 1650 cm^{-1} characteristic of ordered α-helical structure, and a peak near 1675 cm^{-1} characteristic of ordered β-sheet turns.[25, 42] V21C also shows a peak near 1624 cm^{-1}, the β-sheet B2 mode, characteristic of ordered β-sheet peptide bonds. The observance of this peak is surprising since the amide units are

symmetrically distributed in the antiparallel sheet and, thus, no signal is expected according to SFG selection rules.25 The wild type spectrum, in contrast, shows less prominent features indicating less ordering of bonds in the protein film, though the weaker signal was also likely due to the reduced amount of wild type protein immobilized. The dips in the spectra are likely normalization artifacts due to water absorption bands in amide spectral range.

The phases of the amide resonances can be used to describe the protein orientation since phase and orientation are closely related in SFG.43 A quantitative analysis of the phase information is complicated by the complex nature of the SFG signal in ppp polarization with contributions from the incident electric fields both in and out of plane. However, comparison with reference systems allows a qualitative estimate of the orientation. The peaks observed here are in phase with the gold background (the features appear as positive peaks in the spectrum) and are markedly different from those found previously for purely α-helix lysine/leucine model peptides on self-assembled monolayers of alkanethiols on gold, where the spectral features appear as negative peaks in the spectrum and are out phase with the non-resonant gold background.33 In this previous work, lysine-leucine peptides were found to adsorb with the peptide backbone oriented parallel to the surface. Thus, here it can be assumed that the helix structures in the proteins are tilted with respect to the surface normal. For end-on orientations, the α-helix orientation should be more upright in T11C compared V21C, consistent with SFG results in Figure 6 that show a more prominent amide resonance near 1650 cm^{-1} in T11C.

NEXAFS Analysis

Figure 7 shows N K-edge spectra of the protein films acquired at X-ray incidence angles of 20°, 55° and 70°, and the difference between the 20° and 70° spectra. All spectra contain an adsorption edge related to excitation of N 1s electrons into continuum states and additional superimposed characteristic peaks. The spectra are dominated by a broad σ* resonance near 405.2 eV attributed to N-C bonds, and an intense transition at 400.6 eV assigned to amide moieties.28, 44–47 The low-intensity pre edge feature at 398.9 eV is characteristic of amide and side chain N-H and N=C bonds. The difference spectra exhibit a negative dichroism for all three of these resonances.

The dichroism for the N 1s transition into the π* orbital of the peptide bond (400.6 eV) is most prominent and is indicative of protein ordering (Figure 7). While the amide π* intensity comes from both helix and β-sheets and thus cannot be disentangled, we have shown previously that α-helix structures are not expected to contribute significantly to the angle dependence of amide peaks because amide bonds have a broad distribution of orientations.33 The peptide bonds in the antiparallel β-sheet should account for the majority of the observed peak angle dependence and can be used to analyze protein orientation. Average tilt angles of the β-sheets relative to the substrate normal were calculated using the standard theoretical framework28 and assuming the amide π* orbitals are oriented perpendicular to the N-- O plane of the peptide bond. Such analysis yielded tilt angles of 46° for V21C and 43° for T11C β-sheets. The accuracy of these values is ~ ± 5° – the accuracy of the NEXAFS experiment and data analysis procedure. Note also that the β-strands have a distribution of orientations while a uniform distribution has been assumed in the tilt angle calculation.

CONCLUSION

Complementary data from XPS, ToF-SIMS, SFG and NEXAFS indicate opposite end-on orientations for the V21C and T11C variants of Protein G B1 immobilized onto maleimide(ethylene glycol)-functionalized and bare gold substrates. The consistency of the results between different techniques allows a more confident interpretation in terms of the

tethered protein orientation. XPS % N values and ToF-SIMS protein-only peak intensities indicated stronger binding for the cysteine variants than for a cysteine-free wild type, and comparison of reacted and unreacted maleimide ring peaks indicated successful formation of the thiol-maleimide bond. ToF-SIMS PCA and peak ratio analysis indicated enrichment of mass peaks from amino acids located at the end of the protein opposite the cysteine for both mutants on both substrates, with the best results on maleimide obtained at higher salt (1.5 M versus no NaCl) and pH (9.5 versus 7.0). SFG spectra on the maleimide surface indicated ordering of α-helical and β-sheet structures for both mutants, with an upright orientation of the proteins. Polarization dependence of N 1s NEXAFS spectra on the maleimide surface also indicated ordering of amide bonds attributed to anisotropic peptide bonds in the β-sheet, allowing estimates of the average β-sheet tilt angle (~ 45° for both mutants).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Jeremy Brison for ToF-SIMS data acquisition, and Daniel Fischer (NIST) and Cherno Jaye (Hunter College) for providing experimental equipment for NEXAFS spectroscopy and for technical assistance at the synchrotron. NEXAFS studies were performed at the NSLS, Brookhaven National Laboratory, which is supported by the U.S. Department of Energy, Division of Materials Science and Division of Chemical Sciences. This work was supported by NIH grants GM-074511 and EB-002027 (NESAC-BIO).

REFERENCES

(1). Choi J, Kim Y, Oh B. Ultramicroscopy. 2008; 108:1396–1400. [PubMed: 18562117]

(2). Oh B, Lee W, Chun B, Bae Y, Lee W, Choi J. Biosens. Bioelectron. 2005; 20:1847–1850. [PubMed: 15681203]

(3). Lee J, Park H, Jung Y, Kim J, Jung S, Chung B. Anal. Chem. 2007; 79:2680–2687. [PubMed: 17341056]

(4). Persson M, Bülow L, Mosbach K. FEBS Letty. 1990; 270:41–44.

(5). McLean M, Stayton P, Sligar S. Anal. Chem. 1993; 65:2676–2678. [PubMed: 8238945]

(6). Stayton PS, Olinger JM, Jiang M, Bohn PW, Sligar SG. J. Am. Chem. Soc. 1992; 114:9298–9299.

(7). Chi QJ, Zhang J. D.; Andersen JET, Ulstrup J. J. Phys. Chem. B. 2001; 105:4669–4679.

(8). Chen SF, Liu LY, Zhou J, Jiang SY. Langmuir. 2003; 19:2859–2864.

(9). Wang H, Castner DG, Ratner BD, Jiang SY. Langmuir. 2004; 20:1877–1887. [PubMed: 15801458]

(10). Shen Z, Yan H, Zhang Y, Mernaugh R, Zeng X. Anal. Chem. 2008; 80:1910–1917. [PubMed: 18290668]

(11). Wagner MS, Castner DG. Appl. Surf. Sci. 2004; 231:366–376.

(12). Mantus D, Ratner B, Carlson B, Moulder J. Anal. Chem. 1993; 65:1431–1438. [PubMed: 8517550]

(13). Samuel NT, Wagner MS, Dornfield KD, Castner DG. Surf. Sci. Spectra. 2001; 8:163–184.

(14). Tidwell CD, Castner DG, Golledge SL, Ratner BD, Meyer K, Hagenhoff B, Benninghoven A. Surf. Interface Anal. 2001; 31:724–733.

(15). Lhoest J, Wagner M, Tidwell C, Castner DG. J. Biomed. Mater. Res. 2001; 57:432–440. [PubMed: 11523038]

(16). Wagner MS, Castner DG. Appl. Surf. Sci. 2003; 203 PII S0169 -4332(02)00794-8.

(17). Wagner M, Shen M, Horbett T, Castner DG. J. Biomed. Mater. Res. A. 2003; 64:1–11. [PubMed: 12868435]

(18). Xia N, Castner DG. J. Biomed. Mater. Res. A. 2003; 67:179–190. [PubMed: 14517875]

(19). Lhoest J, Detrait E, van den Bosch de Aguilar P, Bertrand P. J. Biomed. Mater. Res. 1998; 41:95–103. [PubMed: 9641629]

(20). Michel R, Pasche S, Textor M, Castner DG. Langmuir. 2005; 21:12327–12332. [PubMed: 16343010]

(21). Xia N, May CJ, McArthur SL, Castner DG. Langmuir. 2002; 18:4090–4097.

(22). Kim Y, Hong M, Kim J, Oh E, Shon H, Moon D, Kim H, Lee T. Anal. Chem. 2007; 79:1377–1385. [PubMed: 17297937]

(23). Shen YR. Nature. 1989; 337:519–525.

(24). Kim J, Somorjai G. J. Am. Chem. Soc. 2003; 125:3150–3158. [PubMed: 12617683]

(25). Chen X, Wang J, Sniadecki J, Even M, Chen Z. Langmuir. 2005; 21:2662–2664. [PubMed: 15779931]

(26). Wang J, Clarke ML, Chen XY, Even MA, Johnson WC, Chen Z. Surf. Sci. 2005; 587:1–11.

(27). Mermut O, Phillips D, York R, McCrea K, Ward R, Somorjai G. J. Am. Chem. Soc. 2006; 128:3598–3607. [PubMed: 16536533]

(28). Stohr, J. NEXAFS Spectroscopy. Vol. Vol. 25. Springer-Verlag; Berlin: 1992.

(29). Okudaira KK, Morikawab E, Hiteb DA, Hasegawac S, Ishiid H, Imamurae M, Shimadae H, Azumaa Y, Meguroa K, Haradaa Y, Saileb V, Sekid K, Uenoc N. J. Electron Spectrosc. Relat. Phenom. 1999; 101–103:389–392.

(30). Petrovykh D, Pérez-Dieste V, Opdahl A, Kimura-Suda H, Sullivan J, Tarlov M, Himpsel F, Whitman L. J. Am. Chem. Soc. 2006; 128:2–3. [PubMed: 16390092]

(31). Lee C, Nguyen P, Grainger D, Gamble L, Castner DG. Anal. Chem. 2007; 79:4390–4400. [PubMed: 17492838]

(32). Lee C, Gong P, Harbers G, Grainger D, Castner DG, Gamble L. Anal. Chem. 2006; 78:3316–3325. [PubMed: 16689532]

(33). Weidner T, Apte J, Gamble L, Castner DG. Langmuir. 2010; 26:3433–3440. [PubMed: 20175575]

(34). Liu X, Jang C, Zheng F, Jürgensen A, Denlinger J, Dickson K, Raines R, Abbott N, Himpsel F. Langmuir. 2006; 22:7719–7725. [PubMed: 16922555]

(35). Clark R, Campbell A, Klumb L, Long C, Stayton PS. Calcif. Tissue Int. 1999; 64:516–21. [PubMed: 10341024]

(36). Wagner M, Tyler B, Castner DG. Anal. Chem. 2002; 74:1824–1835. [PubMed: 11985314]

(37). Wagner MS, Graham DJ, Ratner BD, Castner DG. Surf. Sci. 2004; 570:78–97.

(38). Wagner MS, Graham DJ, Castner DG. Appl. Surf. Sci. 2006; 252:6575–6581.

(39). Moore FG, Becraft KA, Richmond GL. Appl. Spec. 2002; 56:1575–1578.

(40). Bain CD, Davies PB, Ong TH, Ward RN, Brown MA. Langmuir. 1991; 7:1563–1566.

(41). Batson PE. Phys. Rev. B. 1993; 48:2608–2610.

(42). Singh, BR., editor. Infrared Analysis of Peptides and Proteins; American Chemical Society. Washington: 2000.

(43). Ward RN, Davies PB, Bain CD. J. Phys. Chem. 1993; 97:7141–7143.

(44). Cooper G, Gordon M, Tulumello D, Turci C, Kaznatcheev K, Hitchcock AR. J. Electron Spectrosc. Relat. Phenom. 2004; 137:795–799.

(45). Gordon ML, Cooper G, Morin C, Araki T, Turci CC, Kaznatcheev K, Hitchcock AP. J. Phys. Chem. A. 2003; 107:6144–6159.

(46). Zubavichus Y, Zharnikov M, Schaporenko A, Grunze M. J. Electron Spectrosc. Relat. Phenom. 2004; 134:25–33.

(47). Ozawa K, Hasegawa T, Edamoto K, Takahashi K, Kamada M. J. Phys. Chem. B. 2002; 106:9380–9386.

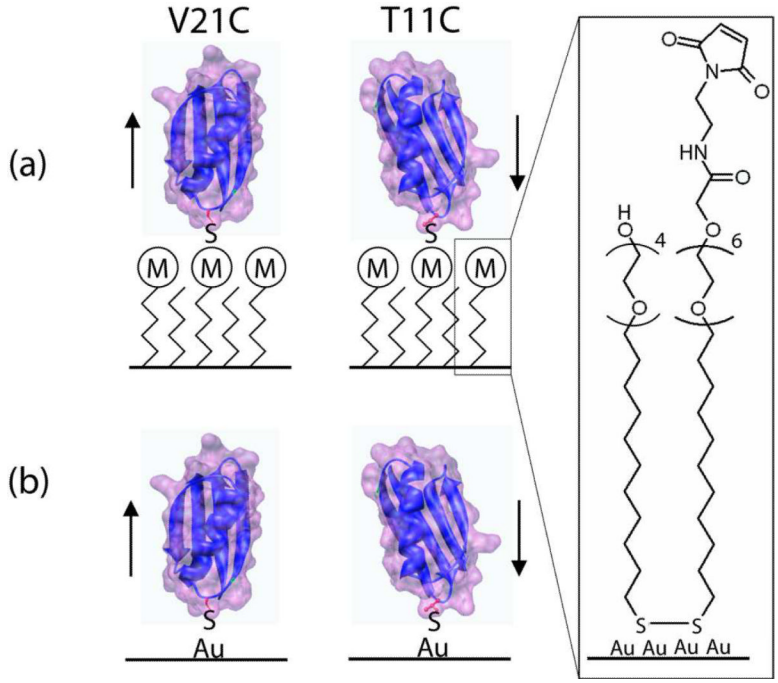

Figure 1.

Protein G1 B1 variants, V21C and T11C, with cysteines introduced at opposite ends of the protein, were immobilized via the cysteine thiol onto (a) maleimide oligo(ethylene glycol)-functionalized gold and (b) bare gold substrates.

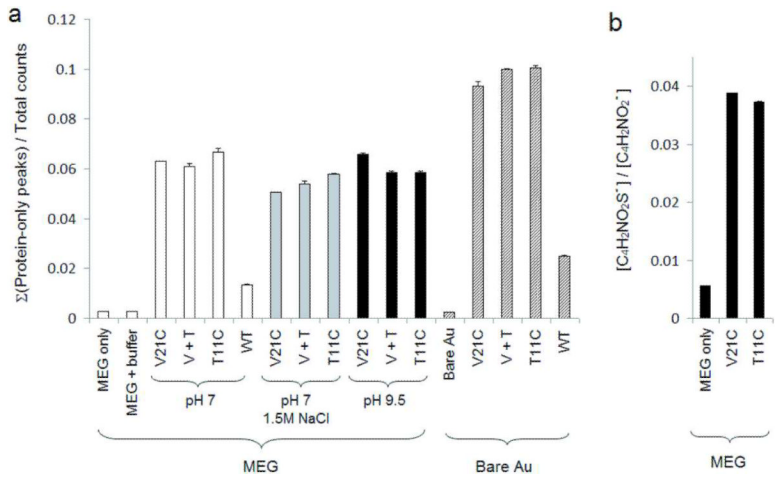

Figure 2.

ToF-SIMS analysis of surface coverage and successful formation of the protein thiol-MEG maleimide bond. (a) The sum of intensities of protein-only peaks was used to assess the relative amount of bound protein. (b) Direct evidence for coupling of the substrate maleimide group with the protein thiol was obtained by comparing the intensity of peaks from reacted and unreacted maleimide rings (reacted: $C_4H_2NO_2S^-$, *m/z* = 127.98; unreacted: $C_4H_2NO_2$, *m/z* = 96.01). The relative intensity of reacted maleimide peaks increased dramatically upon adding the cysteine-containing proteins.

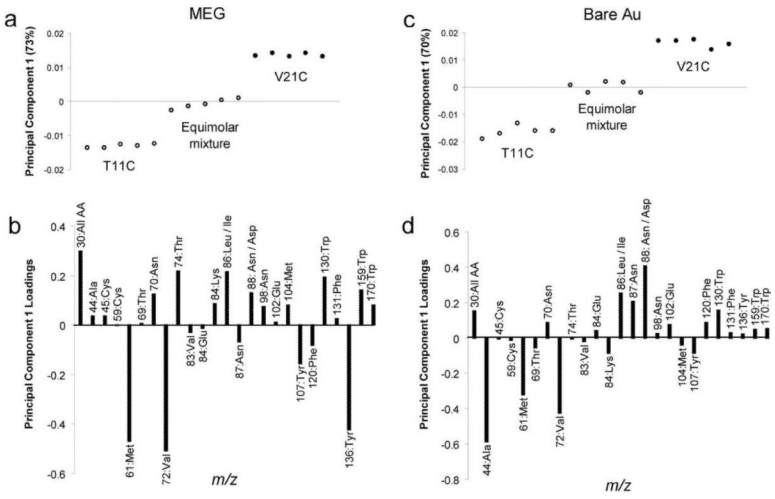

Figure 3.

PCA results for positive ion ToF-SIMS data from immobilization of V21C, T11C and an equimolar mixture of the two onto MEG and bare gold surfaces. Principal component one (PC1) scores (a, c) and loadings (b, d) generated from a set of all protein-only peaks. On both substrates, PC1 accounted for 70–75% of the total variance between spectra, separated V21C and T11C, and gave the equimolar mixture an intermediate score (a, c). Peaks from amino acids located at the C-terminal end of the protein (70 Asn, 86 Leu/Ile, 98 Asn), which are expected to be near the exposed surface of the protein film in V21C, scored positively on both substrates, while peaks derived from the N-terminal end of the protein (107 Tyr, 136 Tyr) scored negatively on MEG (b) and gave mixed results on bare gold (d).

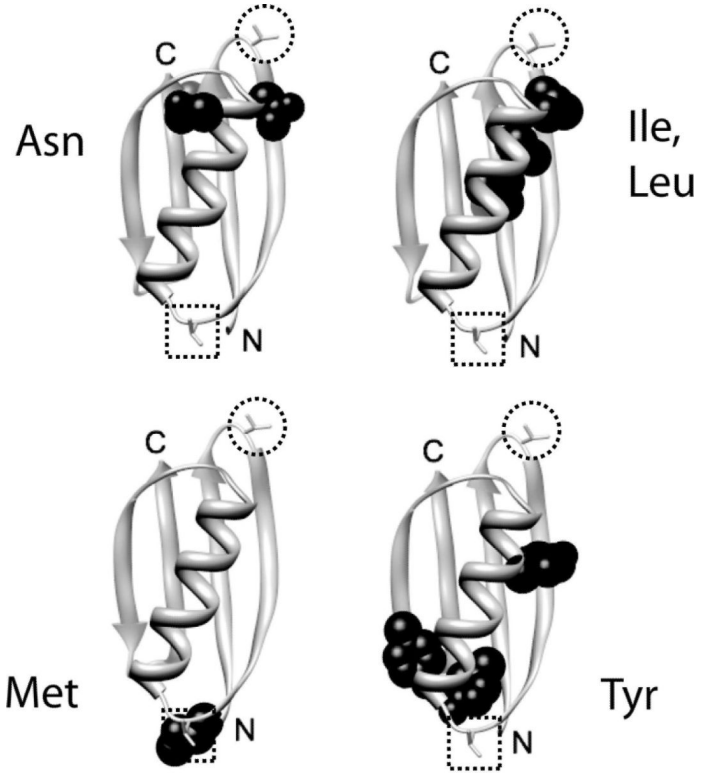

Figure 4.
Amino acids with asymmetric distributions used in ToF-SIMS analysis (shown in black). Isoleucine and leucine produce an identical characteristic positive ion fragment. The locations of the cysteine mutations are marked by dashed circles (for T11C) and squares (V21C).

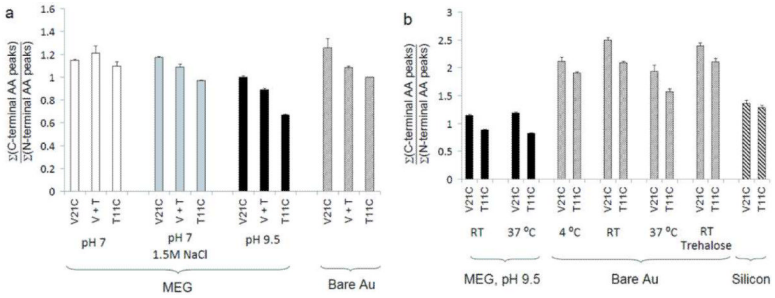

Figure 5.

ToF-SIMS peak intensity ratios. Peak ratios were calculated as the sum of intensities of C-terminal peaks – 70, 87 and 98 from asparagine and 86 from leucine/isoleucine – divided by the sum of intensities of N-terminal peaks – 107 and 136 from tyrosine (rounded mass values). (a) Since the C-terminus is at the opposite end of the protein from the substrate-bound cysteine in V21C, but at the same end in T11C, the trend, V21C > equimolar mixture ("V + T") > T11C, is expected for end-on orientations of the two mutants. On MEG substrates, pH 9.5 buffer gave the best results and was used for subsequent tests of temperature and adding trehalose (b). Nonspecifically immobilized samples – V21C and T11C solutions evaporated onto separate silicon wafer substrates – exhibited similar peak ratios, as expected.

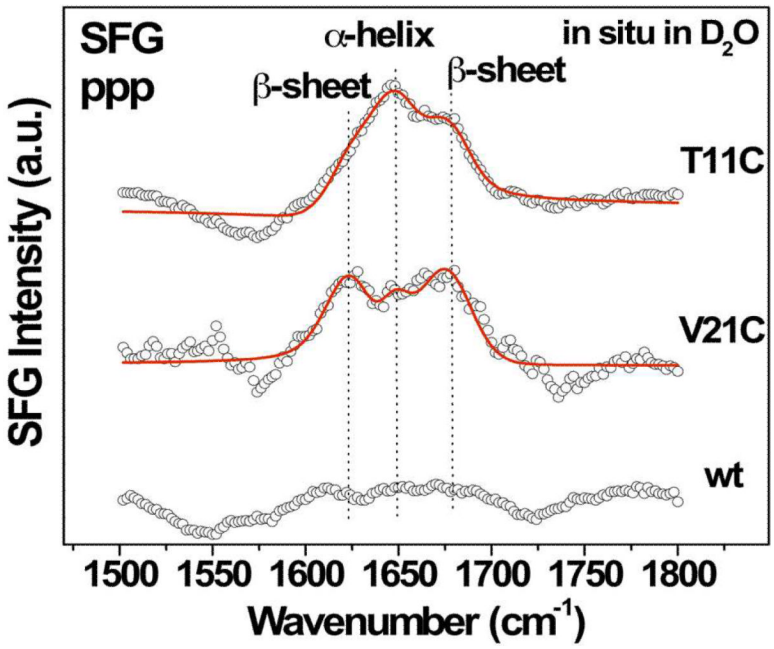

Figure 6.

SFG amide I spectra of Protein G B1 wild type and cysteine mutants on MEG *in situ*. Protein was immobilized in 50 mM sodium carbonate at pH 9.5; buffer was replaced with deuterated water for SFG analysis. The amide N-H peak near 1645 cm^{-1} – prominent in the T11C sample – is characteristic of ordered α-helices, with a phase that indicates an upright helix orientation (orthogonal to the substrate). The amide N-H peaks near 1624 cm^{-1} and 1675 cm^{-1} are characteristic of ordered β-sheet structures.

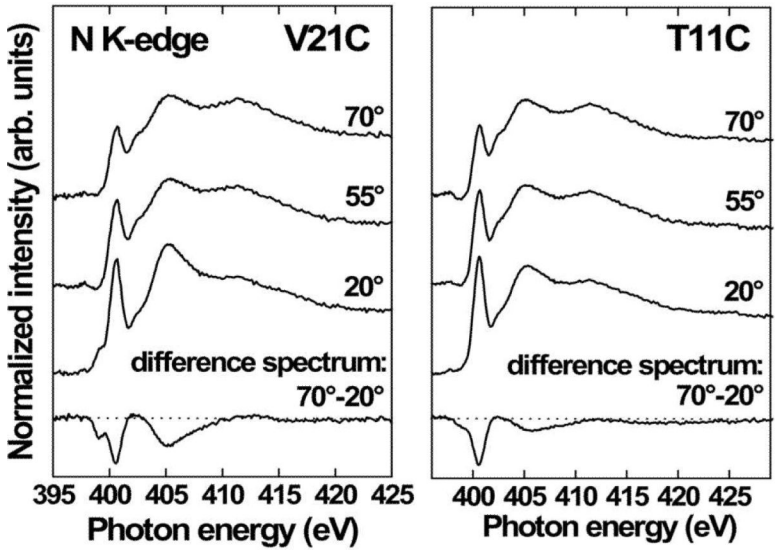

Figure 7.

NEXAFS spectra of the nitrogen *K*-edge for V21C and T11C immobilized onto MEG at pH 9.5, acquired at angles of 20°, 55° and 70°, along with the difference between 20° and 70° spectra. The prominent negative dichroism observed at 400.6 eV in the difference spectra is characteristic of ordered amide bonds and is attributed to anisotropic β-sheet peptide bonds in Protein G B1.

Table 1

Ion fragments used in ToF-SIMS Analysis

Source	Mass	Formula
Positive ion fragments derived from protein only		
All amino acids	30.04	CH_4N^+
Alanine (Ala)	44.05	$C_2H_6N^+$
Cysteine (Cys)	44.98	CHS^+
Cysteine (Cys)	59.00	$C_2H_3S^+$
Threonine (Thr)	69.04	$C_4H_5O^+$
Asparagine (Asn)	70.03	$C_3H_4NO^+$
Valine (Val)	72.08	$C_4H_{10}N^+$
Threonine (Thr)	74.07	$C_3H_8NO^+$
Cysteine (Cys)	76.03	$C_2H_6NS^+$
Glutamic Acid (Glu)	83.05	$C_5H_7O^+$
Lysine (Lys)	84.05	$C_4H_6NO^+$
Leucine (Leu)	84.09	$C_5H_{10}N^+$
Leucine, Isoleucine (Ile)	86.10	$C_5H_{12}N^+$
Asparagine (Asn)	87.06	$C_3H_7N_2O^+$
Asparagine (Asn)	88.04	$C_3H_6NO_2^+$
Asparagine (Asn)	98.03	$C_4H_4NO_2^+$
Methionine (Met)	102.06	$C_4H_8NO_2^+$
Tyrosine (Tyr)	107.05	$C_7H_7O^+$
Phenylalanine (Phe)	120.09	$C_8H_{10}N^+$
Tryptophan (Trp)	130.07	$C_9H_8N^+$
Phenylalanine (Phe)	131.06	$C_9H_7O^+$
Tyrosine (Tyr)	136.09	$C_8H_{10}NO^+$
Tryptophan (Trp)	159.10	$C_{10}H_{11}N_2^+$
Tryptophan (Trp)	170.07	$C_{11}H_8NO^+$
Negative ion fragments used to confirm maleimide-thiol bond		
Unreacted maleimide ring	96.01	$C_4H_2NO_2^-$
Reacted maleimide ring	127.98	$C_4H_2NO_2S^-$
Positive ion fragments from asymmetric amino acids		
Asparagine (Asn)	70.03	$C_3H_4NO^+$
Leucine, Isoleucine (Ile)	86.10	$C_5H_{12}N^+$
Asparagine (Asn)	87.06	$C_3H_7N_2O^+$
Asparagine (Asn)	98.03	$C_4H_4NO_2^+$
Tyrosine (Tyr)	107.05	$C_7H_7O^+$
Tyrosine (Tyr)	136.09	$C_8H_{10}NO^+$

Table 2

XPS Elemental Composition of MEG, Bare Gold and Protein-coated Substrates (atomic %)[a]

	Organic overlayer (Au *4f* signal excluded)[b]				
Sample	**C *1s***	**O *1s***	**N *1s***	**S *2p***	**Au *4f*[c]**
MEG, no protein	72.9 ± 0.7	21.7 ± 0.2	4.1 ± 0.2	1.2 ± 0.1	25.0 ± 0.5
MEG, no protein, exposed to buffer	71.7 ± 0.8	22.7 ± 0.4	4.1 ± 0.1	1.5 ± 0.2	26.0 ± 0.5
MEG, V21C, pH 7.0	69.2 ± 0.5	20.4 ± 0.5	9.1 ± 0.1	1.3 ± 0.1	16.2 ± 0.1
MEG, T11C, pH 7.0	69.5 ± 0.4	19.5 ± 0.0	9.7 ± 0.4	1.3 ± 0.2	14.5 ± 0.4
MEG, V21C + T11C, pH 7.0	68.3 ± 0.8	20.8 ± 0.3	9.7 ± 0.4	1.2 ± 0.0	15.3 ± 0.2
MEG, wild type, pH 7.0[d]	68.6 ± 0.8	22.1 ± 0.4	7.5 ± 0.4	1.8 ± 0.0	17.4 ± 1.2
MEG, V21C, pH 7.0, 1.5 M NaCl	69.9 ± 0.9	21.5 ± 0.4	7.5 ± 0.5	1.1 ± 0.1	18.2 ± 0.3
MEG, T11C, pH 7.0, 1.5 M NaCl	69.4 ± 0.4	21.3 ± 0.2	8.1 ± 0.2	1.2 ± 0.1	16.9 ± 0.4
MEG, V21C + T11C, pH 7.0, 1.5 M NaCl	69.5 ± 0.2	21.3 ± 0.1	8.2 ± 0.1	1.1 ± 0.0	16.7 ± 0.1
MEG, V21C, pH 9.5	69.8 ± 2.8	20.3 ± 0.4	8.7 ± 0.1	1.2 ± 0.1	16.5 ± 0.2
MEG, T11C, pH 9.5	70.1 ± 0.9	19.9 ± 0.2	8.7 ± 0.1	1.3 ± 0.1	17.1 ± 0.4
MEG, V21C, pH 9.5, blocked thiol[d]	70.1 ± 0.4	21.2 ± 0.1	6.2 ± 0.3	2.4 ± 0.2	22.4 ± 0.3
MEG, T11C, pH 9.5, blocked thiol[d]	70.4 ± 0.8	21.3 ± 0.4	6.3 ± 0.1	2.1 ± 0.2	20.0 ± 0.4
Bare gold, no protein	93.9 ± 3.1	6.1 ± 0.2	0	0	64.2 ± 2.9
Bare gold, V21C	66.8 ± 1.4	18.4 ± 0.2	14.3 ± 0.3	0.6 ± 0.1	28.8 ± 1.2
Bare gold, T11C	66.7 ± 0.7	17.9 ± 0.2	14.7 ± 0.3	0.7 ± 0.0	30.0 ± 0.2
Bare gold, V21C + T11C	66.3 ± 1.0	18.5 ± 0.1	14.5 ± 0.4	0.7 ± 0.0	30.3 ± 0.6
Bare gold, wild type[d]	64.9 ± 0.1	20.4 ± 0.2	13.1 ± 0.0	1.6 ± 0.1	36.9 ± 0.2

[a] XPS data for the first set of samples is shown here. Data for the second set is included in Table S1 in the Supporting Information.

[b] The gold substrate Au *4f* signal was excluded to show the elemental composition of the organic overlayer. The remaining signals were re-normalized to 100%.

[c] The Au *4f* signal before re-normalization. Attenuation of this signal was used as an indicator of coverage of the gold substrate.

[d] The N *1s* signal was lower and the substrate Au *4f* signal higher for cysteine-free wild type and mutants with thiols blocked than for unblocked mutants, indicating lower surface coverage in the absence of a reactive thiol.