Investigating the Conformation of Surface-Adsorbed Proteins with Standing-Wave X-Ray Fluorescence

Supporting Information

Ernesto Scoppola,[†] Georgi G. Gochev,[‡,†] Jakub Drnec,[¶] Linus Pithan,[¶] Dmitri Novikov,[§] and Emanuel Schneck[*,‖,†]

[†]Biomaterials Department, Max Planck Institute of Colloids and Interfaces, Potsdam, Germany

[‡]Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, Krakow, Poland

[¶]European Synchrotron Radiation Facility (ESRF), 38000 Grenoble, France

[§]Deutsches Elektronen-Synchrotron (DESY), 22607 Hamburg, Germany

[‖]Physics Department, Technische Universität Darmstadt, Germany

E-mail: emanuel.schneck@pkm.tu-darmstadt.de

X-ray reflectivity of bare multilayer substrates in air

Measurements

SWXF experiments were carried out at the beamline ID03 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France) with an incident beam energy of 7.0 keV. The reflectivity R, i.e., the intensity of the reflected beam relative to the intensity of the incident

beam, was recorded as a function of the scattering vector component perpendicular to the interface,

$$Q_z = \frac{4\pi}{\lambda} \sin(\theta_i), \tag{S1}$$

where θ_i is the incident angle. Measurements were carried out using an angle range of $0° \leq \theta_i \leq 4°$ ($Q_z \leq 0.5$ Å^{-1}, see Eq. S1). The relative resolution, $\delta Q_z / Q_z$, was assumed as 1% during the data reduction procedure, where resolution was taken into account by convolution with a Gaussian function of suitable width.

Data Modeling

Reflectometry curves were analyzed with a fitting procedure based on parameterized volume fraction profiles of all material components, namely sapphire (sa), aluminium oxide (Al$_2$O$_3$) and nickel-carbon (NiC) multilayer. We consider here a sapphire substrate coated with a NiC multilayer and a subsequent Al$_2$O$_3$ layer. In the following, we set $z = 0$ to be the Sa/NiC interface to describe the model, noting that in the main text $z = 0$ is re-defined to coincide with the center-of-mass position of the Si reference layer.

Model Distribution Function

At first, the sapphire volume fraction profile is described with an error function,

$$\Phi_{\mathrm{sa}}(z) = \frac{1}{2} \cdot \left[1 - \mathrm{erf}\left(\frac{-z}{\sqrt{2}\sigma_{\mathrm{Sa}}} \right) \right], \tag{S2}$$

where z denotes the distance to the sapphire/nickel interface and σ_{Sa} the adjustable interfacial roughness. The NiC multilayer is described as an N-fold repetition ($N = 30$) of two slab functions defined as the difference between two error functions:

$$\Phi_{\mathrm{Ni}}(z) = \frac{1}{2} \sum_{i=1}^{N} \left[\mathrm{erf}\left(\frac{z - z_i^{\mathrm{Ni}} - d^{\mathrm{Ni}}/2}{\sqrt{2}\sigma_i^{\mathrm{A}}} \right) - \mathrm{erf}\left(\frac{z - z_i^{\mathrm{Ni}} + d^{\mathrm{Ni}}/2}{\sqrt{2}\sigma_i^{\mathrm{B}}} \right) \right] \tag{S3}$$

$$\Phi_{\mathrm{C}}(z) = \frac{1}{2} \sum_{i=1}^{N} \left[\mathrm{erf}\left(\frac{z - z_i^{\mathrm{C}} - d^{\mathrm{C}}/2}{\sqrt{2}\sigma_i^{\mathrm{B}}} \right) - \mathrm{erf}\left(\frac{z - z_i^{\mathrm{C}} + d^{\mathrm{C}}/2}{\sqrt{2}\sigma_{i+1}^{\mathrm{A}}} \right) \right] \tag{S4}$$

where the index i stands for the i-th repetition in the multilayer. Here, $z_i^{\mathrm{Ni}} = (d^{\mathrm{C}} + d^{\mathrm{Ni}}) \cdot i$ and $z_i^{\mathrm{C}} = (d^{\mathrm{C}} + d^{\mathrm{Ni}}) \cdot i + d^{\mathrm{Ni}}$, where d^{C} and d^{Ni} are the thicknesses of the carbon and nickel slabs, respectively. The sapphire/nickel and carbon/nickel interfacial roughnesses are represented by the parameters σ_i^{A} and σ_i^{B} respectively. Moreover, for $2 \leq i \leq N-1$, these roughnesses are constant such that $\sigma_i^{\mathrm{A}} = \sigma^{\mathrm{A}} = \sigma_{\mathrm{Sa}}$ and $\sigma_i^{\mathrm{B}} = \sigma^{\mathrm{B}}$. At last, the aluminum oxide top layer can be described by mean of a slab function,

$$\Phi_{\mathrm{Al2O3}}(z) = \frac{1}{2} \left[\mathrm{erf}\left(\frac{z - z^{\mathrm{Al2O3}} - d^{\mathrm{Al2O3}}/2}{\sqrt{2}\sigma^{\mathrm{Al2O3}}} \right) - \mathrm{erf}\left(\frac{z - z^{\mathrm{Al2O3}} + d^{\mathrm{Al2O3}}/2}{\sqrt{2}\sigma^{\mathrm{air}}} \right) \right], \tag{S5}$$

where z^{Al2O3}, d^{Al2O3} represent its central position and width. The Al2O3/carbon and Al2O3/air interfacial roughnesses are represented by the parameters σ^{Al2O3} and σ^{air} respectively. For ensuring the continuity of the volume fraction profiles $\sigma^{\mathrm{Al2O3}} = \sigma^{\mathrm{A}}$.

With all volume fraction profiles at hand, the corresponding x-ray SLD profile $\rho(z)$ is then calculated as

$$\rho(z) = \rho_{\mathrm{Sa}} \cdot \Phi_{\mathrm{Sa}}(z) + \rho_{\mathrm{C}} \cdot \Phi_{\mathrm{C}}(z) + \rho_{\mathrm{Ni}} \cdot \Phi_{\mathrm{Ni}}(z) + \rho_{\mathrm{Al2O3}} \cdot \Phi_{\mathrm{Al2O3}}(z), \tag{S6}$$

where ρ_{Sa}, ρ_{C}, ρ_{Ni} and ρ_{Al2O3} are the x-ray SLDs of sapphire, carbon, nickel, and Al2O3, respectively, calculated as[1]

$$\rho = \frac{1}{v_m} \sum_{k} N_k \varrho_k \tag{S7}$$

where the index k identifies the chemical element, N_k its number per volume v, and $\varrho_k^{(x)}$ its complex x-ray scattering length.[1] Both real and imaginary part of the nickel, carbon and Al2O3 SLD were fitted during the modelling process by varying v while keeping conctant the corresponding ϱ_k value calculated with the NIST SLD Calculator.[2] Best fit ρ values are summarized in Table S1 and compared with the theoretical values for idealized bulk mate-

rials.

Theoretical reflectivity curves are calculated by discretizing the SLD profiles into 1–Å–thick layers of constant SLD and subsequent application of Parratt's recursive procedure.[3] In fits to experimental data the best-matching model parameters are obtained by minimization of the total squared difference χ^2 between all theoretical and experimental values. Sampling of the physically-plausible parameter space is performed with a previously validated in-house python routine[4–7] based on the Metropolis Monte Carlo algorithm[8] with random variation of parameters.

In Figure S1 the best reflectivity model is shown together with the experimental data. Finally, the resulting SLD profile was then used for computing the corresponding angle-dependent standing wave intensity distribution as was done earlier[9,10] with a web-based tool.[11]

Table S1: Best-matching values of the model parameters reported in Eq. S2, S3, S4, and S5.

Parameter	Best-matching value [Å]
σ_{Sa}	7.0
d^{Ni}	34.5
d^{C}	25.3
σ^{A}	7.0
σ^{B}	5.6
d^{Al2O3}	49.5
σ^{Al2O3}	7.0
σ^{Air}	2.0
SLD	Best-matching value $[10^{-6}\text{Å}^{-2}]$
ρ_{Sa}	28.25-0.425i
$\rho_{Ni}{}^{a}$	65.68-1.670i
$\rho_{C}{}^{b}$	14.94-0.089i
ρ_{Al2O3}	18.61-0.283i

[a] Idealized bulk value $(67.99\text{-}1.745i)\cdot10^{-6}\text{Å}^{-2}$;
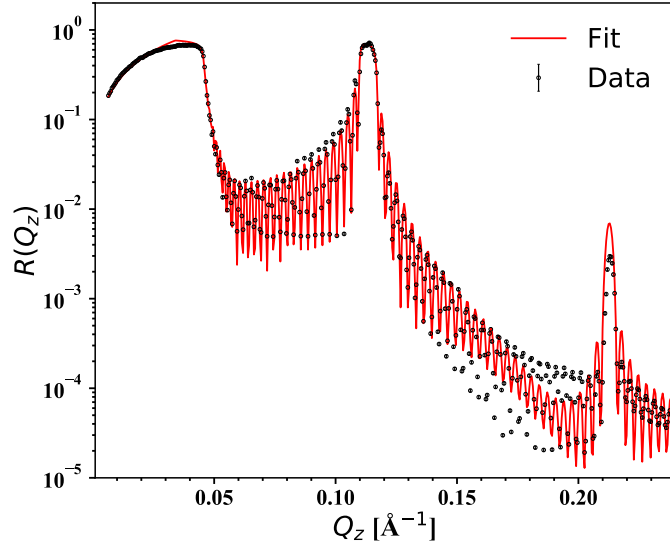[b] Idealized bulk value $(16.17\text{-}0.035i)\cdot10^{-6}\text{Å}^{-2}$.

Figure S1: Best fit reflectivity profile (red line) in comparison with experimental data (black circles). The model takes into account sample over-illumination for $Q_z \leq 0.045$ Å^{-1}.

Si fluorescence under water and in air

Fig. S2 shows the angle-dependent Si fluorescence intensity of OTS-functionalized Al oxide surface under water (left) and in air (right). The fit (solid line) yields center of mass positions of the Si distribution with respect to the surface that are consistent within an uncertainty of $\pm$ 2 Å.
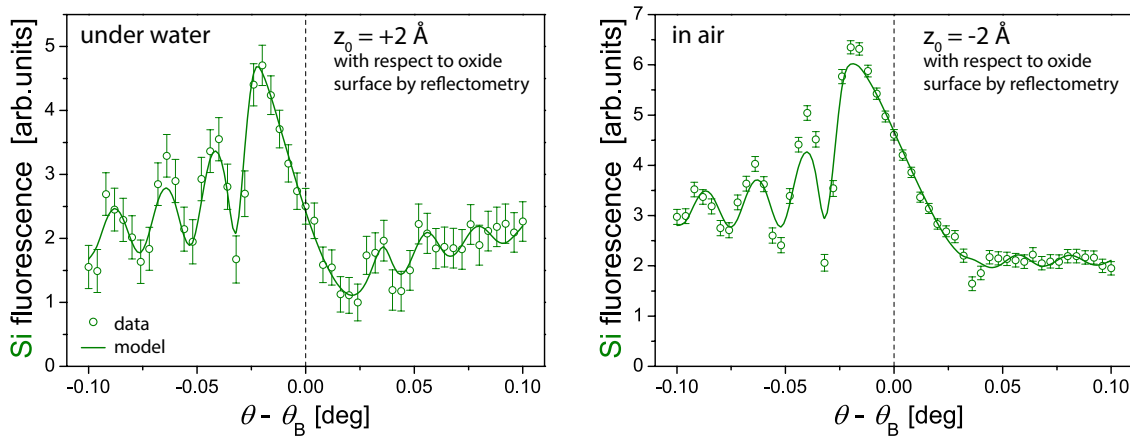

Figure S2: Angle-dependent Si fluorescence intensity of OTS-functionalized Al oxide surface under water (left) and in air (right). Te solid lines indicate the theoretical intensity according to the best-matching model parameters.

Primary structure of β-casein

Fig. S3 shows the primary stucture of β-casein based on the Expasy UniProtKB database for -casein A2-5P (entry: P02666).


```
N-terminus          Domain I                                                    Domain II
1                                               40–41
RELEELNVPGEIVESLSSSEESITRINKKIEKFQSEEQQQ–TEDELQDKIHPFAQTQSLVYPFPGPIPNSLPQNIPPLTQT
                                                        135–136      Domain III
PVVVPPFLQPEVMGVSKVKEAMAPKHKEMPFPKYPVEPFTESQSLTLTDVENLHL–PLPLLQSWMHQPHQPLPPT
                                                    209
VMFPPQSVLSLSQSKVLPVPQKAVPYPQRDMPIQAFLLYQEPVLGPVRGPFPIIV   C-terminus
```
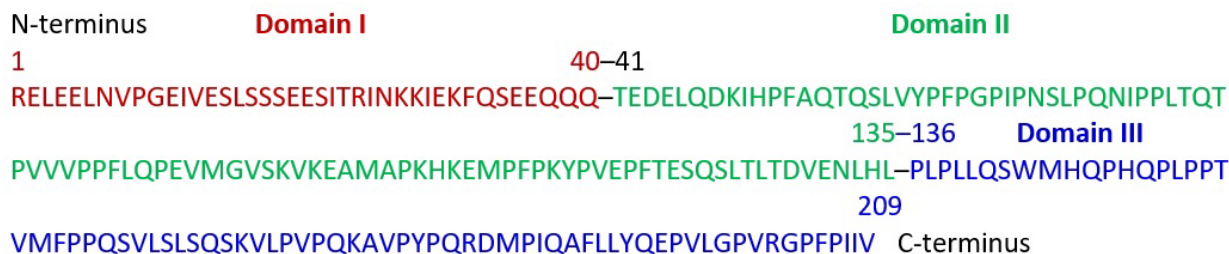
Figure S3: Primary stucture of β-casein.

Hydropathic score vs. amino acid sequence

Fig. S4 shows the hydropathic score vs. amino acid sequence as generated by the Expasy ProtScale tool[12] (database by Kyte and Doolittle[13]), where the localizations of the phosphorylated serine units and the S-containing methionine units are also indicated. The P atoms are localized solely in the N-terminal domain I, while the S atoms are equally distributed over domains II and III.

P and S fluorescence from dry β-casein layers on bare and OTS-functionalized Al oxide

Fig. S5 shows angle-dependent P and S fluorescence intensities from dried β-casein adsorbed onto bare and OTS-functionalized Al oxide surfaces.
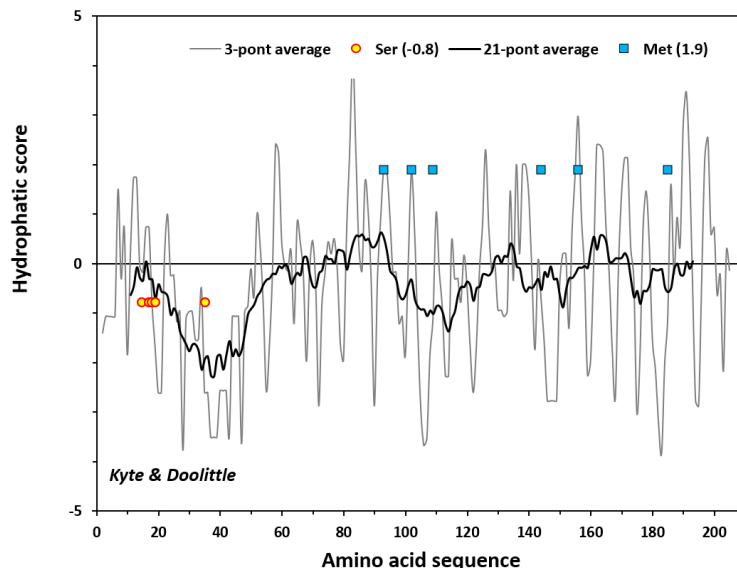
Figure S4: Hydropathic score vs. amino acid sequence. Localizations of the phosphorylated serine units and the S-containing methionine units are indicated with circles and squares, respectively.

Fluorescence background subtraction

In background subtraction approaches, avoiding background under- or over-subtraction is a key aspect. Although the combined thickness h of the water layer and the cover foil are comparatively reproducible ($h \approx 2$ μm), already minor h–variations between the sample and the (protein-free) reference system can significantly alter the intensity scale due to attenuation effects, which exhibit non-linear thickness-dependence. Because of the low incident angle ($\theta \approx 0.9 \pm 0.1$ deg), the resulting long beam path ($s = h/\sin\theta \approx 0.13$ mm) through the layers, and the comparatively low incident beam energy (7 keV), attenuation of the incident beam is more critical than the attenuation of the fluorescence radiation, which reaches the detector in the perpendicular direction with a layer-internal beam path of only ≈ 2 μm. However, the situation is complicated by the energy dependence of the attenuation of the fluorescence radiation. In our subtraction procedure, we therefore introduce a pre-factor, α, that allows correcting for the ensuing slightly different intensity scales of the fluorescence spectra of sample and reference measurements.
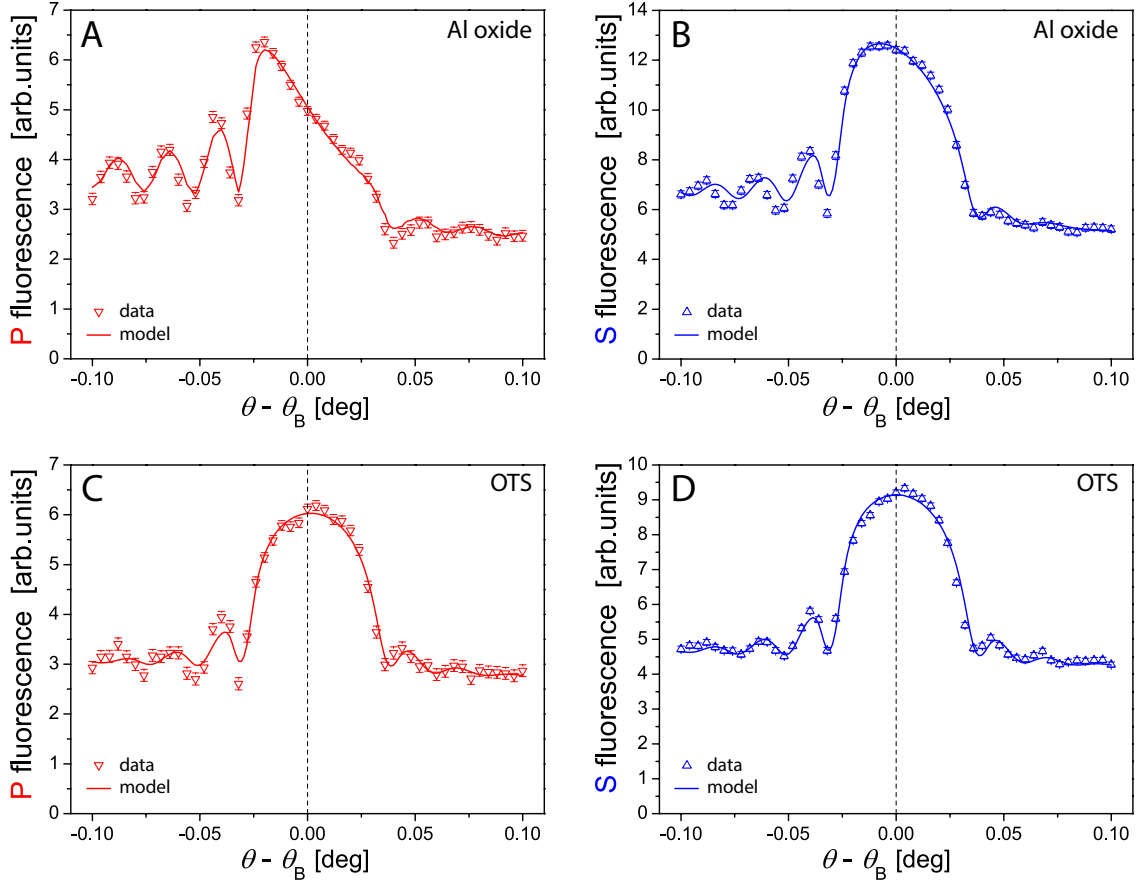
Figure S5: (A and B) Angle-dependent fluorescence intensities emitted by the P (panel A) and S (panel B) atoms, respectively, of dried β-casein adsorbed onto a hydrophilic bare Al oxide surface. (C and D) Angle-dependent fluorescence intensities emitted by the P (panel C) and S (panel D) atoms, respectively, of dried β-casein adsorbed onto a hydrophobic OTS-functionalized surface. The solid lines indicate the simulated fluorescence intensities according to the best-matching values of the model parameters.

$$I(\theta, E) = I_{\mathrm{samp}}(\theta, E) - \alpha I_{\mathrm{ref}}(\theta, E) \tag{S8}$$

, where I_{samp} and I_{ref} are the fluorescence intensities of sample and reference, respectively, and I is the background-corrected intensity. Due to the energy dependence of the attenuation, a constant value of α would be insufficient to perform an accurate background correction along the whole spectrum. Conversely, however, this means that in a narrow energy range around an elemental line of interest a meaningful correction can be performed with a suitable constant choice of α. Since the proteins only contribute P and S lines to

the spectrum, while all other contributions to the spectrum are nominally identical, α has to be chosen such that the baseline on both sides of the P and S lines vanishes. With this criterion, all unknown contributions to the background are eliminated, and we are essentially left with the fluorescence emission of only the proteins. Analogous reasoning applies to the Si fluorescence when a bare Al oxide surface is used as reference for a surface hydrophobized with Si-containing OTS. The subtraction obviously has to be performed angle-wise, so that angle-dependent background intensity modulations are fully accounted for. Fig. S6 exemplarily shows the sample spectral intensity I_{samp} (panel A), the protein-free reference spectral intensity I_{ref} (panel B), and the resulting background-corrected spectral intensity I (panel C) for two incident angles in an energy interval accommodating the P and S lines of the proteins. The plots also show the angle-averaged intensities, which have high statistics. It is seen that both I_{samp} and I_{ref} exhibit pronounced angle dependence regarding their shape and absolute intensity scale. However, after background correction according to Eq. S8 only the P and S peaks are left on a vanishing baseline, with the characteristic angle dependence of their intensities. The peak shapes are well captured with Gaussian fits without baseline, see solid lines. However, the peak intensities were determined through a simple numerical integration of the actual data points within an interval of $\pm\,50$ eV around the central emission line. Note that this procedure is equivalent to Gaussian peak fitting and subsequent analytical integration.

Estimation of parameter uncertainties

In general, both statistical and systematic uncertainties have to be considered for the estimates. Following the book of Bevington et al.,[14] the statistical parameter uncertainty can be derived from the parameter dependence of the χ^2-deviation between the model and the actual data points in the angle-dependent elemental fluorescence intensity plots (Figs. 3-5 in the main text). To this end, a quantitative estimate requires knowledge of the statis-
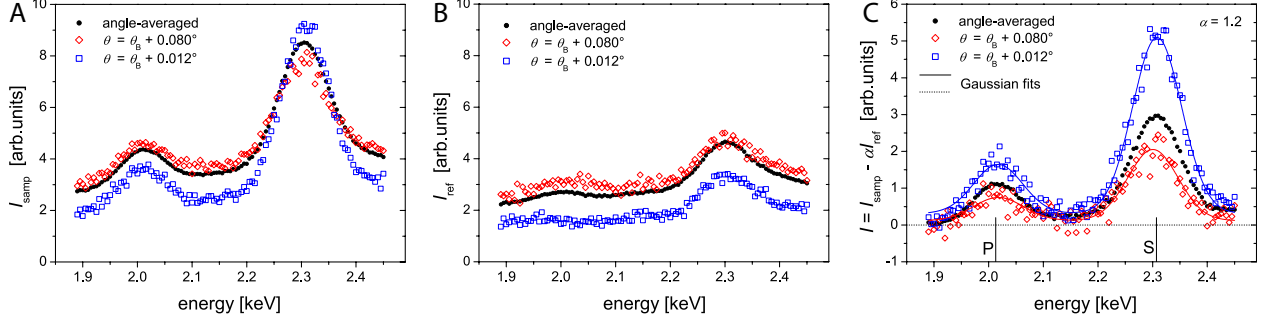
Figure S6: (A) Sample spectra, (B) protein-free reference spectra, and (C) the resulting background-corrected spectra for two incident angles ($\theta = \theta_B + 0.012$ deg, blue symbols and $\theta = \theta_B + 0.080$ deg, red symbols) in an energy interval accommodating the P and S lines of the proteins. Solid lines indicate Gaussian fits without baseline. Black dots indicate the angle-averaged spectra.

tical uncertainties ΔI of the individual intensity data points, which according to Poisson statistics follow from the total count numbers N_{samp} and N_{ref} of the sample and reference measurements, respectively, and from Gaussian error propagation

$$\Delta I = \sqrt{\left(\frac{\partial I}{\partial I_{samp}}\right)^2 (\Delta I_{samp})^2 + \left(\frac{\partial I}{\partial I_{ref}}\right)^2 (\Delta I_{ref})^2} = \sqrt{N_{samp} + \alpha^2 N_{ref}}. \tag{S9}$$

Importantly, the dependence of χ^2 on the parameter of interest has to be probed while all other model parameters are freely varied. This procedure leads to purely statistical uncertainties (two-sigma, 95% confidence interval) of only $\lesssim 1$ Å for both the position z_0 and the width w of the elemental distributions.

However, as pointed out in the main text, also significant systematic uncertainties are associated with both parameters, as they are sensitive to under- or over-subtraction of the reference background of the fluorescence spectra. If we assume that the pre-factor α in Eq. S8 has an uncertainty of ± 0.05, then a systematic uncertainty in w of ≈ 3 Å is obtained. With regard to z_0 uncertainties between 2 and 8 Å are obtained in this way.

By combining statistical and systematic uncertainties, Δ_{stat} and Δ_{sys}, respectively, in an independent fashion, $\Delta^2 = \Delta_{stat}^2 + \Delta_{sys}^2$, we obtain estimated total uncertainties of $\approx 3\text{-}8$ Å for z_0 and of ≈ 4 Å for w. These are listed in Table 1 in the manuscript.

Comparison of background subtraction and full-spectrum-fitting approaches

Background subtraction (BS) and full-spectrum-fitting (FSF) approaches both have their advantages and caveats: In FSF,[9,15–17] the complete description of the intensity often relies on certain assumptions regarding the elemental origin of fluorescence lines and on the spatial distribution of the respective elements. BS,[10] on the other hand, comes with the risk of background under- or over-subtraction. Importantly, FSF and BS lead to consistent results when performed carefully. This is illustrated in Fig. S7, where in order to optimize the analysis procedures a comparison of both approaches was made for the sulfur fluorescence from a protein sample measured at the beamline DIFFABS at synchrotron Soleil in 2014.
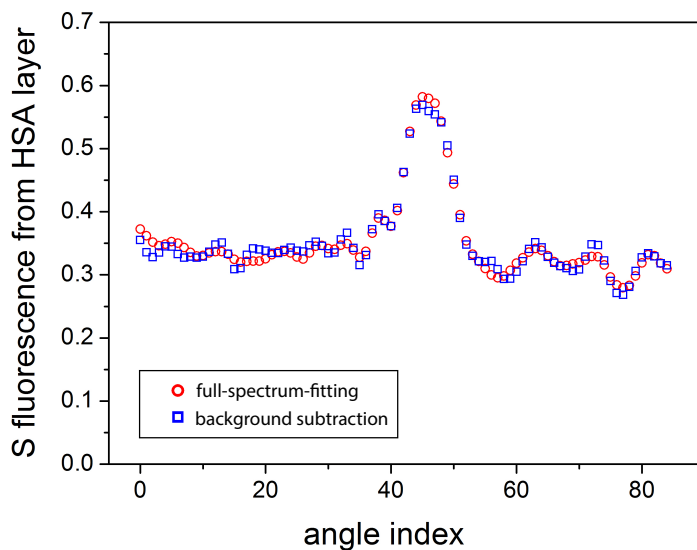

Figure S7: Comparison of angle-dependent protein S fluorescence obtained with the BS and FSF approaches. Sample: Human serum albumin (HSA) adsorbed to the surface of an Al-oxide-terminated Ni/Al multilayer substrate.

Influence of a variation of the parameters z_0 and w on the calculated curves

Figs. S8 and S9 illustrate the influence of a variation of the parameters z_0 and w, respectively, on the calculated curve of the angle-dependent fluorescence intensity. Here, the experimental data points correspond to the P fluorescence by β-casein adsorbed to a bare Al oxide surface under water. It is seen that a variation of the parameters z_0 or w by ± 5 Å results in a substantial deviation from the experimental data points.
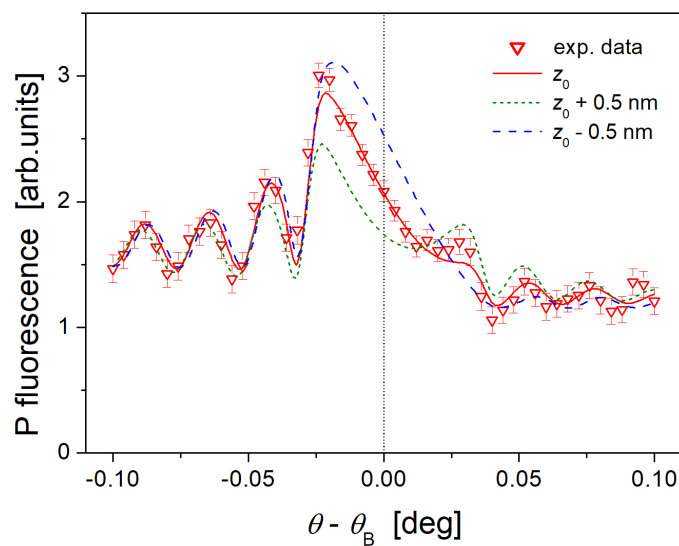

Figure S8: Influence of a variation of the parameter z_0 on the calculated curves. Experimental data: P fluorescence by β-casein adsorbed to a bare Al oxide surface under water.
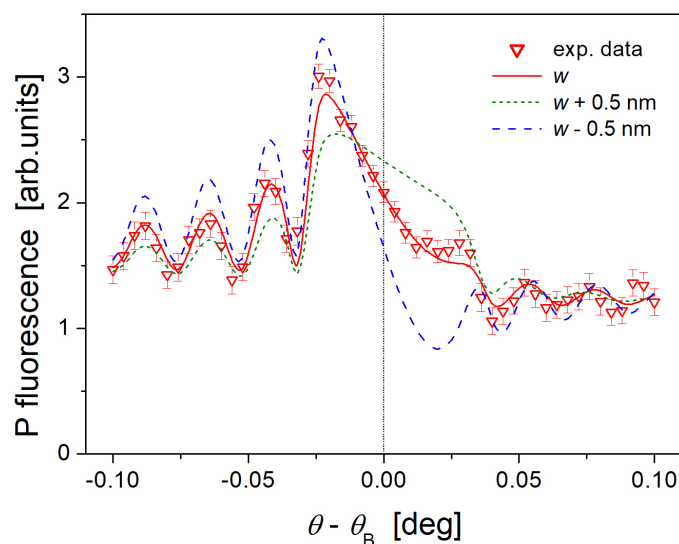
Figure S9: Influence of a variation of the parameter w on the calculated curves. Experimental data: P fluorescence by β-casein adsorbed to a bare Al oxide surface under water.

References

(1) Hayter, J. B.; Mook, H. Discrete thin-film multilayer design for X-ray and neutron supermirrors. *Journal of Applied Crystallography* **1989**, *22*, 35–41.

(2) Kienzle, P. Neutron Activation and Scattering Calculator. https://www.ncnr.nist.gov/resources/activation/.

(3) Parratt, L. G. Surface studies of solids by total reflection of X-rays. *Physical review* **1954**, *95*, 359.

(4) Lolicato, F.; Joly, L.; Martinez-Seara, H.; Fragneto, G.; Scoppola, E.; Baldelli Bombelli, F.; Vattulainen, I.; Akola, J.; Maccarini, M. Nanoparticle–Membrane Interactions: The Role of Temperature and Lipid Charge on Intake/Uptake of Cationic Gold Nanoparticles into Lipid Bilayers (Small 23/2019). *Small* **2019**, *15*, 1970124.

(5) de Santis, A.; Vitiello, G.; Appavou, M.-S.; Scoppola, E.; Fragneto, G.; Barnsley, L. C.;

Clifton, L. A.; Ottaviani, M. F.; Paduano, L.; Krauss, I. R., et al. Not just a flu-idifying effect: omega-3 phospholipids induce formation of non-lamellar structures in biomembranes. *Soft matter* **2020**, *16*, 10425–10438.

(6) Scoppola, E.; Micciulla, S.; Kuhrts, L.; Maestro, A.; Campbell, R. A.; Konovalov, O. V.; Fragneto, G.; Schneck, E. Reflectometry reveals accumulation of surfactant impurities at bare oil/water interfaces. *Molecules* **2019**, *24*, 4113.

(7) Gochev, G. G.; Scoppola, E.; Campbell, R. A.; Noskov, B. A.; Miller, R.; Schneck, E. β-Lactoglobulin Adsorption Layers at the Water/Air Surface: 3. Neutron Reflectometry Study on the Effect of pH. *The Journal of Physical Chemistry B* **2019**, *123*, 10877–10889.

(8) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of state calculations by fast computing machines. *The journal of chemical physics* **1953**, *21*, 1087–1092.

(9) Schneck, E.; Scoppola, E.; Drnec, J.; Mocuta, C.; Felici, R.; Novikov, D.; Fragneto, G.; Daillant, J. Atom-scale depth localization of biologically important chemical elements in molecular layers. *Proceedings of the National Academy of Sciences* **2016**, *113*, 9521–9526.

(10) Schneck, E.; Rodriguez-Loureiro, I.; Bertinetti, L.; Marin, E.; Novikov, D.; Konovalov, O.; Gochev, G. Element-specific density profiles in interacting biomembrane models. *Journal of Physics D: Applied Physics* **2017**, *50*, 104001.

(11) Stepanov, S.; Forrest, R. Fitting dynamical X-ray diffraction data over the World Wide Web. *Journal of Applied Crystallography* **2008**, *41*, 958–962.

(12) https://web.expasy.org/protscale/, accessed Oct 10, 2020.

(13) Kyte, J.; Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology* **1982**, *157*, 105–132.

(14) Bevington, P. R.; Robinson, D. K.; Blair, J. M.; Mallinckrodt, A. J.; McKay, S. Data reduction and error analysis for the physical sciences. *Computers in Physics* **1993**, *7*, 415–416.

(15) Padmanabhan, V.; Daillant, J.; Belloni, L.; Mora, S.; Alba, M.; Konovalov, O. Specific ion adsorption and short-range interactions at the air aqueous solution interface. *Physical review letters* **2007**, *99*, 086105.

(16) ben Jabrallah, S.; Malloggi, F.; Belloni, L.; Girard, L.; Novikov, D.; Mocuta, C.; Thiaudière, D.; Daillant, J. Electrolytes at interfaces: accessing the first nanometers using X-ray standing waves. *Physical Chemistry Chemical Physics* **2017**, *19*, 167–174.

(17) Malloggi, F.; Jabrallah, S. B.; Girard, L.; Siboulet, B.; Wang, K.; Fontaine, P.; Daillant, J. X-ray Standing Waves and Molecular Dynamics Studies of Ion Surface Interactions in Water at a Charged Silica Interface. *The Journal of Physical Chemistry C* **2019**, *123*, 30294–30304.