Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

Microscopic Mechanism of Specific Peptide Adhesion to Semiconductor Substrates**

Michael Bachmann, Karsten Goede, Annette G. Beck-Sickinger, Marius Grundmann, Anders Irbäck, and Wolfhard Janke*

anie_201000984_sm_miscellaneous_information.pdf

Supplementary Information

Michael Bachmann, Karsten Goede,* Annette G. Beck-Sickinger, Marius Grundmann, Anders Irbäck, and Wolfhard Janke

E-mail: goede@physik.uni-leipzig.de

Modeling and Computer Simulation

Hybrid peptide–silicon model. The characteristic properties of HF treated Si(100) surfaces in de-ionized water as described in the Manuscript effectively enter into the definition of our hybrid model which serves as the basis of our analysis and interpretation of the specificity of peptide adhesion on these interfaces. We conclude that the key role of water is the slowing down of the oxidation process of the Si(100) surface, but for the actual binding process its influence is rather small (up to screening effects). In particular, we do not expect that stable water layers form between adsorbate and substrate. Our model contains all peptide atoms, while the substrate is simplified and consists only of atomic layers with surface specific atomic density. The substrate and its surface structure itself is fixed and thus its energy is not considered in the model. Therefore, the energy of a single peptide with conformation $\mathbf{X}$ (where dihedral backbone and side-chain angles are the degrees of freedom) and interacting with the substrate, whose surface structure is characterized by the Miller index (hkl), is generally written as

$$E(\mathbf{X}) = E_{\text{pep}}(\mathbf{X}) + E_{\text{pep-sub}}^{\text{Si}(hkl)}(\mathbf{z}). \tag{1}$$

*To whom correspondence should be addressed

Here, $\mathbf{z} = (z_1, z_2, \ldots, z_N)$ is the perpendicular distance vector of all N peptide atoms from the surface layer of the substrate. The effect of the surrounding solvent is implicitly contained in the force field parameters.

The peptide is described by a simplified all-atom model,

$$E_{\mathrm{pep}}(\mathbf{X}) = E_{\mathrm{ev}}(\mathbf{X}) + E_{\mathrm{loc}}(\mathbf{X}) + E_{\mathrm{hb}}(\mathbf{X}) + E_{\mathrm{hp}}(\mathbf{X}) \tag{2}$$

which consists of intrinsic excluded volume repulsions E_{ev} between all atoms, a local potential E_{loc} which represents the interaction among neighboring NH and CO partial charges, hydrogen bonding energy E_{hb}, and the interaction between hydrophobic side chains, E_{hp}. The individual contributions in Eq. (2) have been described in detail elsewhere.[1–3]

Table 1: Atomic van der Waals radii $\sigma = \sigma_{\mathrm{Si}}, \sigma_i$ and energy depths $\varepsilon = \varepsilon_{\mathrm{Si}}, \varepsilon_i$ used in the simulations.

atom	σ [Å]	ε [kcal/mol]
Si	2.0[7]	0.05[8]
H	1.2[9]	0.04[9]
C	1.8[7,9]	0.05[9]
O	1.5[7,9]	0.08[9]
N	1.6[7,9]	0.09[9]

The interaction of the peptide with the substrate is modeled in a coarse-grained way, i.e., each peptide atom feels the mean field of the atomic substrate layers. The atomic density of these layers is dependent on the surface characteristics, i.e., it depends on the crystal orientation (the Miller index *hkl*) of the substrate at the surface. We make the following assumptions for setting up the model: (i) According to the discussion of Si(100) surface properties in de-ionized water, the Si(100) surface is considered to be hydrophobic. This has the effect that it is not favorable for water molecules to reside between the adsorbed peptide and the substrate. Furthermore, polarization effects between sidechains and substrate are not expected. (ii) Since dangling bonds on the Si(100) surface are probably saturated by covalent bonds to hydrogen atoms (due to the HF etching process), we assume that covalent bonds between peptide and surface atoms are not formed.

Thus, the surface is also considered to be uncharged.[4] (iii) Si dimers sticking off the substrate are not considered. This and the hydration effect are expected to weakly screen the peptide from the substrate. Based on these assumptions, we use a generic noncovalent Lennard-Jones approach for modeling the interaction between peptide atoms and surface layer,[5,6]

$$E_{\text{pep-sub}}^{\text{Si}(hkl)}(\mathbf{z}) = 2\pi\rho^{\text{Si}(hkl)} \times \sum_{i=1}^{N} \varepsilon_{i,\text{Si}}\sigma_{i,\text{Si}}^2 \left[\frac{2}{5}\left(\frac{\sigma_{i,\text{Si}}}{z_i}\right)^{10} - \left(\frac{\sigma_{i,\text{Si}}}{z_i}\right)^{4}\right], \tag{3}$$

where $\rho^{\text{Si}(hkl)}$ is the atomic density of the Si(hkl) surface layer. Si has diamond structure and since we assume the surface to be ideally planar, $\rho^{\text{Si}(100)} \approx 0.068\text{Å}^{-2}$. The noncovalent interaction between the peptide atoms and the Si substrate is parametrized by force-field parameters $\varepsilon_{i,\text{Si}} = \sqrt{\varepsilon_i \varepsilon_{\text{Si}}}$ and $\sigma_{i,\text{Si}} = \sigma_i + \sigma_{\text{Si}}$. It thus depends on the energy depths ε_i and van der Waals radii σ_i of the individual atoms. The parameter values used in the simulations are listed in Table Table 1.

Monte Carlo computer simulation. For the computer simulations of the hybrid model, a multiple-thread variant of the multicanonical Monte Carlo method, embedded into the BONSAI (**B**io-**O**rganic **N**ucleation and **S**elf-**A**ssembly at **I**nterfaces) package,[10] was employed.

The canonical partition function at temperature T can be written as

$$Z = \int_{E_{\text{min}}}^{\infty} dE\, g(E)\, e^{-E/RT}, \tag{4}$$

where $g(E) = e^{S(E)/R}$ is the density of states that connects (microcanonical) entropy S and energy E. Therefore, all informations regarding the phase behavior of the system – typically governed by the competition between entropy and energy – is encoded in $g(E)$. Consequently, for a detailed global analysis of the phase behavior, a precisely estimated density of states is extremely helpful. Unfortunately, the density of states covers many orders of magnitude in the phase transition regimes, so that its estimation requires the application of sophisticated generalized-ensemble Monte Carlo methods.

Multicanonical sampling[11] allows the estimation of $g(E)$, in principle, within a single simulation. The idea is to increase the sampling rate of conformations being little favored in the free-energy landscape and, finally, to perform a random walk in energy space. This is achieved in the simplest way by setting $T = \infty$ and introducing suitable multicanonical weights $W_{\mathrm{muca}}(E) \sim g^{-1}(E)$ in order to sample conformations X according to a transition probability

$$\omega(X \to X') = \min\left[e^{[S(E(X))-S(E(X'))]/R}, 1\right], \tag{5}$$

where $S(E(X))/R = -\ln W_{\mathrm{muca}}(E(X)) = \ln g(E(X))$.

The implementation of multicanonical sampling is not straightforward as the multicanonical weights $W_{\mathrm{muca}}(E)$ are obviously unknown *a priori*. Therefore, starting with $W^{(0)}_{\mathrm{muca}}(E) = \mathrm{const}$, the weights have to be determined by an iterative procedure until the multicanonical histogram h_{muca} is almost "flat", i.e., if the estimate for the density of states after the nth run, $\hat{g}^{(n)}(E)$, satisfies

$$\hat{g}^{(n)}(E) W^{(n-1)}_{\mathrm{muca}}(E) \approx \mathrm{const}. \tag{6}$$

in the desired range of energies. An efficient, error-weighted estimation method for the multicanonical weights is described in detail in Refs.[12,13]

Eventually, if Eq. (6) is reasonably satisfied, the multicanonical weights $W^{(n)}_{\mathrm{muca}}(E) = [\hat{g}^{(n)}(E)]^{-1}$ are then used in a final long production run, where all quantities of interest are measured and stored in a time-series file. The canonical expectation value of any quantity O at temperature T is then obtained from the multicanonical time series of length M by reweighting,

$$\langle O \rangle = \frac{\sum_{t=1}^{M} O(\mathbf{X}_t) e^{-E(\mathbf{X}_t)/RT} W^{-1}_{\mathrm{muca}}(E(\mathbf{X}_t))}{\sum_{t=1}^{M} e^{-E(\mathbf{X}_t)/RT} W^{-1}_{\mathrm{muca}}(E(\mathbf{X}_t))}, \tag{7}$$

where t is the multicanonical Monte Carlo "time" step (or sweep).

The exemplified result for the microcanonical entropy $S(E)$ obtained in the simulation for pep-

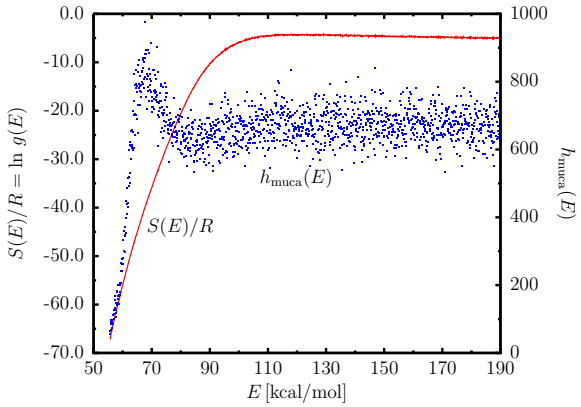

Figure 1: Microcanonical entropy $S(E)/R = \ln g(E)$ (up to an unimportant constant) and multicanonical histogram $h_{\mathrm{muca}}(E)$ for the peptide sequence S1.

tide sequence S1 (see Manuscript) and the corresponding, almost flat multicanonical histogram h_{muca} from the same simulation are plotted in Fig. Figure 1. This shows that the density of states could be estimated with high accuracy (more than 25 orders of magnitude) and the flatness of the multicanonical histogram signals that also the entropically strongly suppressed conformations were sampled with high statistics which is necessary to reasonably analyze the thermodynamics of adsorption.

In our simulations, conformational updates included rotations about single dihedral backbone and sidechain torsion angles, as well as rigid body rotations and translations. A simulation box of dimension $[50\text{Å}]^3$ with periodic boundary conditions parallel to the substrate was used. In perpendicular direction mobility is restricted by the Si substrate at $z = 0$ and a steric wall at $z = z_{\mathrm{max}} = 50\text{Å}$, where the atoms experience hard-wall repulsion.

We have also performed consistency and validity checks of the multicanonical results in independent replica-exchange (parallel tempering)[14,15] Monte Carlo simulations.

Peptide synthesis

The peptides were synthesized by automated multiple solid phase peptide synthesis (Syro, Multisyntech, Bochum, Germany) using the Wang resin to obtain a peptide acid (30 mg, resin loading

0.6 mmol/g) and the fluorenyl-9-methoxycarbonyl (Fmoc)/*tert.* butyl strategy. Fmoc-amino acids (tenfold excess) were introduced by double coupling procedures (2 × 36 min) using in situ activation with diisopropylcarbodiimide and 1-hydroxybenzotriazol. The Fmoc removal was carried out with 40% piperidine in DMF for 3 min, 20% piperidine for 7 min and finally 40% piperidine for 5 min. The obtained peptides were cleaved by a mixture of trifluoroacetic acid/thioanisol/thiocresol (90/5/5, v/v) for 3 hours. Afterwards the peptides were precipitated from ice-cold diethyl ether, collected by centrifugation and washed four times. Purification of the peptides was achieved by preparative high-performance liquid cromatography (HPLC) on a RP C-18 column (Waters, 300 × 25 mm, 5 μm) with a gradient of 20%–70% B in A (A = 0.1% trifluoroacetic acid in water; B = 0.08% trifluoroacetic acid in acetonitrile) over 45 min and a flow of 15 mL/min. All relevant fractions were collected and further analyzed by electrospray ionization mass spectrometry (ESI-MS) on an API 3000 PE Sciex (Canada, Toronto) and by analytical reversed-phase HPLC on a Vydac RP18-column (4.6 × 250 mm; 5 μm / 300 Å) using linear gradients of 10–60% B in A over 30 min and a flow rate of 0.6 mL/min. The achieved purity of the peptide was ≥95%.

Experiments

Sample preparation. (100)-grown Si substrate pieces with an average surface size of (5 × 5) mm^2 have been prepared by wafer cleaving (wafers from Korth, Altenholz, Germany). Native oxides and other particles covering the surfaces were removed by etching for 1 min in an ammonium fluoride solution (87.5% NH_4F: H_2O, 12.5% HF: H_2O) followed by a water rinse and distilled water bath.[16] This etch does not attack the cleaned semiconductor surfaces. Si substrates showed typical hydrophobic properties after etching, which indicates a clean Si surface. We have investigated the clean semiconductor surfaces by atomic-force microscopy (AFM) to estimate the cleanness and flatness of the respective surfaces. The maximum particle coverage was 0.2%, which is well below the peptide adhesion coefficients (PAC) discussed in the text. After a possible sample tilt was leveled, the maximum root-mean square (rms) roughness of the clean (100) sample

surfaces was 0.37 nm, while for most samples the roughness rms value was well below 0.30 nm. These values are typical for freshly etched semiconductor surfaces. Suitable clean and flat sample pieces have been exposed to a diluted watery solution of the respective peptide within seconds after etching. The peptide concentration in this Tris-buffered saline (Carl Roth, Karlsruhe, Germany; pH 7.6) was 1 μg/mL. With the molar mass 1335 g/mol for any of the four peptide sequences, this equals a solution molarity of 0.75 μM. At this very low value, no major inter-molecular aggregation in solution is expected for these short, polar side-chain dominated peptides.[16,17] The dwell was 2 h. The saline solution in distilled water was used to further minimize peptide cluster lumping in solution. Samples have been subsequently exposed to a water rinse and short distilled-water bath to remove unbound particles. Peptide particles which still appear on the surface after the wash can be expected to have formed a strong binding to the surface when the semiconductor was still in solution. This sticking to the surface holds for the majority of clusters.[16,17]

Measurements. After a drying time of 6 h in air, sample surfaces have been investigated by a Dimension 3000 AFM in combination with a Nanoscope IIIa (Digital Instruments, now Veeco, Woodbury NY, U.S.) operating in tapping mode. The AFM probe was n^+ silicon with a 123-μm cantilever and a spring constant of 59 N/m driven near its resonance frequency of 380 kHz. Scan rates were of the order of 5.0–0.15 μm/s, depending on the image size of 10–0.5 μm. Very similar images have been obtained with other probes (226 μm, 188 kHz, 45 N/m). PAC values have been obtained by performing a grain analysis on (5×5) μm^2 images using the SPIP program (version 1.9223, Image Metrology A/S, Lyngby, Denmark). Images were leveled using a first-order plane fit when necessary to remove a sample tilt. This granted for setting the minimum detection height in the grain analysis to 0 nm above the average surface height, thus allowing each cluster to be detected. A calibration of these PAC-on-Si values was achieved by relating them to the respective PAC values on GaAs (100) substrates from the same peptide solution. This ensured errors well below 0.1 (on the cPAC scale). AFM images in this manuscript have been prepared with the Gwyddion 2.10 free software.

References

(1) A. Irbäck, B. Samuelsson, F. Sjunnesson, S. Wallin, *Biophys. J.* **2003**, *85*, 1466–1473.

(2) A.Irbäck, S. Mohanty, *Biophys. J.* **2005**, *88*, 1560–1569.

(3) A. Irbäck, S. Mohanty, *J. Comput. Chem.* **2006**, *27*, 1548–1555.

(4) J. Tien, A. Terfort, G. M. Whitesides, *Langmuir* **1997**, *13*, 5349–5355.

(5) W. A. Steele, *Surf. Sci.* **1973**, *36*, 317–352.

(6) R. Hentschke, *Macromol. Theory Simul.* **1997**, *6*, 287–316.

(7) S. S. Batsanov, *Inorg. Mater.* **2001**, *37*, 871–885.

(8) H. Heinz, H. Koerner, K. L. Anderson, R. A. Vaia, B. L. Farmer, *Chem. Mater.* **2005**, *17*, 5658–5669.

(9) T. A. Halgren, *J. Am. Chem. Soc.* **1992**, *114*, 7827–7843.

(10) M. Bachmann, *The BONSAI Project: Simulation of bio-organic nucleation and self-assembly at interfaces*. The peptide model (2) used in this package was introduced in Ref. 1 and is also contained in the protein folding and aggregation simulator (PROFASI),[3] which is freely available at http://cbbp.thep.lu.se/activities/profasi.

(11) B. A. Berg, T. Neuhaus, *Phys. Rev. Lett.* **1992**, *68*, 9–12.

(12) W. Janke, *Histograms and all that*, in *Computer Simulations of Surfaces and Interfaces*, NATO Science Series, II. Mathematics, Physics and Chemistry - Vol. *114*, Proceedings of the NATO Advanced Study Institute, ed. by B. Dünweg, D. P. Landau, and A. I. Milchev (Kluwer, Dordrecht, **2003**), pp. 137-157.

(13) B. A. Berg, *Fields Inst. Commun.* **2000**, *26*, 1–24.

(14) K. Hukushima, K. Nemoto, *J. Phys. Soc. Jpn.* **1996**, *65*, 1604–1608.

(15) C. J. Geyer, *Markov chain Monte Carlo maximum likelihood*, in *Computing Science and Statistics*, Proceedings of the 23rd Symposium on the Interface, ed. by E. M. Keramidas (Interface Foundation, Fairfax Station, **1991**), 156–163.

(16) K. Goede, P. Busch, M. Grundmann, *Nano Lett.* **2004**, *4*, 2115–2120.

(17) K. Goede, M. Grundmann, K. Holland-Nell, A. G. Beck-Sickinger, *Langmuir* **2006**, *22*, 8104–8108.