

## Differences in Solution Behavior among Four Semiconductor-Binding Peptides

Simon Mitternacht,<sup>†</sup> Stefan Schnabel,<sup>‡</sup> Michael Bachmann,<sup>†,‡</sup> Wolfhard Janke,<sup>‡</sup> and Anders Irbäck<sup>\*,†</sup>

Computational Biology & Biological Physics Group, Department of Theoretical Physics, Lund University, Sölvegatan 14A, SE-223 62 Lund, Sweden, and Institute for Theoretical Physics and Centre for Theoretical Sciences (NTZ), University of Leipzig, Augustusplatz 10/11, D-04109 Leipzig, Germany

Received: November 15, 2006; In Final Form: March 1, 2007

Recent experiments have identified peptides that adhere to GaAs and Si surfaces. Here, we use all-atom Monte Carlo simulations with implicit solvent to investigate the behavior in aqueous solution of four such peptides, all with 12 residues. At room temperature, we find that all four peptides are largely unstructured, which is consistent with experimental data. At the same time, we find that one of the peptides is structurally different and more flexible, as compared to the others. This finding points at structural differences as a possible explanation for differences in adhesion properties among these peptides. By also analyzing designed mutants of two of the peptides, an experimental test of this hypothesis is proposed.

### Introduction

The advancing progress in manipulating proteins and non-biological macromolecules and materials at the nanometer scale opens up possibilities for constructing novel hybrid materials with potential applications in bionanotechnology.<sup>1,2</sup> An important development in this direction is the identification of proteins that can bind to specific compounds. Over the past decade, genetic engineering techniques have been successfully employed to find peptides with affinity for, for example, metals,<sup>3,4</sup> semiconductors,<sup>5</sup> and carbon nanotubes.<sup>6</sup> However, the mechanisms by which peptides bind to these materials are not completely understood; e.g., it is unclear what role conformational changes play in the binding process.

Here, we report atomic-level simulations of the solution behavior of four 12-residue peptides whose adhesion properties to (100) surfaces of GaAs and Si crystals were studied in recent experiments.<sup>5,7,8</sup> The main quantity measured in the experiments was the peptide adhesion coefficient (PAC), defined as the percentage of surface coverage, after drying and washing of the samples that were originally in contact with the peptide solution. This quantity was measured by AFM for the different peptide–substrate combinations<sup>7,8</sup> and was found to show a clear dependence on both peptide and substrate (see below).

How the binding occurs in these peptide–surface systems is unclear. However, although the bound peptides were found to form clusters,<sup>8</sup> it seems unlikely that the peptides aggregate before binding to the surface, because the hydrophobicity of the peptides studied is low and the peptide concentration was low, in the nanomolar range. A more accurate description is probably that the peptides bind one by one, a process that, in principle, can occur in two fundamentally different ways. One possibility is a docking behavior, in which the peptides bind to the surface without undergoing any major conformational change. This scenario assumes that the peptides have a stable structure in solution and that this structure matches the structure

of the surface, for example, with respect to polarization. The second variant is that the peptide is unstructured before binding occurs. Although the bound peptide structure need not be unique, the process would then have similarities with coupled folding–binding,<sup>9</sup> which can be an efficient mode of binding when compared to docking.<sup>10,11</sup> Measurements of circular dichroism (CD) spectra suggest that all four studied peptides are largely unstructured in solution,<sup>8</sup> thus favoring the second type of binding over docking.

Recent studies have found that the adhesion propensity of peptides to various surfaces can be in part explained in terms of adhesion properties of their constituent amino acids;<sup>12,13</sup> however, the amino acid composition alone cannot explain the PAC values obtained experimentally for the four peptides studied here. In fact, two of these peptides share exactly the same amino acid composition, but still have quite different adhesion properties. In order to explain the adhesion properties, it might thus be necessary to take structural characteristics into account.

The aim of our study is to get a more detailed picture of the behavior in aqueous solution of these peptides and to look for possible structural differences not seen in the CD analysis.<sup>8</sup> A perfect model for folding simulations does not exist. It is worth noting, however, that the model we use,<sup>14,15</sup> despite a simplified energy function, is capable of folding both  $\alpha$ -helical and  $\beta$ -sheet peptides without changing any model parameters.<sup>15</sup>

Three of the peptides we study have previously been simulated<sup>16</sup> using the ECEPP/3 force field.<sup>17</sup> This study found only minor differences in folding behavior between these peptides. To further elucidate the structural properties of these peptides at room temperature, we here perform simulations using an alternative model, which has given realistic results for the stability and its temperature dependence for the peptides that it was able to fold.<sup>15</sup>

Simulating the actual binding of the peptides to the surface is more challenging due to uncertainties about the precise form of the peptide–surface interactions and their dependence on solvation effects.<sup>18,19</sup> Nevertheless, such simulations have been performed for gold-binding peptides.<sup>20</sup> The phase structure for

\* Corresponding author. Phone: +46 46 2223493. Fax: +46 46 2229686. E-mail: anders@thep.lu.se.

<sup>†</sup> Lund University.

<sup>‡</sup> University of Leipzig.

**TABLE 1: The Four Sequences Studied and Their PAC Values for Adsorption to (100) GaAs and Si Surfaces (from Goede et al.<sup>8</sup>)<sup>a</sup>**

label	sequence	PAC, %	
		GaAs	Si
S1	AQNPSDNNHTH	25	1
S2	AQNPSDNNATA	14	3
S3	TNHDHNSAPNTQ	17	15
S4	AQAPSDAATHH	21	6

<sup>a</sup> S2 is a double His → Ala mutant of S1; S3, a random permutation of S1; and S4, a triple Asn → Ala mutant of S1.

chain adsorption to attractive surfaces has been investigated using lattice models for polymers<sup>21–23</sup> and peptides.<sup>24</sup> Simplified statistical–mechanical models have also been used to study molecular recognition of patterned surfaces<sup>25–30</sup> and conformational changes of proteins adsorbed to a solid surface.<sup>31,32</sup>

## Model and Methods

**Peptides Studied.** The four peptides we study are listed in Table 1, where also PAC values for (100) GaAs and Si surfaces can be found. The sequence S1 was selected from a huge library of 12-mers for adhesion to GaAs.<sup>5</sup> Its very poor propensity to adhere to Si is noteworthy. The sequence S2 is obtained from S1 by exchanging two histidines for alanines. This double mutation leads to a reduced PAC for GaAs and a slightly increased PAC for Si. The peptide adhering best to the Si surface is S3, which is a random permutation of S1. The Si PAC is a factor of 15 higher for S3 than for S1, despite that their amino acid composition is the same. The last sequence, S4, is derived from S1 by replacing three asparagines with alanines. This change results in a slightly reduced GaAs PAC and an increased Si PAC.

**Peptide Model.** The model we use contains all atoms of the peptide chain, including H atoms, but no explicit water molecules. It assumes fixed bond angles, bond lengths, and peptide torsion angles (180°) so that each amino acid has the Ramachandran angles  $\phi$ ,  $\psi$  and a number of side chain torsion angles as its degrees of freedom. Here, a brief presentation of the energy function will be given. Detailed descriptions of the parametrization of the geometry<sup>14</sup> and the different energy terms<sup>15</sup> can be found elsewhere.

The energy function consists of four terms,

$$E = E_{\text{ev}} + E_{\text{loc}} + E_{\text{hb}} + E_{\text{hp}} \quad (1)$$

The first term,  $E_{\text{ev}}$ , represents excluded volume effects and is of the form

$$E_{\text{ev}} = \kappa_{\text{ev}} \sum_{i < j} \left[ \frac{\lambda_{ij}(\sigma_i + \sigma_j)}{r_{ij}} \right]^{12} \quad (2)$$

where the sum is over all atom pairs. The parameters  $\sigma_i$  are atomic radii, and  $\lambda_{ij}$  is a scale factor, which is 1.0 for pairs connected by three covalent bonds and 0.75 otherwise.

The second term represents an interaction between neighboring NH and CO partial charges along the backbone. It is given by

$$E_{\text{loc}} = \kappa_{\text{loc}} \sum_I \left[ \sum_{\substack{i=N,H \in I \\ j=C,O \in I}} \frac{q_i q_j}{r_{ij}} \right] \quad (3)$$

where the outer sum is over all amino acids, and the  $q_i$  are partial charges.

The H bond contribution,  $E_{\text{hb}}$ , consists of two parts: backbone–backbone bonds and backbone–side chain bonds,

$$E_{\text{hb}} = \epsilon_{\text{hb}}^{(1)} \sum_{\text{bb-bb}} u(r_{ij}) \nu(\alpha_{ij}, \beta_{ij}) + \epsilon_{\text{hb}}^{(2)} \sum_{\text{bb-sc}} u(r_{ij}) \nu(\alpha_{ij}, \beta_{ij}) \quad (4)$$

where  $r_{ij}$  denotes the HO distance;  $\alpha_{ij}$ , the NHO angle; and  $\beta_{ij}$ , the HOC angle. The function  $u(r)$  is given by

$$u(r) = 5 \left( \frac{\sigma_{\text{hb}}}{r} \right)^{12} - 6 \left( \frac{\sigma_{\text{hb}}}{r} \right)^{10} \quad (5)$$

and the angular dependence is

$$\nu(\alpha, \beta) = \begin{cases} (\cos \alpha \cos \beta)^{1/2} & \text{if } \alpha, \beta > 90^\circ \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

The last energy term,  $E_{\text{hp}}$ , represents an effective hydrophobic attraction and has the form

$$E_{\text{hp}} = - \sum_{I < J} M_{IJ} C_{IJ} \quad (7)$$

where the sum is over all pairs of nonpolar amino acids. The  $M_{IJ}$  ( $\geq 0$ ) are constants that determine the strength of attraction between amino acids,  $I$  and  $J$ .  $C_{IJ}$  is a geometric factor and a measure of the degree of contact between two side chains. It is defined as

$$C_{IJ} = \frac{1}{N_I + N_J} \left[ \sum_{i \in A_I} f(\min r_{ij}^2) + \sum_{j \in A_J} f(\min r_{ij}^2) \right] \quad (8)$$

where  $A_I$  denotes a predefined set of  $N_I$  side chain atoms for residue  $I$ . The function  $f(x)$  is given by  $f(x) = 1$  if  $x < A$ ,  $f(x) = 0$  if  $x > B$ , and  $f(x) = (B - x)/(B - A)$  if  $A < x < B$  [ $A = (3.5\text{\AA})^2$  and  $B = (4.5\text{\AA})^2$ ].

In this model, folding is mainly driven by hydrogen bonding and the effective hydrophobic force. The sequences studied here, S1–S4, are all only weakly hydrophobic, which makes hydrogen bonding the dominating driving force.

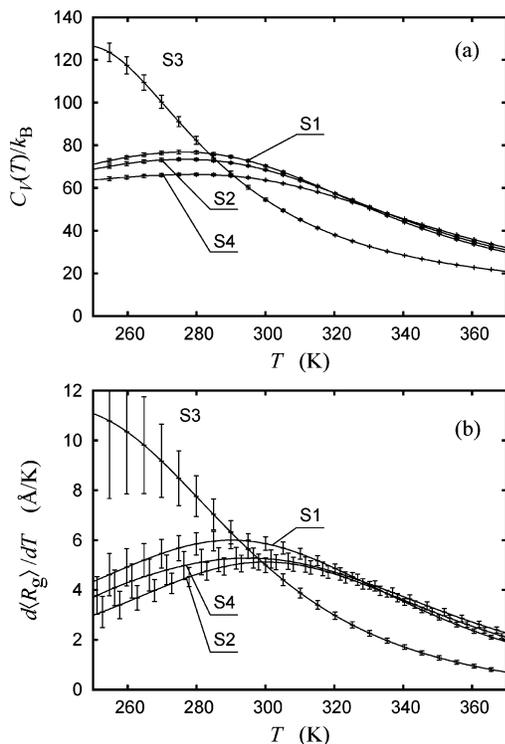
**Simulation Method.** To investigate the solution behavior of the peptides S1–S4, we perform simulated-tempering<sup>33,34</sup> simulations with eight temperatures in the range 275–369 K (274.9, 286.7, 299.0, 311.8, 325.2, 339.1, 353.6, and 368.8 K), and some reference runs at a constant temperature of 1000 K. The conformational updates we use are rotations of single backbone and side chain torsion angles and a semilocal backbone update, biased Gaussian steps (BGS),<sup>35</sup> which updates seven or eight consecutive angles in a manner that keeps the rest of the molecule approximately fixed. In the simulated-tempering runs, these updates are called in different proportions at different temperatures with more BGS at lower temperatures. At 299 K, the fractions of attempted single-angle backbone moves, side chain moves, and BGS are 0.29, 0.51, and 0.20, respectively.

Our simulations are carried out using the software package PROFASI,<sup>36</sup> which is a C++ implementation of the above model. Each simulation comprises  $10^9$  elementary MC steps.

The results of our simulations are analyzed using multihistogram techniques.<sup>37</sup> All statistical uncertainties quoted are  $1\sigma$  errors obtained by the jackknife method.<sup>38</sup>

## Results and Discussion

**Overall Structure and Temperature Dependence.** Cooperative structural activity is typically signaled by a peak in the



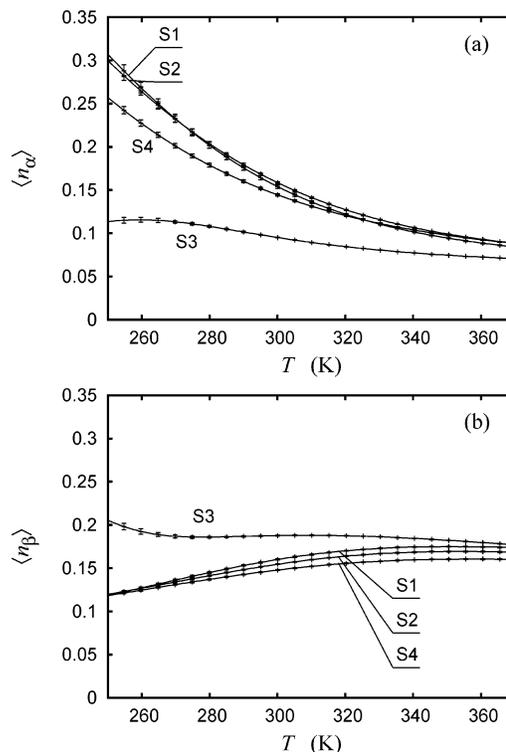
**Figure 1.** Temperature dependence of (a) the specific heat  $C_V = d\langle E \rangle/dT$  and (b)  $d\langle R_g \rangle/dT$ , for the sequences S1–S4.  $R_g$  is the radius of gyration (calculated over all non-H atoms).

statistical fluctuations of system relevant quantities, such as the energy. Figure 1 shows how the specific heat,  $C_V = d\langle E \rangle/dT = (\langle E^2 \rangle - \langle E \rangle^2)/k_B T^2$ , and the temperature derivative of the radius of gyration,  $d\langle R_g \rangle/dT$ , vary with temperature for the sequences S1–S4 ( $\langle \dots \rangle$  denotes a Boltzmann average). The qualitative behavior of the three sequences S1, S2, and S4 is virtually identical. For all three sequences, the specific heat exhibits a broad peak with maximum around 280 K. The  $d\langle R_g \rangle/dT$  curves show a similar broad peak, although the statistical errors are larger, and the maximum is slightly shifted toward higher temperature.

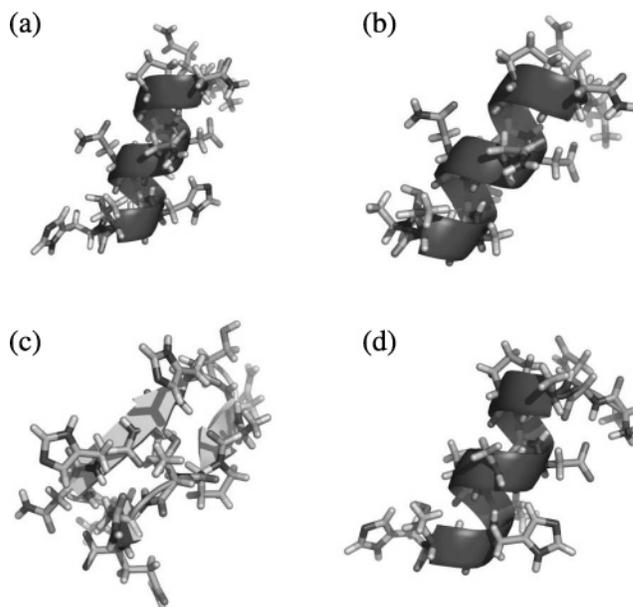
In the temperature regime where these peaks occur, it turns out that the secondary-structure content of these three sequences changes relatively rapidly. As the temperature decreases, the  $\alpha$ -helix content,  $\langle n_\alpha \rangle$ , increases, whereas the  $\beta$ -strand content,  $\langle n_\beta \rangle$ , decreases slightly, as can be seen from Figure 2. These results indicate that the structures with lowest energy are  $\alpha$ -helical for S1, S2, and S4. It should be noted, however, that the  $\alpha$ -helix content remains small,  $<0.25$ , all the way down to 273 K.

The sequence S3 shows a markedly different behavior. Neither  $C_V$  nor  $d\langle R_g \rangle/dT$  has a maximum within the temperature range studied; both quantities increase monotonically with decreasing temperature (see Figure 1). Furthermore, the  $\beta$ -strand content remains larger than the  $\alpha$ -helix content at low temperature for this sequence (see Figure 2). The  $\beta$ -strand content does not decrease with decreasing temperature, and the  $\alpha$ -helix content increases much less than for the other sequences.

Figure 3 shows typical low-energy conformations for the four different sequences, as obtained by simulated annealing.<sup>39</sup> As one might expect from the temperature dependence of the  $\alpha$ -helix and  $\beta$ -strand contents, the structure is  $\alpha$ -helical for S1, S2, and S4. However, the  $\alpha$ -helix does not span the entire chain, but rather, the region between residues 3 and 12. That the beginning of the sequence is not  $\alpha$ -helical is not unexpected,

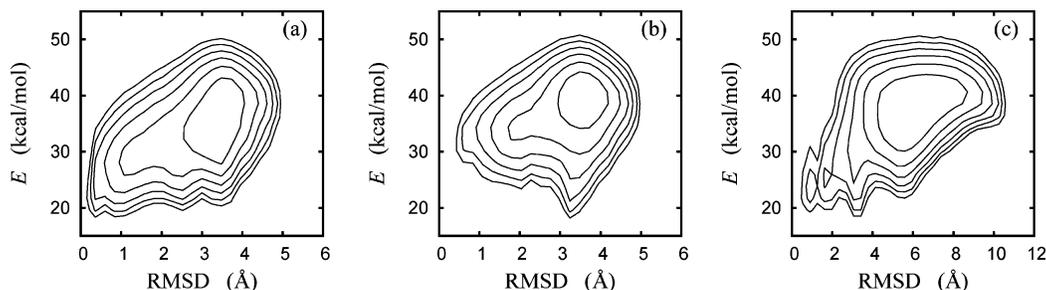


**Figure 2.** Temperature dependence of (a) the  $\alpha$ -helix content  $\langle n_\alpha \rangle$  and (b) the  $\beta$ -strand content  $\langle n_\beta \rangle$ , for the sequences S1–S4. We define a residue as  $\alpha$ -helical if its Ramachandran angles  $\phi$  and  $\psi$  satisfy  $\phi \in (-90^\circ, -30^\circ)$  and  $\psi \in (-77^\circ, -17^\circ)$ , and  $n_\alpha$  denotes the fraction of the 10 inner residues that are  $\alpha$ -helical. Similarly,  $n_\beta$  is the fraction of the 10 inner residues with Ramachandran angles satisfying  $\phi \in (-150^\circ, -90^\circ)$  and  $\psi \in (90^\circ, 150^\circ)$ .



**Figure 3.** Typical low-energy conformations for (a) S1, (b) S2, (c) S3, and (d) S4. These structures were obtained as the lowest-energy structures in 10 simulated annealing runs for each sequence, starting from random conformations. In each run, the temperature was decreased geometrically from 369 to 0.7 K in 100 steps. At each temperature, 100 000 elementary MC steps were performed. Drawn with PyMOL.<sup>40</sup>

because there is a proline at position 4. The lowest-energy structure we find for S3 is a  $\beta$ -hairpin. Its turn is at residues 6 and 7. The second strand of the  $\beta$ -hairpin, spanning residues 8–12, is not perfect but broken in the vicinity of the proline at position 9.



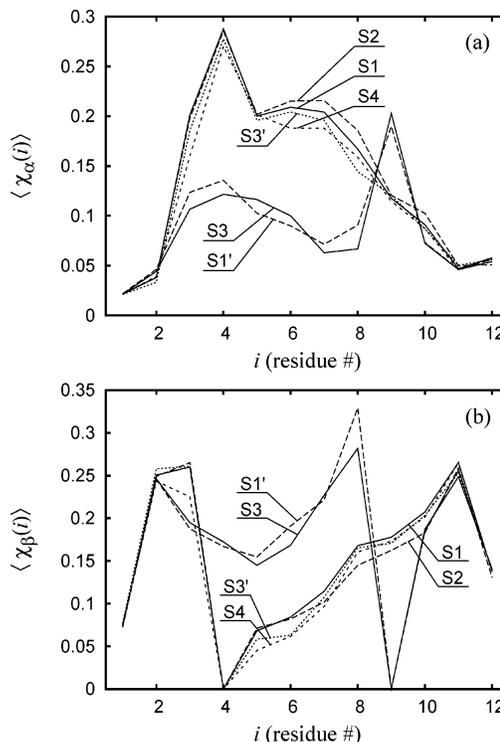
**Figure 4.** Free energies  $F(\Delta, E)$  calculated as functions of rmsd,  $\Delta$ , and energy,  $E$ , for S1 and S3 at  $T = 299$  K. The reference structure is either an  $\alpha$ -helix or a  $\beta$ -hairpin (see text). The contours are spaced at intervals of  $1k_B T$ . Contours more than  $6k_B T$  above the minimum free energy are not shown. The free energy  $F(\Delta, E)$  is defined by  $P(\Delta, E) \propto \exp(-F(\Delta, E)/k_B T)$ , where  $P(\Delta, E)$  is the joint probability distribution of  $\Delta$  and  $E$  at temperature  $T$ . (a) rmsd from the  $\alpha$ -helix for S1 (calculated over residues 5–12). (b) rmsd from the  $\alpha$ -helix for S3 (residues 1–8). (c) rmsd from the  $\beta$ -hairpin for S3 (all residues). Note that the  $x$  scale is different in (c).

It must be stressed that the states illustrated in Figure 3 are only weakly populated at room temperature, as is evident from the secondary-structure contents shown in Figure 2. Our results are thus consistent with the CD analysis of the solution behavior of these peptides,<sup>8</sup> at room temperature and pH 7.6, which suggests that they all are largely unstructured.

Our conclusion that the  $\alpha$ -helix content, at low temperature, is higher than the  $\beta$ -strand content for S1 and S2 is in agreement with a previous study of S1–S3 based on the ECEPP/3 force field;<sup>16</sup> however, in that study, the sequence S3 was found to be  $\alpha$ -helical, as well. Furthermore, the  $\alpha$ -helix content of S1 and S2 was significantly higher as compared to what we find and to what is indicated by the CD results.<sup>8</sup>

Having studied the overall structure and the temperature dependence, we now turn to a more detailed structural description at  $T = 299$  K, which is close to where the CD measurements were taken.<sup>8</sup> This discussion will focus mainly on S1 and S3, because the double mutant S2 and the triple mutant S4 show a behavior very similar to that of S1.

**Structural Characterization at  $T = 299$  K.** To further elucidate the structure and free-energy landscape of these peptides, we analyze root-mean-square deviations (rmsd) from suitable reference structures (calculated over backbone atoms). We first consider an  $\alpha$ -helical reference structure. The N-terminal part of S1 is rather flexible due to a proline at position 4. Similarly, the C-terminal part of S3 is flexible due to a proline at position 9. To reduce noise, we omit these tails when calculating rmsd. The reference structure used is an  $\alpha$ -helix with eight residues. With rmsd calculated this way, we study the free energy  $F(\Delta, E)$  as a function of rmsd,  $\Delta$ , and energy,  $E$ , at 299 K. Figure 4a and b shows contour plots of  $F(\Delta, E)$  for S1 and S3. For both sequences, the free-energy minimum is at an rmsd of  $\approx 3.4$  Å, which is approximately the average value for random structures, as obtained from control runs at 1000 K. This finding supports the conclusion that both S1 and S3 are largely unstructured at 299 K. A clear local free-energy minimum corresponding to an  $\alpha$ -helix is missing for both sequences. For S1, there is, however, a valley from the global minimum in the direction of low rmsd and low energy, and there is a small but significant fraction of  $\alpha$ -helical conformations with  $\Delta \approx 1$  Å and relatively low energy. For S3, there is a valley in the same direction, but it is less pronounced, and conformations with a  $\Delta$  as small as 1 Å are rare. There is also a second valley for S3, where the lowest populated energies are found. The appearance of this second valley, where  $\Delta > 3$  Å, is not unexpected, given that the lowest-energy structure found for S3 is a  $\beta$ -hairpin (see Figure 3c). Figure 4c shows  $F(\Delta, E)$  for S3 when this  $\beta$ -hairpin is taken as the reference structure. A local minimum with  $\Delta \approx 1$  Å and low energy can be found,



**Figure 5.** Secondary-structure profiles for S1–S4, S1' (AQNTSDN-NPHTH), and S3' (TNHPHSNADTNQ) at  $T = 299$  K. The lines are only guides to the eye. The statistical errors are small, 0.003 or smaller, and have, for clarity, been omitted. (a) The probability that residue  $i$  is in the  $\alpha$ -helix state,  $\langle \chi_\alpha(i) \rangle$ , against  $i$ . (b) The probability that residue  $i$  is in the  $\beta$ -strand state,  $\langle \chi_\beta(i) \rangle$ , against  $i$ .

but it is very weakly populated. The dominating global minimum corresponds to unstructured conformations. In fact, the average rmsd from the  $\beta$ -hairpin for random S3 conformations, as obtained from a control run at 1000 K, is  $\sim 6$  Å, which is approximately where the global minimum is found at  $T = 299$  K.

Next, we examine how the  $\alpha$ -helix and  $\beta$ -strand contents (as defined in the caption of Figure 2) vary along the chains. Let  $\chi_\alpha(i) = 1$  if residue  $i$  is in the  $\alpha$ -helix state and  $\chi_\alpha(i) = 0$  otherwise so that  $\langle \chi_\alpha(i) \rangle$  is the probability of finding residue  $i$  in the  $\alpha$ -helix state, and let  $\chi_\beta(i)$  denote the corresponding function for the  $\beta$ -strand state. Figure 5 shows  $\langle \chi_\alpha(i) \rangle$  and  $\langle \chi_\beta(i) \rangle$  against  $i$  for S1–S4 at  $T = 299$  K. The low-energy conformations of S1, S2, and S4 shown in Figure 3 contain an  $\alpha$ -helix starting near position 3 and ending at the C terminus. The  $\alpha$ -helix probability profile in Figure 5a reveals that the stability of this  $\alpha$ -helix is not uniform along the chain; its N-terminal part is most stable, whereas the stability decreases

**TABLE 2: The Fraction  $\lambda_\alpha$  ( $\lambda_\beta$ ) of Conformations That Have at Least One Continuous  $\alpha$ -Helix ( $\beta$ -Strand) Segment of Length 3 or More**

T (K)	S1			S3		
	275	299	369	275	299	369
$\lambda_\alpha$	0.21	0.12	0.03	0.06	0.04	0.01
$\lambda_\beta$	0.03	0.03	0.04	0.04	0.04	0.03

significantly toward the C terminus. For S3, it can be seen from Figure 5b that the  $\langle\chi_\beta(i)\rangle$  values are similar in the two regions that make the strands of the  $\beta$ -hairpin in Figure 3c. An exception is the proline at position 9, for which  $\langle\chi_\beta(i)\rangle$  is strictly 0 (proline has a fixed  $\phi = -65^\circ$  in the model, which falls outside the  $\phi$  interval in our  $\beta$ -strand definition). We also note that the two end residues tend to be unstructured for all four sequences, with relatively small values of both  $\langle\chi_\alpha(i)\rangle$  and  $\langle\chi_\beta(i)\rangle$ .

From the single-residue probabilities  $\langle\chi_\alpha(i)\rangle$  and  $\langle\chi_\beta(i)\rangle$ , one cannot tell whether the formation of secondary structure is cooperative. To study that for S1, S2, and S4, we calculate the helix–helix correlation coefficient for *neighboring* residues at  $T = 299$  K, as defined by

$$r_{ii+1}^{(\alpha)} = \frac{C_{ii+1}^{(\alpha)}}{\sqrt{C_{ii}^{(\alpha)} C_{i+i+1}^{(\alpha)}}} \quad (9)$$

where

$$C_{ij}^{(\alpha)} = \langle\chi_\alpha(i)\chi_\alpha(j)\rangle - \langle\chi_\alpha(i)\rangle\langle\chi_\alpha(j)\rangle \quad (10)$$

For all three peptides, we find that the largest  $r_{ii+1}^{(\alpha)}$  values occur in the region from  $i = 4$  to  $i = 9$  and are in the range 0.3–0.5. These values indicate that helix formation is a rather weakly cooperative process for these peptides. Consequently, the free-energy barrier to helix formation should be low, a conclusion that is in line with the results shown in Figure 4a. For S3,  $r_{ii+1}^{(\alpha)}$  is  $\approx 0.3$  or smaller for all  $i$ . The analogous strand–strand correlation coefficient,  $r_{ii+1}^{(\beta)}$ , defined in terms of  $\chi_\beta(i)$ , is smaller than 0.25 for all  $i$  for all four sequences.

Another way of analyzing secondary-structure correlations is to look at the typical lengths of unbroken  $\alpha$ -helix and  $\beta$ -strand segments. Specifically, we calculate the fraction of conformations, at fixed  $T$ , that have at least one unbroken  $\alpha$ -helix ( $\beta$ -strand) stretch with 3 residues or more, which we denote by  $\lambda_\alpha$  ( $\lambda_\beta$ ). Table 2 shows  $\lambda_\alpha$  and  $\lambda_\beta$  for S1 and S3 at three different temperatures. For S1 at  $T = 299$  K, we find that  $\lambda_\alpha = 0.12$ . This result can be compared with what one would expect if the  $\chi_\alpha(i)$  were independent random variables with  $i$ -dependent individual distributions, given by Figure 5a. In this uncorrelated case, it turns out that one would find  $\lambda_\alpha = 0.04$ . This comparison shows that the correlations are significant but not very strong. For S3, we find that  $\lambda_\beta = 0.04$  at  $T = 299$  K. A calculation analogous to that for S1 shows that  $\lambda_\beta = 0.04$  is precisely what one would expect in the absence of correlations. Hence, we find that secondary-structure correlations are very weak for S3.

Finally, it is also instructive to identify the backbone H bonds that are most likely to occur. We consider an H bond formed if its energy is less than  $-\epsilon_{\text{hb}}^{(1)}/3$ . For S1, we find that the bonds NH(Asp6)–CO(Asn3) and NH(Asn7)–CO(Asn3) occur in  $\approx 38$  and  $\approx 34\%$  of the conformations, respectively, at  $T = 299$  K, whereas no other backbone H bond has a frequency of occurrence above 15%. These results confirm that the  $\alpha$ -helix seen in low-energy conformations for S1 is most stable in its N-terminal part. Note also that in our simulations, this helix

often starts with a fork-like H bonding; the CO(Asn3) group acts as an acceptor for two bonds. For S3, there is only one backbone H bond that occurs in more than 15% of the conformations at  $T = 299$  K; namely, NH(Asn11)–CO(Ala8) with a frequency of occurrence of  $\approx 21\%$ . The paucity of H bonds underscores the notion that this peptide is highly flexible.

**Two Other Sequences.** Why do we find a different behavior for S3? A major reason is the different position of the proline; the proline residue with its special geometry is at position 9 in the sequence S3, but at position 4 in S1, S2, and S4. To gauge the importance of the proline location, we repeated the same calculations for a variant of S3, S3', with Asp4 and Pro9 interchanged. We find that the behavior of S3' closely resembles that of S1, S2, and S4. As an example, we show in Figure 5 the  $\alpha$ -helix and  $\beta$ -strand probability profiles for S3'. The S3' profiles are nearly identical to those for S1, S2, and S4. In the reshuffling of S1 to get S3, the change of proline position thus seems particularly important.

We also studied the sequence obtained by interchanging Pro4 and Thr9 in S1, which we call S1'. We find that this transposition of S1 leads to a behavior similar to that of S3, as is illustrated by Figure 5, which confirms the importance of the position of the proline.

Neither S1' nor S3' has, to our knowledge, been studied experimentally.

## Conclusions

We have investigated the solution behavior of four synthetic peptides, S1–S4, that experimentally have been shown to exhibit specific adhesion properties to (100) GaAs and Si semiconductor substrates. We find that S1, the double mutant S2, and the triple mutant S4 all show a very similar behavior with respect to structure as well as thermodynamics. At room temperature, these peptides are largely unstructured but have a small but significant  $\alpha$ -helix content. For S3, which is a random permutation of S1, we find a different behavior. S3 is more flexible than the other three peptides, with a very small content of both  $\beta$ -strand and  $\alpha$ -helix structure. The lowest-energy structure we find for S3 is not  $\alpha$ -helical but a  $\beta$ -hairpin.

In the experiments, S1–S4 showed good adhesion to GaAs, especially S1. The main difference between the peptides was that S3, in contrast to the other three, adhered well to Si, too.

The experimentally observed<sup>8</sup> difference between the peptides S1 and S3 shows that their adhesion properties cannot be predicted from the amino acid composition alone. Our results suggest that, although S1–S4 all are predominantly unstructured, there is a clear difference in structural preferences between S3 and the other three peptides due to different proline positions. To what extent this difference actually explains the different adhesion properties of the peptides remains to be seen. A possible test of this would be to do experiments on the sequence S3', which in our model shows a solution behavior similar to that of S1, S2, and S4. It would be very interesting to see whether the adhesion properties of S3' resemble those of S1, S2, and S4, with proline at the same position as in S3', or whether they resemble those of S3, with 83% sequence identity to S3'.

**Acknowledgment.** This work was in part supported by the German Science Foundation (no. JA 483/24-1) and the Swedish Research Council. M.B. was supported by a research fellowship from the German Science Foundation (no. BA 2878/1-1). We also acknowledge support from the DAAD-STINT Personnel

Exchange Programme and the John von Neumann Institute for Computing (NIC), Forschungszentrum Jülich (no. hlz11).

## References and Notes

- (1) Sarikaya, M.; Tamerler, C.; Jen, A. K.-Y.; Schulten, K.; Baneyx, F. *Nat. Mater.* **2003**, *2*, 577–585.
- (2) Gray, J. J. *Curr. Opin. Struct. Biol.* **2004**, *14*, 110–115.
- (3) Brown, S. *Nat. Biotechnol.* **1997**, *15*, 269–272.
- (4) Sano, K.-I.; Shiba, K. *J. Am. Chem. Soc.* **2003**, *125*, 14234–14235.
- (5) Whaley, S. R.; English, D. S.; Hu, E. L.; Barbara, P. F.; Belcher, A. M. *Nature* **2000**, *405*, 665–668.
- (6) Wang, S. Q.; Humphreys, E. S.; Chung, S.-Y.; Delduco, D. F.; Lustig, S. R.; Wang, H.; Parker, K. N.; Rizzo, N. W.; Subramoney, S.; Chiang, Y.-M.; Jagota, A. *Nat. Mater.* **2003**, *2*, 196–200.
- (7) Goede, K.; Busch, P.; Grundmann, M. *Nano Lett.* **2004**, *4*, 2115–2120.
- (8) Goede, K.; Grundmann, M.; Holland-Nell, K.; Beck-Sickinger, A. *G. Langmuir* **2006**, *22*, 8104–8108.
- (9) Dyson, H. J.; Wright, P. E. *Curr. Opin. Struct. Biol.* **2002**, *12*, 54–60.
- (10) Shoemaker, B. A.; Portman, J. J.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8868–8873.
- (11) Gupta, N.; Irbäck, A. *J. Chem. Phys.* **2004**, *120*, 3983–3989.
- (12) Willett, R. L.; Baldwin, K. W.; West, K. W.; Pfeiffer, L. N. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7817–7822.
- (13) Peelle, B. R.; Krauland, E. M.; Wittrup, K. D.; Belcher, A. M. *Langmuir* **2005**, *21*, 6929–6933.
- (14) Irbäck, A.; Samuelsson, B.; Sjunnesson, F.; Wallin, S. *Biophys. J.* **2003**, *85*, 1466–1473.
- (15) Irbäck, A.; Mohanty, S. *Biophys. J.* **2005**, *88*, 1560–1569.
- (16) Gökoğlu, G.; Bachmann, M.; Çelik, T.; Janke, W. *Phys. Rev. E.* **2006**, *74*, 041802.
- (17) Némethy, G.; Gibson, K. D.; Palmer, K. A.; Yoon, C. N.; Paterlini, G.; Zagari, A.; Rumsey, S.; Scheraga, H. A. *J. Phys. Chem.* **1992**, *96*, 6472–6484.
- (18) Cormack, A. N.; Lewis, R. J.; Goldstein, A. H. *J. Phys. Chem. B* **2004**, *108*, 20408–20418.
- (19) Liang, T.; Walsh, T. R. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4410–4419.
- (20) Braun, R.; Sarikaya, M.; Schulten, K. *J. Biomater. Sci., Polym. Ed.* **2002**, *13*, 747–757.
- (21) Vrbova, T.; Whittington, S. G. *J. Phys. A* **1996**, *29*, 6253–6264.
- (22) Bachmann, M.; Janke, W. *Phys. Rev. Lett.* **2005**, *95*, 058102.
- (23) Bachmann, M.; Janke, W. *Phys. Rev. E.* **2006**, *73*, 041802.
- (24) Bachmann, M.; Janke, W. *Phys. Rev. E.* **2006**, *73*, 020901(R).
- (25) Muthukumar, M. *J. Chem. Phys.* **1995**, *103*, 4723–4731.
- (26) Bratko, D.; Chakraborty, A. K.; Shakhnovich, E. I. *Chem. Phys. Lett.* **1997**, *280*, 46–52.
- (27) Golumbfskie, A. J.; Pande, V. S.; Chakraborty, A. K. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11707–11712.
- (28) Bogner, T.; Degenhard, A.; Schmid, F. *Phys. Rev. Lett.* **2004**, *93*, 268108.
- (29) Kriksin, Y. A.; Khalatur, P. G.; Khokhlov, A. R. *J. Chem. Phys.* **2005**, *122*, 114703.
- (30) Moghaddam, M. S.; Chan, H. S. *J. Chem. Phys.* **2006**, *125*, 164909.
- (31) Zhdanov, V. P.; Kasemo, B. *Proteins* **1998**, *30*, 168–176.
- (32) Castells, V.; Yang, S.; Van Tassel, P. R. *Phys. Rev. E.* **2002**, *65*, 031912.
- (33) Marinari, E.; Parisi, G. *Europhys. Lett.* **1992**, *19*, 451–458.
- (34) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776–1783.
- (35) Favrin, G.; Irbäck, A.; Sjunnesson, F. *J. Chem. Phys.* **2001**, *114*, 8154–8158.
- (36) Irbäck, A.; Mohanty, S. *J. Comput. Chem.* **2006**, *27*, 1548–1555. The PROFASI package is freely available at <http://cbbp.thep.lu.se/activities/profasi>.
- (37) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195–1198.
- (38) Miller, R. G. *Biometrika* **1974**, *61*, 1–15.
- (39) Kirkpatrick, S.; Gelatt, C. D., Jr.; Vecchi, M. P. *Science* **1983**, *220*, 671–680.
- (40) DeLano, W. L. The PyMOL Molecular Graphics System (2002) on World Wide Web <http://www.pymol.org>.