

Genetically engineered gold-binding polypeptides: structure prediction and molecular dynamics

ROSEMARY BRAUN¹, MEHMET SARIKAYA² and KLAUS SCHULTEN^{1,*}

¹ *Beckman Institute and Department of Physics, University of Illinois, Urbana, IL 61801, USA*

² *Materials Science and Engineering, University of Washington, Seattle, WA 98195, USA*

Received 21 November 2001; accepted 5 March 2002

Abstract—The biological control of inorganic crystal formation, morphology, and assembly is of interest to biologists and biotechnologists studying hard tissue growth and regeneration, as well as to materials scientists using biomimetic approaches for the control of inorganic material fabrication and assembly. Biomimetics requires an accurate understanding of natural mechanisms at the molecular level. Such understanding can be derived from the use of metal surfaces to study surface recognition by proteins together with combinatorial genetics techniques for the selection of suitable peptides. Polymerization of these peptides produces engineered polypeptides large enough to encode their own folding information with low structural complexity, while enhancing binding affinity to surfaces. The low complexity of such polypeptides can aid in analyses, leading to modeling and eventual manipulation of the structure of the folded polypeptide. This paper presents structure predictions for gold-binding protein sequences, originally selected by combinatorial techniques. Molecular dynamics simulations lasting 5 ns were carried out using solvated polypeptides at the gold surface to assess the dynamics of the binding process and the effects of surface topography on the specificity of protein binding.

Key words: Molecular dynamics; gold; protein design; surface binding; biomimetics.

INTRODUCTION

Many biological hard tissues (bones, dentin, enamel, spicules, particles, and spines) contain proteins in addition to inorganic minerals [1]. Proteins are essential in the formation of biological materials through their control of nucleation, growth, morphogenesis, phase organization, and distribution [2–4]. Genetic engineering techniques (cell-surface display [5] and phage display [6]) have led to the isolation of simple proteins with specific binding affinity to selected practical inorganics.

*To whom correspondence should be addressed. E-mail: kschulte@ks.uiuc.edu

Using these proteins as building blocks opens avenues to the engineering of materials with novel properties.

A recent example of this genetic approach is furnished by gold-binding proteins (GBPs). These proteins were selected for the ability to bind to gold in the presence of high salt concentrations, conditions under which other proteins do not exhibit gold binding [5, 7]. During the selection processes, it was found that the gold-binding polypeptides bind more strongly to gold after the surface is treated with HF to remove surface impurities, indicating that the GBPs recognize the native gold surface rather than a partially removed contaminant. Additionally, GBPs preferentially bind gold over chromium, demonstrating substrate specificity [5].

The GBP sequences contain multiple repeats of a 14–30 amino acid sequence [5, 7]. The repeating polypeptides retain their binding properties as part of other proteins (e.g. alkaline phosphatase) if they contain a sufficient number of repeats, and affinity for gold increases with the number of repeats. Interestingly, none of the GBP sequences contains cysteine, which is known to form a covalent bond to gold [8]. All the GBP sequences to date have contained methionine. The experimental literature is inconclusive as to whether methionine can form a covalent bond to gold [9, 10]; however, because the protein was released from the surface in the presence of a detergent, and because other polypeptides containing methionine did not bind to gold, it is unlikely that the methionine sulfur contributes to the binding. Whereas many proteins and self-assembled monolayers bind to gold via disulfide bonds [11–13], the nature of GBP binding is thought to differ from this well-known thiol linkage and offer a new avenue for protein–gold surface interaction. The GBP sequences are all found to be rich in serine and threonine, and physisorption of these polar side-chains onto gold may account for the observed binding.

In addition to their binding characteristics, GBPs alter gold crystal morphology under ambient conditions [7]. In the presence of GBP, gold formed large, flat hexagonal crystals displaying the {111} face. Such crystals were not seen to form in the presence of control proteins that do not bind to gold; using the well-known Faraday approach, flat crystals only form under boiling or highly acidic conditions [14]. A possible cause of flat-gold formation may be the preferential binding onto the Au{111} surface, obscuring it from further gold accretion. It is not readily apparent how the GBP adheres to gold, nor why the {111} surface would be preferred to, for example, the more sparsely populated {211} face.

Knowledge of the structure of GBP could provide insight into the mechanism by which GBP binds to gold. Molecular dynamics simulations of the predicted structure, which can give a detailed description of the dynamics over the course of several nanoseconds, could shed light on the interaction of GBP with various crystal surfaces. These simulations could also elucidate the role of water in the GBP–gold interaction.

subsequently removed, as well as waters between the GBP and Au surface. The complete system consisted of 594 protein atoms, 1457 (672) Au atoms for the {111} ({211}) surface, and approximately 13 000 water atoms, bringing the total to approximately 15 000 atoms.

Molecular dynamics simulations were carried out using the program NAMD2 [20], with the CHARMM26 force field [21]. Additional terms were necessary in order to model the metal surface. It is known that a single water molecule physisorbs onto a metal surface with an energy of between 7 and 15 kcal/mol, which while stronger than hydrogen bond energy (5 kcal/mol), is almost as large as the interaction energy between water molecules in bulk water under ambient conditions (19.8 kcal/mol) [22]. At the interface, a compromise between the forces leading to physisorption and solvation can be expected. Typically, the interfacial structure of the water decays to bulk-like properties within a few solvent diameters; on the metal side, the electron density and surface structure decay to bulk-like arrangements at yet shorter distances [23]. The interaction between the metal (which is rigid but with a diffuse electronic structure) and the adsorbates (which exhibit incoherent particle motion and localized electronic structure) may be represented by a model potential in cases where the metal's properties are not the focus of interest. The overall potential between a molecule and a metal surface, U_{surf} , may be represented [24] by a Lennard–Jones 10–4 potential,

$$U_{\text{surf}} = 2\pi \sum_{l=0}^{\infty} n_l \epsilon \sigma^2 \left[\frac{2}{5} \left(\frac{\sigma}{z + d_l} \right)^{10} - \left(\frac{\sigma}{z + d_l} \right)^4 \right], \quad (1)$$

where d_l is the separation between the first and l th lattice planes, n_l is the number density of atoms in plane l , and z is the perpendicular separation of the adsorbate particle from the first ($l = 0$) lattice plane. The parameters σ and ϵ are the Lennard–Jones 6–12 potential parameters describing the interaction between the adsorbate and surface atoms. Because there is no significant contribution from below the second lattice plane, only the first two terms of the sum need to be considered. The corrugation of the more sparsely populated (e.g. {211}) planes may be mimicked by making the d_l 's sinusoidal functions [23]. For Au, σ and ϵ are 1.8474 Å and 0.0390 kcal/mol, respectively [25].

In addition to the metal–adsorbate interactions, it is of interest to consider the electrostatic potential contribution due to the interaction of molecular point charges with the induced image charge ‘beneath’ the metal plane. For a smooth surface, this contribution is given by

$$U_{\text{surf}}^{\text{img}}(z_1, \dots, z_n) = \frac{\epsilon_m - \epsilon_s}{\epsilon_m(\epsilon_m + \epsilon_s)} \left[\sum_i \frac{q_i^2}{4z_i} + \frac{1}{2} \sum_{\substack{i,j \\ i \neq j}} \frac{q_i q_j}{|\vec{r}_{ij} + 2\vec{n}z_j|} \right], \quad (2)$$

where the first term describes the interaction of a charge q_i at a distance z_i above the image plane (generally defined to be the plane in which the nuclei of the atoms

in the topmost lattice plane lie) with its own image charge and the second term describes the interaction between a charge and the images of all other charges. \vec{r}_{ij} denotes the separation between real charges, \vec{n} is the unit vector perpendicular to the image plane; and the ϵ 's are the dielectric constant for the surface ($\epsilon_s = \infty$) and the medium ($\epsilon_m = 80$, water), respectively. However, for small distances z above the gold plane, the Lennard–Jones contribution dominates the image-charge contribution, whereas for large z , neither term contributes significantly. Furthermore, it has been observed that for a neutral many-particle system, the sum of charge–image charge interactions is close to zero [26, 27]; such cancellation has been noted even with monolayer films of water. Thus, it is reasonable to exclude the charge–image charge terms from the Hamiltonian.

Molecular dynamics simulations using force fields of this type (a smoothed potential representing the metal surface and no charge-image contribution) have been shown to reproduce accurately the experimentally measurable surface concentrations of water [28] and hydrocarbons [29]. Explicitly including the polarizability has been shown not to affect the interfacial structure [28], since the induced dipole moment of a water molecule in the polarizable models tends to point in the same direction as the permanent moment. Additionally, molecular mechanics using the above model force field has been used to calculate the minimum energy monolayer of hydrocarbon chains on MoSe₂; the resulting structure agrees well with the observed STM structure [29]. The utility of this force-field model in predicting the behavior of organic molecules in contact with Au [12, 27, 29] suggests that this force-field model accurately represents the atomic interactions.

Molecular dynamics simulations were carried out, at time steps of 1 fs, using the program NAMD2 [19] in conjunction with the Au potential for both the {111} and {211} planes, with the gold atoms held fixed to speed up computation. Each system was simulated with periodic boundary conditions, with full electrostatics computed using the Particle Mesh Ewald (PME) [30] method with a grid spacing of the order of 1 Å or less. Each system was energy-minimized using the Powell algorithm and then heated for 2 ps under Langevin dynamics at a temperature of 300 K and with a damping coefficient of 10 ps⁻¹. Each system was then equilibrated for 1 ns at constant pressure and temperature. Pressure was maintained at 1 atm using the Langevin piston method [31], while temperature coupling was enforced by velocity reassignment every 2 ps. The systems were then simulated for 4 ns further in the NpT ensemble using Langevin dynamics/Langevin piston at a temperature of 300 K and a damping coefficient of 10 ps⁻¹. Analysis of trajectories and energetics was performed using X-PLOR [19] and VMD [32].

RESULTS

Secondary structure prediction was carried out on three known GBP sequences, necessitated by lack of experimental structural data. For GBP1, there was agreement between several prediction methods indicating an antiparallel β -sheet, with each

repetition of the sequence contributing another strand to the β -sheet. For GBP2, the results for various methods differed, with some indicating helical regions where others predicted β -strands or coils. For GBP3, the neural network predictions did not arrive at any specific structure other than a random coil, and the similarity searches matched sequences of a globular character. Due to the ambiguity of GBP2 and the indeterminate structure for GBP3, we have confined our detailed studies to GBP1. Three repeats of the GBP1 14-amino acid sequence were mapped onto the antiparallel β -sheet structure suggested by the prediction tools; each repetition of the sequence contributes an additional strand to the β -sheet and is believed to contribute linearly to the binding potential based on the predicted structure and experimental results [33]. The predicted structure was minimized and checked for bad angles, contacts, etc.

The predicted GBP1 structure was placed on both $\{111\}$ and $\{211\}$ gold surfaces large enough to accommodate it and fully solvated; water molecules between the GBP and Au surface were removed. Because of the aforementioned symmetry of the structure, the GBP can be flipped by 180° without consequence. The system in its pre-equilibrium configuration is shown in Fig. 1. The stability of the predicted structure in contact with the gold surfaces can be assessed by analyzing its deviation from the initial structure over the course of the equilibration. Figure 1 depicts the root mean square deviation (RMSD) of the C_α atoms of the protein during equilibration. After 500 ps, the protein stabilizes at approximately 3 Å RMSD from its pre-equilibrium structure. Because the initial structure was a predicted, rather than a crystallographic one, such deviation is expected. Four molecular dynamics simulations lasting 4 ns were carried out on each plane. No major conformational changes were observed. Each strand of the GBP adsorbs onto the $\{111\}$ Au plane with an energy of ~ 9 kcal/mol, roughly twice the adsorption energy of water onto

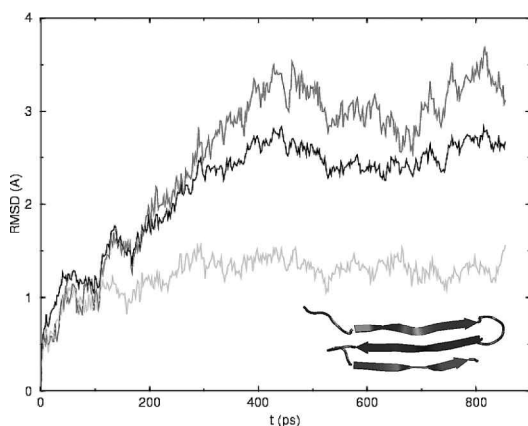


Figure 1. RMSD of C_α atoms during the equilibration relative to the predicted starting structure (black). The protein is stable after 500 ps. Polar residues (green) exhibit a small RMSD compared with the fluctuations observed in the hydrophobic residues (red). The inset depicts the initial β -sheet configuration of GBP. (This figure is published in colour on <http://www.vspub.com/jcots/JBS>)

the {111} plane [28]. It is observed that the main contributions to the adsorption energy come from the polar amino acids rather than the hydrophobic ones (Fig. 2). Visual examination of the GBP position on the gold surfaces (Fig. 3) reveals that while the protein is composed equally of polar and hydrophobic residues, the atoms in close contact with the gold surface are primarily those from the polar side-chains. Additionally, the polar residues exhibited a much smaller deviation from their initial position than the nonpolar residues (Fig. 1), indicating stability with respect to the fixed gold lattice.

The adsorption energy of GBP onto the {211} plane is much less favorable than that of GBP onto the {111} plane (Fig. 2). The difference in lattice geometry between the {111} and the {211} plane, i.e. the two-fold higher density of atoms of the {111} surface, is expected to result in a two-fold reduction of the potential. However, the observed reduction in the adsorption energy on the {211} plane cannot be fully accounted for by the lower density of gold atoms. The presence of water in the {211} surface corrugations is believed to hinder interaction of the GBP with the Au surface, thereby contributing to face-specific binding. Although water molecules within 3 Å of both GBP and the Au surface were removed at the beginning of the simulation, over the course of 4 ns water molecules diffused between the GBP and Au surface. While the main contribution to the water molecule count depicted in Fig. 4 is from water molecules at the edges of GBP, the difference between the Au{111} and Au{211} surface is primarily due to diffusion of water underneath GBP in the corrugations of the {211} plane. A sample snapshot revealing the differences in water placement is shown in Fig. 5. The specific interaction of a single polypeptide with a crystallographic surface could be studied experimentally

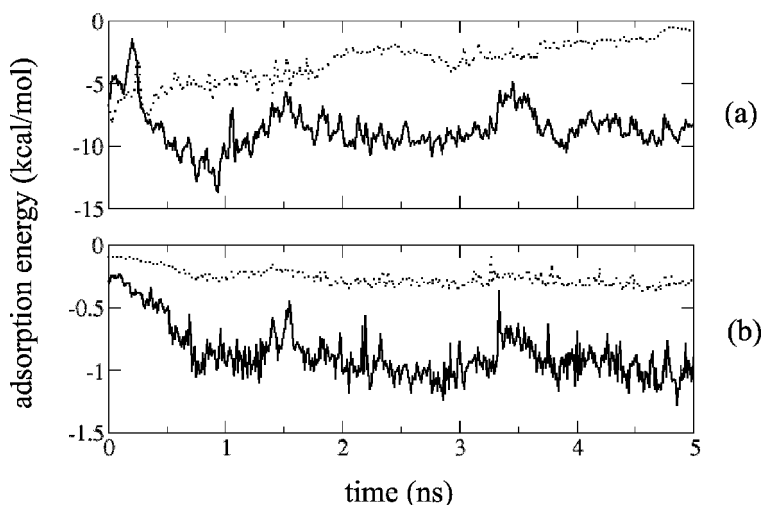


Figure 2. Adsorption energies. In (a), GBP on Au{111} is shown as a solid line, while GBP on Au{211} is shown as a dotted line. In (b), the solid line shows the average contribution to the adsorption energy on the {111} plane from Ser and Thr; the dashed line shows the average contribution from all other residues.

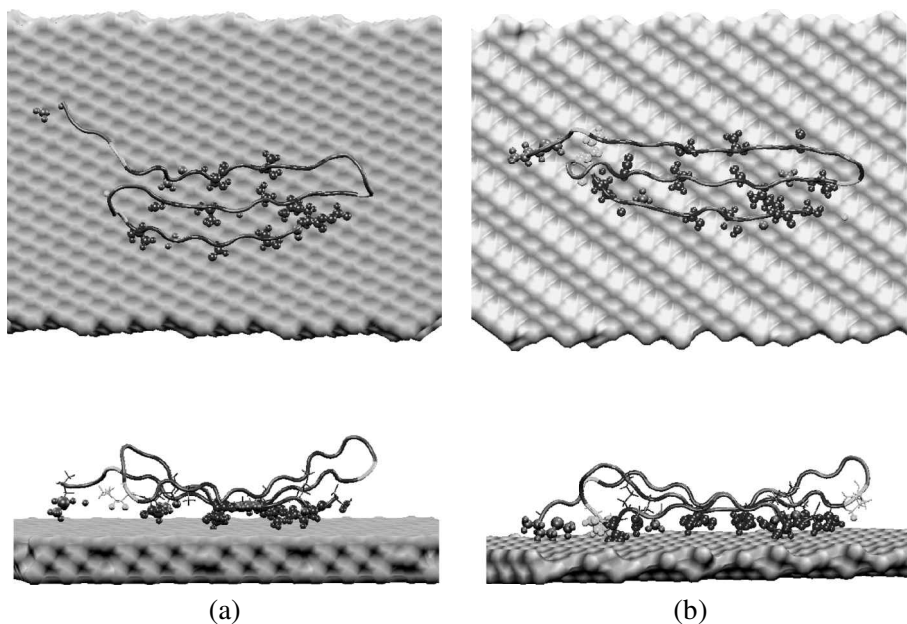


Figure 3. GBP on (a) the {111} and (b) the {211} Au surface, viewed from above (top) and edged-on (bottom). Atoms near the surface are shown as spheres. Coloring corresponds to residue type: polar residues are highlighted in blue; charged are in green; and hydrophobic are in red. (This figure is published in colour on <http://www.vspub.com/jconts/JBS>)

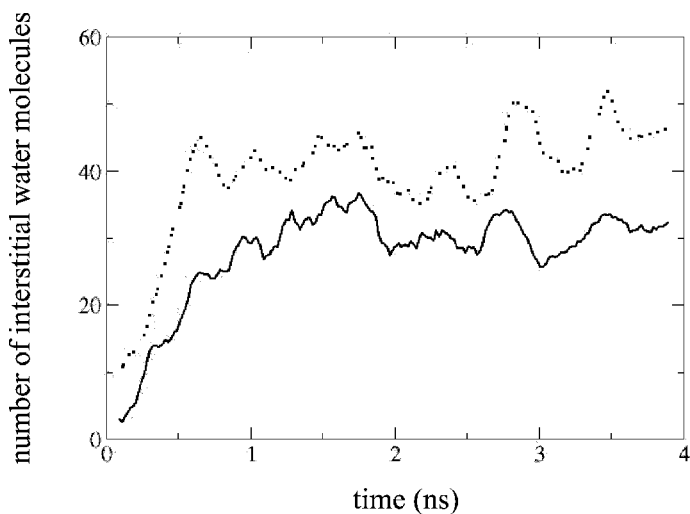


Figure 4. Number of water molecules within 3 \AA of both GBP and the Au plane. The data are smoothed by taking a running average over a 20 ps window. Data for the {111} plane are shown as a solid line; data for the {211} plane are shown as a dotted line.

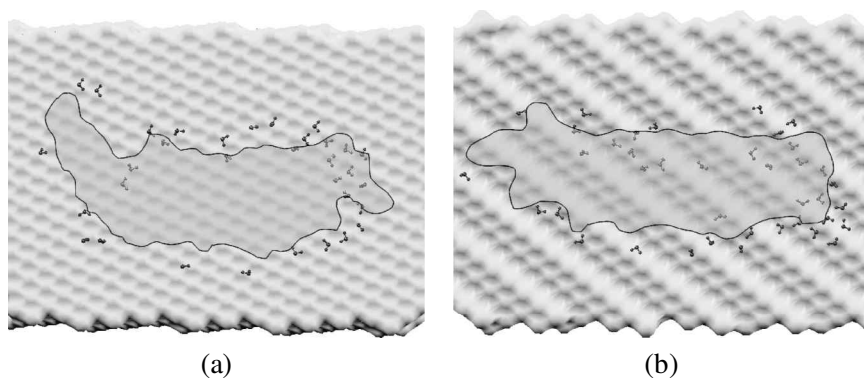


Figure 5. Water molecules between GBP and Au after 2 ns. The dark shadow (outlined) represents the GBP. Water molecules beneath the GBP are highlighted in blue. Fewer are seen on the {111} surface (a) than in the corrugations of the {211} surface (b). (This figure is published in colour on <http://www.vspub.com/jconts/JBS>)

using AFM equipped with a functionalized tip and with modeling studies using steered molecular dynamics [34] simulations in which GBP is pulled across the gold surfaces under full solvation.

CONCLUSIONS

The predicted structure and molecular dynamics simulations of the engineered polypeptide selected for its gold-binding characteristics explain the nature of the GBP–gold interaction as well as differences of the interactions between the Au{111} and Au{211} crystal surfaces. The results of the model predict a periodic structure of hydroxyls on the surface of the antiparallel β -sheet that is conducive to binding to the Au lattice. This model is similar to the electrostatic binding of organic matrix proteins to CaCO_3 surfaces [4] and the binding of small antifreeze proteins with tandem repeats onto ice crystals [35, 36]; in both cases, repetitive charged groups commensurate with the lattice spacing have been interpreted to be the cause of crystal binding. Results of the simulations of GBP1 in contact with gold surfaces show that the protein is stable. The close contacts to the gold surface come from the side-chains of the polar amino acids, consistent with the experimental observation that the proteins selected for their ability to bind to gold are richer in serine and threonine than those which fail to bind tightly to gold. Since experimental observation [33] indicates that GBP1 binding is not influenced by high salt concentrations, it may be reasonably assumed that the structural model derived from the simulation is representative of GBP1 binding at varying salt concentrations.

The nature of mineral–protein interaction is of great importance in many fields of biology and materials sciences, including biomineralization and the inhibition or promotion of ice formation in organisms, and in the assembly of functional

inorganics using polypeptides. Engineered polypeptides have great potential as model systems in fundamental studies of protein interaction with inorganic crystals. These simple proteins may serve as a conceptual model for the protein control of mineral formation in hard-tissue growth and tissue engineering, and as molecular erector sets in materials nano-assembly [37, 38]. Although spectroscopic work remains to be performed, extensive modeling as demonstrated here provides a first glimpse of the engineered protein conformation and binding to inorganic crystals, a significant step towards molecular biomimetics via the exploitation of recognition capabilities and interactions found in biological systems [38].

Acknowledgements

We wish to thank Ferenc Molnar and James Phillips (University of Illinois), Stanley Brown (University of Copenhagen), and Daniel Heidel (University of Washington) for discussions. The work of RB and KS was supported by the National Institutes of Health (NIH PHS 5 P41 RR05969), the Roy J. Carver Charitable Trust, and NRAC grant MCA935028. The work of MS was supported by ARO-DURINT.

REFERENCES

1. H. A. Lowenstam, *Science* **211**, 1126 (1981).
2. S. Mann, *Nature* **332**, 119 (1988).
3. M. Sarikaya, *PNAS* **96**, 14183 (1999).
4. G. Fallini, S. Albeck, S. Weiner and L. Addadi, *Science* **271**, 67 (1996).
5. S. Brown, *Nature Biotechnol.* **15**, 269 (1997).
6. S. R. Whaley, D. S. English, E. L. Hu, P. F. Barbara and A. M. Belcher, *Nature* **405**, 665 (2000).
7. S. Brown, M. Sarikaya and E. A. Johnson, *J. Mol. Biol.* **299**, 725 (2000).
8. P. Fenter, A. Eberhardt and P. Eisenberger, *Science* **266**, 1216 (1994).
9. R. Brizzolara, J. Boyd and A. Tate, *J. Vac. Sci. Technol. A* **15**, 773 (1997).
10. J. A. Cuff, M. E. Clamp, A. S. Siddiqui, M. Finlay and G. Barton, *Bioinformatics* **14**, 892 (1998).
11. M. Tarek, K. Tu, M. Klein and D. Tobias, *Biophys. J.* **77**, 964 (1999).
12. Z. Zhang and T. Beck, *Langmuir* **12**, 1227 (1995).
13. J. Qian, R. Hentschke and W. Knoll, *Langmuir* **13**, 7092 (1997).
14. J. Turkowitch, P. C. Stevenson and J. Hillier, *Trans. Faraday Soc. Discuss.* **11**, 55 (1951).
15. A. Strong and B. Moore, *Chem. Commun.* 473 (1998).
16. K. T. Simons, R. Bonneau, I. Ruczinski and D. Baker, *Proteins* **37** (Suppl 3), 171 (1999).
17. *QUANTA 97*, Molecular Simulations, Burlington, MA (1997).
18. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Crystallogr.* **26**, 283 (1993).
19. A. T. Brünger, *X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR*. Yale University Press, New Haven (1992).
20. L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan and K. Schulten, *J. Comp. Phys.* **151**, 283 (1999).
21. A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. W. E. Reiher, B. Roux, M. Schlenkrich,

- J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorcikiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B* **102**, 3586 (1998).
22. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
23. J. Shelley and D. Berard, *Rev. Comput. Chem.* **12**, 137 (1998).
24. R. Hentschke, *Macromol. Theory Simul.* **6**, 287 (1997).
25. A. Rappe, C. Casewitt, K. Colwell, W. Goddard and W. Skiff, *J. Am. Chem. Soc.* **114**, 10024 (1992).
26. J. Shelley, G. Patey, D. Berard and G. Torrie, *J. Chem. Phys.* **107**, 2122 (1997).
27. J. Qian, R. Hentschke and W. Knoll, *Langmuir* **13**, 7092 (1997).
28. J. Shelley and D. Berard, *Rev. Comput. Chem.* **12**, 137 (1998).
29. R. Hentschke, *Macromol. Theory Simul.* **6**, 287 (1997).
30. T. A. Darden, D. M. York and L. G. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).
31. S. E. Feller, Y. H. Zhang, R. W. Pastor and B. R. Brooks, *J. Chem. Phys.* **103**, 4613 (1995).
32. W. F. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics* **14**, 33 (1996).
33. M. Sarikaya, R. Humbert and S. Brown, unpublished results.
34. S. Izrailiev, S. Stepaniants, M. Balsera, Y. Oono and K. Schulten, *Biophys. J.* **72**, 1568 (1997).
35. C. A. Knight, C. C. Cheng and A. L. DeVries, *Biophys. J.* **59**, 409 (1991).
36. Y.-C. Liou, A. Tocilj, P. Davies and J. Zhongchao, *Nature* **406**, 322 (2000).
37. C. Mirkin and T. Taton, *Nature* **405**, 626 (2000).
38. P. Ball, *Nature* **409**, 413 (2001).