IMAGING THE MOLECULAR DIMENSIONS AND OLIGOMERIZATION OF PROTEIN MOLECULES AT THE SOLID/LIQUID INTERFACE BY SURFACE ORIENTED MOLECULAR SIZING (SOMS) MICROSCOPY

By

Mark Joseph Waner

A DISSERTATION

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

1998

ABSTRACT

IMAGING THE MOLECULAR DIMENSIONS AND OLIGOMERIZATION OF PROTEIN MOLECULES AT THE SOLID/LIQUID INTERFACE BY SURFACE ORIENTED MOLECULAR SIZING (SOMS) MICROSCOPY

By

Mark Joseph Waner

The structure and behavior of proteins at the solid/liquid interface is of great scientific interest. It has application both to fundamental biochemical understanding, as well as to biotechnological purposes. Interfaces play a critical role in many physiological processes. The mechanism of protein adsorption to surfaces is not very well understood. The current model put forth in much of the literature assumes a two step model. In the first step of this model the protein collides with the surface and adsorbs if its energy is sufficient to overcome the free energy of desorption of surface adsorbed solvent. The second step is often assumed to involve significant conformational change of the secondary and tertiary structure of the protein or enzyme, akin to denaturation. This unfolding of the protein would tend to indicate that loss of function would occur concomitantly, but studies have found very little loss in activity upon adsorption for a number of different protein systems.

The recent development of the atomic force microscope (AFM) offers another tool for the examination of protein structure at liquid/solid interfaces. For atomically flat crystals the AFM has been used to determine atomic positions to < 1 Å resolution. In the case of samples with topographic features larger than atoms, the probe tip of the AFM 'convolutes' with the size and shape of surface features. This has hindered the use of AFM for molecular level structural determination of proteins at the liquid/solid interface.

The work presented in this dissertation covers the development of the surface oriented molecular sizing (SOMS) technique which makes use of the angstrom height resolution of the AFM and a physically based mathematical framework for the analysis of the height distribution of adsorbed protein molecules. The surface adsorption and orientation (SAO) model is developed using statistical thermodynamics to model the expected height distributions for molecules adsorbed on a surface.

The SOMS technique will be shown to be viable through studies of ferritin and concanavalin A (Con A) at the water/mica interface. Using this technique we are able to determine both the three-dimensional size and the oligomerization state of the adsorbed molecules. This technique will then be utilized for the examination of denaturation of Con A at the interface, by a number of mechanisms. Further, the structural and orientational changes in Con A as a function of pH will also be presented. The final chapter of this dissertation will present an extension of these studies to the deposition and structure of Con A thin films on mica.

DEDICATION

To my parents, James and Carol, for all their love, support and encouragement. I wouldn't have made it through this without you.

ACKNOWLEDGEMENTS

I have had the great pleasure to work with and have support from many people during my graduate work at M.S.U. First, however, I must express my gratitude to those who helped me get here in the first place. The encouragement of my parents and undergraduate advisors, Dr. Nick Baumgartner and Dr. David Ewing, was integral to my decision to pursue the Ph.D. Here I would also like to mention Dr. George Leroi, for his advice and support, which began with my first visit to the department as a prospective graduate student in 1991.

I would like to thank Dr. Marcos Dantus for his insight, support and time. He has supported not only my development as a researcher, but also as an educator. Our collaborator, Dr. Melvin Schindler has been of considerable help in the development of the science that is presented in this dissertation and has provided valuable advice and critical insight. The initial work in our group on this project began with Martha Gilchrist, who helped greatly in my learning of the field. I have also had considerable assistance and support from Dr. James Geiger. He has provided insightful reviews of this manuscript and our Journal of Physical Chemistry publication, as well as allowing me the use of instrumentation and equipment in his laboratory and his biochemical expertise.

I would also like to acknowledge the other professors in the department who have been both excellent teachers and mentors. I would especially like to thank Dr. Paul Hunter, who has been a great help in my development as a teacher. I would also like to thank Dr. Katharine C. Hunt, who has been a role model both as a top-notch teacher and

researcher. I would also like to thank Dr. Gerald T. Babcock for his support during my studies, and his excellent example as a teacher and researcher.

The strong support staff in the department of chemistry have also played a key role during my stay here. The glass, machine and electronics shops have provided first class expertise. The business office and the secretarial staff have always done an outstanding job in making sure that I could get my job done on time, and their efforts are greatly appreciated.

Last, but certainly not least, I would like to thank my future wife, Kristina Henson for her love, support and understanding as I finished my research and wrote this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFM	atomic force microscope or microscopy
Con A	Concanavalin A
HEPES	N-2 hydroxyethylpiperazine N'-2 ethanesulfonic acid
HPLC	high-performance liquid chromatography
NMR	nuclear magnetic resonance
NSOM	near-field scanning optical microscope or microscopy
PSPD	position sensitive photodiode
PZT	lead zirconium titanate
QCM	quartz crystal microbalance
SANS	small angle neutron scattering
SAXS	small angle x-ray scattering
SAO	surface adsorption and orientation model
SEM	scanning electron microscope or microscopy
SFA	surface force apparatus
SOMS	surface oriented molecular sizing
SPM	scanned probe microscope or microscopy
STM	scanning tunneling microscope or microscopy
TEM	transmission electron microscope
XRD	x-ray diffraction

INTRODUCTION

In this dissertation I will discuss the work I have done over the past two and a half years, under the direction of Dr. Marcos Dantus. My interest has been focused on the better understanding of the physical characteristics of protein molecules at the liquid/solid interface. A large part of my work has been the development of a new technique for the direct observation of tertiary structure and oligomerization state of protein molecules at the solid/liquid interface using the atomic force microscope (AFM).

The first chapter of this work is devoted to a description of the current understanding of protein behavior at liquid/solid interfaces, and some of the questions that remain to be adequately understood or investigated. I will then discuss the use of the AFM as a tool for investigating biomolecular structures on surfaces and its limitations for accurate lateral dimensions of biomolecules. Chapter two will lay out the theoretical framework for the surface oriented molecular sizing (SOMS) technique and describe how this approach is used to overcome one of the shortcomings of the AFM, and the other related scanned probe microscopes (SPM). The experimental methods and materials will be described in chapter three, along with an explanation of our data analysis procedure. Chapter four will give the experimental results for SOMS measurements of ferritin and concanavalin A (Con A). The system which has been the focus of most of my work is Con A, and most of chapter four is devoted to better understanding its tertiary and quaternary structure at the water/mica interface as it adsorbs molecule by molecule. The

final chapter contains results which use the AFM, and the understanding gained by the results of chapter four, to examine the structure of Con A as it adsorbs to form monolayer films.

Chapter 1 : Perspective on the Field of Study

1.1 Proteins at interfaces

The behavior of proteins at the solid/liquid interface has been of great scientific interest, as it has application to fundamental biochemical understanding, as well as to more commercial purposes. Interfaces play a critical role in many physiological processes. Many processes involving proteins occur at or near interfaces, such as between cell membranes and the cytoplasm. For example, the cell adhesive protein fibrinogen is regarded as causing thrombosis when it adsorbs to many biomaterials [1]. On the more practical side is the interest in gaining understanding of protein structure and function at interfaces for the development of biological assays, biocompatible materials for implants and surgical instruments and even materials which might reduce contact lens fouling [2-4]. Recently there have been ACS symposia devoted specifically to the studies of proteins at interfaces [2,4].

The adsorption of a protein molecule from solution onto a solid surface is a complex process [5,6]. Current literature tends to favor a two step model for the adsorption process, illustrated schematically in Figure 1 [4]. The first step involves the

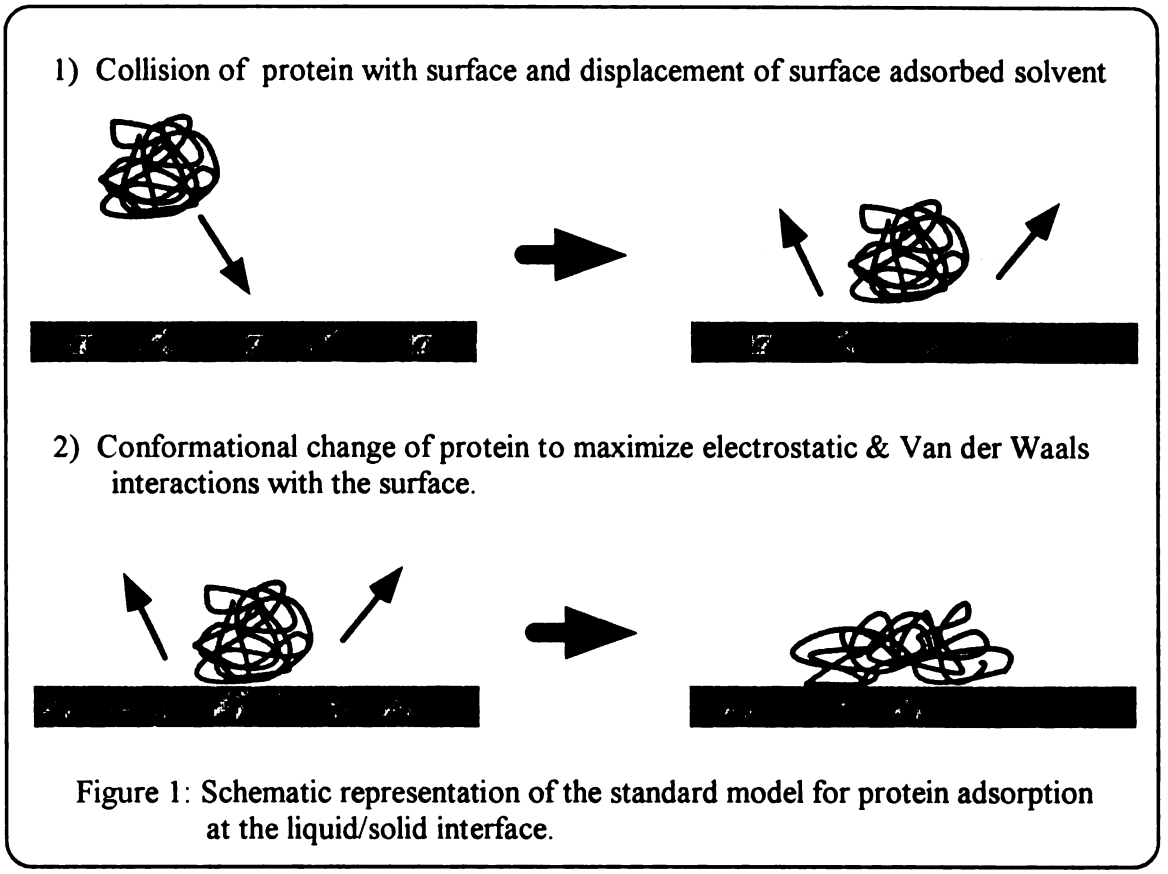


Figure 1: Schematic representation of the standard model for protein adsorption at the liquid/solid interface.

collision of the protein, which is translating by Brownian motion, with the surface. If the energetic benefit of the removal of surface adsorbed water is greater than the thermal energy of the protein molecule, then it remains on the surface, otherwise it returns to solution [6]. Once the protein is adsorbed, the second step is thought to involve a conformational change that maximizes electrostatic interactions, local interactions between electron donor moieties in the protein with electron acceptors on the surface and maximization of van der Waals forces [7]. It is often assumed that this conformational change involves fairly substantial denaturation. There is now some experimental evidence that 'compact' globular proteins will conserve their native structure when adsorbed, and that only some 'soft/flexible' proteins suffer some degree of denaturation upon adsorption, especially on hydrophobic surfaces [8]. The second step, therefore,

involves primarily an orientation process which is thermodynamically driven and yields proteins in the most energetically favorable orientation.

Much of the work in this field has been focused primarily on three areas of study: a) kinetics of the adsorption process, b) monolayer and 2D crystal growth and c) structural transitions of proteins. The bulk of the studies, up to this time, have investigated the kinetics and thermodynamics of the adsorption process and the formation of monolayer films. The study of isolated, individual protein molecules at interfaces is even more limited.

While there is much interest in the structure of biomolecules at interfaces, publications in this area are relatively few. Much of the reason for this is that the number of techniques available to quantitatively measure structure and structural changes of individual protein molecules at interfaces is very limited. Table I lists the tools most commonly used for the determination of protein structure today, as well as information about the limits of each technique. Of the techniques listed, only scanning and transmission electron microscopy (SEM and TEM, respectively), small angle x-ray and neutron scattering (SAXS and SANS, respectively), and AFM are readily utilized for the study of protein structure at interfaces. X-ray diffraction (XRD), Electrophoresis, HPLC and MS are applicable for protein structure determination in pure form or in solution. Although NMR spectroscopy primarily has been used for low molecular weight proteins in solution, new developments which increase the intensity of observed NMR signals may extend its use to the study of proteins at interfaces [9,10].

Table I: Some common techniques for the determination of protein structure.

Technique	Resolution	Concentration	Requirements/Limitations	Analysis
X-ray diffraction	Å	~100 μm crystal (~1μg)	Crystalline samples	Complex
Electrophor-esis,HPLC	N/A	≥ 0.5 mg/ml	MW measurement only	Simple
NMR	Å	≥ 1 mg/ml	Low MW, soluble proteins	Complex
SEM, TEM	Å-nm	≥ 0.02 μg/ml	≥ 300 kD, dehydrated and coated	Simple
SAXS, SANS	Å-nm	≥ 1 mg/ml	Synchotron radiation source	Complex
AFM	1 Å-20 nm	≥ 0.02 μg/ml	artifacts induced by imaging	Simple*

A further limitation of most of the techniques (SEM, TEM and AFM excluded) is that they require fairly concentrated samples of protein in order to obtain good signal to noise. The SEM and TEM both require high vacuum and conductive samples, which means samples must be chemically fixed, dehydrated and coated with graphite or gold. The only technique that has the potential to observe individual protein molecules at solid/liquid interfaces, under conditions approaching something close to physiological, is the AFM.

1.2 Atomic force microscopy (AFM)

Since its recent introduction [11], atomic force microscopy has been used extensively for imaging biological molecules [12-16]. The AFM is one member of the

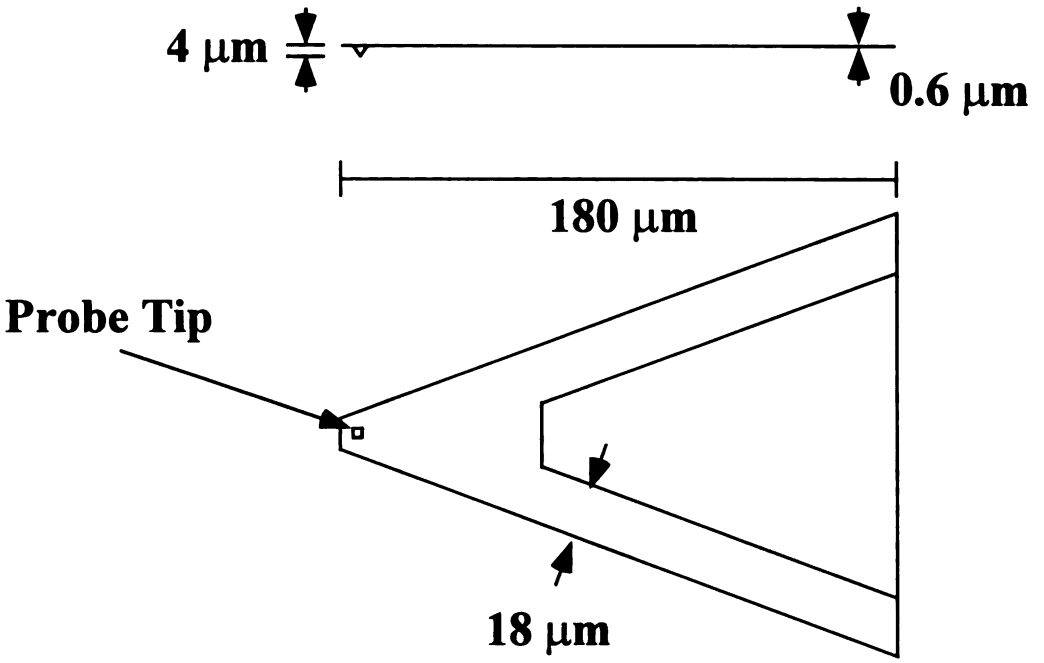


Figure 2: Scale drawing of Park Scientific Microlever type A cantilever

scanning probe microscope (SPM) family which also includes the scanning tunneling microscope (STM) and the near-field scanning optical microscope (NSOM). All of these instruments rely on raster scanning a very small (~10-100 nm) probe tip over the surface of a sample. In the case of the AFM the probe tip (~10-50 nm radius of curvature) is mounted on a cantilever arm (see Figure 2). The cantilever bends in response to the forces which the probe tip senses as it comes close to the sample surface. This interaction force between the sample and the probe tip may be an attractive or a repulsive force. At large tip sample separation there is effectively no interaction force. As the probe and sample are brought closer together (~ 5-15 nm) there is initially an attractive van der Waals interaction force. As the tip-sample separation becomes smaller the interaction force becomes repulsive because the electron clouds of the sample and the probe come

close enough for electrostatic forces to dominate. The interaction is analogous to the potential energy surface for a simple diatomic such as H_2. The analogy is not completely sound, however, due to the fact that in the AFM case there is a discontinuity due to the spring instability of the cantilever (see Figure 3, number 2). In addition the shape of the force potential for the AFM differs from that of isolated atom case due to the effect of geometry and the presence of large numbers of atoms. One of the reasons that the AFM was not developed until the mid 1980's was that the microfabrication techniques necessary for the production of small, sharp probe tips mounted on sensitive cantilevers were still being developed by the electronics industry.

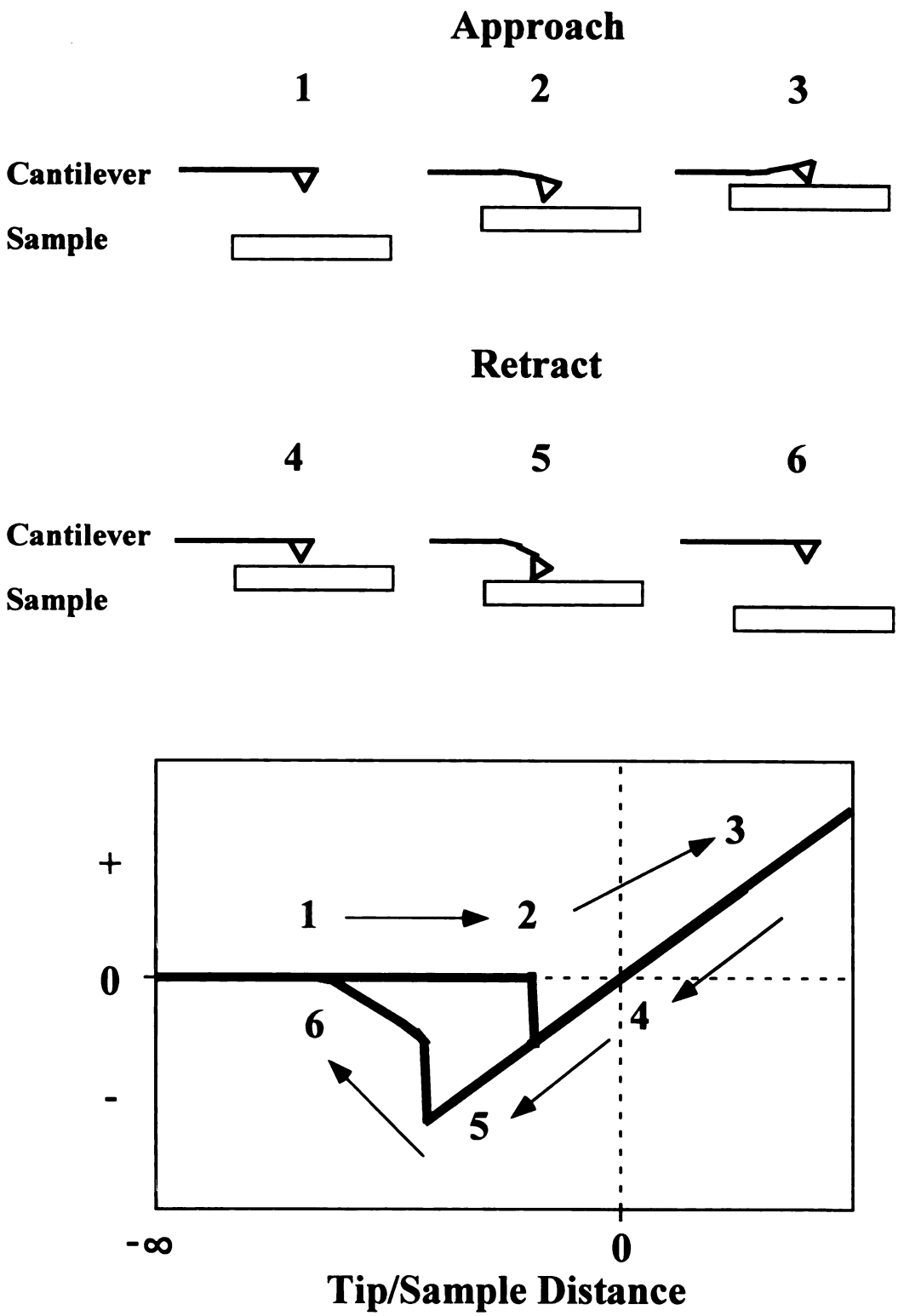
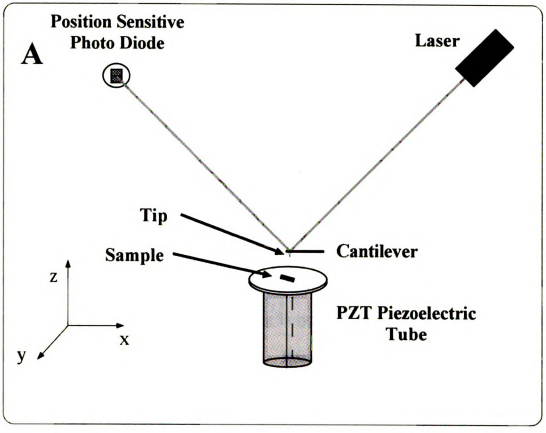


Figure 3: Schematic force vs. distance curve for AFM in air

A schematic diagram of an AFM is given in Figure 4. The key components are the piezoelectric scanner, the cantilever with integrated probe tip and the diode laser with position sensitive photodiode. The piezoelectric scanner is typically a lead-ziconate-titanate (PZT) material. A tube design is often used since it offers x,y and z motion in one device. In this design an electrode is deposited on the inside of the tube and four electrodes are deposited on the outer surface of the tube. Lateral motion of the tube is achieved by control of the voltages applied to opposite outer electrodes, relative to the inner electrode. The z motion is controlled by application of voltage to all four of the outer electrodes and the inner electrode. The diode laser and position sensitive photodiode (PSPD) system allows the measurement of the tip-sample force. The bending of the cantilever causes a deflection of the laser beam. Initially the laser, cantilever and PSPD are aligned such that the beam bounces off the end of the cantilever and onto the center of the two elements of the PSPD. The two elements of the PSPD are referred to as segments A and B. In order to minimize the effect of noise and fluctuations in the photodiodes the AFM typically makes use of the so-called error signal for the measurement of the cantilever deflection or tip-sample force. The error signal is given by:

$$Err = \frac{A - B}{A + B}$$

Where A and B are the voltages measured for each of the photodiode segments.



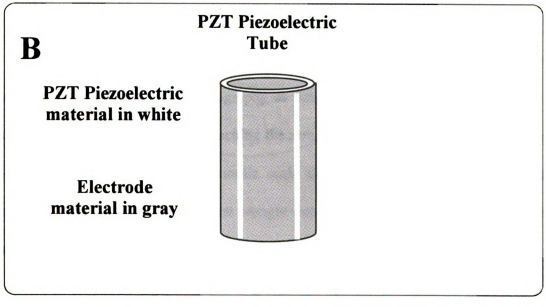


Figure 4: Schematic diagram of A) the atomic force microscope (AFM) and B) PZT scanner tube

When one speaks of using an AFM, they might be speaking of any one of a number of different modes of imaging. The common modes of imaging are: constant and variable force contact modes, non-contact mode and intermittent contact. Other related techniques include lateral force and magnetic force microscopy.

The standard mode of imaging is the constant force, contact mode of imaging. In this method the tip and sample are brought into repulsive force 'contact' and a computer controlled feedback loop adjusts the z position of the scanner in order to maintain constant force. The image formation is based on plotting the z position of the scanner at each x and y point. This mode of imaging is typically used for obtaining topographic information for a wide variety of samples. Because of the sharp tips and the applied repulsive forces this mode of imaging is often not suitable for very soft, weakly adsorbed samples, unless performed under liquid where the adhesion force is negligible.

The variable force contact mode (also known as error mode of imaging) is used when one wishes to image the topography of lattice scale features such as O^{-2} ions in mica. In this situation one is approaching the noise limit of the instrument. In order to get information at the lattice or atomic scale the AFM requires an atomically flat substrate. The tip and sample are first brought into contact. After a successful approach in a region of the sample which is flat one turns off the force feedback loop. The image formation in this mode is given by plotting the error signal at each point of the raster scan in x and y. Figure 5 presents an image of the hexagonal (001) mica lattice taken in the error mode.

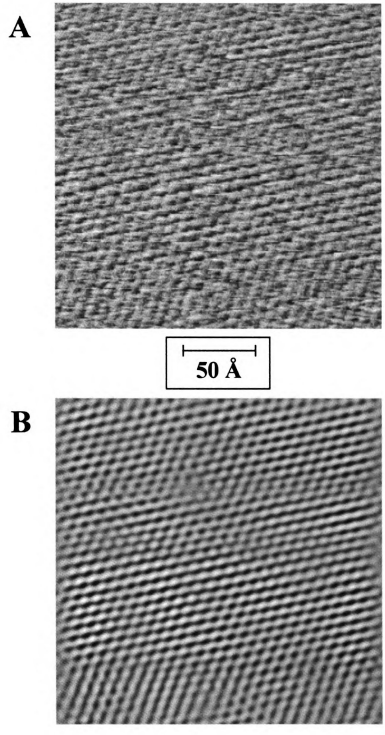


Figure 5: Error mode AFM image of the mica lattice. The raw image is shown in A), while B) is the same image after application of a Fourier filter.

The non-contact mode of imaging typically utilizes the attractive regime of the force vs. distance curves. In this mode of imaging the cantilever is mounted on a small piezoelectric actuator that is used to oscillate the cantilever near its resonant frequency. The modulation amplitude is typically on the order of tens to hundreds of angstroms. The force interactions between the tip and sample are monitored either by measuring the changes in frequency or amplitude of the cantilever oscillation. Because the forces being monitored are long-range attractive, this method of imaging applies forces on the order of piconewtons between the tip and sample. While this mode of imaging has somewhat lower imaging resolution, it often provides a very good way to image soft, elastic or weakly bound samples without inducing damage.

The intermittent contact mode (also called tapping mode) of imaging is similar to non-contact in that the cantilever is oscillated near its resonant frequency. In this case, however, the amplitude is increased and the tip-sample separation is reduced. The sample is brought close enough to the probe tip so that at the minimum of the cantilever oscillation, the tip and sample make contact. Because the tip and sample make only intermittent contact, there is less worry about damage to the sample. Also, in the ideal case the next point in the raster scan is not reached until the tip has been lifted from the surface. That means that the lateral forces, which arise from the motion of the scanning in contact mode, are eliminated.

The AFM technique allows the examination of biological specimens without the need for chemical fixation/dehydration or growth of crystals, which is normally

associated with the high-resolution techniques of SEM/TEM or XRD [12,17,18]. In fact, samples may be imaged with AFM under liquids at physiological concentrations and pH [19-22]. A number of groups have utilized the high resolution (~1 Å) of the AFM for examination of 2D crystallization of protein molecules [11,18-21]. Others, especially Hansma, Bustamante and Henderson, have studied isolated proteins [17,22-28] or DNA molecules [14,20,23-26] and their complexes [27]. Recently AFM has been used by the Marchant group to obtain valuable structural information for the multimeric Von Willebrand factor (MW~260 kD, monomer) [28-30].

In the case of low molecular weight globular proteins and other small three-dimensional structures, however, AFM has been limited by an inherent loss of resolution for lateral measurements (10-20 nm, at best). The AFM is routinely used to obtain angstrom resolution images of crystalline and SAM surfaces [16,31,32]. This angstrom level resolution, however, is only possible for atomically flat samples with highly ordered periodic features. When imaging sample features which are not highly ordered and rough, on the scale of the probe tip, lateral resolution is limited by the finite size of the probe tip (~10-50 nm radius of curvature) as well as the sample being examined [30,33-35] (see Figure 6). Figure 6 shows an ideal probe tip with radius of curvature of 20 nm and the path it would follow when scanning a small spherical object of radius 6.5 nm. While the lateral information for the sample object is lost, the height of the object is measured with a high degree of accuracy. The problem of tip convolution in AFM imaging manifests itself in a few common ways. Three examples of tip convolution are illustrated in Figure 7. In Figure 7A an image of tobacco mosaic virus is presented, along

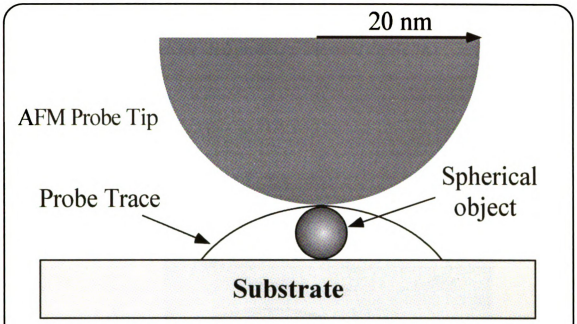


Figure 6: Lateral tip/sample 'convolution' in AFM

with a cross sectional line scan taken at the area highlighted by the white line. The tobacco mosaic virus is a cylindrical rod with a radius of 180 Å. The line scan shows that the height is measured to be 180 Å, but the lateral dimension is measured to be 350 Å. Another example of a tip convolution manifestation is presented in Figure 7B. In this case the shape of the probe tip deviates from the ideal spherical shape. This type of tip produces an image with protrusions in the surface that show a repeating shape (e.g. note the triangular shape that appears in each of the lightly shaded features). A third manifestation of the tip convolution effect is presented in Figure 7C. In this image each raised surface feature seems to be repeated. This is the result of a so-called double tip, in which the probe tip has been damaged and there are effectively two probe tips that can interact with the sample surface. While Figures 7B and C highlight artifacts due to bad

probe tips, it is important to remember that this convolution effect leads to limitations for lateral measurements even for ideal probes (as shown in Figure 7A).

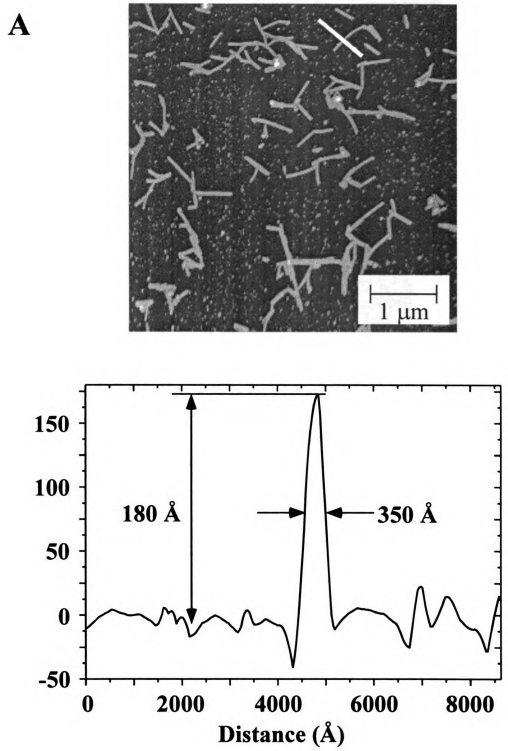


Figure 7A: Image of tobacco mosaic virus on mica illustrating the effect of lateral convolution of the AFM probe tip.

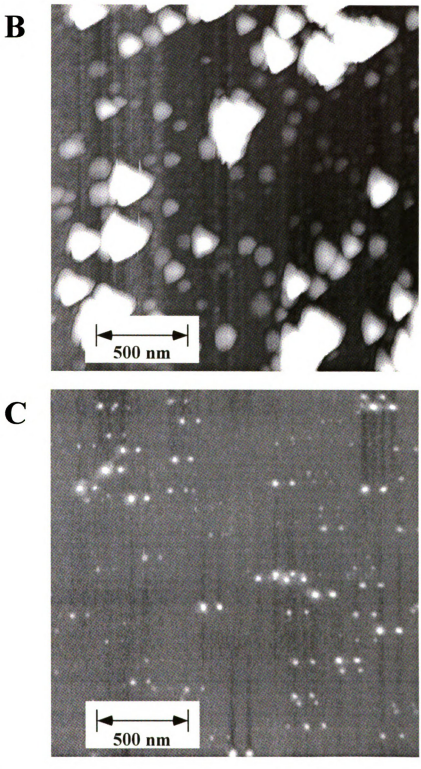


Figure 7 B & C: Effect of tip convolution problem when the probe tip is B) non-spherical or when there is a C) double tip.

This 'convolution' effect occurs regardless of imaging conditions (vacuum, ambient or under liquid) and has limited the applicability of the AFM for the study of biomolecules at interfaces. Some groups have approached this problem by making sharper probe tips [35] which has led to commercially available tips with radius of curvature as low as 10 nm, whereas many of the initially available tips were 40-50 nm in radius. Still others have examined ways to first characterize the shape of the probe tip and then mathematically 'deconvolute' their AFM data [34-36]. This is still problematic for our purposes since the probe tips are fragile enough that they may change size or shape, even over a short period of use. Further, this 'convolution' effect is not a true mathematical convolution of two functions rather it involves a shadowing effect of the tip as it interacts with the sample. Because of this, the 'deconvolution' of data leads to the determination of an upper limit to the lateral dimensions. The approach which is taken in the current work is to take advantage of the angstrom level resolution of the AFM for measurements of height, which is essentially independent of the probe size (please see Chapter 2.3 for more detail).

1.3 Proposal of SOMS microscopy

This dissertation presents a new approach for making very high resolution (<5 Å) three-dimensional measurements of protein molecules at liquid/solid interfaces, using atomic force microscopy (AFM). Such measurements are of particular biological relevance since the preponderance of biological activities of enzymes and their binding properties occur on membrane or cytoskeletal surfaces. To avoid the loss of resolution

that is normally introduced into AFM measurements as a result of probe tip geometry we have devised a measurement strategy that measures the heights of an ensemble of individual protein molecules that have adsorbed to the atomically flat surface of mica. A scale model of an ellipsoid (of dimensions 30 x 45 x 75 Å) circumscribed by a parallelepiped and adsorbed to a mica surface in three different orientations is presented in Figure 8. Also shown in Figure 8 is a rendering of an AFM probe tip (drawn to scale) positioned over one of the protein molecules which illustrates that the height of the molecule is accurately probed by the AFM even though the lateral dimensions will be overestimated by the convolution effect.

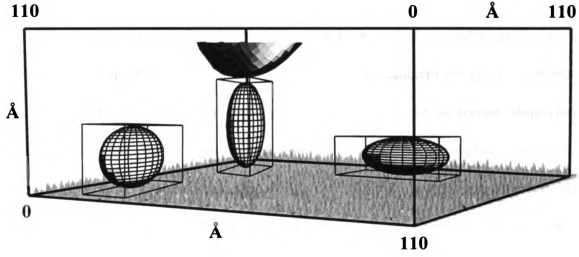


Figure 8: Scale model of AFM probe tip scanning ellipsoidal particles adsorbed to mica in three different orientations

The technique has been termed Surface Oriented Molecular Sizing (SOMS) because it utilizes the statistical thermodynamic distribution of orientations for adsorbed molecules to determine the three dimensional structure of proteins. Further, the statistical thermodynamic nature of the surface adsorption process can be used to estimate the surface adsorption energy for different protein orientations at the interface. The use of the projected height of individual adsorbed protein molecules as the defining parameter for molecular dimensions makes this approach minimally dependent on probe tip geometry and utilizes the angstrom resolution of the AFM in the z direction. Indeed, it is advantageous to use a broader tip for making these measurements, thereby minimizing the applied pressure between the tip and protein and decreasing the chance of underestimating the height of a molecule due to the digitization inherent in the microscope (see section 2.2). Furthermore, this technique is generally applicable to contact and intermittent contact modes of imaging under ambient or liquid conditions. The relative ease, speed, and resolution with which these molecular dimension measurements may now be performed make this technique particularly useful for examining multiple conformational equilibria of adsorbed protein subunits and the changes in shape and assembly state of multi-subunit proteins resulting from oligomerization and aggregation at liquid/solid interfaces. Such changes may be systematically examined as a function of protein concentration and ligand or allosteric activator binding. Such measurements could help to determine whether single amino acid substitutions affect protein activity as a consequence of an induced change in the oligomerization state of the mutant protein or as a result of a chemical change in the active site. The studies in this dissertation focus primarily on the structure of

concanavalin A (Con A) molecules at the water/mica interface, although results for ferritin are also presented.

1.4 References

1. Feng, L.; Andrade, J. D. Structure and Properties of Fibrinogen. In *Proteins at Interfaces II, Fundamentals and Applications*; Horbett, T. A., Brash, J. L., Eds.; American Chemical Society: Washington, 1995; Vol. ACS Symposium Series 602; pp 66.

2. Horbett, T. A.; Brash, J. L. Proteins at Interfaces: Current Issues and Future Prospects. In *Proteins at Interfaces, Physicochemical and Biochemical Studies*; Brash, J. L., Horbett, T. A., Eds.; American Chemical Society: Washington, 1987; Vol. ACS Symposium Series 343; pp 1.

3. Norde, W. *Adv. Coll. Inter. Sci.* **1986**, 25, 267.

4. Brash, J. L.; Horbett, T. A. Proteins at Interfaces: An Overview. In *Proteins at Interfaces II, Fundamentals and Applications*; Horbett, T. A., Brash, J. L., Eds.; American Chemical Society: Washington, 1995; Vol. ACS Symposium Series 602; pp 1.

5. Elwing, H.; Askenda, A.; Ivarsson, B.; Nilsson, U.; Welin, S.; Lundstrom, I. Protein Adsorption on Solid Surfaces: Physical Studies and Biological Model Reactions. In *Proteins at Interfaces, Physicochemical and Biochemical Studies*; Brash, J. L., Horbett, T. A., Eds.; American Chemical Society: Washington, 1987; Vol. ACS Symposium Series 343; pp 468.

6. Nadarajah, A.; Lu, C. F.; Chittur, K. K. Modeling the Dynamics of Protein Adsorption to Surfaces. In *Proteins at Interfaces II, Fundamentals and Applications*; Horbett, T. A., Brash, J. L., Eds.; American Chemical Society: Washington, 1995; Vol. ACS Symposium Series 602; pp 181.

7. van Oss, C. J.; Wu, W.; Giese, R. F. Macroscopic and Microscopic Interactions Between Albumin and Hydrophilic Surfaces. In *Proteins at Interfaces II, Fundamentals and Applications*; Horbett, T. A., Brash, J. L., Eds.; American Chemical Society: Washington, 1995; Vol. ACS Symposium Series 602; pp 80.

8. Yan, G.; Li, J.-T.; Huang, S.-C.; Caldwell, K. D. Calorimetric Observations of Protein Conformation at Solid-Liquid Interfaces. In *Proteins at Interfaces II, Fundamentals and Applications*; Horbett, T. A., Brash, J. L., Eds.; American Chemical Society: Washington, 1995; Vol. ACS Symposium Series 602; pp 256.

9. Tilton jr., R. F.; Kuntz, I. D. *Biochemistry* **1982**, 21, 6850.

10. Miller, K. W.; Reo, N. V.; Schoot Uiterkamp, A. J. M.; Stengle, D. P.; Stengle, T. R.; Williamson, K. L. *Proc. Nat. Acad. Sci. USA* **1981**, 78, 4946.

11. Binnig, G.; Quate, C. F.; Gerber, C. *Phys. Rev. Lett.* **1986**, 56, 930.

12. Lal, R.; John, S. A. *Am. J. Physiol.* **1994**, 266, C1.

13. Butt, H.-J.; Downing, K. H.; Hansma, P. K. *Biophys. J.* **1990**, 58, 1473.

14. Bustamante, C.; Vesenka, J.; Tang, C. L.; Rees, W.; Guthold, M.; Keller, R. *Biochemistry* **1992**, 31, 22.

15. Bustamante, C.; Rivetti, C. *Ann. Rev. Biophys. Biomol. Struct.* **1996**, 25, 395.

16. Louder, D. R.; Parkinson, B. A. *Anal. Chem.* **1994**, 66, 84R.

17. Schnyder, T.; Engel, A.; Lustig, A.; Wallimann, T. *J. Biol. Chem.* **1988**, 263, 16954.

18. Dykstra, M. J. *A Manual of Applied Techniques for Biological Electron Microscopy*; Plenum: New York, 1993.

19. Weisenhorn, A. L.; Drake, B.; Prater, C. B.; Gould, S. A. C.; Hansma, P. K.; Ohnesorge, F.; Egger, M.; Heyn, S. P.; Gaub, H. E. *Biophys. J.* **1990**, 58, 1251.

20. Hansma, H. G.; Vesenka, J.; Siegerist, C.; Kelderman, G.; Morret, H.; Sinsheimer, R. L.; Elings, V.; Bustamante, C.; Hansma, P. K. *Science* **1992**, 256, 1180.

21. Putman, C. A. J.; Van der Werf, K. O.; De Grooth, B. G.; Van Hulst, N. F.; Greve, J. *Appl. Phys. Lett.* **1994**, 64, 2454.

22. Weisenhorn, A. L.; Maivald, P.; Butt, H.-J.; Hansma, P. K. *Phys. Rev. B* **1992**, 45, 11226.

23. Benzanilla, M.; Manne, S.; Laney, D. E.; Lyubchenko, Y. L.; Hansma, H. G. *Langmuir* **1995**, 11, 655.

24. Hansma, H. G.; Sinsheimer, R. L.; Li, M. Q.; Hansma, P. K. *Nucleic Acids Res.* **1992**, 20, 3585.

25. Henderson, E. *Nucleic Acids Res.* **1992**, 20, 445.

26. Shaiu, W.-L.; Vesenka, J.; Jondle, D.; Henderson, E.; Larson, D. D. *J. Vac. Sci. Technol. A* **1993**, 11, 820.

27. Wyman, C.; Rombel, I.; North, A. K.; Bustamante, C.; Kustu, S. *Science* **1997**, 275, 1658.

28. Marchant, R. E.; Lea, A. S.; Andrade, J. D.; Bockenstedt, P. *J. Coll. Inter. Sci.* **1992**, 148, 261.

29. Siedlecki, C. A.; Lestini, B. J.; Kottke-Marchant, K.; Eppel, S. J.; Wilson, D. L.; Marchant, R. E. *Blood* **1996**, 88, 2939.

30. Eppel, S. J.; Zypman, F. R.; Marchant, R. E. *Langmuir* **1993**, 9, 2281.

31. Alves, C. A.; Smith, E. L.; Porter, M. D. *J. Am. Chem. Soc.* **1992**, 114, 1222.

32. Bottomley, L. A. *Anal. Chem.* **1998**, 70, 425R.

33. Allen, M. J.; Hud, N. V.; Balooch, M.; Tench, R. J.; Seikhaus, W. J.; Balhorn, R. *Ultramicroscopy* **1992**, 42B, 1095.

34. Markiewicz, P.; Goh, M. C. *Langmuir* **1994**, 10, 5.

35. Keller, D. J.; Franke, F. S. *Surf. Sci.* **1993**, 294, 409.

36. Wilson, D. L.; Dalal, P.; Kump, K. S.; Benard, W.; Xue, P.; Marchant, R. E.; Eppell, S. J. *J. Vac. Sci. Technol. B* **1996**, 14, 2407.

Chapter 2: Conceptual and Theoretical Framework

2.1 Introduction to the Surface Oriented Molecular Sizing technique

The loss of lateral resolution that is normally introduced into AFM measurements as a result of probe tip geometry has limited the use of the AFM for the measurement of macromolecular structure. In order to better examine protein structure with the AFM we have devised a measurement strategy that determines the heights of an ensemble of individual protein molecules that have adsorbed to the atomically flat surface of mica (see Figure 8). The technique has been termed Surface Oriented Molecular Sizing (SOMS) because it utilizes the statistical thermodynamic distribution of orientations for adsorbed molecules to determine their three dimensional structure.

In order to model the distribution of molecular orientations we needed to develop a theoretical treatment of the data. The first model used in our understanding [1] was based on the statistical orientation process that parallelepipeds experienced when being dropped onto a surface. This model made the simplification that orientation and adsorption of protein molecules might behave similarly to boxes being dropped on a surface. Large numbers of parallelepipeds, scaled to the relative dimensions of protein molecules were dropped onto a surface landing with a statistically determined distribution. It was found that most of the parallelepipeds landed such that they projected their shortest dimension from the surface, or in other words they tend to land such that the side of the box with the largest surface area is in contact with the substrate [1].

Although this initial idea was a good place to start, the extrapolation of the 'box' model to macromolecular objects, such as proteins, falls short because it fails to account for the chemical properties of the protein and the surface and the thermodynamic orientation process occurring at the surface. The model of protein adsorption presented here takes into account both statistical, as well as thermodynamic properties of the protein and surface. In this chapter the development of a surface adsorption and orientation (SAO) model which treats the orientation of protein molecules during adsorption as a statistical thermodynamic phenomenon is discussed.

2.2 Surface Adsorption and Orientation (SAO) model

In the absence of any localized recognition sites in the protein molecule, which would tend to lead to a non-statistical orientation of molecules, the ensemble of molecules is expected to assume a thermodynamically driven distribution of orientations. A surface adsorption and orientation (SAO) model is proposed, which treats this interfacial protein orientation as an optimization of free energy. Although in the case of the current work there is a lack of specific recognition sites on the protein molecules, the effect of chemical recognition between the surface and the protein molecules will be discussed. It will be shown that the SAO model can also be used to provide valuable physical insight for systems that do have preferential orientation due to some recognition between protein and substrate.

The model assumes each protein can be enclosed by a parallelepiped of dimensions a, b and c, where $a \leq b \leq c$ (Figure 9 A). The gross structure of many globular proteins is often approximated by an ellipsoid. A parallelepiped is compatible with ellipsoidal models in the limit of well-rounded edges. Furthermore, the measured height is independent of the degree to which the edges are rounded. The most highly favored interaction between the protein molecules and the hydrophilic mica surface results in orientations that maximize the surface area of contact. Therefore, a protein orientation having side bc adsorbed on the mica (where b and c are the largest dimensions by definition) is considered to be the most energetically favorable, yielding a maximum probability for measured heights corresponding to dimension a (see Figure 9 B).

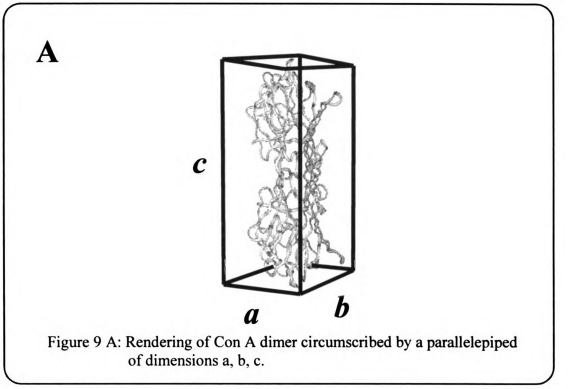


Figure 9 A: Rendering of Con A dimer circumscribed by a parallelepiped of dimensions a, b, c.

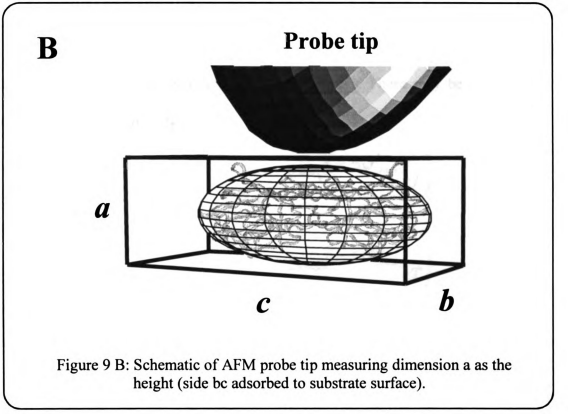


Figure 9 B: Schematic of AFM probe tip measuring dimension a as the height (side bc adsorbed to substrate surface).

The optimization of the orientation free energy, after the initial adsorption step, is assumed to be dependent on the protein surface in contact with the substrate. The free energy of orientation for each surface is given by:

$$\Delta G_{orient}^{bc} \cong \Delta G_{bc} A_{bc} \tag{1}$$

$$\Delta G_{orient}^{ac} \cong \Delta G_{bc} A_{ac} \tag{2}$$

$$\Delta G_{orient}^{ab} \cong \Delta G_{ab} A_{ab} \tag{3}$$

Where ΔG_{bc} is the free energy of orientation per unit area of the protein in contact with the substrate (in this case for side *bc*), and A is the surface area of the side identified by the subscript.

The probability of orientation such that height *a*, *b* or *c* are measured depends on the different free energies of orientation, equations 1-3. These may be estimated by a Boltzmann distribution using:

$$P(a) \cong \frac{2\exp\left(-\Delta G_{orient}^{bc} / kT\right)}{Q} \tag{4}$$

$$P(b) \cong \frac{2\exp\left(-\Delta G_{orient}^{ac} / kT\right)}{Q} \tag{5}$$

$$P(c) \cong \frac{2\exp\left(-\Delta G_{orient}^{ab} / kT\right)}{Q} \tag{6}$$

Where the factor of two arises from the symmetry of the parallelepiped, k is Boltzmann's constant, T is the temperature and Q is the canonical partition function. To this point the model does not make any assumptions about the presence or absence of particular recognition sites between the protein and the surface, since the free energy of each face of the protein is independent.

If the system lacks any specific chemical recognition sites then all the free energies of orientation per unit area are equal (*ie.* $\Delta G_o \equiv \Delta G_{bc} = \Delta G_{ac} = \Delta G_{ab}$). Under this condition equations 4-6 may be further simplified by taking the ratios of the probabilities of measuring each orientation.

$$\frac{P(a)}{P(b)} = \exp\left(\frac{[A_{ac} - A_{bc}]\Delta G_0}{kT}\right) \tag{7}$$

$$\frac{P(b)}{P(c)} = \exp\left(\frac{[A_{ab} - A_{ac}]\Delta G_0}{kT}\right) \tag{8}$$

$$\frac{P(a)}{P(c)} = \exp\left(\frac{[A_{ab} - A_{bc}]\Delta G_0}{kT}\right) \tag{9}$$

The free energies of orientation have been substituted by their corresponding expression from equations 1, 2 and 3. These equations make it possible to predict the probability of each protein orientation on the surface, given the dimensions of the protein molecule. Conversely, it is also possible to start with the experimentally observed dimensions (*a, b* and *c*) and their probabilities (*P(a)*, *P(b)* and *P(c)*) and obtain a value for ΔG_0 using Equations 7-9. The magnitude of ΔG_0 determines the relative probabilities of measuring

each of the three main dimensions of a molecule. Figure 10 illustrates the theoretical height distribution for a parallelepiped of the same dimensions as Con A dimer for values of ΔG_0 differing by a factor of two.

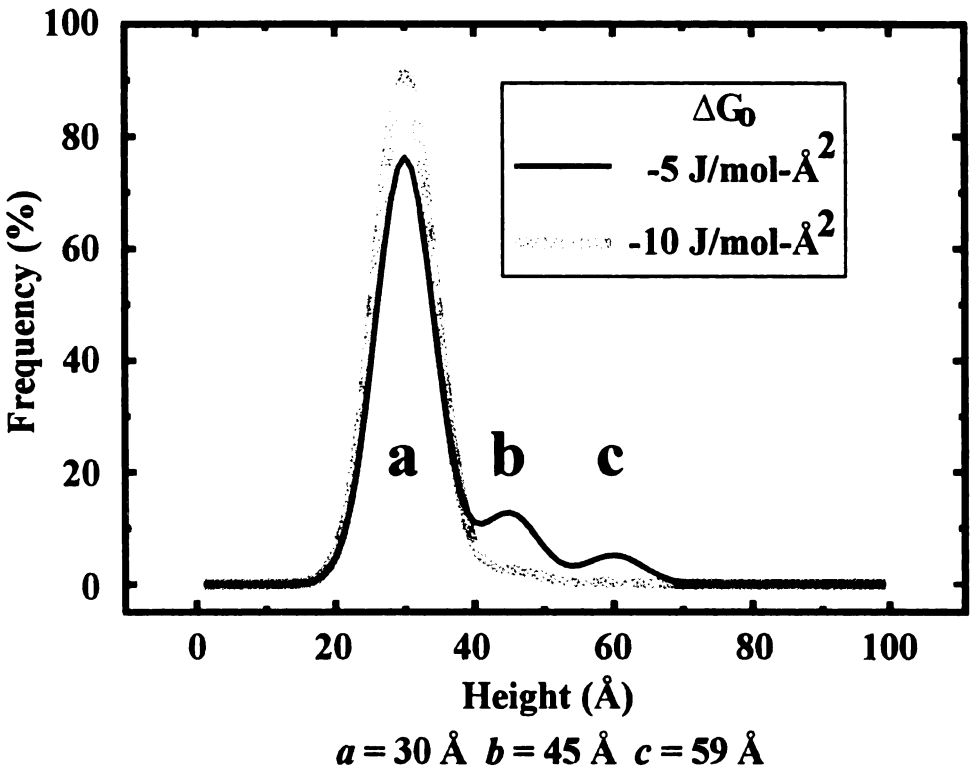


Figure 10: Height distribution predicted by SAO model, for given dimensions

The SAO model, therefore, provides a physically based mathematical framework to analyze the height distribution profile obtained from AFM measurements. This sensitivity of the observed distribution to the dimensions of the molecule used makes it possible to also determine the relative amounts of each oligomeric species present (*e.g.* monomer, dimer, tetramer). This model is useful for the determination of tertiary and quaternary structural information for protein molecules at the liquid/solid interface. It is possible to determine this structural information regardless of other physical data, such as molecular weight or XRD structure determination, so the technique could be quite useful

for the analysis of proteins which have not yielded to structural determination by other techniques.

Given a large enough sampling, on the order of 10^3 individual measurements, a distribution of measured heights is obtained which is characteristic of the molecular dimensions of the protein. This has been verified using a model height distribution and different numbers of randomly obtained data points using Mathematica 3.0 (see Appendix 2). For a system with dimensions a= 30 Å, b= 45 Å, c= 60 Å (similar to Concanavalin A which is the focus of this work) a theoretical height distribution was calculated using a sum of three Gaussian functions, of width and relative areas consistent with the experimental data to be presented in chapter 4. Then the observed height distributions for different numbers of random measurements of the ideal distribution are calculated (see Figure 11). It is quite apparent that 100 measurements is insufficient to elucidate the three-dimensional structure of this system, but by 500-1000 measurements a reasonable approximation of the actual distribution is obtained. The sum of squares was calculated for the random sampling of points compared to the ideal height distribution. The results of this calculation are plotted in Figure 12. It follows the expected logarithmic dependence on the number of data points sampled.

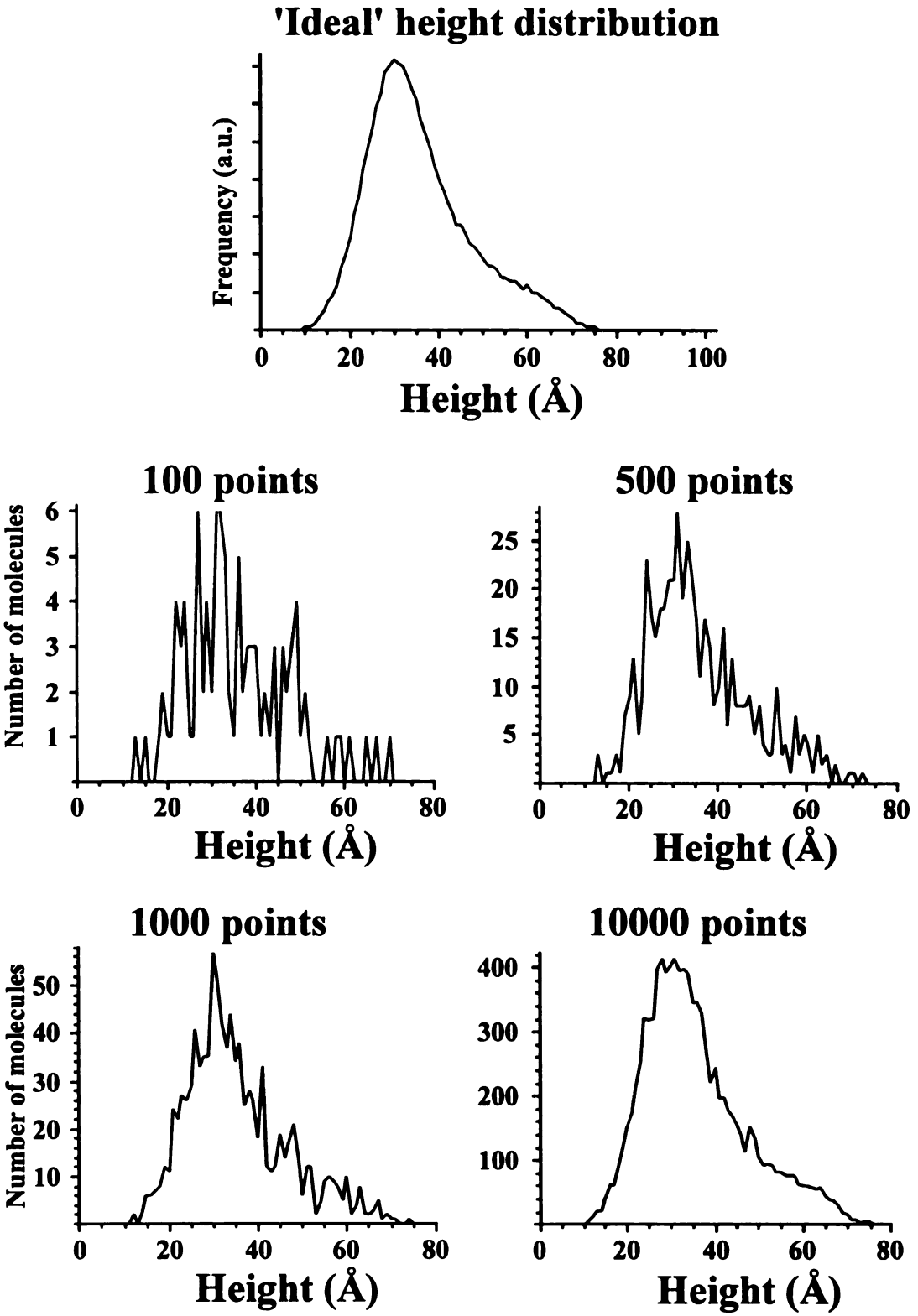


Figure 11: Experimental data modeled from an ideal height distribution

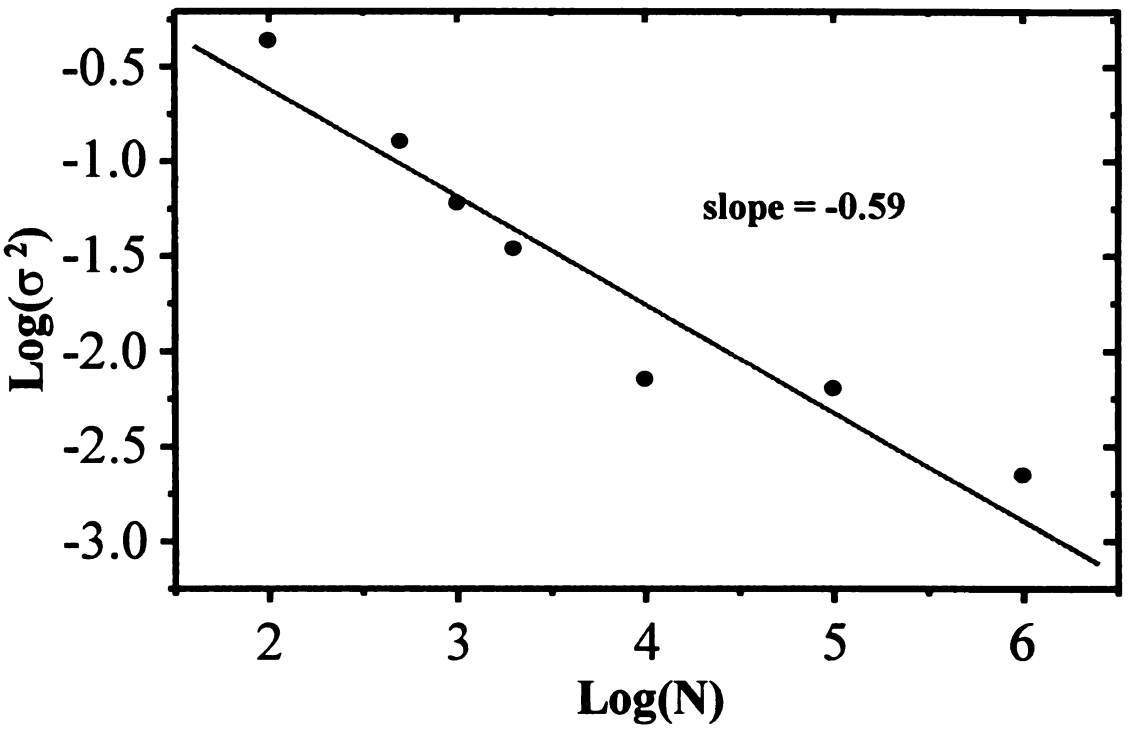


Figure 12: Log-log plot of variance versus number of measurements. See Appendix 2 for calculation.

Most proteins are asymmetrical or oligomeric, therefore for a given protein, different dimensions are obtained corresponding to different protein orientations (and/or oligomers) on the substrate. In order to interpret the height distribution and decompose it into the dimensions of the protein, the SAO model is used to predict how molecules of specific dimensions are expected to adsorb on the substrate. With this theoretical framework it is possible, not only to predict the way that a single oligomer of a protein will adsorb, but also to model the adsorption of samples containing mixtures of different oligomeric species.

2.3 Other considerations

In order to accurately measure the heights of protein molecules, there are a couple of other considerations which need to be addressed. The first is that the measurement may be highly influenced by the nature of the protein and the pressure applied to the protein by the tip during the imaging process. The second is that there is a possibility of underestimation of the actual molecular height due to the digitization of the imaging. Each of these issues will be addressed by experiments in chapter 4, but the theoretical issues are discussed here.

The AFM image contrast is based on the force interactions between the probe tip, typically made from silicon or silicon nitride, and the sample. The forces applied during the AFM imaging process are typically in the range of ~100 pN to hundreds of nN. While these forces may seem quite small, one must consider that this force is applied over a very small area at any one time. Therefore, the applied pressure can be quite large. The applied pressure for a 1 nN force and a 50 nm radius surface area is 1.27×10^5 N/m^2, which is a rather significant pressure. It has been shown that the measurement of colloidal gold is insensitive to tip/sample applied forces in standard AFM imaging conditions [2]. The structure of globular proteins, however, is not insensitive to these forces. In this range of applied pressures one needs to be concerned about the possibility of compression of the protein structure. Gekko et al. have studied the compressibility of globular proteins containing a variety of structural types [3]. This work has shown that compressibility scales with the amount of α-helical structural elements, while proteins with more β-sheet character are rather incompressible. The Con A protein is primarily ß-

sheet in nature and has very little α-helical character (see Figure 21 in Chapter four), and therefore is expected to be quite incompressible. This will be addressed by experiments examining the effect of imaging force on the observed height distribution, as well as by measurements made on succinyl Con A, which has been shown to exist only in the dimeric form [4].

Because the raster scanning process of the AFM is digitized, one must be concerned with the possibility that the probe tip is not necessarily making a measurement at every location of the sample. For example, if a particle of 32 Å diameter is on a flat substrate (see Figure 13) and one acquires a 2 μm x 2 μm image digitized to 256 x 256 pixels it is possible to underestimate the true height of the sample since the probe doesn't necessarily interact with the apex of the particle. In this case, each scan line is 7.8 nm away from the previous line. If a particle happens to lie between two consecutive scan lines and the radius of curvature of the probe tip is 20 nm, the probe tip will not interact with the apex of the particle. This problem depends on the digitization, probe tip size and the size of the sample features of interest.

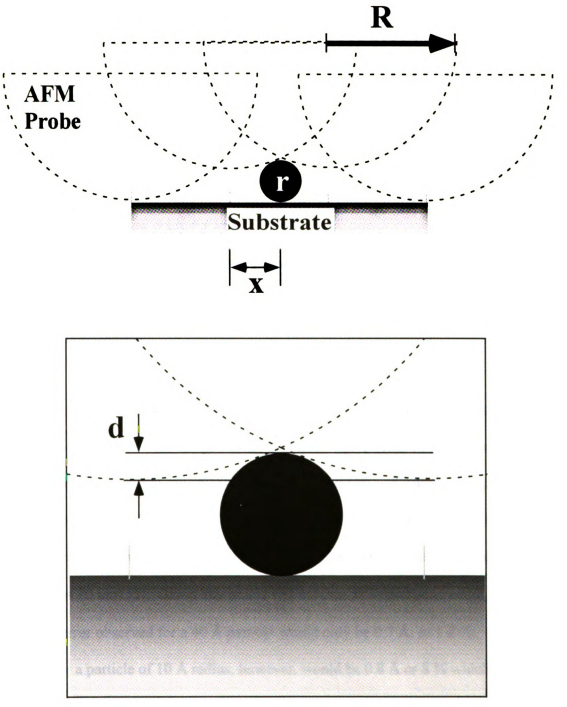


Figure 13: AFM height measurement discrepancy due to digitization

In the worst case scenario there is a possibility of completely missing a sample feature which is actually present. Of more importance for the current work and the conditions chosen for this work is that one may see the feature, but the height measurement will be underestimated since the probe tip does not necessarily interact with the highest feature of the sample particle. The underestimation of the height, or discrepancy *d*, is given by:

$$d = (R + r) - \sqrt{(R + r)^2 - X^2} \qquad (10)$$

where R is the radius of curvature of the probe tip, r is the radius of a spherical object on the surface and X is the horizontal distance between the probe and spherical object at the point of closest contact.

In order to estimate the potential discrepancy in our height measurements, 5 nm colloidal gold particles were imaged using one probe tip. Because of the tip convolution effect, the observed lateral dimensions of a spherical object will reflect not only the size of the object, but also the probe tip. Using the approach of Markiewicz and Goh, the calculated radius of curvature of the tip was 24 nm (the manufacturer's nominal value was given as 20 nm) [5]. Given this radius and a 2 μm, 512 pixel scan range, the maximum error observed for a 40 Å particle would only be 0.7 Å, or 1.8 %. The error expected for a particle of 10 Å radius, however, would be 0.8 Å or 8 % which is significant. The error associated with the height measurement of very small particles or molecules (<20 Å) can become significant. It should also be noted that the error expected for the smaller particles is within the range of the surface roughness of the mica.

2.4 References

1. Gilchrist, M. Macromolecular size determination by scanning force microscopy. M. S., Michigan State University, 1996.

2. Vesenka, J.; Manne, S.; Giberson, R.; Marsh, T.; Henderson, E. *Biophys. J.* **1993**, 65, 992.

3. Gekko, K.; Hasegawa, Y. *Biochemistry* **1986**, 25, 6563.

4. Gunther, G. R.; Wang, J. L.; Yahara, I.; Cunningham, B. A.; Edelman, G. M. *Proc. Nat. Acad. Sci. USA* **1973**, 70, 1012.

5. Markiewicz, P.; Goh, M. C. *Langmuir* **1994**, 10, 5.

Chapter 3: Experimental methods

3.1 Materials and instrumentation

Standards. Colloidal gold particles were obtained from Sigma (St. Louis, MO) and used without further purification. Three sizes of gold particles were used for the AFM height calibration: 50 ± 8 Å (3.8×10^{13} particles/ml), 90 ± 12 Å (6.0×10^{12} particles/ml), and 180 ± 8 Å (6.7×10^{11} particles/ml). The given dimensions and standard deviations were determined by the manufacturer using TEM. Tobacco Mosaic Virus (cylindrical shape; diameter 18 nm) (American Type Culture Collection) was also used to check the microscope height calibration.

Proteins. Cationized Ferritin was obtained from Polysciences (Warrington, PA) as a 20 mg/ml solution. Concanavalin A (Con A, *Canavalia ensiformis*) was obtained from Boehringer-Mannheim (Indianapolis, IN) as lyophilized solid. Succinyl Concanavalin A was obtained from Sigma (St. Louis, MO) as a 95 % protein lyophilisate. Proteins used in this work were used as obtained from the commercial source.

SDS-PAGE gel electrophoresis. Gels were prepared using the following: Sodium dodecyl sulfate (SDS), electrophoretic grade, obtained from Boehringer-Mannheim (Indianapolis, IN), Bis-Acrylamide (30-40%) purchased from Bio-Rad (Hercules, CA), TEMED (electrophoresis grade) obtained from Fisher Biotech (Springfield, NJ) and ammonium persulfate acquired from Gibco BRL (Germany). Tris base (Boehringer-Mannheim ,Indianapolis, IN) was used as the buffer for gel preparation and sample

buffer. A Mini-Protean II gel electrophoresis apparatus from Bio-Rad (Hercules, CA) was used for pouring and running the gels. Kaleidoscope pre-stained molecular weight standards (7.1-200 kD) for SDS-PAGE gel electrophoresis were obtained from Bio-Rad (Hercules, CA). Typically a 12% gel was used. The recipe for gel formation and solutions for electrophoresis was followed, with the following exception. When preparing the gels, the monomer solution was not degassed prior to adding TEMED and ammonium persulfate to initiate crosslinking.

Other chemicals. Three different buffer solutions were used for the experiments found in this work. The three were based on: HEPES obtained from Sigma (St. Louis, MO), magnesium acetate 99% from Fluka Chemika (Buchs, Switzerland) and Tris from Boehringer-Mannheim (Indianapolis, IN). The $MgCl_2$ was ACS grade from Columbus Chemical Industries (Columbus, OH). The water used for the preparation of solutions was purified with a MilliQ Type I MilliPore water purification unit (19 MΩ, resistivity). Methyl α-D mannopyranoside was obtained as 99% solid from Sigma (St. Louis, MO) and used without further purification. Muscovite mica used in this work came from two sources. High quality research grade sheets were a gift of Dr. Richard Schwendeman and natural samples were obtained from Ward's Natural Science Establishment. Highly oriented pyrolytic graphite (HOPG) was a gift of Dr. Thomas J. Pinnavaia.

Instrumentation. The microscope used for this work was an Autoprobe CP scanning probe microscope (Park Scientific Instruments). Two types of piezoelectric tube scanners (Park Scientific Instruments) were used for this work. One of these was a

five μm high resolution scanner, which had a maximum lateral range of five microns and z range of one micron. The second scanner used for some of the studies in chapter 5, was a 100 micron scanner, which was equipped with a proprietary Scan Master feature which corrects for non-linearities in the piezoelectric tube. The 100 μm scanner had a z range of about 7 μm.

The UV-vis spectrometer used for this work was an ATI Unicam dual beam, scanning spectrophotometer. It was used both in scanning and fixed wavelength mode, using air for the background determination, and the given solvent as the reference sample. The concentrations of Con A and succinyl Con A stock solutions were determined using A_{280} with an extinction coefficient of 1.37 ml mg^{-1} cm^{-1} [1].

3.2 Sample preparation

Measurements of protein height require an atomically flat substrate (less than 5Å corrugations). In most cases we found surface roughness better than 5 Å over areas of tens of microns squared. Mica cleaves very easily along the (001) plane. Along this direction of the crystal are alternating layers of aluminosilicate and potassium ions. The (001) plane resulting from the cleavage process consists of a layer of oxygen atoms. This exposed plane is quite hydrophilic and contains a significant negative surface charge, in part due to the loss of some endogenous potassium counter ions during the cleavage process. Treating a freshly cleaved surface of mica with a 5 mM $MgCl_2$ solution results in the replacement of some of the K^+ ions with magnesium, thereby reducing the negative

surface charge. The magnesium treated surface has been demonstrated to enhance adsorption of many macromolecules [2]. It is believed that the divalent Mg ions contain sufficient charge density and size to adequately replace the lost potassium ions [3]. We have examined the effect of the magnesium treatment on Con A adsorption, and found no significant changes. These results will be presented in chapter four.

Proteins were dissolved in 10 mM buffer solution and diluted to the working concentration just prior to application of the sample to the substrate. Twenty five microliters of protein solution is deposited onto the mica. Twenty minutes later, the surface is gently rinsed twice with 200 μl portions of MilliQ water and allowed to dry for an hour and a half to 8 hours at 21°C and 35-50% relative humidity. At sufficiently low protein concentrations (~10-400 ng/ml for Con A) this preparation yields a dispersed population of individual protein molecules adsorbed on the magnesium treated mica surface (Figure 14). The concentration of protein, in each set of experiments has been chosen to provide a large population of individual protein molecules and to avoid the seeding of aggregates and monolayers. Overall, the protein coverage for SOMS measurements is always kept below 1%.

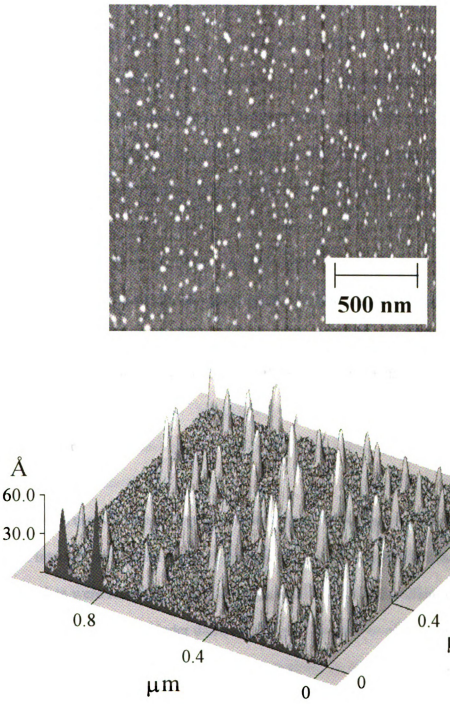


Figure 14: AFM images of Con A molecules on mica, shown in standard and three-dimensional renderings.

The concentration of stock solutions prepared from the lyophilized Con A (Boehringer-Mannheim) was found to vary, even when the same mass of lyophilisate and volume of buffer were used. It was not unusual to observe up to an order of magnitude difference from preparation to preparation. It is believed that this is due to an inhomogeneous mixing of the Con A with the residual sodium chloride during the freeze-drying process. This problem was avoided by careful measurement of concentration by A_{280} and by homogenization of the lyophilisate with a spatula prior to weighing out of sample.

In Chapter 5 data examining the orientation and oligomerization state of Con A molecules as they aggregate to form monolayers on a surface, will be presented. The experimental details for these experiments will be discussed in chapter 5.

3.3 AFM imaging

Height measurements may be obtained from AFM studies carried out by contact mode or intermittent contact mode in ambient or under liquid conditions. For this study the AFM data were recorded in contact mode under controlled ambient conditions (~35-50% humidity and 21°C). Under these conditions a monolayer of water exists at the mica surface [4] and the protein molecules are expected to remain hydrated. This limited hydration is expected to stabilized the adsorbed protein molecules, and will be discussed further in chapter four. It was found that the protein molecules adhered to the mica well enough for measurements only when the humidity was less than ~50%. At higher

humidities, multilayers of water at the surface will increase the total attractive capillary forces between probe and sample [4-7]. Humidities of less than 30% were avoided for two reasons. First, at much lower humidity the protein may become dehydrated, leading to possible structural deformations. Also, at humidities below 30%, static repulsive forces between mica and the probe tip become large enough that the long range forces dominate the tip-sample interactions. The influence of protein hydration on structure has been previously studied by scanning tunneling microscopy (STM) [8,9]. The largest structural changes were found to occur as the humidity dropped below 30%. Unfortunately those measurements were not able to provide quantitative structural information because the contrast mechanism (tunneling current) is highly dependent on the amount of solvent present.

For the measurements presented, sharpened Microlever cantilevers (Park Scientific Instruments) having a spring constant of approximately 0.05 N/m and an integrated silicon nitride probe with a radius of curvature of approximately 20 nm, according to the manufacturer, were used as probes (see Figure 2). The piezoelectric scanner was calibrated for vertical measurements using Tobacco Mosaic Virus (TMV) and/or colloidal gold particles. Horizontal measurements were calibrated using a two-dimensional grating with one μm periodicity in x and y.

AFM images were acquired in constant force mode. In this mode of imaging, the computer control includes a feedback loop that measures the cantilever deflection (proportional to the A-B/A+B signal from the position sensitive photodiode) and moves

the scanner up or down in the z direction, so as to maintain a constant applied tip-sample force. The minimum applied force required for stable imaging was found to occur within the electrostatic repulsive regime and was constant for all images. This net force was ~0.1 nN (according to the ProScan 1.5 software from Park Scientific Instruments), which may vary by as much as a factor of two from cantilever to cantilever due variability in the manufacturing process. The total force applied between the tip and the cantilever under our imaging conditions, also contains a large attractive or capillary force (~10^{-9} - 10^{-6} N, depending on the ambient humidity [10]) in addition to the overall repulsive force. The effect of the tip-sample applied force on the AFM measurements for both ferritin and Con A will be presented in chapter four.

The scan size for all the data scans presented, was 2.0 x 2.0 μm and 512 x 512 pixels, which was imaged at a linear scan rate of 1.5 Hz (3 μm/s along the fast-scan direction). This relatively slow scan rate was chosen in order to minimize the lateral forces which occur between the tip and sample during the raster scanning process. The lateral forces might be further minimized through the use of intermittent or non-contact modes of imaging, or by imaging under liquid.

3.4 Data Analysis

The raw images from the microscope contain line to line offset variation in contrast due to the scaling used by the ProScan data acquisition software (Park Scientific Instruments) and instrumental drift. Often there is also a slope to the image background.

This occurs when a sample is not normal to the *x-y* plane of the scanner (see Figure 15 A). In addition, there may also be a parabolic curvature to the image data along the scan direction. This artifact arises, in large part, due to rocking motion of the scanner tube during its lateral motion (see Figure 15 B). Each of these artifacts can be compensated for by using a 1^{st} or 2^{nd} order polynomial subtraction made to each scan line of the image. This background correction is performed using Pro Scan image analysis software from Park Scientific Instruments. The net effect of this process is that the average background level of each scan line is brought to the same value.

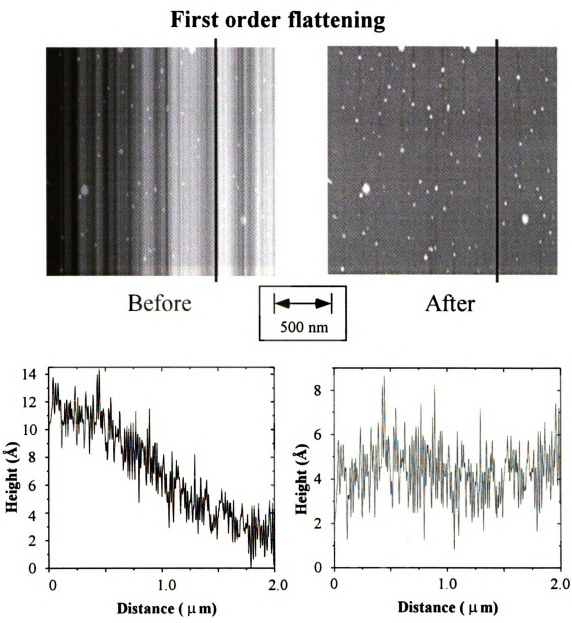


Figure 15 A: Example of image requiring first order flattening

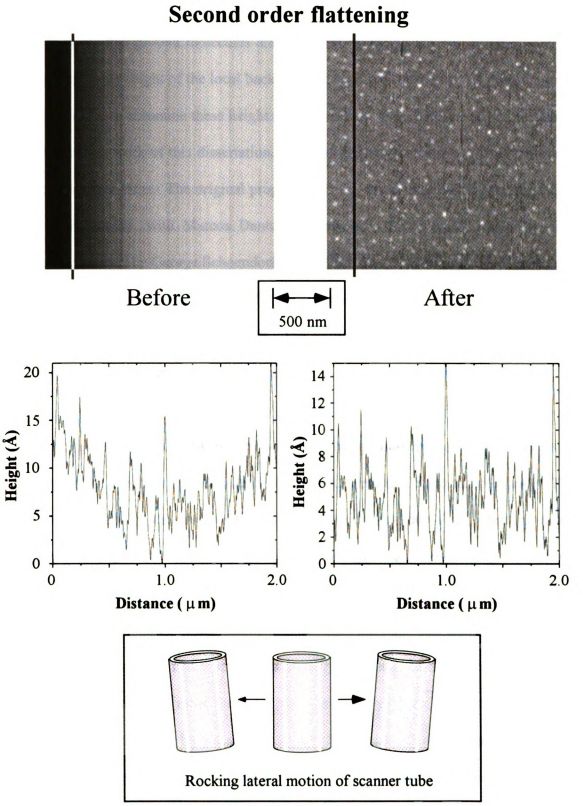


Figure 15 B: Example of image requiring second order flattening

Each image contained approximately 10-50 molecules per square micron. The heights of individual protein molecules are measured by taking the maximum height of the molecule minus height of the local background. A custom data analysis program was designed in order to automate these height measurements. Two versions of this program were used for the work of this dissertation. Both of the programs made use of the same measurement algorithms. The original program was written in Quick Basic 4.0 for DOS by Dr. Neil Bowlby, with Marcos Dantus and myself. The current version of the software was written by George Schoendorff and myself in Visual Basic 5.0 for Windows and made use of a subroutine (supplied by Park Scientific Instruments) which opens the hierarchical data format (HDF) image files. The Visual Basic source code for the current version of the program may be found in Appendix 2.

The HDF analysis software first loads the image into memory and prepares a histogram of the height values for each pixel in the image. If the sample has < 1% surface coverage (as is the case for the work presented in chapter four), then the most frequent pixel height value is a good first approximation to the average background level of the image. This value is used to determine the threshold value to be used for finding the local maxima contained in the image. The smallest particles that could be accurately measured were between 6 and 9 angstroms in height, so in general the threshold value used for analysis was the most frequent pixel height plus 6 to 9 angstroms. Local maxima may be confined to be within a given area by the specification of a radius value. For the analysis of the images in this work, I used a radius of 12 pixels. Twelve pixels in x and y corresponds to a distance of 47 nm, or more than twice the nominal radius of

curvature of the probe tip. In order to eliminate large protein aggregates from consideration, the local maxima search algorithm may be given a cutoff value as well.

The local maxima search finds not only isolated particles, but also anomalous features of the background and large aggregates. In order to eliminate these from the data set an interactive peak selection routine has been built in. In this interactive mode, a magnified gray scale image of the area around each local maximum (±12 pixels in x and y) is displayed one at a time. Determination of the height of the maximum, relative to the local background requires a reproducible measurement of the local background. To this end, a histogram of pixel heights is calculated based on the pixels within 40 pixels along the y direction and 1 pixel along the x direction. Preference towards the y scan direction is given since all the images for this work were acquired with the fast scan direction along the y axis of the scanner. The scanner noise and drift are minimized along the scan direction since those points were accumulated within a short time of one another. The most frequent pixel value for the region is taken to be a good estimate of the local background. Finally, the measured height of the particle of interest is calculated by taking the maximum height of the particle minus the most frequent local pixel height. Each maximum may be accepted or rejected on a case-by-case basis, by the program user. In the work of this dissertation, particles were rejected if they were of irregular shape (i.e. out-of-round), if the local histogram was anomalous (i.e. non Gaussian, indicating significant noise) or if a particle is much smaller or larger laterally than other similar particles in the image. Using these criteria, particles that are composed of large aggregates or contain anomalously high or low maxima are eliminated.

Each data set, generally consisting of more than a thousand individual height measurements, is tabulated in 3 Å bins to produce a height distribution histogram. The choice of bin size is based primarily on the substrate roughness (which is generally 3-5 Å) and is a reasonable, unbiased method of 'smoothing' the data. Within the 5 Å margin of error claimed, the results are independent of this bin size choice. A combination of Gaussian functions is then used to fit the height distribution data.

$$y = \frac{A}{w\sqrt{\pi/2}} \exp\left[-2\left(\frac{x - x_c}{w}\right)^2\right] \qquad (11)$$

This fitting was done iteratively using constraints on the three parameters: center position x_c, width w and amplitude A. In all cases, the amplitudes of the functions were allowed to vary freely. The center positions of the functions were allowed to vary independently. The widths of the functions were constrained to be between 8 and 13 Å. In the case of the Con A analysis, the functions centered at 59 Å (having the lowest amplitude), have the greatest uncertainty and their positions were fixed while the other parameters were optimized. The centers of the Gaussian functions are taken as the measured molecular dimensions, while the area under the curves are proportional to the probability with which a given molecular orientation is measured (see Chapter 2). Consistency in the fitting parameters was checked by multiple repetitions of the experiments on different days. The results presented here conform to this strict criterion, with variations in measured molecular dimensions of less than 5 Å.

The final analysis consists of two different procedures: a.) determination of molecular dimensions, and b.) assignment of quaternary structure (when different oligomerization states are possible). In order to assign the observed dimensions to the macromolecule, a model which takes into account the orientations of molecules at interfaces is required (see Chapter 2). Protein dimensions obtained by this method may be compared, for consistency, with other biophysical data available for the protein under investigation, such as molecular weight. For example, the volume of each protein may be estimated from the dimensions as measured by SOMS. Using the molecular weight of the protein under study, a volume, based on a density of ~1 g/cm^3, may then be calculated. In the case of the Con A results of this dissertation, the data are shown to be consistent with the structures determined by x-ray diffraction (XRD).

3.5 References

1. Yariv, J.; Kalb, A. J.; Levitzski, A. *Biochim. Biophys. Acta* **1968**, 165, 303.

2. Apell, H. J.; Colchero, J.; Linder, A.; Marti, O. Investigation of the Na, K-ATPase by SFM. In *STM and AFM in Biology*; Marti, O., Amrein, M., Eds.; Academic Press, Inc.: San Diego, 1993; pp 282.

3. Vesenka, J.; Guthold, M.; Tang, C. L.; Keller, D.; Delaine, E.; Bustamante, C. *Ultramicroscopy* **1992**, 42-44, 1243.

4. Hu, J.; Xiao, X.-D.; Ogletree, D. F.; Salmeron, M. *Science* **1995**, 268, 267.

5. Bustamante, C.; Vesenka, J.; Tang, C. L.; Rees, W.; Guthold, M.; Keller, R. *Biochemistry* **1992**, 31, 22.

6. Edstrom, R. D.; Meinke, M. H.; Yang, R.; Elings, V.; Evans, D. F. *Biophys. J.* **1990**, 58, 1437.

7. Thundat, T.; Zheng, X.-Y.; Chen, G. Y.; Warmack, R. J. *Surf. Sci. Lett.* **1993**, 294, L939.

8. Leggett, G. J.; Davies, M. C.; Jackson, D. E.; Roberts, C. J.; Tendler, S. J. B.; Williams, P. M. *J. Phys. Chem.* **1993**, 97, 8852.

9. Parker, M. C.; Davies, M. C.; Tendler, S. J. B. *J. Phys. Chem.* **1995**, 99, 16155.

10. Bustamante, C.; Rivetti, C. *Ann. Rev. Biophys. Biomol. Struct.* **1996**, 25, 395.

Chapter 4: Application of SOMS

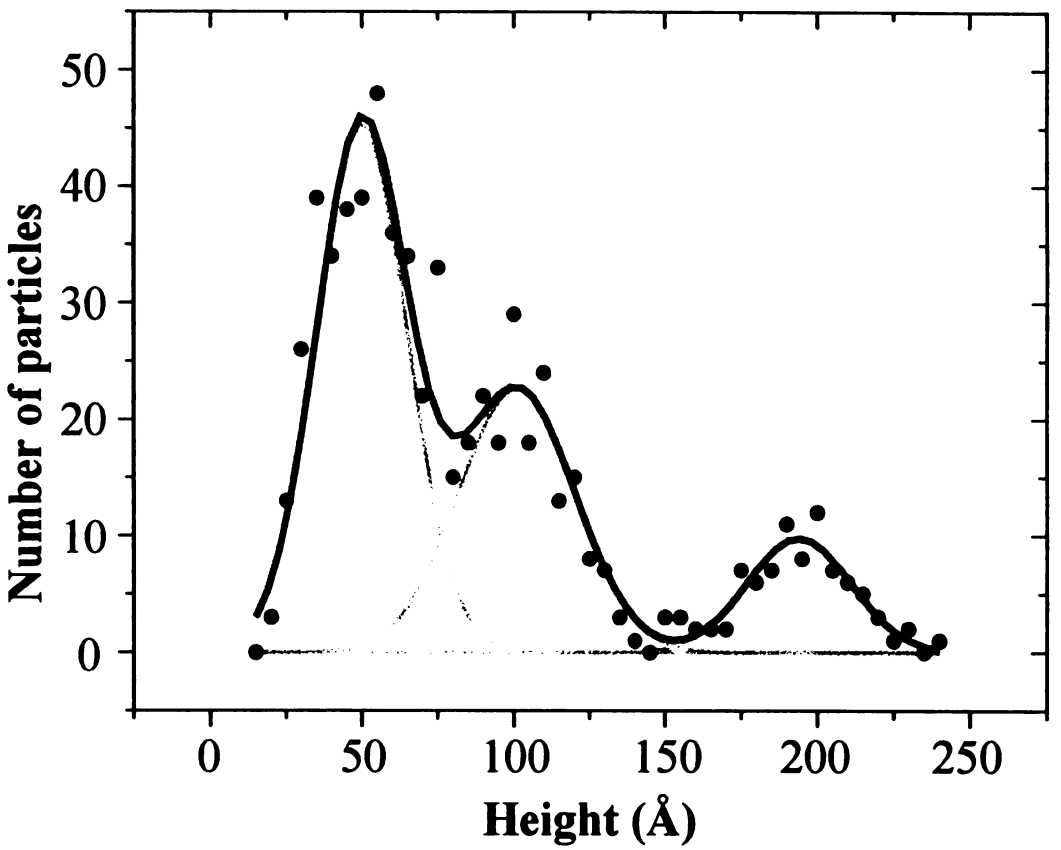
This chapter presents measurements of colloidal gold, ferritin and Con A at the water/mica interface. These samples will be used to show that the SOMS technique is capable of providing valuable new information about the structure, oligomerization and orientation of single macromolecules at interfaces. The data obtained from this technique will also be shown to provide important information regarding the properties of Con A at the liquid/solid interface

4.1 Analysis of gold and ferritin at the liquid/solid interface

4.1.1 Measurement of colloidal gold at the water/mica interface

One of the concerns with using the AFM to measure the heights of protein molecules at the liquid/solid interface is that the force applied during imaging might distort the structure of the protein molecule. While this issue will be examined later in this chapter, initially, however a proof of principle for the height measurement accuracy using an incompressible standard will be presented. Colloidal gold has been shown to be a useful, and incompressible standard for AFM measurements [1]. In addition, since colloidal gold is spherical one should observe only one orientation for each particle size. For these reasons, 5, 9 and 18 nm gold particles were measured as a proof of principle for the SOMS approach. The height distribution for the three sizes of colloidal gold on mica is plotted in Figure 16. Figure 16 also gives a tabulation of heights and standard deviations measured by the manufacturer (using TEM) and by the AFM. Note that for

the data in this figure the microscope was calibrated based on measurements of the 5 nm gold colloid and Tobacco Mosaic virus (cylinder with 18 nm diameter). The manufacturer data is based on the measurement of 100 particles, while the AFM measurements are based on measuring the number of particles given in Figure 16.



TEM	AFM	Particles
5.0 ± 0.8	5.0 ± 0.8	353
9.0 ± 1.2	10.1 ± 1.6	207
18.0 ± 0.8	19.4 ± 2.0	87

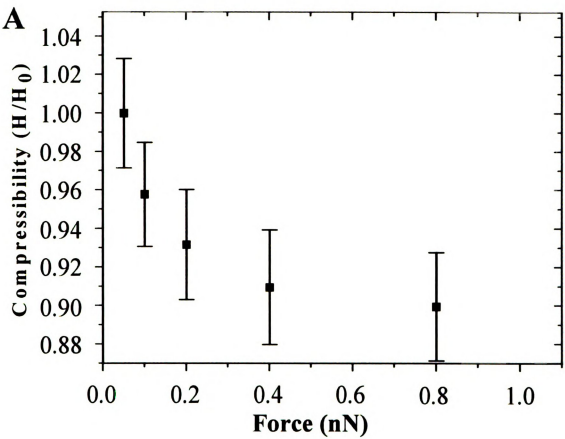
measurements are in units of nm

Figure 16: Height distribution measured by AFM for three different sizes of colloidal gold on mica.

It is clear that the agreement between the AFM and TEM measurements is quite good, but the AFM measurements are overestimating the heights of the larger particles. The AFM measurements also show a greater standard deviation for the larger particles. There are a few lines of reasoning that could explain these discrepancies. First, the calibration of the scanner was probably off by a small amount. Given the results of this study the microscope was re-calibrated using the data to determine a three-point calibration. Second, the number of 18 nm particles measured by the AFM was much less than for the other two sizes. The measurement of fewer particles for the larger colloids was due to the fact that the adsorption of the particles was found to vary inversely with the size of the particle. Finally, it has been observed that larger colloidal particles sometimes vary from the ideal spherical shape. Because the larger particles adhere weakly to the mica and may posses some ellipticity, it might be expected that the AFM could have more trouble with their measurement. The agreement of the AFM and TEM data suggests that the error in the measurement is mostly in the size distribution of the particles themselves, and not as a result of the measurement technique.

Vesenka et al. have suggested that colloidal gold makes a good incompressible standard for SPM measurements [1]. This assertion was tested using the AFM and 5 nm colloidal gold samples. As a means of measuring this compressibility samples were imaged at different forces and the heights for individual gold particles on the surface were measured. The data was then plotted as the height measured for a given force, referenced to the height measured for the minimum applied force. The data for the first trial of this experiment is presented in Figure 17 A, while a representative image from

this data set is given in Figure 17 B. It is readily seen that there is 10% compression going from a force of 0.05 nN to 0.8 nN, which is quite unexpected. Vesenka et al. observed no significant compression of 5-20 nm gold for forces as high as 225 nN [1]. The experiment was repeated the following day with significantly different results. The results of this second trial are given in Figure 18 A and B. Note that the variation of the measured height is much less pronounced. The reason for this inconsistency in the data is found by visual inspection of the images from each trial. The particles observed in Figure 17 B are broader (laterally) than those of Figure 18 B. The reason for this difference is a good example of a tip convolution artifact. A new, unused probe tip was used for the second trial, while a previously used tip was used during trial one. It is likely that the older tip had some adsorbed material stuck to it, which caused the interaction region of the tip to be larger. Further, it appears that this adsorbed material is probably protein material. It is likely that the compressibility behavior observed in the first experiment (see Figure 17) is characteristic of the adsorbed protein and not the gold.



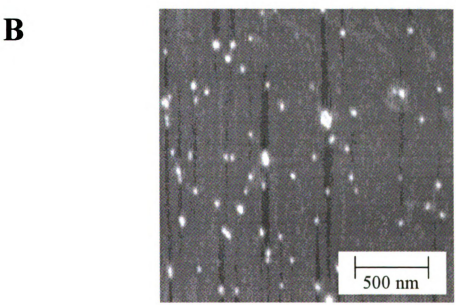
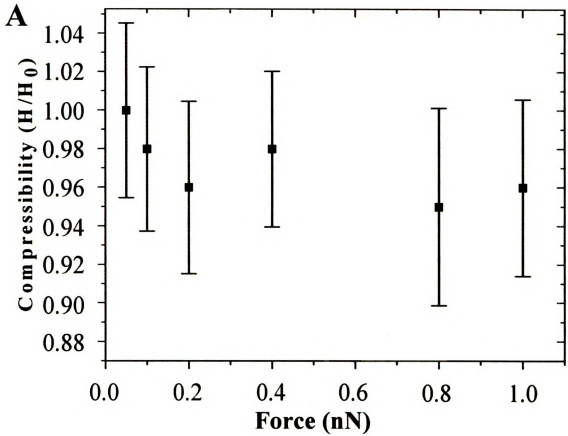


Figure 17: A) Compressibility of 5 nm colloidal gold on mica as a function of force.
B) AFM image from this data set.



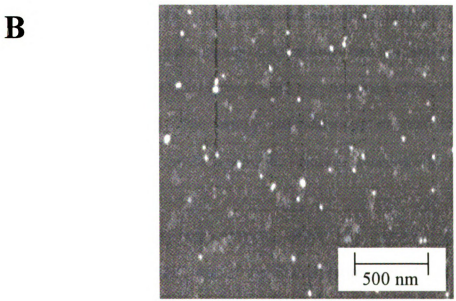


Figure 18: A) Plot of compressibility as a function of force for 5 nm colloidal gold using a new probe tip. B) AFM image from this data set.

4.1.2 SOMS analysis of ferritin at the water/mica interface

The SOMS approach for the determination of macromolecular dimensions was first tested using the protein ferritin (cationized, horse spleen). Ferritin is an important iron storage protein that is ubiquitous in fungi, plants and animals [2]. It has been widely studied and characterized by a variety of techniques, including electron microscopy and XRD [2,3]. The basic structure of this protein consists of an apoferritin shell composed of 24 polypeptide chains organized as a 120 Å diameter, hollow sphere containing six 10 Å channels that allow the exchange of iron (stored as ferric oxyhydroxide) [2]. Apoferritin has a molecular weight of 450 kD, but when its core is filled with the maximum number of iron ions (4500) it has a mass of 900 kD. The large size and high iron content of this protein has allowed the use of electron microscopy for its localization within cells and for the determination of structural characteristics. The compact structure and rigidity of the apoferritin shell makes ferritin structure quite resistant to denaturation [4]. In fact, as part of the processing of commercial ferritin it experiences temperatures of 80°C [4]. Feder and Giaver have studied the adsorption of ferritin to surfaces using electron microscopy [4]. The tertiary and quaternary structure of ferritin has been recently reported by Yau and Zhou using the scanning tunneling microscope (STM) [5].

Since ferritin is large, robust, stable and exhibits multiple oligomeric organizations, it is an ideal test of the SOMS technique. A 20 μg/ml solution of cationized horse spleen ferritin was applied to a magnesium treated piece of mica, as described in Chapter 3. The height distribution obtained from the AFM measurements is

given in Figure 19. Three Gaussian populations, centered at 100 Å, 116 Å and 135 Å, were fit to the data. The dominant population at 116 Å (62% of data) is attributed to the presence of intact ferritin molecules, which have been observed by electron and scanning tunneling microscopies to be 120 Å in diameter [3,5]. The population at 135 Å (21 % of data) is consistent with the presence of ferritin trimer, which was reported by Yau and Zhou to have a 130 Å by 380 Å size. The presence of the population centered at 100 Å (17% of data) may be partially denatured or dehydrated ferritin molecules. It should be noted, however, that there have been reports of ferritin structures, measured by electron microscopy, between 75 and 125 Å. In their studies, Feder and Giaever measured the diameter of ferritin adsorbed onto a carbon surface, to be 101±8 Å [4]. The measurement of a 100 Å population in the current studies is, therefore consistent with previous observations of ferritin. The study done by Feder and Giaever was done using an electron microscope, so the ferritin they observed was dehydrated. The population at 100 Å in the current data may very well represent a subset of the molecules which may have dehydrated to a greater extent when adsorbed to the mica surface.

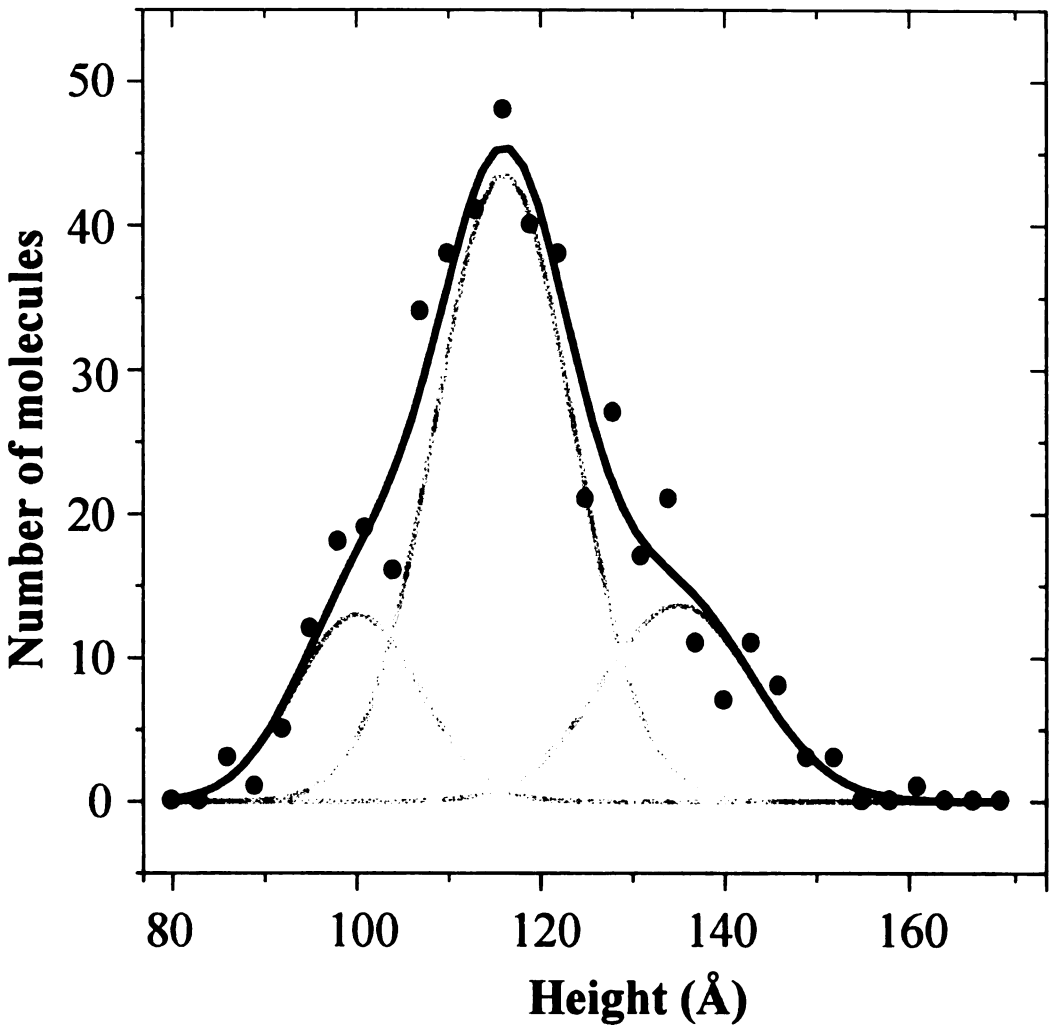


Figure 19: Height distribution measured for ferritin molecules at the water/ mica interface.

The results for ferritin presented to this point are consistent with previous work on the ferritin system, using STM, TEM and XRD. These data are also consistent with the proposal of the SOMS technique and the SAO model. Although the intact monomer of ferritin is spherical, the trimer has been observed to be non-spherical with dimensions of 130 Å by 380 Å [5]. The data of this section show that the trimer orients such that its shortest dimension projects from the substrate surface, which is consistent with the SAO model.

It has been predicted that the compact and rigid structure of the apoferritin core should make ferritin quite stable at the liquid/solid interface [4]. The above results indicate that the native structure of ferritin is conserved upon adsorption at the water/mica interface. The 24 polypeptide chains which make up the apoferritin shell are composed of 4 α-helical structural units between 35 and 42 Å long as well as a fifth short helix, that lies along a channel of the apoferritin shell [2]. Gekko and Hasegawa have examined the compressibility of globular proteins as a function of structural characteristics [6]. They found that proteins containing a significant fraction of α-helical character were in general more compressible. Although the primary structural characteristic of the ferritin polypeptides are α-helices, ferritin should be fairly incompressible due to its compact structure. This hypothesis was tested by measuring the height distribution of ferritin as a function of imaging force. The results of this study are presented in Figure 20. As in the gold compressibility study, the results are plotted as the ratio of the height measured compared to the height measured at the minimum imaging force, as a function of imaging force. The ferritin molecules are seen to undergo

compression of about 10% in going from an imaging force of 0.05 nN to 1.2 nN. This leads to an error of about 10-12 Å when using a force that's an order of magnitude greater than that used for the SOMS experiments of this dissertation. For a structurally robust protein like ferritin it is clear that contact mode AFM at ambient conditions is capable of making very precise measurements of structure, as long as the applied force is kept to a minimum.

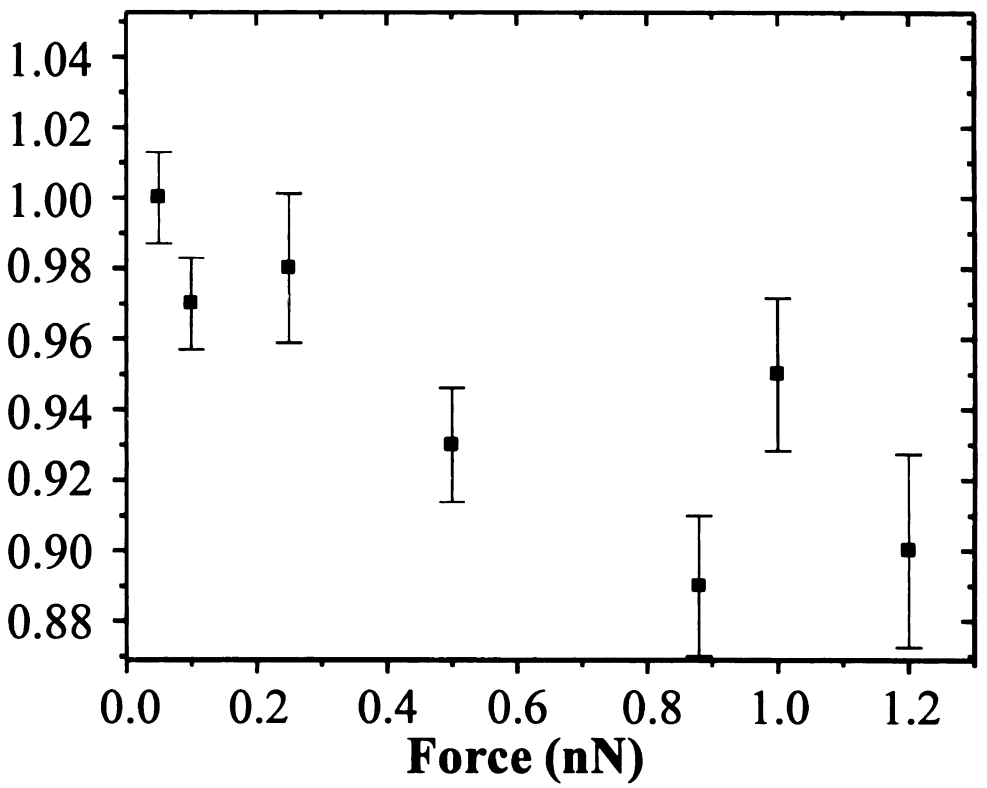


Figure 20: Plot of compressibility as a function of force for ferritin molecules adsorbed at the water/mica interface.

The SOMS results presented for colloidal gold and ferritin have been shown to be consistent with previous studies of these systems. These results indicate that the height measurements of small particles and macromolecules provide a useful way to obtain structural information with the AFM. While the SOMS approach avoids the probe tip

convolution problem that reduces the lateral resolution of the AFM for measurements of these systems, one must also be careful to watch for changes in the average lateral dimensions of particles that may indicate changes in the probe tip characteristics. The height measurements made by the AFM appear to be unaffected by the imaging force near the minimum imaging force required for contact mode imaging of these systems. The SOMS approach allowed the assignment of different oligomeric species of ferritin molecules adsorbed at the water/mica interface. While different oligomerization states were observed for ferritin, the monomeric units of the protein are spherical in shape. This made ferritin a good first test, but doesn't provide a very stringent test of the SOMS technique. The measurement of a protein with an ellipsoidal shape will be investigated in the next section.

4.2 Analysis of Concanavalin A (Con A) at the liquid/solid interface.

4.2.1 Background information for Con A

Concanavalin A (Con A) is a lectin, obtained from the Jackbean (*Canavalia ensiformis*), that binds mannose, glucose, and glycoconjugates containing these saccharides [7-9]. The strength of the interaction between Con A and saccharides has been likened to approach that of an antibody interaction, which has led to interest in the use of Con A adsorbed to surfaces for use in glucose sensor devices [10]. The native protein exists as a tetramer, composed of four identical subunits with a molecular weight of 25,500 kDa each [9]. In solutions at pH ≥7.0, Con A is normally observed as a

tetramer [9]. At pH ≤5.5, however the dimeric form is favored [11]. Gordon and Young have reported the dissociation of Con A into dimers as a consequence of dilution and ionic strength at a concentration ≤5 μg/ml and a pH ≥7.0 [12]. According to x-ray diffraction studies, the molecular dimensions of the dimer are approximately 30 Å x 45 Å x 75 Å, while those of the tetramer are approximately 60 Å x 70 Å x 70 Å (Protein DataBank, 1CN1, [13]) (see Figure 21).

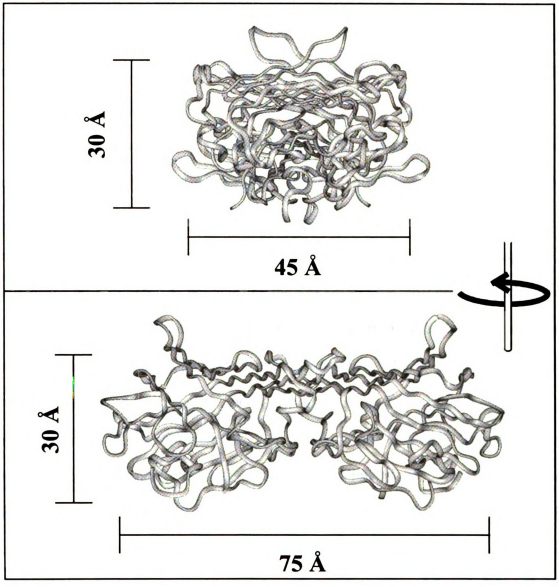


Figure 21: Ribbon rendering of the Con A dimer, based on XRD structure of Shoham et al. [BNL Protein DataBank, 1CN1]

4.2.2 SOMS analysis of Con A at the water/mica interface

Preparations of Con A solutions with concentrations in the range of ~10-400 ng/ml yielded dispersed individual protein molecules upon adsorption on magnesium treated mica surfaces, as described in the experimental section (see Chapter 3, Figure 14, for example). Analysis of several such images results in the height distribution shown in Figure 22. This distribution is the result of measuring the heights of 2000 individual protein molecules. The height distribution is best fit by four Gaussian functions centered at 18, 28, 42 and 59 Å (see Table II). Inspection of the XRD structure of the Con A dimer (see Figure 21), indicates that the observed height distribution corresponds to individual Con A dimers. As predicted by the SAO model the most prevalent orientation is the one in which the largest surface area is in contact with the mica substrate, projecting the 30 Å dimension perpendicular to the surface. The 28 and 42 Å dimensions are in excellent agreement with the x-ray structure (see Figure 21). The 59 Å measurement is significantly shorter than the 75 Å expected for Con A dimers projecting their longest axis normal to the mica surface, but it is quite close to the 60 Å dimension of the tetramer. The assignment of this population of molecules to the dimer will be verified later in this chapter by measurements made on succinylated Con A. Results to be presented later in this section will also show that the presence of the 18 Å population, which varies from ~8-35% of the total population, can be attributed to denatured Con A.

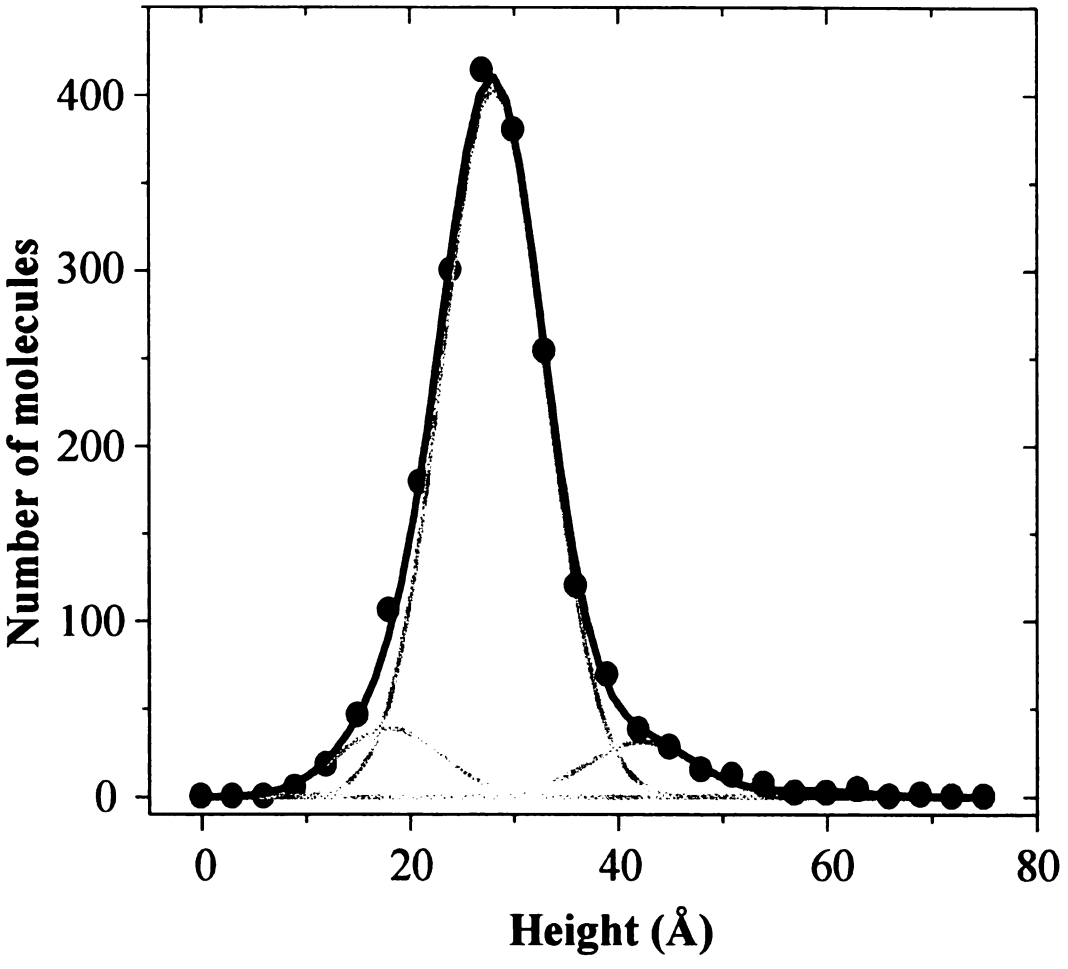


Figure 22: Height distribution measured for Con A molecules at the water/ mica interface.

Table II: SOMS data obtained for Con A at the water/mica interface.

	Height (Å)	Width (Å)	Area[a]	Probability (%)[b]
*	18	10	489	
a	28	10	5079	91.8
b	42	10	409	7.4
c	59	10	45	0.8
		$\Delta G_0 = -9.4 \pm 2.0$ J/mol-Å^2		

[a]These measurements correspond to the area under the Gaussian function and are proportional to the population of molecules with the given height . [b]The probability excludes the particles found to have a dimension of 18-19 Å, which are determined to arise from the presence of denatured protein molecules. The values of ΔG_0 given in Tables II and III are the average of three measurements, using equations 7-9 (see Chapter 2).

As in the case of ferritin, there was concern that the imaging forces applied to the protein molecules might cause distortion of their structure. Experiments were performed at the minimum imaging force, nominally 0.1 nN, and at forces as high as 1 and 2 nN. Samples were first imaged at the minimum force for stable contact mode imaging, then the force was increased and the sample was re-imaged. The results of this study are presented in Table III.

Table III: SOMS data for Con A at the water/mica interface as a function of imaging force.

0.1 nN 1068 pts.

center	width	Area	% of total	rel. %[a]
18	9	932	28.5	
28	12	2094		89.5
45	9	215		9.2
59	12	30		1.3

ΔG_0=-8.6 J/mol Å^2

1.5 nN 414 pts.

center	width	Area	% of total	rel. %[a]
18	9	449	34.9	
28	12	789		94.0
45	8	45		5.4
59	12	5		0.6

ΔG_0=-10.1 J/molÅ^2

Summary of the three trials

center	**width**	**% of total**	**rel. %[a]**
18	**9**	**29 ± 6**	
28	**12**		**91 ± 2**
45	**9**		**8 ± 2**
59	**12**		**0.9 ± 0.3**

ΔG_0=-9.4 ± 0.7 J/mol Å^2

2.0 nN 432 pts.

center	width	Area	% of total	rel. %[a]
18	8	297	22.7	
28	12	919		90.9
45	9	83		8.2
59	12	9		0.9

ΔG_0=-9.5 J/mol Å^2

[a]Relative percent is the area under the curve for the population of one molecular orientation relative to the population of all orientations $a \leq b \leq c$

The height distributions show no statistically significant differences when data is acquired at an imaging force between 0.1 nN and 2 nN. This result is consistent with the work of Gekko and Hasegawa. In their study of the structural factors which influence compressibility for globular proteins they found a general trend that compressibility scales with the amount of α-helical character of the protein [6]. They observed that proteins that contain ß-sheets tend to be less compressible than those containing primarily α-helices. The ß –structure of Con A is perhaps its most apparent feature when examining the XRD structures. The monomer unit consists primarily of two antiparallel sheets that together account for over half the amino acids in the sequence [9]. In addition, Gekko and Hasegawa found a correlation between the presence of certain residues and an increased or decreased compressibility [6]. A low compressibility was associated with the Asparagine, Glycine, Serine and Threonine, while a higher compressibility was associated with the presence of Leucine, Glutamic acid, Phenylalanine and Histidine [6]. Analysis of the sequence of Con A finds that 33 % of the amino acids are Asparagine, Glycine, Serine or Threonine, while 18 % are of Leucine, Glutamic acid, Phenylalanine or Histidine (see Table IV). This result further indicates that Con A is not expected to exhibit a large compressibility.

Table IV: Analysis of the composition of the Con A sequence and the potential influence on the compressibility of Con A based on Gekko & Hasegawa [6].

Residue	Number	% of total
ALA	19	8
ARG	6	2
ASN	12	5
ASP	20	8
GLN	5	2
GLU	7	3
GLY	16	7
HIS	6	2
ILE	15	6
LEU	18	8
LYS	12	5
MET	2	1
PHE	11	5
PRO	11	5
SER	31	13
THR	19	8
TRP	4	2
TYR	7	3
VAL	16	7

Decrease Compressibility	Number	% of total
ASN	12	5
GLY	16	7
SER	31	13
THR	19	8
	sum	**33**

Increase Compressibility	Number	% of total
GLU	7	3
LEU	18	8
HIS	6	2
PHE	11	5
	sum	**18**

In order to verify the presence of dimers at the interface, and to rule out the possibility that Con A was denaturing at the mica surface, two experiments were carried out. The first involves imaging of succinylated Con A, which is known to exist only as a dimer, over a wide range of pH [14]. The second experiment was to determine the height distribution for Con A which has been chemically denatured.

Results from the succinylated Con A samples (1576 individual height measurements) show a very similar height distribution to that obtained from Con A (see Figure 23 and Table V). The three main populations in the height distribution are centered at 29, 45 and 59 Å, corresponding closely (within 3 Å) to those found for Con A. These results give further evidence that Con A exists as a dimer and not a tetramer at the water/mica interface, under these conditions. This result is consistent with the observations of Gordon and Young, that Con A at a concentration of <5 μg/ml dissociates to form the dimeric species [12]. This experiment allows an unambiguous assignment of the oligomeric structure of Con A at the interface. The 59 Å population may be attributed to the dimer oriented such that side *ab* is in contact with the surface, projecting *c*=75 Å away from the surface. This dimension is not projected at 0° inclination, however, which can be reasoned by inspection of the XRD structure of Figure 21. Since the ends of the molecule are quite rounded, one might expect some angle of inclination relative to the surface. A simple calculation indicates an inclination of about 38° from normal.

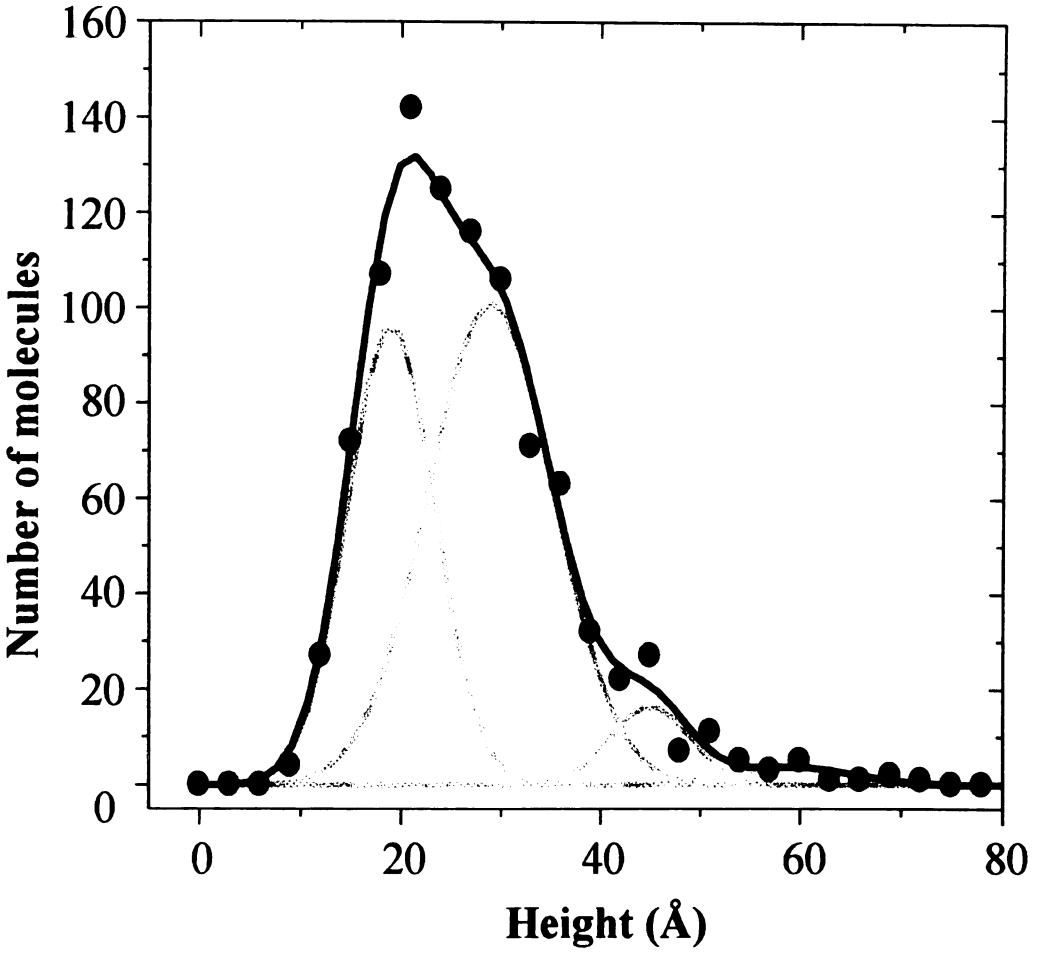


Figure 23: Height distribution measured for Succinyl Con A molecules at the water/mica interface.

Table V: SOMS data for succinyl Con A at the water mica interface

	Height (Å)	Width (Å)	Area[a]	Probability (%)[b]
*	19	9	1056	
a	29	13	1597	87.7
b	45	8	164	9.0
c	59	13	59	3.2

$\Delta G_0 = -6.2 \pm 0.2$ J/mol-Å^2

[a, b]See notes in Table II

The results of this experiment support the assertion that the SOMS approach has the capability and sensitivity to measure changes in the oligomerization state of proteins at the liquid/solid interface. There is a further population of molecules centered at 19 Å, which is similar to that observed at 18 Å for Con A, but in this case accounting for 37% of the total molecular orientations measured. The assignment of this population will be made in the following experiment.

To investigate the possibility of denaturation of Con A at the water/mica interface Con A was treated with 2% sodium dodecyl sulfate (SDS) for 2-3 min. at 90-95°C prior to dilution and preparation of samples. As discussed in Chapter one, the prevailing model of protein adsorption involves a denaturation step after the initial adsorption [15]. The height distribution from the SDS treated Con A samples, which is the result of 1983 individual height measurements, is shown in Figure 24 (see also Table VI). Note that the major population for this sample is centered at 18 Å, leading to the assignment of the 18-19 Å population to denatured Con A. These results indicate that ~60-70% of the protein molecules were denatured when treated with SDS at 90-95°C. The rest of the distribution

shows populations at 29 Å, 46 Å and 59 Å in excellent agreement with the previous two sets of data (within 3 Å). This indicates that some of the Con A molecules were not denatured, or possibly underwent refolding upon dilution of the SDS during sample preparation. This result allows us to assign the peak at 19 Å in the succinylated Con A samples to the presence of ~37% denatured protein in the sample. The degree of denaturation for the succinylated Con A is slightly higher than the typical ~30% denaturation observed for Con A on mica. This may be a result of the succinylation process or improper handling during production.

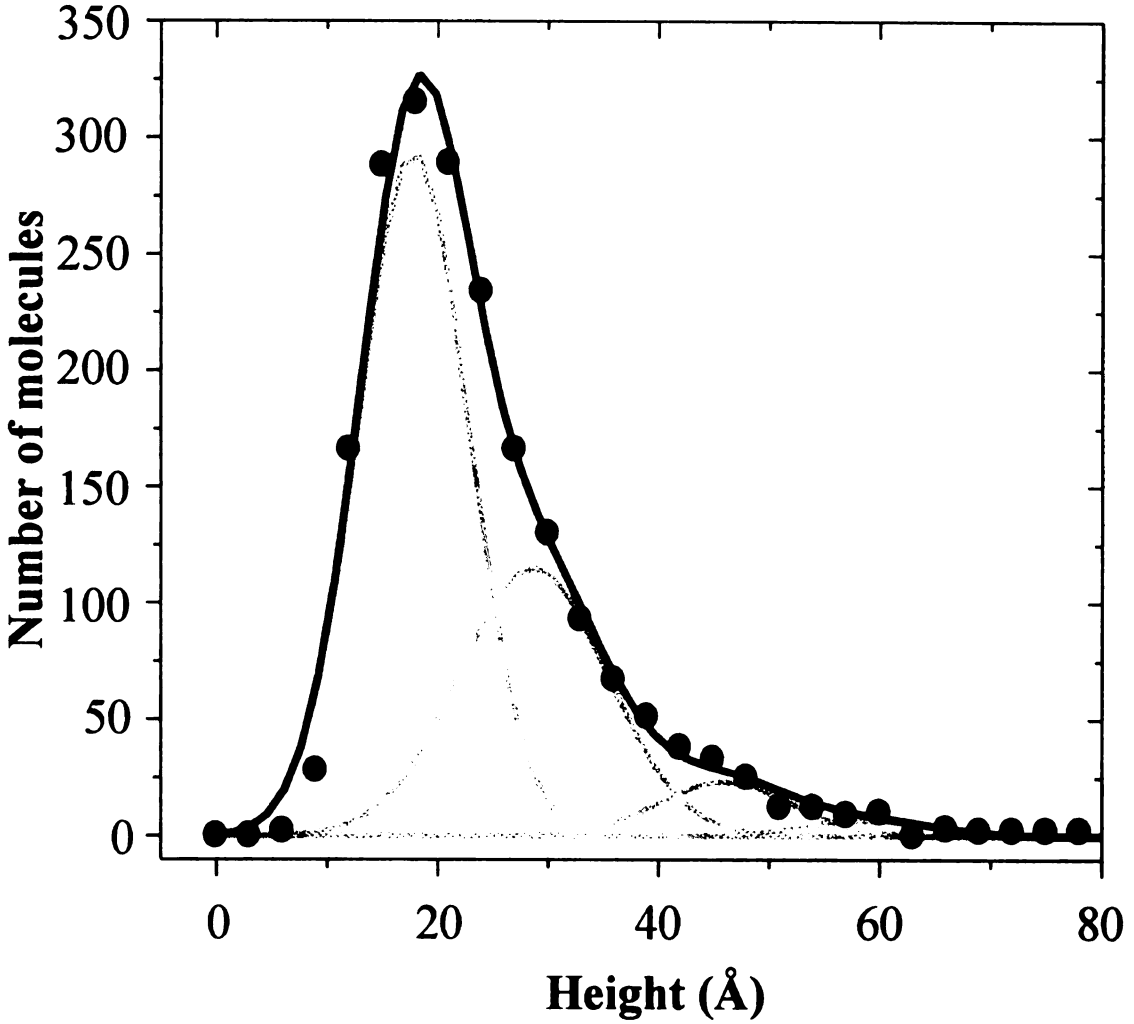


Figure 24: Height distribution measured for SDS treated Con A molecules at the water/mica interface.

Table VI: SOMS data for SDS denatured Con A at the water/mica interface

	Height (Å)	Witdh (Å)	Area[a]	Probability (%)[b]
*	18	10	3679	
a	29	13	1938	81.2
b	46	12	350	14.7
c	59	12	100	4.2

[a, b]See notes in Table II

Further analysis of the results by the SAO model can be used to estimate the free energy of orientation per unit area, ΔG_0, for Con A and succinylated Con A on Mg^{+2} treated mica. Using equations 7-9 (see Chapter 2) and values (height and probability) from Table II, we calculate average values of -9.4±2.0 J/mol- $Å^2$ and -6.2±0.2 J/mol- $Å^2$, for Con A and succinylated Con A, respectively. The quoted uncertainty is the standard deviation of the three values calculated from equations 7-9. Day to day variability of as much as 40 % in the calculated value of ΔG_0 has been observed. Much of this derives from the fact that the 42-45 Å or 59 Å populations are quite small relative to the 28-30 Å population. Because of the exponential dependence in equations 7-9, a small difference in the 59 Å population can make a large difference in the calculated ΔG. This variability in the calculated free energy will be discussed again later in this chapter.

The approximate surface area of each face of the protein can be calculated from the measured dimensions. With these values the free energy of orientation for each of the three different orientations is calculated (see Table II). The most stable of these being ~ -20 kJ/mol when the Con A molecules have their largest surface oriented towards the

mica. These values are larger than kT (2.5 kJ/mol), indicating this orientation is relatively stable. Differences in the free energy for each orientation determine the observed height distribution. The values for Con A and succinylated Con A are similar, as expected, however, I believe that there may be some small difference related to the chemical modification of amino acid side chains by the succinylation process. Succinylation involves the modification of amines in the protein with a concomitant decrease in positive charge. For Con A, an average of 10 sites per subunit are succinylated [14], which may cause changes in the interaction with the mica surface.

In conclusion, angstrom resolution measurements from individual protein molecules at the liquid/solid interface have been observed by a new approach to AFM. The measurement of populations centered at 28, 42 and 59 Å, indicate that Con A adsorbs onto mica substrates as a dimer. These molecules tend to orient such that the smallest dimension is projected normal to the surface. This observation is consistent with the SAO model proposed in Chapter 2. These measurements are in very good agreement with the x-ray structures determined by Weisgerber et al. and Shoham et al. [13,16]. The height distribution observed for Con A at the water/mica interface was shown to be unaffected by imaging forces between 0.1 and 2.0 nN. Although the tetrameric form of Con A is normally observed in solution, the solutions from which samples were prepared were ~10-400 ng/ml at a pH=7.3, which according to Gordon and Young are expected to contain primarily the dimeric species [12]. This dimeric state is conserved upon adsorption of the Con A molecules to the mica surface. The presence of the dimeric oligomerization state was further confirmed with measurements of succinylated Con A.

Application of the SAO model to these measurements gives approximate values for the free energy of orientation for these molecules. The possibility of protein denaturation caused by the adsorption process was addressed by measurements of Con A treated by the detergent SDS. This question will be examined more closely in the following section.

4.2.3 Examination of Con A denaturation at the water/mica interface

It was shown in the previous section that denatured Con A at the water/mica interface is indicated by the presence of a population of molecules with a height of 18-19 Å. The treatment of Con A with SDS and heat was chosen for the previous study primarily because it is known to effectively denature most protein systems. Additionally SDS is known to strongly bind to most polypeptides, and the thinking was that upon dilution of the sample in buffer much of the SDS-protein interactions might remain. In this way we sought to try to prevent refolding of the Con A before or possibly during the adsorption process. In order to test this idea a series of experiments were performed to further examine the denaturation of Con A at the liquid/solid interface.

Comparison of the degree of denaturation typically observed for Con A or succinyl Con A with that observed for Con A treated with SDS at 90-95°C, shows a factor of two or more in the population of molecules at 18-19 Å. Although significant, this result was not as dramatic as anticipated. One possibility is that upon dilution of the denatured stock solution or during the actual adsorption process, some of the Con A

molecules may undergo refolding. Con A samples which were denatured by various methods were prepared and measured in order to test this hypothesis.

The first of these experiments involved treating the Con A with 6 M urea. It was expected that if refolding were occurring, it might be more likely to happen with urea, which doesn't bind to the polypeptide backbone as strongly as SDS. The height distributions obtained were quite similar to that of the SDS treated Con A (see Figure 25). The amount of denatured Con A, as determined by taking the area of the population centered at 17-18 Å compared to the total area, was found to be 43-44%. These numbers are based on 2 sets of data run on two different days, containing 435 and 1220 individual molecular measurements. Clearly the amount of denaturation observed for the urea treated Con A is higher than untreated Con A, but much less than SDS treated protein. This indicates that some refolding of the protein may have occurred during the dilution process. During the dilution of the sample, the initial concentration of ~6 M urea was reduced by three orders of magnitude.

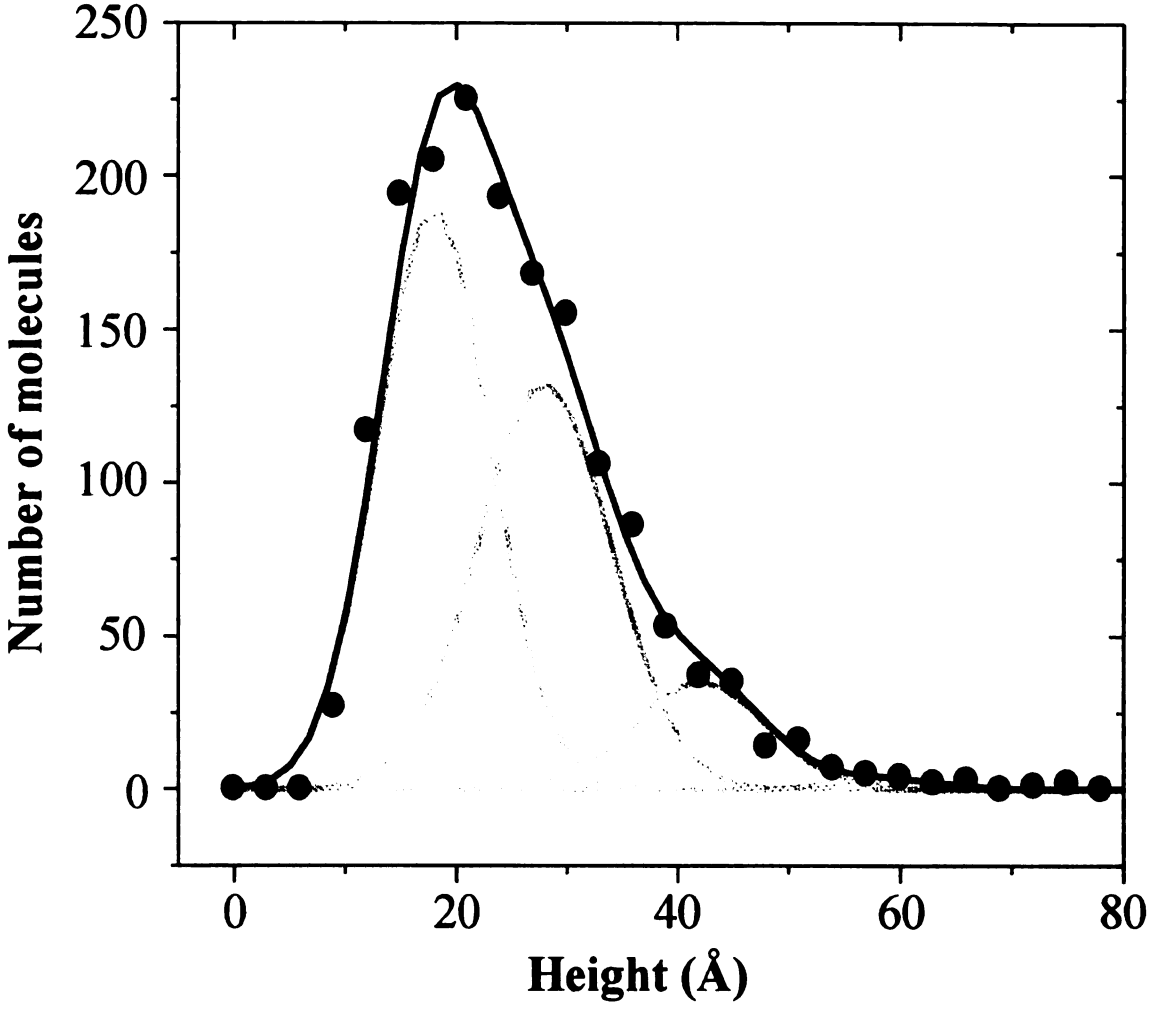


Figure 25: Height distribution measured for Con A molecules treated with 6 M urea, prior to dilution and adsorption at the water/mica interface.

To examine the effect of heat on the denaturation of Con A, aliquots of Con A in pH=7.3 HEPES stock solution were treated at 90-95°C for 3-5 minutes. This solution was then diluted to working concentration and applied to magnesium treated mica. These samples produced SOMS data that were indistinguishable from the SDS treated samples, which were also heated to 90-95°C. In this case two sets of samples were examined. The height distributions showed 65-74% of the molecules to have heights centered between 14 and 18 Å.

Upon SDS/heat or temperature denaturation conditions, approximately twice as many Con A molecules appear to be denatured at the water/mica interface, compared to untreated molecules. The results from the urea treated samples indicate that Con A undergoes significant refolding, prior to or during the adsorption process, although it most likely occurs prior to the adsorption process.

As mentioned in Chapter one, the prevailing model for protein adsorption at the liquid/solid interface is thought to involve two steps [15]. The first step involves the collision of a protein molecule with the surface and its layer of adsorbed water, while the second is thought to involve some conformational change of the protein to maximize its interaction with the surface. The results to this point have shown that less than one third of the Con A molecules adsorbed at the water/mica interface appear to be significantly denatured. In contrast to this case, approximately two thirds of the adsorbed molecules at the interface are denatured when treated with 2% SDS and/or 90-95°C temperatures. Because there is ≤33% denaturation in the case of untreated Con A adsorbed to mica,

there is reason to believe that this is consistent with the second step of the prevailing protein adsorption model (see Chapter one, Figure 1). To examine this more closely samples of Con A on mica were prepared and imaged as before, and then were aged under ambient conditions (21°C and 35-50% relative humidity).

After 24 hours the samples were imaged, producing height distributions that were not significantly changed from that of freshly prepared samples (1.5 – 8 hours old). It is a reasonable assumption that after 24 hours, equilibrium of the layer of surface adsorbed water with the protein and the atmosphere had been reached. Even when imaged after being adsorbed to the mica surface for 48 hours, there was still not significantly more denaturation observed. After approximately 72 hours, however, a significant change in the amount of denatured protein was observed. Prior to aging the samples (prepared on two different days), they were found to consist of 28-29% denatured molecules. After three days' aging on the surface 61 –73% of the molecules measured were found to orient with heights centered at 18 Å. There are two probable explanations for this observation. First, it is possible that the denaturation of Con A due to its adsorption to the mica surface is a rather slow process, and this experiment indicates a long time scale for the second step of the prevailing protein adsorption model for Con A adsorption to mica. This is a reasonable hypothesis given that after 24-48 hours the samples did not show significantly more denaturation. The second possibility is that the Con A is denaturing due to its being kept at 21°C for an extended period. It is also possible that both of these scenarios are contributing. In either case, however, it is clear that the majority of Con A molecules at the water/mica interface conserve their native dimeric structure, even when stored for 24

hours at ambient conditions. This observation has potential application toward the design of saccharide sensor devices, or Con A/saccharide based bioassays.

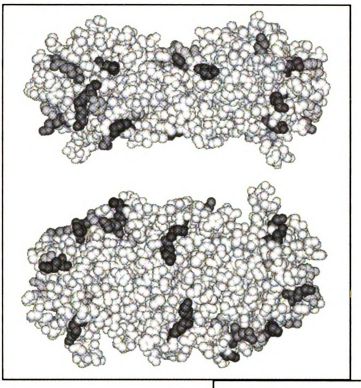
4.2.4 Effect of pH on the structure and adsorption of Con A molecules at the water/mica interface

The structural characteristics of Con A in solution have been studied under a variety of conditions. One important factor that influences the structure of Con A in solution is pH. The effect of various pH conditions on the structure and function of Con A have been studied by a number of workers [11,17-21]. Some of the key findings of this work have been that at acid pH, between 2 and 5.5, the dimeric organization is favored [11], and that between a pH of 8 and 9 there is an irreversible structural transition [17,18,22]. Using circular dichroism (CD) and optical rotatory dispersion (ORD), Zand et al. observed Con A to approach a random coil form at a pH of 9.1 [17].

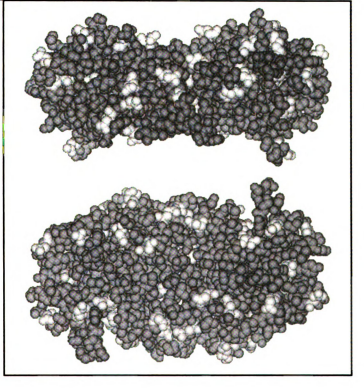
The structural changes in Con A as a function of pH may in part be caused by changes in the ionic character of the amino acid residues. The isoelectric point of Con A has been reported to be in the range of 4.5-8 [9,11]. However, a value of 5-5.5 is typical of the accepted value used in recent literature [23,24] and is consistent with a calculation of pI=5.00 based on the amino acid sequence [25]. The studies of this dissertation have, up to this point, looked at Con A deposited on the mica surface from a HEPES buffer at pH=7.3. At this pH the protein will have a net negative charge, although 11% of the residues posses some positive character. At a pH of 4.6, however, 89% of the residues

have positive charge, leading to a net positive charge for the molecule. A CPK structure for the Con A dimer, based on XRD data [13] is presented in Figure 26. In this figure the positively charged amino acids appear shaded.

Observation of Figure 26 indicates that there is a fairly even distribution of the positively charged amino acids along the surface of the Con A dimer at both pH values. This observation is consistent with the simplifying assumption made in deriving the SAO model (see Equations 7-9 in Chapter two). The simplifying assumption was that no face of the protein contained a unique surface recognition site, which allowed the definition of $\Delta G_o \equiv \Delta G_{bc} = \Delta G_{ac} = \Delta G_{ab}$. Since no face of the protein has significantly more charge than another, the charge would not contribute to making one face of the protein inherently more favorable for adsorption to the negatively charged mica surface. An important prediction of the SAO model is that changes in the free energy of the adsorption process will lead to changes in the observed height distribution. Figure 10 gives theoretically predicted height distributions for a system with the dimensions observed in our experiments for Con A at two different values of the free energy of orientation (given in units of energy per unit area). It is apparent that the SAO model predicts that a more favorable ΔG_o will produce a higher probability of finding molecules oriented with their shortest dimension projected normal to the surface. To test this prediction, a careful study of the height distributions for Con A at the water/mica interface as a function of pH were carried out.



pH = 7.3



pH = 4.6

Figure 26: Positively charged amino acids in Con A at two pH values. Positive residues are shaded.

For this study, solutions of Con A were prepared in 0.01 M HEPES at pH=7.3 and 0.01 M Acetate at pH=4.6. The concentrations were checked, prior to dilution, to verify that they were equal. To insure good comparison between the measurements, samples at each pH were run at the same time (i.e. prepared and imaged the same day). The results presented here are the result of three separate trials. The results from Con A deposited from a HEPES buffer at pH=7.3 consist of data sets containing 715, 761, 990 and 1047 molecular measurements made on four samples prepared on three different days. The results for Con A deposited from the acetate buffer at pH=4.6 are taken from data sets containing 379, 585 and 892 individual measurements made on three samples prepared on three different days. The measured height distributions, representing the sum of all the trials, are presented in Figure 27. For purposes of calculating an average and standard deviation, however, each data set was fit to four Gaussian functions independently. The results obtained from the Gaussian fits to the data are summarized in Table VII.

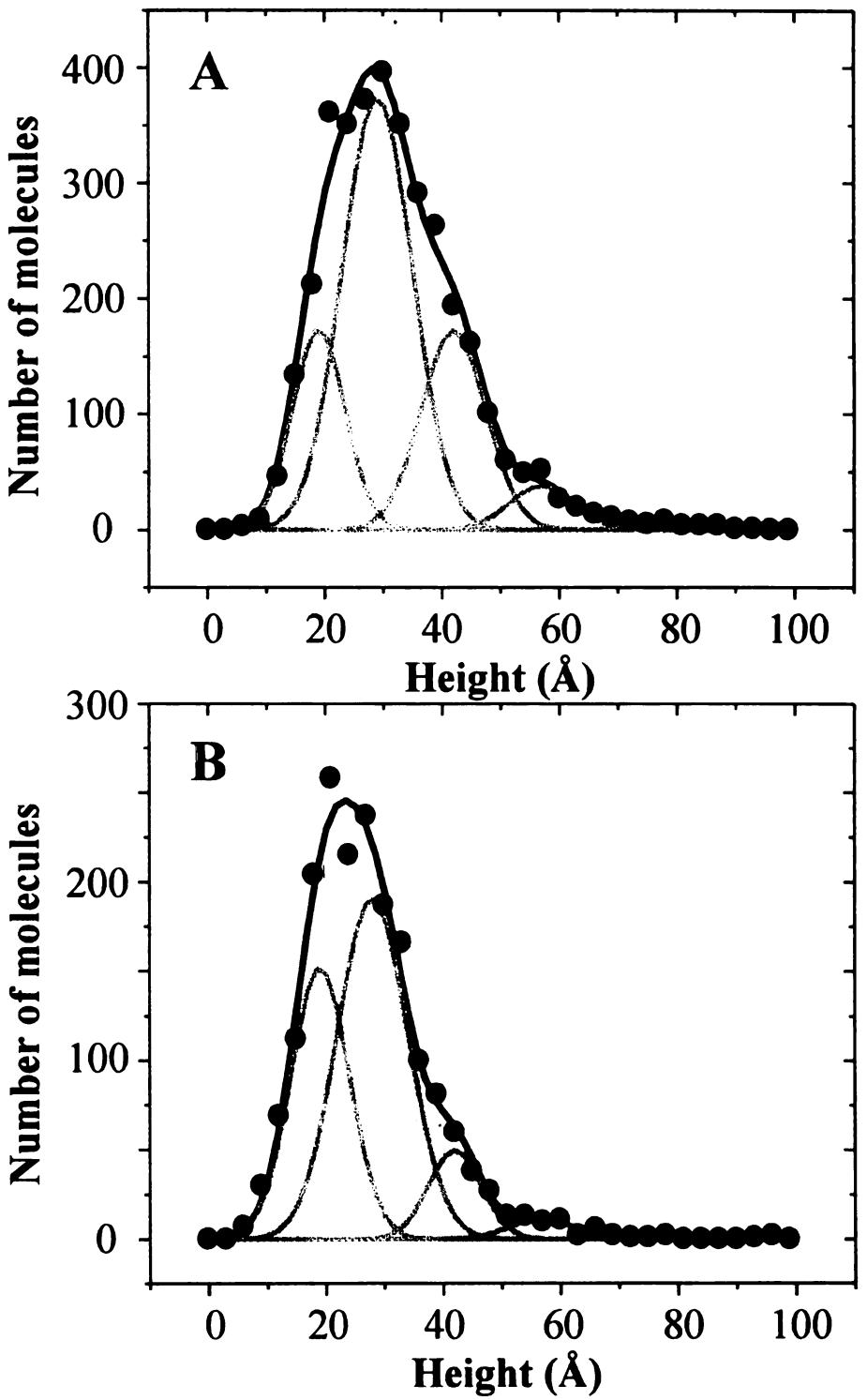


Figure 27: Height distributions measured for Con A molecules deposited from solutions at A) pH=7.3 and B) pH=4.6.

Table VII: SOMS data for Con A at the water/mica interface deposited from solutions at two different pH values.

A) pH=7.3

	Height (Å)	Width (Å)	Probability (%)*	% denatured
	19	9		18 ± 5
a	29	12	71 ± 9	
b	43	11	24 ± 5	
c	59	11	5 ± 4	

B) pH=4.6

	Height (Å)	Width (Å)	Probability (%)*	% denatured
	18	9		34 ± 5
a	28	12	84 ± 14	
b	42	9	13 ± 11	
c	59	11	3 ± 3	

*See note b in Table II

Examination of the fitting parameters is in close agreement with the widths and positions presented in Tables II, III, V and VI. It is clear that there was more denaturation of the Con A which had been deposited from the acetate buffer at pH=4.6, compared to that of Con A in HEPES at pH=7.3. This behavior has not been previously observed for Con A. The studies of Auer and Schilz have examined Con A in the acidic pH range [19]. In this work they observed the release of Mn^{+2} and Ca^{+2} ions from Con A beginning at a pH around 6.0 and increasing to a maximum by pH≈3 [19]. The loss of the metals from Con A results in a conformational change of the amino acids near the metal binding sites. While this change is observable in the XRD structures of the native and apoprotein [13,26], the change in the gross molecular size of the dimer is less than 5

Å. The transition from native Con A to the demetallized form has been likened to an "unlocking" of the conformation [13,26]. This is consistent with the observation of more denatured protein for samples adsorbed to mica at pH=4.6, and may explain why more denaturation occurs during the adsorption process.

The other striking feature of the results is that there is a tendency for Con A molecules to more strongly favor the orientation that places side bc on the mica surface, when deposited from a pH=4.6 solution. This preferential orientation of the Con A dimer is likely driven by a more favorable interaction of the positively charged Con A dimer with the net negative charge of the mica surface. This change in the distribution over orientations is predicted by the SAO model, proposed in chapter 2. From the results of Table VII, the free energy of orientation is calculated from equations 1-3 and 7 of chapter 2. The results of this calculation are presented in Table VIII.

Table VIII: Free energy of orientation values for Con A at the water/mica interface deposited from solutions at two different pH values.

A) pH=7.3

ΔG_0	-3.4	J/mol Å^2
Std. Dev.	0.7	J/mol Å^2
ΔG_{bc}	-8.5	kJ/mol
ΔG_{ac}	-5.8	kJ/mol
ΔG_{ab}	-4.2	kJ/mol

B) pH=4.6

ΔG_0	-6.5	J/mol Å^2
Std. Dev.	2.8	J/mol Å^2
ΔG_{bc}	-16.0	kJ/mol
ΔG_{ac}	-10.9	kJ/mol
ΔG_{ab}	-7.7	kJ/mol

Note that the standard deviation in the probability of measuring dimension c (given in Table VII) is on the order of the average, indicating significant uncertainty in this population. Therefore, the results presented in Table VIII are based solely on the P(a)/P(b) ratio (see equation 7, chapter 2). The exclusion of the equations which consider the population of molecules with dimension c (i.e. 59 Å) is reasonable, since the variability of the P(a)/P(c) and P(b)/P(c) is quite large. The data of Table VIII clearly show an increase in the free energy of orientation (by a factor of nearly two), upon going from a pH of 7.3 to 4.6. When the data of Table VIII A and Tables II and III are compared one sees a fairly large discrepancy in the calculated free energy of adsorption. In the earlier data the Con A deposited from a HEPES buffer at pH=7.3 gave ΔG_0 ~9 J/mol $Å^2$, however the data presented in this section for Con A under the same conditions gives ΔG_0 ~3.5 J/mol $Å^2$. This variability will be discussed further in section 4.3.

The effect of high pH on Con A structure at the water/mica interface was also examined. It has been observed that Con A undergoes irreversible loss of its β-structure when it is exposed to basic media in the pH range 8-9 transition [17,18,22]. A stock solution of Con A in borate buffer (pH=10) was prepared. Results for the analysis of three height distributions, containing 709, 1133 and 1962 molecular measurements, are presented in Table IX.

Table IX: A) SOMS data for Con A at the water/mica interface, as deposited from a pH=10 solution. B) Free energy of orientation values for Con A at the water/mica interface deposited from a solution at pH=10

A)

	Height (Å)	Witdh (Å)	Probability (%)*	% denatured
	17	8		51 ± 13
a	28	11	71 ± 7	
b	42	11	23 ± 7	
c	59	10	6 ± 2	

B)

ΔG_0	-4.1	J/mol Å^2
Std. Dev.	0.8	J/mol Å^2

ΔG_{bc}	-10.2	kJ/mol
ΔG_{ac}	-6.8	kJ/mol
ΔG_{ab}	-4.9	kJ/mol

The data for Con A deposited from a pH=10 solution show significantly more denaturation than Con A deposited from a pH=7.3 solution. The percentage of denatured molecules lies somewhere between that of Con A treated with 6 M urea (see section 4.2.3) and the SDS and/or heat treated Con A. This observation is consistent with the CD studies of Zand et al. that observed Con A to approach a random coil form at a pH of 9.1 [17]. The presence of an irreversible structural transition, occurring somewhere between a pH of 8 and 9 has also been noted by Pflumm et al. [18,22].

As seen from Table IX and Table VII-A, the relative probabilities for the three molecular orientations of the Con A dimer at the water/mica interface, are not significantly different for pH=7.3 and pH=10. While the free energy of orientation per unit area observed for the pH=10 sample is slightly higher than that of the Con A at pH=7.3 (see Tables VIII and IX), there is only a slight difference in the free energy of orientation for each protein face. This is in part due to the slight differences in the average Gaussian peak positions for the data sets. This negligible difference in the free energy, as compared to the large change in going from pH=7.3 to pH=4.6, is not too surprising. At a pH of 7.3 Con A has a net negative charge, but at a pH lower than 5 the charge on the protein becomes positive. In the present case, the protein has become more negatively charged going from pH=7.3 to pH=10, though the change in charge does not seem to drastically affect the free energy of the adsorption process onto the net negative mica surface.

Despite the fact that the surface was treated with Mg^{+2} prior to the adsorption of protein, there is reason to believe that the net charge at the surface remains somewhat negative. This is supported by the data of this section, but this was tested independently by running a sample of Con A at pH=7.3 on mica with and without the magnesium treatment. In this experiment 643 measurements were made for Con A on untreated mica and 696 measurements of protein on Mg^{+2} treated mica. The height distributions obtained were statistically identical, matching the results presented in Table VII A. The expectation was that if the negative surface charge is significantly reduced by the presence of the positive magnesium ions, then the free energy of orientation for

negatively charged Con A would favor the 30 Å population more strongly. Based on the results of this section, it is reasonable to assume that the surface charge of the mica remains a net negative.

This brings up another interesting feature of the data presented in this section. It was anticipated that the adsorption of Con A from the pH=4.6 solution would progress at a faster rate than that at pH=7.3-10, where the molecule maintains a net negative charge. In fact there was no increase in the number of molecules deposited at a pH of 4.6 vs. a pH of 7.3. Analysis of 7-8 images, acquired on the same day, showed that for Con A deposited from Acetate buffer (pH=4.6) 76 ± 3 molecules/μm^2 while 73 ± 6 molecules/μm^2 were observed for images of the pH=7.3 deposited samples. The uncertainties given are ±1 σ.

4.3 Discussion and conclusions

Height measurements of colloidal gold particles of different sizes (i.e. 5, 9 and 18 nm) made by the AFM in contact mode at ambient conditions showed that these particles are accurately measured by this technique. The gold particles were observed to be quite incompressible, showing no significant change in structure within an order of magnitude of the minimum force required for stable imaging.

The first application of the SOMS technique was the spherical protein ferritin. The analysis of the height distributions obtained for ferritin showed the presence of three

populations of molecules at the water/mica interface. The dominant population, centered at 116 Å is attributed to the native spherical 24-mer, which has been observed by x-ray and electron microscopy to be 120 Å [3,4]. The second most prevalent species, centered at 135 Å is attributed to the presence of a trimer of the native 24-mer, which has been previously observed by STM [5]. The remaining population at 100 Å is consistent with dehydrated ferritin, as observed in vacuum by electron microscopy [4].

Measurements of individual Con A and succinyl Con A molecules at the water/mica interface has shown that the Con A dimer is the prevalent structure. The measurements obtained (29 x 45 x 59 Å) for these dimers are in good agreement with the dimensions determined by x-ray diffraction studies. Further, the results for Con A, succinyl Con A and SDS treated ConA are self-consistent, and reproducible. In addition to the structural information obtained, the proposed model also allows us to predict a value for the free energy of adsorption for Con A at the water/mica interface and the orientation of the adsorbed proteins.

The denaturation of Con A was examined using the SOMS technique. Typically about 30% or less of the Con A molecules at the water/mica interface were found to be denatured. The assignment of molecules with heights of 17-19 Å to the presence of denatured molecules was made based on the observation of Con A treated with SDS and heated to 90-95°C. This chemical denaturation produced samples with 60-70% of the observed molecules to have heights centered at 17-19 Å. The presence of a population of denatured molecules for untreated Con A at physiological pH is consistent with the

prevailing two step model of protein adsorption (see Figure 1). Measurements of Con A that was adsorbed to the mica surface and then aged from one to two days showed that the amount of denatured protein remains nearly constant. After three days, however, we found that ~60-70% of the Con A had unfolded on the surface. It is interesting to note that Barnes et al. found that the useful lifetime for their Con A based piezoelectric glucose sensor was three days [10]. In their work Con A was adsorbed onto siloxane modified silver electrodes, and the loss of sensor activity was attributed to bacterial attack [10]. If bacterial attack is the reason for the inactivation of the Con A, we should have observed the relatively large bacteria with the AFM. In fact, we saw no indication of bacterial growth with the AFM, which indicates that the structure of Con A at the water/mica interface is quite stable over the range of 24-48 hours. After that the protein may denature simply because it cannot maintain its structure without additional hydration or that the half life of the second step of the protein adsorption model is quite long. Con A treated with 6 M urea gave a height distribution containing only ~40% denatured molecules, indicating that some of the molecules may have undergone refolding as the urea concentration was diluted by three orders of magnitude.

The SOMS model was further tested by comparing Con A deposited onto mica from buffers of different pH. In this study samples of Con A in 0.01 M HEPES (pH=7.3) and in 0.01 M acetate (pH=4.6) or 0.01 M borate (pH=10) were prepared in parallel and imaged within a few hours of one another. At the lower pH, where Con A possesses a net positive charge, the SOMS analysis showed that the free energy of adsorption was more favorable than for Con A adsorbed from solutions at pH between 7.3 and 10, where Con

A maintains a net negative charge. This observation is consistent with adsorption onto a negatively charged surface, as predicted by the SAO model presented in Chapter two.

It was expected that the adsorption to the negatively charged mica surface would be faster for Con A at a pH=4.6 than for pH=7.3. This expectation was based on the fact that 4.6 is below the isoelectric point of Con A (pI=5-5.5). Analysis of the number of molecules observed at each of the two pH values showed that the number of proteins adsorbed per square micron lie within the same range (i.e. 73-78 molecules/μm^2 and 68-78 molecules/μm^2, respectively). The adsorption process appears to be limited by diffusion of the molecules to the surface, rather than any electrostatic forces present. The likely explanation for this observation is that in the dilute solutions used (~10-400 ng/ml), there is good shielding of charge by the solvent. The issue of rate of adsorption as a function of protein charge will be further addressed in Chapter five.

While these results are shown to be qualitatively consistent with the SAO model, there are large discrepancies observed in the free energies of orientation predicted. The results presented in section 4.2.2 predict a value of -9.4 ± 2 J/mol $Å^2$ while the later samples of Con A in HEPES (see section 4.2.4) gave a value of -3.4 ± 0.7 J/mol $Å^2$. There are a couple of reasons for this discrepancy. First of all the data of section 4.2.2 were for samples prepared using mineral quality mica (Ward's Natural Science Establishment, NY), while the later data was taken using high quality mica sheets (gift of Dr. Richard Schwendeman). A second reason for this discrepancy lies in the way the SAO model calculates the value of ΔG_0. The ratios of the probabilities of observing each

molecular orientation are exponentially related to the free energy. Since the number of measurements observed at 59 Å is much smaller than the number of observations made of molecules centered at ~30 or ~45 Å, there is much more uncertainty in its probability. If one ignores the ΔG_0 calculated from the ratio P(b)/P(c) (as was done in section 4.2.4) the standard deviation in the calculated value is usually less by a factor of two. If the value of ΔG_0 for Table II is re-calculated one obtains a value of -8.3 ± 1 J/mol Å^2, which is still more than twice the value obtained for Con A on the optical quality mica. Although it may seem that sticking with the natural mica would have been better, given the lower amount of denatured Con A observed, it should be noted that the samples prepared using this substrate showed a wider range of variability. This variability included measuring some samples which contained ~10-20% denatured protein, with ΔG_0~-9 J/mol Å^2, as well as others with 30% denatured molecules and ΔG_0~-3-4 J/mol Å^2. The Con A in HEPES samples deposited onto the high quality mica gave ~30% denatured protein and ΔG_0~-3-4 J/mol Å^2. A likely explanation for these observations is that some areas of the natural mica may have been more positively charged (possibly due to the presence of excess metal cations) which would explain the increased population of 28-30 Å molecules. This type of shift in the height distribution with increasing positive charge was shown for Con A adsorbed to mica from the acetate buffer at pH=4.6.

The studies of Con A at the liquid/solid interface presented here not only serve to validate the SAO model and the SOMS technique, but they also give new insight into the structure and oligomerization of this lectin. This study marks the first such study of individual Con A molecules at the liquid/solid interface. The results have important

implications to those interested in designing Con A based sensors and assays. The strong binding of saccharides by Con A has been of interest for use in glucose sensors [10,27,28]. Further the use of Con A as a model system for antibody-antigen interactions [27], and the stability of Con A at the liquid/solid interface observed in this work indicates that Con A may be useful in the design of new bioassays.

We conservatively estimate that the accuracy of this technique is 5 Å or less. This estimate is related to a number of factors. First of all there is 3-5 Å roughness in the mica substrates which we use. Second, although the ultimate resolution of the AFM in the z dimension is 1 Å or less, we are not optimizing our imaging conditions for single point measurements of height, rather we trading accuracy for acquisition time. For the purposes of three-dimensional protein oligomerization state and orientation, the accuracy (< 5 Å) is sufficient.

The technique presented here provides a means to measure the molecular dimensions for the complete size range of monomeric and oligomeric proteins adsorbed to a surface (natural or synthetic). These measurements may then be utilized to determine the oligomerization state of the protein. Probe-protein interactions, mainly lateral and vertical pressure, can cause the molecule to appear shorter [29]. With regard to protein-substrate interaction, there is an attractive interaction between the polar surface of the protein molecule and the charged mica surface, possibly deforming the molecule and causing it to flatten against the mica surface. A study using lysozyme has shown that it retains its activity when adsorbed onto mica and imaged with AFM, so these

deformations may not be extensive enough to alter biological activity [30]. Our measurements confirm the conservation of native structure for Con A and ferritin, when compared with available x-ray data.

A thermodynamic model was derived to understand the orientation of proteins at the liquid/solid interface. Large deviations from this SAO model by adsorbed proteins would imply a preferred orientation that is presumably related to recognition between a protein domain and the surface. The extent of deviation from the orientation model may therefore be useful in defining preferred molecular orientation for oligomers or multiple equilibria for dissociated subunits of oligomers. In these cases one is able to extend the simple model presented here, by allowing each face of the protein to have a different value for ΔG_0 (Equations 1-3). In this way a quantitative value may be assigned to the energy difference associated with each orientation.

An examination of the preferred orientation of the molecules also provides the opportunity to predict the organization of individual subunits within the protein oligomer and potentially evaluate the conformational flexibility of dissociated individual subunits. Measurements with angstrom resolution may be performed on functional complexes that are associated with a surface, a more common biological interaction within and between cells. These results present an approach which extends the analytical power and the resolution of AFM to provide direct physical characterization (i.e. dimensions, orientation and free energy of adsorption) of biomolecules at interfaces, and is minimally dependent on the geometry of the probe tip. This approach provides the capability to

examine properties of protein structure and oligomerization at liquid/solid interfaces using picomole amounts of protein. Further, it offers the ability to distinguish between the possible quaternary forms and their interconversions resulting from the binding of ligands, substrates, or allosteric effectors. Such measurements can be most useful in studies of enzyme regulation, DNA/RNA organization, gene regulation, protein biosynthesis and chaperonin function, bacterial attachment and adhesion, and the effects of post-translational modification on protein organization and activity.

4.4 References

1. Vesenka, J.; Manne, S.; Giberson, R.; Marsh, T.; Henderson, E. *Biophys. J.* **1993**, 65, 992.

2. Munro, H. N.; Linder, M. C. *58* **1978**, 2.

3. Massover, W. H. *Micron* **1993**, 24, 389.

4. Feder, J.; Giaever, I. *J. Coll. Inter. Sci.* **1980**, 78, 144.

5. Yau, S.-T.; Zhou, Y. *Mod. Phys. Lett. B* **1995**, 9, 187.

6. Gekko, K.; Hasegawa, Y. *Biochemistry* **1986**, 25, 6563.

7. *Concanavalin A as a Tool*; Bittiger, H.; Schnebli, H. P., Eds.; John Wiley & Sons:, 1976.

8. Sharon, N.; Lis, H. *Science* **1989**, 246, 227.

9. Reeke, G. N., Jr.; Becker, J. W.; Cunningham, B. A.; Wang, J. L.; Yahara, I.; Edelman, G. M. Structure and Function of Concanvalin A. In *Concanavalin A*; Chowdhury, T. K., Weiss, A. K., Eds.; Plenum Publishing Corp.: New York, 1975.

10. Barnes, C.; D'Silva, C.; Jones, J. P.; Lewis, T. J. *Sens. Act. B* **1991**, 3, 295.

11. Liener, I. E. Isolation and Properties of Concanavalin A. In *Concanavalin A as a Tool*; Bittiger, H., Schnebli, H. P., Eds.; John Wiley & Sons:, 1976; pp 2.

12. Gordon, J. A.; Young, R. K. *J. Biol. Chem.* **1979**, 254, 1932.

13. Shoham, M.; Yonath, A.; Sussman, J. L.; Moult, J.; Traub, W.; Kalb, A. J. *J. Mol. Biol.* **1979**, 131, 137.

14. Gunther, G. R.; Wang, J. L.; Yahara, I.; Cunningham, B. A.; Edelman, G. M. *Proc. Nat. Acad. Sci. USA* **1973**, 70, 1012.

15. Brash, J. L.; Horbett, T. A. Proteins at Interfaces: An Overview. In *Proteins at Interfaces II, Fundamentals and Applications*; Horbett, T. A., Brash, J. L., Eds.; American Chemical Society: Washington, 1995; Vol. ACS Symposium Series 602; pp 1.

16. Weisgerber, S.; Helliwell, J. R. *J.Chem. Soc., Faraday Trans.* **1993**, 89, 2667.

17. Zand, R.; Agrawal, B. B. L.; Goldstein, I. J. *Proc. Nat. Acad. Sci. USA* **1971**, 68, 2173.

18. Pflumm, M. N.; Beychok, S. *Biochemistry* **1974**, 13, 4982.

19. Auer, H. E.; Schilz, T. *Int. J. Peptide Protein Res.* **1984**, 24, 462.

20. Lis, H.; Sharon, N. *Ann. Rev. Biochem.* **1986**, 55, 35.

21. Sharon, N.; Lis, H. *FASEB J.* **1990**, 4, 3198.

22. Pflumm, M. N.; Wang, J. L.; Edelman, G. M. *J. Biol. Chem.* **1971**, 246, 4369.

23. Lvov, Y.; Ariga, K.; Ichinose, I.; Kunitake, T. *Thin Solid Films* **1996**, 284-285, 797.

24. Gallinet, J.-P.; Gauthier-Manuel, B. *Eur. Biophys. J.* **1993**, 22, 195.

25. Swiss Institute of Bioinformatics. Compute pI/MW Tool. Available http://expasy.hcuge.ch/ch2d/pi_tool.html, June 1998.

26. Reeke, G. N. j.; Becker, J. W.; Edelman, G. M. *Proc. Nat. Acad. Sci. USA* **1978**, 75, 2286.

27. Janata, J. *J. Amer. Chem. Soc.* **1975**, 97, 2914.

28. Kremer, F. J. B.; Engbersen, J. F. J.; Kruise, J.; Bergveld, P.; Starmans, D. A. J.; Feijen, J.; Reinhoudt, D. N. *Sensors and Actuators B* **1993**, 13-14, 176.

29. Radmacher, M.; Fritz, M.; Cleveland, J. P.; Walters, D. A.; Hansma, P. K. *Langmuir* **1994**, 10, 3809.

30. Radmacher, M.; Fritz, M.; Hansma, H. G.; Hansma, P. K. *Science* **1994**, 265, 1577.

Chapter 5: Study of Con A Orientation at the Liquid/Solid Interface: From Isolated Molecules to Thin Films

5.1 Introduction to Con A film formation

While the majority of papers dealing with protein structure at liquid/solid interfaces deal with measurements of monolayer or multilayer films, and 2D crystals, very little is known about the mechanisms of adsorption or the microstructure of thin protein films. The data presented thus far in this dissertation have shown that the dimeric oligomerization state of Con A is favored at the water/mica interface, and that Con A molecules tend to orient such that the surface area in contact with the mica is maximized. One of the issues that hasn't been addressed is whether the Con A simply dissociated to the dimeric form in solution as a result of concentration, pH or ionic strength [1], or whether the interfacial chemistry favors the dimeric form. Would tetramers of Con A be observed as the molecules adsorb to the interface from a concentrated solution? If so, how would they orient? In order to investigate these issues a series of experiments was designed to examine Con A molecules at the water/mica interface as deposited from more concentrated protein solutions.

The interest in thin films of Con A dates back to the mid-seventies. Janata published a study of an 'immunoelectrode', composed of a PVC coated Pt wire that was subsequently layered with Con A [2]. This device was shown to be sensitive to the binding of yeast mannan. The description of this device as an 'immunoelectrode' was

justified in that Con A-saccharide interactions, while not a true antibody-antigen interaction, are often used as a model for ligand binding and antibody-antigen interactions [2]. From the mid-eighties to the present there have been a number of studies of Con A associated in thin films, studied by a variety of techniques [3-10].

The findings of these studies are indicative of one of the problems in the proteins at interfaces field, namely that there is significant variability in conditions and results, and many unanswered questions. Afshar-Rad et al. deposited films of Con A from a pH=5-5.5 buffer onto mica [3]. Using the surface force apparatus, they observed an apparent layer thickness of 38 Å, consistent with the presence of dimers [3]. Gallinet and Gauthier-Manuel also used a surface force apparatus and found 12.5 Å, 23 Å, 39 Å and 46 Å layers within their multilayer films on mica [6]. The thickness of expelled layers in their study was dependent on the pH of the deposition solution used. The same year Safar et al. prepared Con A films on glass and mica [7]. Using TEM they investigated the morphology, finding rough (±400 Å), thick films (455 ± 72 Å), but did not determine the number of layers present [7]. Using CD and FTIR spectroscopies Safar et al. found that Con A retained its β-structure in the films they deposited [7]. Lvov et al. deposited Con A onto polyethyleneimine cation layers. Using a quartz crystal microbalance (QCM) they determined the mean layer thickness to be 57 Å ±10% when depositing from pH=7 solution and 17 Å ±10% depositing from a pH=5.6 solution [8]. Lvov et al. viewed this result as consistent with studies that have found preference for the dimer at pH~5.6, but tetramer formation at pH~7[11,12]. Fitzpatrick et al. deposited Con A films onto mica at pH=7 and examined them using x-ray photoelectron spectroscopy (XPS) [10].

The film thickness was determined by monitoring the N 1s signal as a function of angle and x-ray energy to determine a depth profile [10]. The measured thickness for films varied between 14 and 41 Å, but the average thickness quoted was 21 ± 4 Å (based on the angle resolved depth profile) [10]. Although they did not make measurements of film thickness or molecular orientations, both Barnes et al. and Haas and Mohwald, note that they had difficulty in controlling the state, stability and evenness of Con A coatings on quartz and phospholipid layers, respectively [3,4]. The data and preparation conditions for each of these studies is summarized in Table X. It is quite apparent from the variability in the measurements of these studies that there is not a clear picture of the structure and growth of Con A films.

Table X: Summary of literature results for Con A thin films deposited on surfaces.

Ref.	**Method**	**Concentration**	**Deposition time**	**pH of solution**	**substrate**	**Thickness (Å)**
[6]	SFA	40 μg/ml	60 min.	3.9	Mica	12.5
[3]	SFA	Not specified	2 hr.	5.5	Mica	38
[8]	QCM	1 mg/ml	20-30 min.	5.6	[a]PEI^+ on Ag	57 (±10%)
[6]	SFA	40 μg/ml	60 min.	6	Mica	39
[8]	QCM	1 mg/ml	20-30 min.	7	[a]PEI^+ on Ag	17 (±10%)
[10]	XPS	1-10 μg/ml	17-19 hr.	7	Mica	21 ± 4
[6]	SFA	40 μg/ml	60 min.	7.4	Mica	23, 46

[a] poly (ethyleneimine)$^+$ modified Ag electrode on QCM

A better understanding of the structure and molecular orientation of Con A thin films is paramount to the fundamental understanding of protein-protein and protein-surface interactions. The results of Chapter four present a well-characterized picture of isolated Con A molecules at the liquid/solid interface. A logical next step would be to ask questions about how this early stage of adsorption might progress as the adsorption continues toward formation of monolayer and thin films. Additionally the potential commercial application of Con A-saccharide based sensors and bioassays, makes Con A a good candidate for this study.

Results presented in Chapter four show that Con A maintains much of its native dimeric structure at the water/mica interface and appears to be little affected by the imaging forces applied by contact mode AFM. As the Con A molecules begin to aggregate at the interface the problem of tip convolution will preclude the determination of individual molecular positions and orientation, but the height of local domains will be accessible. While techniques such as neutron reflection and ellipsometry are capable of determining overall film thickness, the AFM allows one to make local morphology and topography with angstrom height resolution and nanometer lateral resolution.

The approach to the problem of making high-resolution measurements of protein films is essentially the same as in the isolated molecule case (i.e. SOMS). Though it is hindered by reduced lateral resolution, the AFM can measure the height of regions of the film to 1 Å accuracy, limited by the substrate flatness and roughness. The height

distribution obtained for images of sub-monolayer protein films will reflect populations of heights corresponding to both the mica substrate and the protein molecules that make up the film. The heights of various regions of the film may then be measured relative to the background. This approach may be extended to complete monolayer coverage as well. As will be shown, the height of the mica substrate may be determined either by producing sub-monolayer films or by physically cutting away a small section of the film from the surface with the AFM.

This chapter will present studies of Con A at the water/mica interface as monolayer films are formed. Data for the early stages of growth, as well as morphology of complete monolayers will be presented. These results will be compared to the observed orientation and structure of individual Con A molecules presented in Chapter four, and to the results of previous studies for Con A films at liquid/solid interfaces. Further, the implications to understanding the film growth mechanism and the design of sensor/assay devices will be discussed as well.

5.2 Experimental and theoretical details

The materials and instrumentation are the same as given in Chapter three, however, the preparation and imaging of samples, as well as the experimental analysis and theory are somewhat different from that presented in Chapters two and three. A brief description of the additional experimental details will be given in this section.

Preparation of Con A Films. 25 µl of solution containing ~120 µg/ml Con A in 0.01 M buffer (HEPES at pH=7.3 or Acetate at pH=4.6), was applied directly to mica, which was freshly cleaved and treated with magnesium ions (as described in Chapter three). Samples were incubated at 21°C for the duration of the deposition time and then rinsed with two 200 µl portions of milli Q water. The deposition time was controlled using a wristwatch and was varied between 1 min. and 60 min. The samples were 'dried' (see Chapter three) at 21°C and 35-50% RH. Films were imaged beginning 1 hour after preparation and were discarded after 12 hours.

AFM imaging of Con A at the liquid/solid interface. Two different types of Microlever cantilevers (Park Scientific Instruments) were used. Measurements were made using either sharpened Microlevers, having a spring constant of approximately 0.05 N/m with an integrated silicon nitride probe (20 nm radius of curvature), or Microlevers with a spring constant of 0.5 N/m containing 50 nm radius of curvature tips. The higher spring constant was desirable for this application since high forces ~10 nN were useful when scraping away portions of the film for the measurement of the background mica height. It was also found that somewhat lower vertical forces, ~1.5 nN and higher lateral forces were capable of cutting away sections of Con A film. The 5 µm and 100 µm scanners were calibrated for vertical measurements with colloidal gold particles (5 and 9 nm) (Sigma, St. Louis). Horizontal measurements are calibrated using a two-dimensional grating with 1 µm periodic spacing in both the x and y directions.

AFM images were acquired in the constant force mode, as discussed in Chapter three. As the films grow to the point of nearly complete coverage, a measurement of film thickness is not made directly, since the mica background is not known. In cases of monolayer or near monolayer coverage a small area was scanned at high force and at a high scan rate. The high scan rate was used in order to cause increased lateral forces between the probe tip and the protein film during the raster scanning process. After repeated scanning, the forces were sufficient to scrape away the Con A molecules from the mica surface (see Figures 33,36,37). The molecules that are scraped away deposit at the edges of the scan area, or in some cases lead to fouling of the probe tip with adsorbed proteins.

Analysis of Con A thin films: Analysis of the Con A thin films was similar to that of the isolated molecules in that we focused on the height distribution. In the case of the films, however, we do not have the lateral resolution necessary to determine the orientation of each molecule in the film. As discussed in Chapter one, the lateral resolution is limited by the size and shape of the probe tip, as well as the sample morphology. In order to get an idea of the best lateral resolution to be expected a scale model was drawn (see Figure 28). As can be seen from Figure 28, the lateral resolution for the protein films will depend on both the film and the probe. Given ideal probe tip size and geometry (i.e. Microlever spherical probe tip with radius of curvature of 20 nm), one could expect lateral resolution of 11 nm for a 30 Å thick discontinuous film. A 60 Å film would have lateral resolution limited to 14 nm, while an 80 Å thick film would limit it to 16 nm. This means that for the Con A films one might expect to be able to

determine morphology for local areas of 11 x 11 nm to 16 x 16 nm. The broader, unsharpened Microlever probe tips, having a radius of curvature of 50 nm, would have resolution limitations of 32 x 32 nm to 56 x 56 nm.

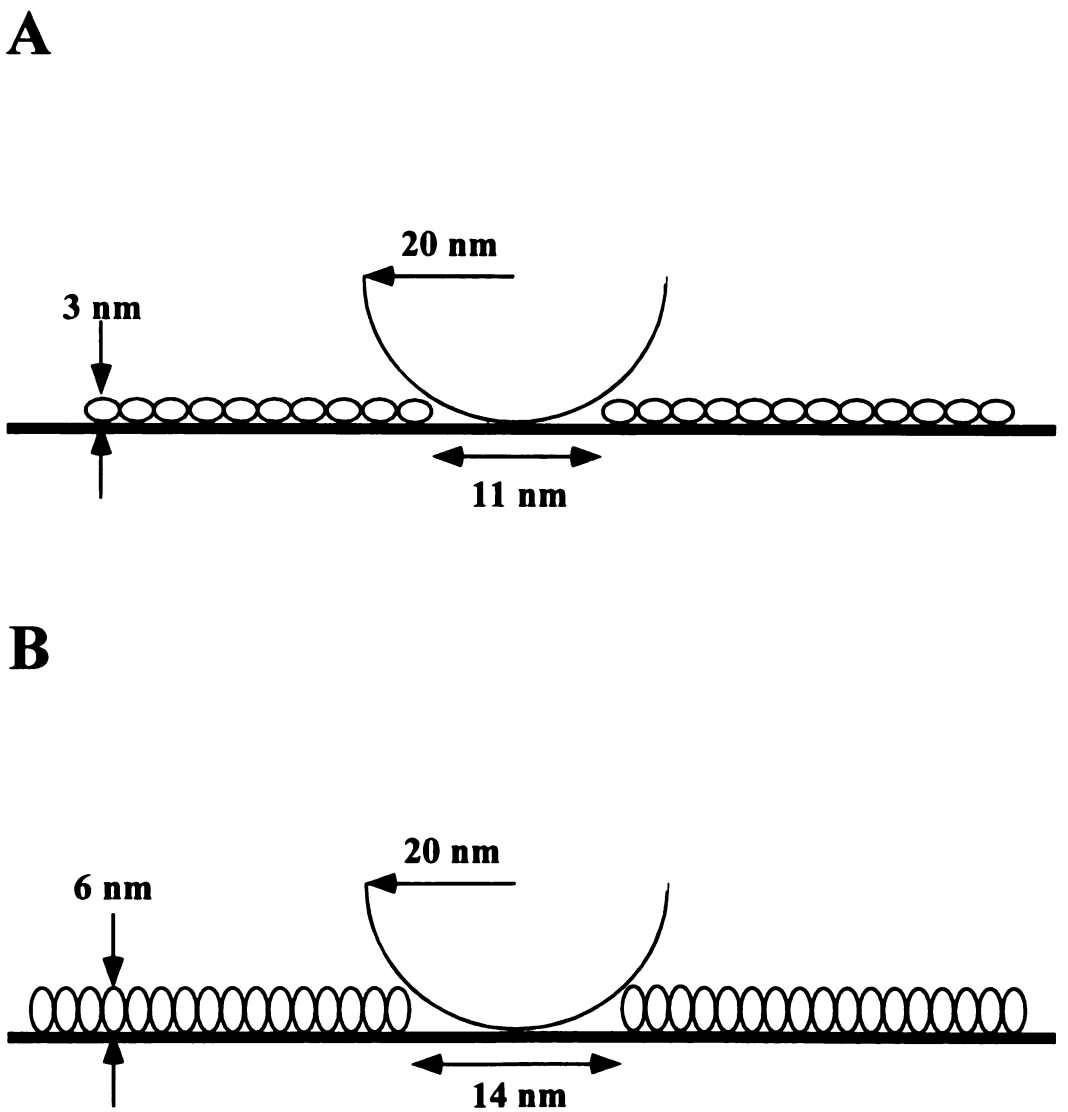


Figure 28: Model illustrating the AFM lateral resolution expected for measurement of thin films on mica.

The same considerations used for the SOMS analysis, namely that the height resolution is on the order of a few angstroms or better and that this height information conveys the molecular orientation information will be used for analysis of the AFM images. In the case of isolated molecules we construct a height distribution from measurements of individual molecular heights. In the case of thin films, where the heights for individual molecules cannot be determined, a slightly different approach needs to be taken to examine the distribution of molecular orientations. Therefore, the height of each pixel in the image (262144 pixels total) was placed into 2 Å bins. For images in which a large area of the protein film was scraped away, or one of a discontinuous film (see Figures 29-37), a distribution of heights containing information about the mica surface and the adsorbed protein molecules is obtained. The population of pixels at lower height values corresponds to the mica substrate, while the higher distributions correspond to the different orientations of molecules in the film. The data were fit with Gaussian functions, in the same way as the height distributions of Chapter four. The orientations of the protein molecules were deduced by taking the heights relative to the mica substrate.

5.3 Results

5.3.1 Growth and orientation of Con A films from pH=7.3 solutions

The first set of experiments involved deposition of Con A from a HEPES buffer (pH=7.3) onto magnesium treated mica. Deposition time was varied between 1 minute and 1 hour. Figures 29-33 give the AFM images and pixel height histograms for these

films. After 1 minute, the Con A molecules have begun to form islands on the mica, but the surface coverage is quite low (see Figure 29). The height histogram given in figure 29 B shows that there are at least three populations of heights for this image. The population centered ~150 Å, which accounts for 53 ± 10 % of the image, corresponds to the height of the mica substrate or background level. One should note that the height values obtained by the constant force mode of imaging are relative heights, due to changes made in the height of the PZT scanner in response to changes in force. If the height of the mica substrate is normalized to 0 Å (which will be done with all of the results given in this chapter), we see that the other two populations are centered at 12 ± 3 Å and 32 ± 2 Å (see Table XI). Surprisingly these values are quite close to the 18-19 Å and 28-30 Å populations observed for ensembles of single Con A molecules on mica (see Chapter four). These data are consistent with the presence of Con A dimers and unfolded protein. The samples of Chapter four were prepared from solutions that were ~10-400 ng/ml Con A. It is known for Con A that the dimeric form is favored when the concentration is below 5 μg/ml [1]. The solutions used to prepare the samples of this section were twenty times more concentrated than this, and were at a pH of 7.3, which is also favorable for the formation of Con A tetramers [11]. This observation of heights corresponding to Con A dimers and denatured protein suggests that the interface, and/or intermolecular forces between Con A molecules are favoring one orientation on the mica surface. The orientation of Con A dimers, which projects the 30 Å dimension normal to the surface is consistent with the observations of Chapter four, however in the case of aggregates of Con A at the interface we observed no 42-45 Å or 59 Å heights corresponding to the other dimensions observed for Con A dimers at the interface. The

other striking feature of this data is that the population of heights centered at 12 ± 3 Å occurs with the same frequency as the population at 32 ± 2 Å. In Chapter four we saw that it was common to observe about 30% of all Con A molecules, deposited from a pH=7.3 solution, as denatured (population centered at 18-19 Å). In the present data there is nearly a fifty-fifty mix of 12 ± 3 Å and 32 ± 2 Å. Much of this 12 Å population may be an artifact of the imaging process (i.e. tip 'convolution'). When the image is scaled such that the population lying 12 Å above the background is highlighted one sees that much of these areas are accounted for by the edges of the islands, where tip convolution is expected. There are, however, some areas which are large enough not to be a tip convolution artifact, but rather to the presence of denatured protein. Figure 30 renders a 1 x 1 μm region of the image from Figure 29 as a contour plot. In Figure 30 the range of the contours is set to be equivalent to the three populations observed in the height distribution of Figure 29. Also in this figure a 15 nm square is given to illustrate the approximate lateral resolution expected for this film. It is clear that many of the 12 Å regions are halos around 30 Å islands, but there are other 12 Å tall regions that are significant. This indicates that either the interface, or the intermolecular forces between Con A molecules is somehow favoring a change in the native dimer conformation.

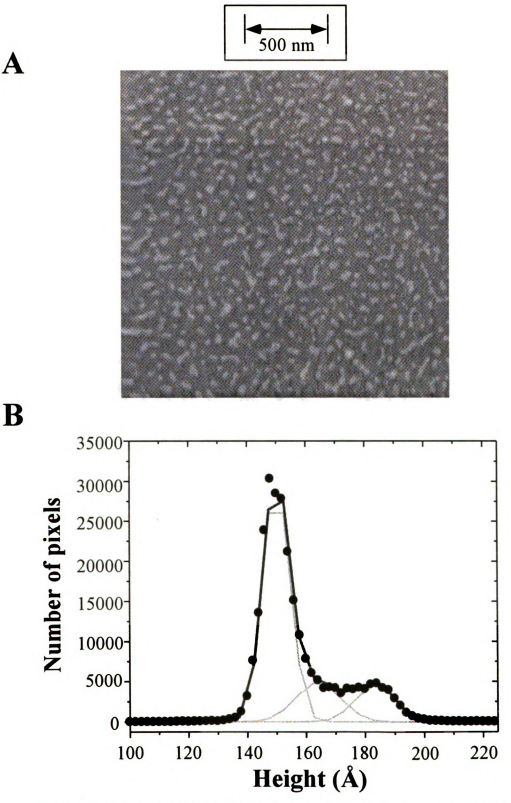


Figure 29: 1 min. deposition of Con A in HEPES on mica. A) AFM image B) Height histogram.

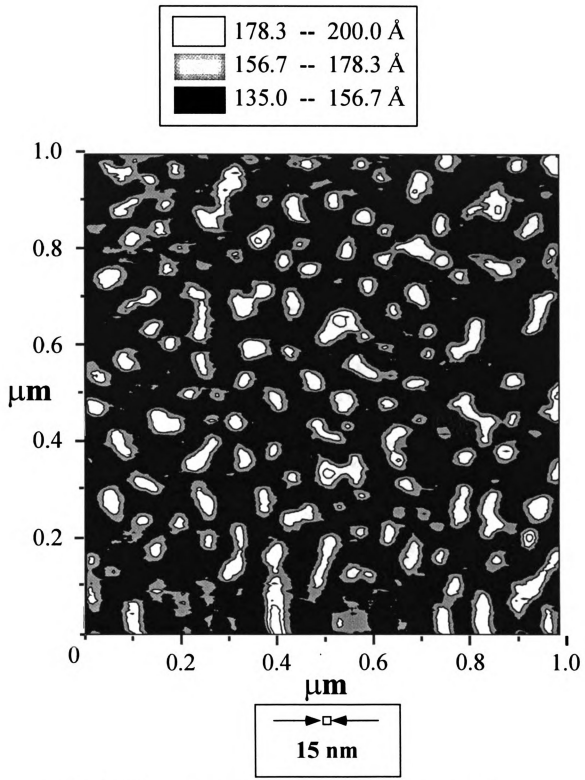


Figure 30: Portion of image from Figure 29, rendered as a contour plot. Approximate lateral resolution is represented by a square drawn to scale of image.

Figure 31 gives a representative AFM image and height histogram for a 5 minute deposition of Con A (HEPES buffer, pH=7.3) onto mica. Again there are three populations observed in the height distribution, corresponding to the mica substrate, and different populations of Con A heights. One should note that there are a few places on the image where there are fairly high heights (i.e. pixels show up as white). The number of pixels associated with these areas are much less than 1% of the total number of pixels and are not apparent in the histogram. It appears that these areas are anomalous, and may be areas where there has been deposition of aggregates, which formed in solution and not on the surface, or possibly due to local differences in the mica surface. The data obtained from this figure are summarized in Table XI. The numbers are quite similar to those observed for the 1 minute deposition, however, there is more surface coverage in this case. Note again that there is a fifty-fifty distribution of 13 ± 1 Å and 30 Å heights, consistent with denatured and dimeric Con A, respectively.

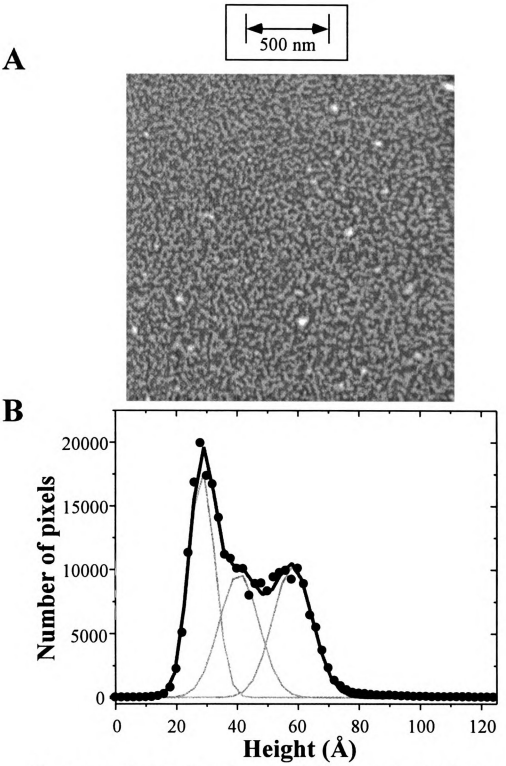


Figure 31: 5 min. deposition of Con A in HEPES on mica. A) AFM image B) Height histogram.

Figure 32 presents representative data for a 10 minute deposition of Con A. The major observation here is that only two distinct populations of heights are observed. The film is still submonolayer and growing by aggregation of the islands seen after 1 minute and 5 minutes. Again, while there are a few anomalous aggregates, the height distribution contains only heights corresponding to the mica substrate and a population of Con A at 18 ± 3 Å. These data were consistent between samples prepared on the same day, as well as ones prepared 1 month apart. This seems to indicate some kind of collapse of the Con A dimer on the surface, or perhaps an imaging artifact. In either case, it is likely that the cause is a cooperative intermolecular force between Con A molecules as it forms a layer at the interface. Because the constant force mode of imaging was used, the height can be anomalously high or low due to large differences in the interaction potential of the probe tip and the sample. In the present case, it is possible (though improbable, as will be discussed later) that as the surface coverage of Con A increases that the probe sample interaction changes leading to a lower measured height. For this to be the case, the interaction potential energy would have to decrease (become less repulsive), which would cause the feedback mechanism to underestimate the true height. This is likely not the case, however, and will be addressed in the following set of experiments.

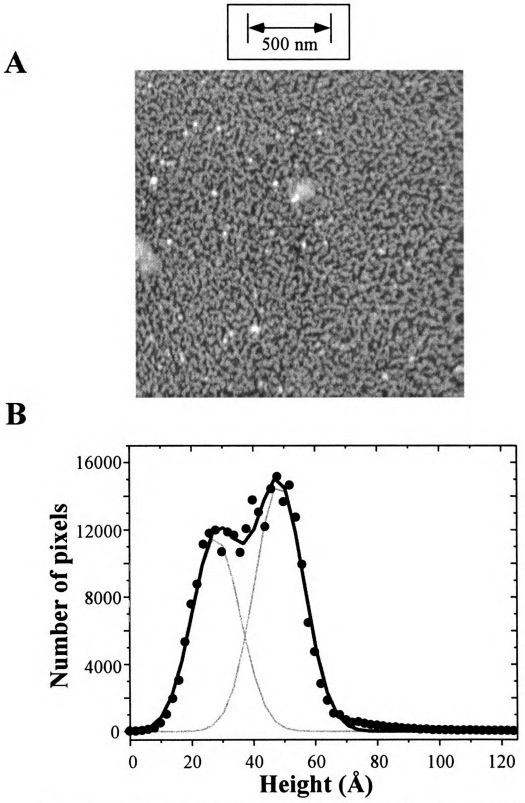


Figure 32: 10 min. deposition of Con A in HEPES on mica. A) AFM image B) Height histogram.

Figure 33 illustrates representative data obtained for a Con A film which was deposited over a 1 hour period of time. In this case it is apparent that there has been complete coverage of the surface with the Con A film. In this case we needed to physically scrape away a small section of the film with the AFM probe, in order to determine the height of the mica background (as described in the experimental section of this Chapter). The height distribution plotted in Figure 33 B shows the presence of at least three different populations of heights. The relative heights obtained, and their relative probabilities are given in Table XI. The two height populations, corresponding to the Con A molecules, are observed at 47 ± 3 Å (87 ± 3%) and 67 ± 6 Å (8 ± 3%). It is quite interesting that after the formation of a continuous film, that the unfolded conformation is not observed, nor is the 30-32 Å high dimeric conformation. There are a couple of possibilities that could explain these observations. One of the explanations is that under these conditions the tetrameric oligomerization state has formed at the interface. The dimensions of the Con A tetramer are 60 Å x 70 Å x 70 Å (Protein DataBank, 1CN1, [13]), which could explain the 67 ± 6 Å population. The 67 Å population is relatively uncommon, only ~8% of the film. It is possible that the large population at 47 ± 3 Å is tetramer which has denatured. Another explanation of the data, which is consistent with the earlier observations, is that a complete monolayer of denatured Con A dimers 18 Å high was produced (consistent with the observation of the 18 Å population after 10 minute deposition). Once this layer formed, a second layer of Con A dimers may have seeded on it. These dimers adsorbing onto the already deposited layer of Con A would likely favor the orientation that maximized interactions with the film (i.e. 30 Å height). This would account for the 47 ± 3 Å population.

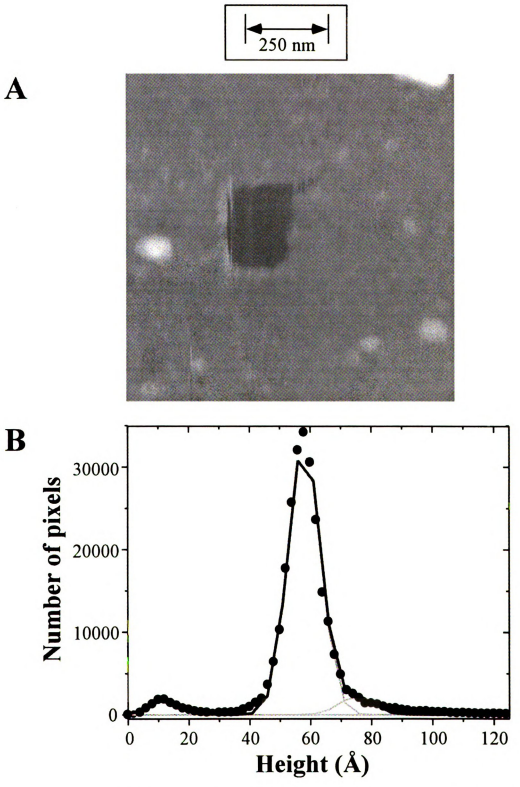


Figure 33: 60 min. deposition of Con A in HEPES on mica. A) AFM image B) Height histogram.

Table XI: Data for Con A films deposited in HEPES buffer (pH=7.3).

Deposition time **# of samples**	**Thickness** **(Å)**	**Relative** **probability (%)**
1 min.	0	53 ± 10
5 samples	12 ± 3	28 ± 8
	32 ± 2	20 ± 3
5 min.	0	41 ± 5
2 samples	13 ± 1	30 ± 1
	30 ± 0	29 ± 4
10 min.	0	37 ± 11
5 samples	18 ± 3	63 ± 11
60 min.	0	5*
3 samples	47 ± 3	87 ± 3
	67 ± 6	8 ± 3

* Background in this data is due to an area of protein that was physically scraped away, as discussed in the experimental section

5.3.2 Growth and orientation of Con A films at acidic pH

This section presents data obtained for the deposition of Con A from a 0.01 M acetate buffer (pH=4.6) onto magnesium treated mica. Figures 34-37 present

representative AFM images and the corresponding pixel height distributions for deposition times ranging from 1 minute to 1 hour.

Figure 34 presents data obtained for a deposition time of 1 minute. It is quite apparent that after only one minute there is significant growth of the Con A film. It has already progressed to the stage of island coalescence. The height distribution of Figure 34 B clearly shows the presence of at least three distinct populations of heights in the image. The data, which include other trials of the experiment, are summarized in Table XII. The two populations observed are centered at 15 ± 2 Å and 36 ± 2 Å, which are within 5 Å of those observed during the growth of films from the pH=7.3 HEPES buffer. It is interesting that the population centered at 36 ± 2 Å is observed 3-4 times more often than that at 15 ± 2 Å. It appears that the Con A is again present as the dimeric and denatured state, however there is a much higher preference for the native dimer in the energetically most favorable orientation. This is consistent with the observations of Chapter four, in which the orientation projecting the 30 Å height was highly favored at acidic pH.

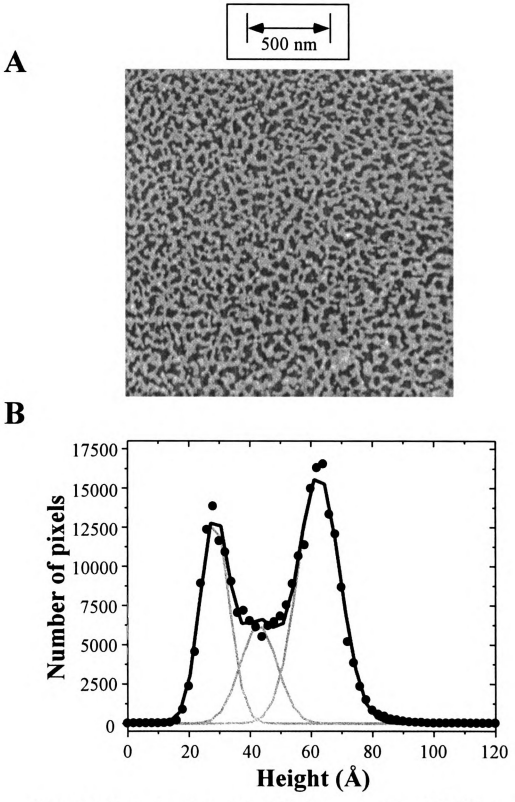


Figure 34: 1 min. deposition of Con A in Acetate on mica. A) AFM image B) Height histogram.

Figure 35 presents representative data of Con A deposited for 5 minutes on mica. It is clear that the film growth has nearly completed the first monolayer. The height distribution (see Figure 35 B) again shows three populations, the relative heights of which are presented in Table XII. Similarly to the data of the previous section (see Figure 32 and Table XI), the height distribution shows a predominance of the 18 Å, or denatured, population. In this case, however, there is a small (~6% of the pixels in the image) population centered at 34 Å, which correlates well with the presence of the Con A dimer.

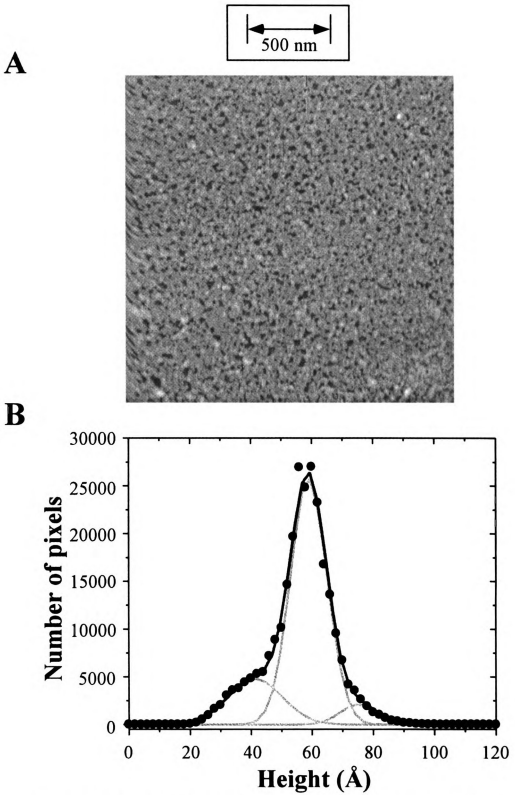


Figure 35: 5 min. deposition of Con A in Acetate on mica. A) AFM image B) Height histogram.

Figure 36 presents data obtained for a 10 minute deposition on mica. In this case the layer was nearly continuous and required a physical scraping away of a small area, for the determination of the mica background. The data for these samples is summarized in Table XII. Besides the population due to the mica background, there are populations centered at 41 and 55 Å. The predominant population is that centered at 41 Å. This population is similar in height to that observed for the one hour deposition of Con A from the HEPES buffer (pH=7.3), and might be explained in the same way.

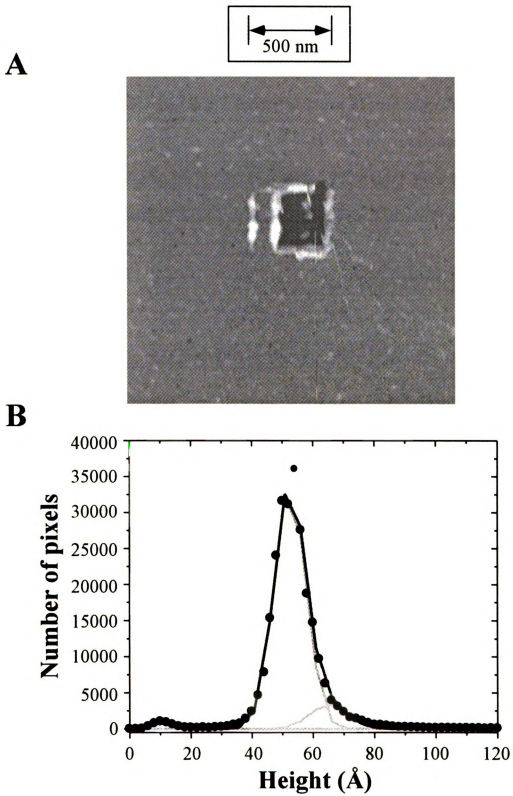
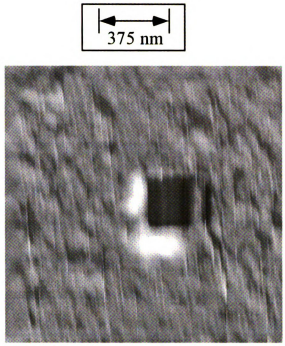


Figure 36: 10 min. deposition of Con A in Acetate on mica. A) AFM image B) Height histogram.

The data obtained for a 1 hour deposition of Con A from the acetate buffer (pH=4.6) is given in Figure 37. Figure 36 A is an AFM image of a 1.5 x 1.5 μm area showing a small area of the film scraped away. There are four populations necessary to obtain a reasonable fit to the data. The three populations corresponding to the film are centered at 47 ± 8 Å, 68 ± 5 Å and 92 ± 1 Å. The first two populations correspond quite well with those observed for the Con A in HEPES film (1 hour deposition) of the previous section. The 92 Å population is likely due to the presence of a multilayer film.

A



B

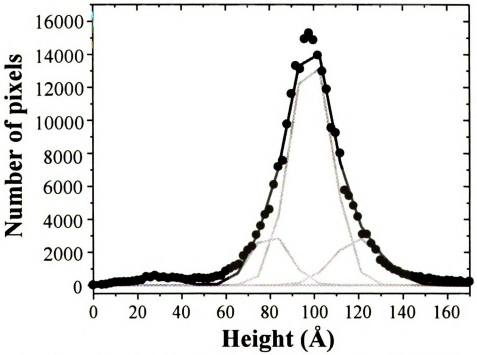


Figure 37: 60 min. deposition of Con A in Acetate on mica. A) AFM image B) Height histogram.

Table XII: Data for Con A films deposited in Acetate buffer (pH=4.6).

Deposition time **# of samples**	**Thickness** **(Å)**	**Relative** **probability (%)**
1 min.	0	26 ± 4
4 samples	15 ± 2	16 ± 3
	36 ± 2	58 ± 6
5 min.	0	20
1 sample	18	74
	34	6
10 min.	0	3 ± 1*
2 samples	41	86 ± 2
.	55 ± 1	11 ± 1
60 min	0	4 ± 2*
2 samples	47 ± 8	16 ± 4
	68 ± 5	62 ± 9
	92 ± 1	18 ± 3

* Background in this data is due to an area of protein that was physically scraped away, as discussed in the experimental section.

5.3.3 Kinetics of the deposition process at physiological and acidic pH

Examination of Figures 29-37, and Tables XI and XII clearly shows that the rate of adsorption of the Con A to the mica surface is faster at the lower pH, where the protein molecules are positively charged. In Chapter four it was noted that the rate of adsorption did not appear to change with pH, for deposition of isolated molecules from very dilute solutions, however, under the present conditions the pH has a significant effect on the rate. For very dilute solutions, it is hypothesized that the electrostatic interactions between the protein molecules and the mica surface are shielded by the solvent. In the case of more concentrated protein solutions, the electrostatic interactions between the protein and the substrate appear to become important. This may be simply a fact of the larger number of molecules which would be in close proximity to the surface at any given instant, as compared to the case of the solutions of Chapter four. In the case of the dilute solutions the rate of adsorption is likely limited by the Brownian motion, or diffusion of the protein molecules toward the surface. In the present case, however, one of the dominant rate determining factors appears to be electrostatic attraction.

The kinetics of the adsorption process were investigated by plotting the amount of amount of substrate surface (i.e. 100% - % surface coverage) present as a function of deposition time. It was found that a plot of the natural logarithm of the % uncoated surface observed versus time gave a straight line (see Figure 38). This indicates that this adsorption process is first order or pseudo-first order. The negative slope of the plot

gives the first order rate constant in units of min^{-1}. The data were fit by linear regression, and only the submonolayer films were included in the plot.

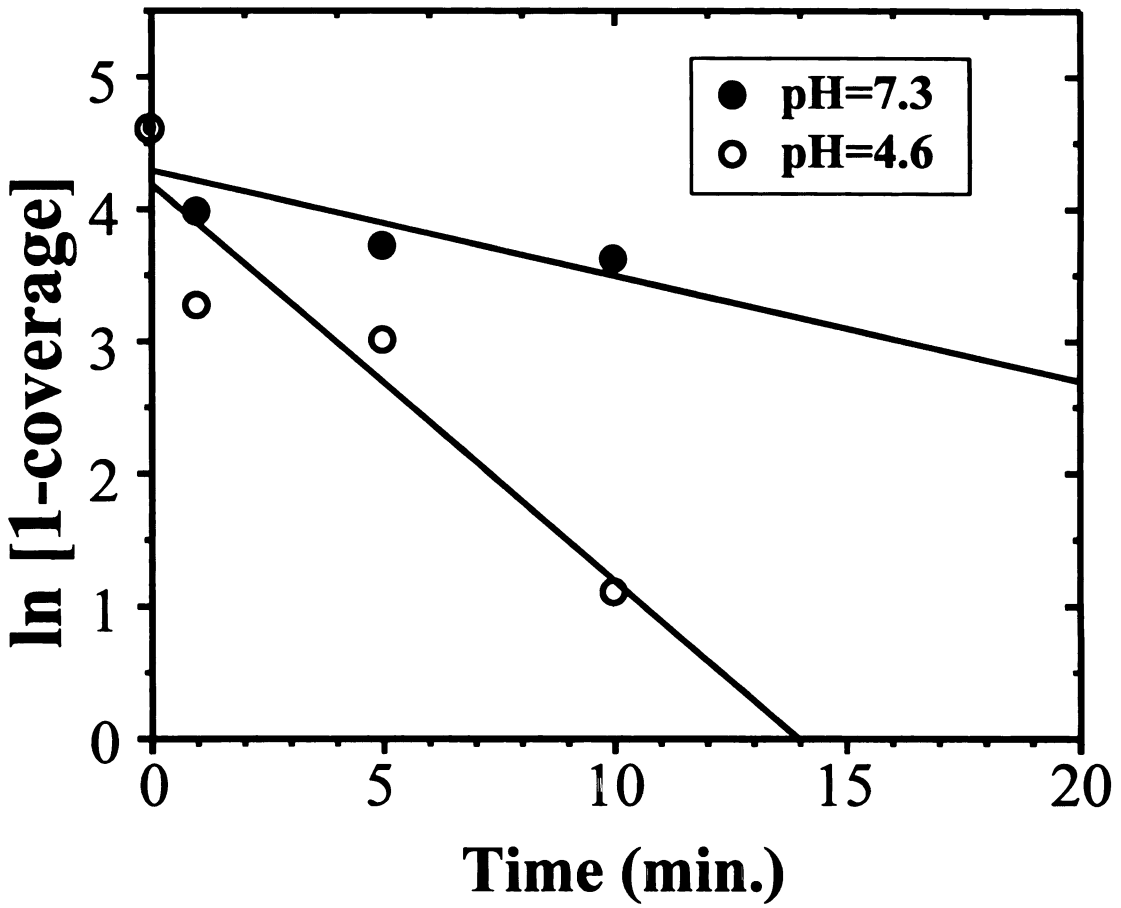


Figure 38: First order kinetics plot of Con A data at two different pH values.

The rate constant for deposition from the pH=7.3 HEPES solution was found to be 0.08 min^{-1}, while the rate constant for the pH=4.6 acetate buffer was 0.30 min^{-1}. This is a difference by nearly a factor of four in the effective rate of film growth.

5.3.4 Morphology of Con A thin films at the water/mica interface

It has been previously reported that thin films of Con A are quite rough [4,5,7]. The current study has found them to be significantly rougher than what would be expected if the molecules were forming ordered arrays on the mica substrate. The Proscan (Park Scientific Instruments) image analysis program was used to calculate RMS and average roughness for the continuous thin films deposited on mica. The results of this analysis are given in Table XIII. Typically values of 3-5 Å of average roughness can be expected for mica, over the typical 2 x 2 μm scans acquired with the AFM. The roughness data presented in Table XIII was calculated for 0.25-1 μm areas, where the film was seen to be continuous and not containing any large aggregates on the film surface. It is clear that the films are typically rougher than the mica substrate, and that the films deposited from the acetate buffer are significantly rougher than those deposited from HEPES.

Table XIII: RMS and Average roughness for Con A thin films.

Solvent:	**HEPES**	**pH=7.3**	**Acetate**	**pH=4.6**
Roughness:	*RMS*	*Average*	*RMS*	*Average*
(Å)	7	5	12	9
σ	2	1	2	1

In order to get a better idea of the film morphology, cross-sectional slices of the film image were plotted. Representative line profiles, for films deposited from the different buffers, are plotted in Figure 39. The data presented in these figures has been normalized, such that the minimum height value of the line is set to be zero. The films deposited from the HEPES buffer tend to have heights that vary between 0 and 30 Å, while those deposited from acetate typically vary from 0 to 60 or 70 Å.

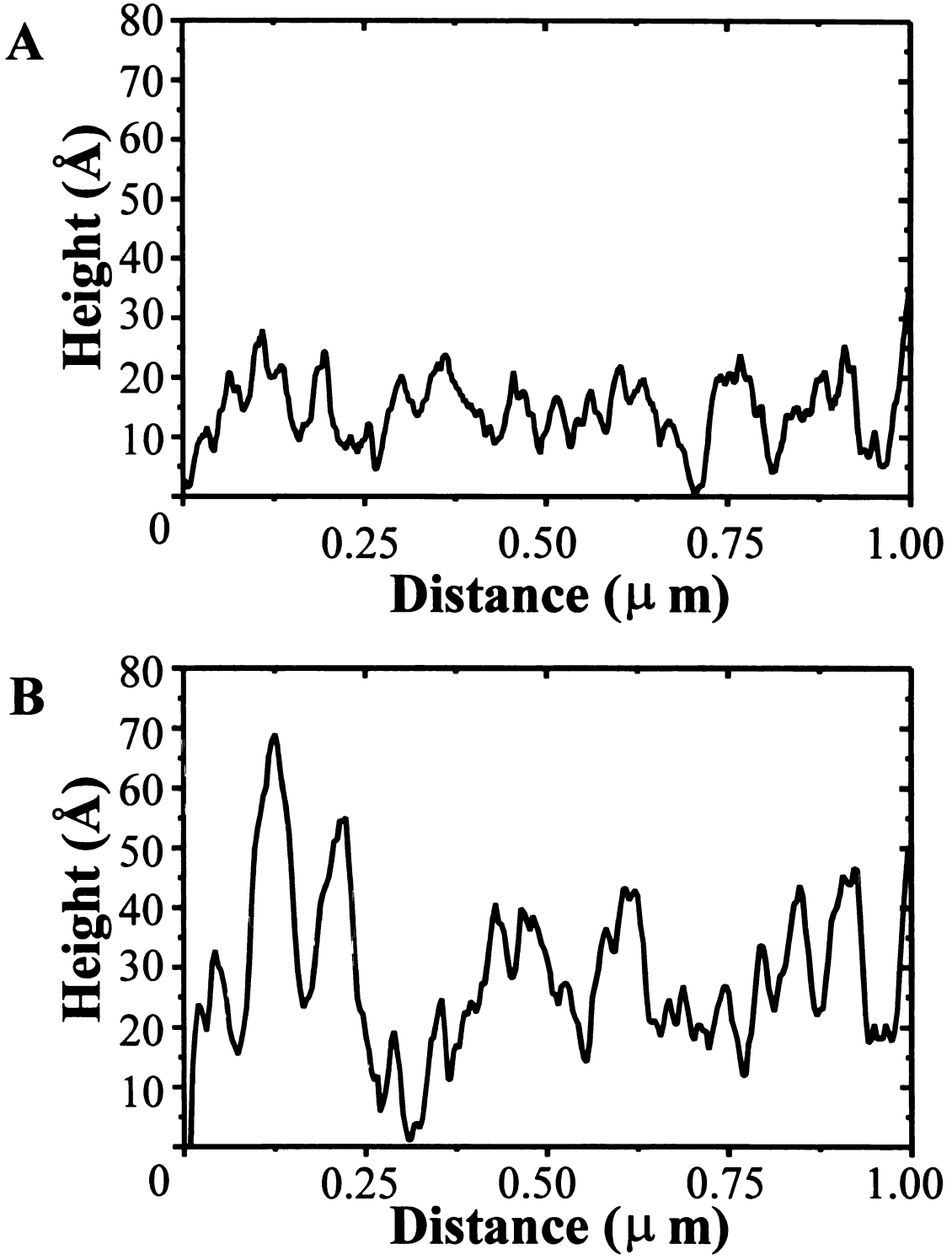


Figure 39: Line profiles taken from AFM images of Con A thin films, deposited at A) pH=7.3 and B) pH=4.6.

The origin of the differences in the morphology of these films can be understood in part by consideration of the difference in the deposition kinetics between the two buffer solutions used. As discussed in section 5.3.4 Con A deposits onto mica at a significantly faster rate from a buffer at pH=4.6 than it does from a solution at pH=7.3. Because the molecules are depositing at a much faster rate from the pH=4.6 solution, there is less time to maximize interactions with the surface and adjacent Con A molecules. This may partly explain why these films appear rougher than those deposited at a slower rate.

Another possible explanation for the morphological differences may in part be due to the formation of multi-layer films. The data presented in Tables XI and XII show that films formed from one hour depositions show features at approximately 47 Å and 67 Å heights in both buffer systems used. The film deposited from the acetate buffer, however, has another feature in its height distribution, appearing at 92 Å. Additionally there is a difference in the relative proportions of each orientation. In the case of the HEPES system the 47 Å heights account for ~87 % of the film, while in the acetate system only ~16 % of the film has this orientation. This inverse proportionality also appears in the relative amount of the 67 Å populations. This data is consistent with the faster kinetics observed, and the formation of multi-layers of Con A on the mica surface, which appears to account for the differences in the observed film morphology.

5.4 Discussion

The data presented in section 5.3.1 and 5.3.2 examined the distribution of orientations for Con A molecules at the water/mica interface as the growth of thin films progressed. Multiple heights were measured, even at very early times in the film formation, where islands of the protein molecules had not yet grown together to form a continuous layer. The heights which are observed at the various deposition times are consistent with previous studies of Con A film thicknesses [3,6,8,10]. In the previous studies of Con A films adsorbed to solid supports, the techniques used (i.e. SFA, QCM, XPS) provided only average thickness of layers. The current study has used the AFM to examine layer thickness during film formation, providing not just average thickness, but also showing the presence of multiple conformations or orientations of Con A at the interface.

The growth of the Con A films appeared to go through stages. The three stages observed in this study are schematically illustrated in Figure 40. In this figure Con A dimers are represented by ellipsoids, with dimensions corresponding to the dimensions of the native and denatured form of the Con A dimer.

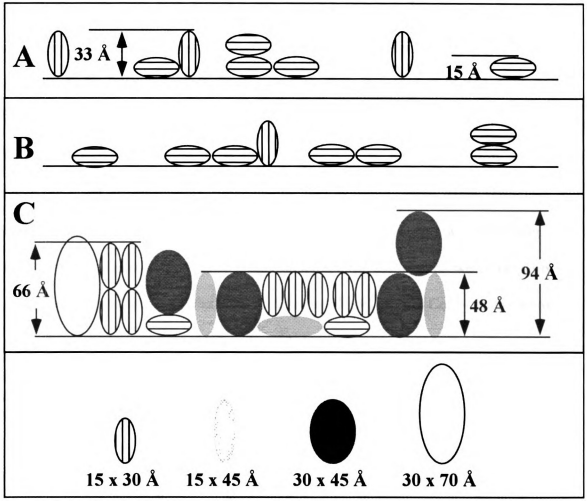


Figure 40: Possible orientations of Con A molecules at the water/mica interface, as a function of time: A) early growth, B) intermediate growth, C) continuous film (as discussed in the text).

In the first stage (see Figure 40A) there was formation of two predominant orientations or conformations observed at 12-15 Å and 30-36 Å. Although the relative amount of each of these orientations was nearly 50%, much of the 12-15 Å population can be attributed to the tip convolution artifact. This is consistent with the presence of Con A dimers, as was observed in Chapter four for Con A deposited from very dilute solutions. The observation of dimer species was expected under acidic conditions, but the presence of dimers deposited from a HEPES buffer at pH=7.3 was surprising. It is well known that in solution Con A favors the tetrameric organization when the pH is 7 or above, while the dimeric species is favored below pH=6 [11,12]. Con A is also known to form dimers when it is diluted below 5 μg/ml [1], as was the case in the solutions used in Chapter four, but observation of dimer species under the present conditions is not generally predicted.

The second stage (see Figure 40B) of film growth involves a change in the orientation of the Con A molecules, where the predominant species have a height of about 18 Å. This is consistent with the presence of unfolded or denatured Con A (see Chapter four).

The third stage of growth (see Figure 40C) occurs when the protein molecules have formed a continuous monolayer, and/or they have begun to seed multi-layers. The predominance of areas with heights centered around 47 Å and 67 Å could be due to a few possible orientations of Con A dimers on the surface, as illustrated in Figure 39C. The 67 Å measurement, however may also be indicative of the formation of Con A tetramer at

the interface. The current results cannot distinguish between the different possible modes of Con A surface orientation for these films. This could be examined in the future by careful studies of saccharide binding affinity as a function of film growth. Since the location of the metal centers (a.k.a. sugar binding sites) are well known, it may be possible to determine the orientation of the Con A molecules, and possibly whether there are tetrameric Con A molecules present. These results indicate that Con A tends to exist at the water/mica interface as a dimer. It appears that this variation from the tetrameric form of Con A, which is typically observed for Con A solutions at neutral pH, is a result of the chemistry of the interface affecting the stability of the tetramer.

The kinetics of the deposition process is dramatically affected by the pH of the deposition solution. The rate of deposition appears to be almost four times faster when deposition takes place from an acetate buffer at pH=4.6, as opposed to the HEPES buffer at pH=7.3. This is in contrast to the results presented in Chapter four, where the rate of deposition from dilute solutions $\sim 10^2$ ng/ml was not significantly different for the pH=7.3 and pH=4.6 solutions. The likely explanation for this is that in dilute solution, the electrostatic interactions between the protein and surface are shielded sufficiently that the deposition process is diffusion limited. In the case of more concentrated solutions (i.e. $\sim 10^2$ μg/ml) the deposition process is not diffusion limited, and the electrostatic attraction between the positively charged Con A (at pH=4.6) and the negative surface charge of the mica substrate increases the rate of deposition.

This difference in the rate of adsorption with pH can also begin to explain the difference in the surface morphology of the films. The films produced by deposition from the acetate buffer system were significantly rougher than those deposited from HEPES buffer. The faster deposition may not allow sufficient time for the Con A molecules to maximize their orientation on the surface before more molecules can come in and deposit.

The study of Con A thin films by AFM has not been previously published, although this system has been interrogated and characterized by a number of other techniques. This study links together what was learned about the orientation and structure of individual Con A molecules at the water/mica interface (see Chapter four and Appendix 1), and what has been determined by previous studies of the bulk properties of Con A thin films. It is hoped that this information can be used to better understand the function of and to design better saccharide sensing devices, and possibly develop new Con A-saccharide based bioassays. Furthermore, the approach used in this study demonstrates that valuable information can be obtained from AFM data, despite the lateral tip convolution problem.

5.5 References

1. Gordon, J. A.; Young, R. K. *J. Biol. Chem.* **1979**, 254, 1932.

2. Janata, J. *J. Amer. Chem. Soc.* **1975**, 97, 2914.

3. Afshar-Rad, T.; Bailey, A. I.; Luckham, P. F.; MacNaughtan, W.; Chapman, D. *Biochim. Biophys. Acta* **1987**, 915, 101.

4. Barnes, C.; D'Silva, C.; Jones, J. P.; Lewis, T. J. *Sens. Act. B* **1991**, 3, 295.

5. Haas, H.; Mohwald, H. *Thin Solid Films* **1989**, 180, 101.

6. Gallinet, J.-P.; Gauthier-Manuel, B. *Eur. Biophys. J.* **1993**, 22, 195.

7. Safar, J.; Roller, P. P.; Ruben, G. C.; Gajdusek, D. C.; Gibbs, C. J. j. *Biopolymers* **1993**, 33, 1461.

8. Lvov, Y.; Ariga, K.; Ichinose, I.; Kunitake, T. *Thin Solid Films* **1996**, 284-285, 797.

9. Kremer, F. J. B.; Engbersen, J. F. J.; Kruise, J.; Bergveld, P.; Starmans, D. A. J.; Feijen, J.; Reinhoudt, D. N. *Sensors and Actuators B* **1993**, 13-14, 176.

10. Fitzpatrick, H.; Luckham, P. F.; Eriksen, S.; Hammond, K. *J. Coll. Inter. Sci.* **1992**, 149, 1.

11. Reeke, G. N., Jr.; Becker, J. W.; Cunningham, B. A.; Wang, J. L.; Yahara, I.; Edelman, G. M. Structure and Function of Concanvalin A. In *Concanavalin A*; Chowdhury, T. K., Weiss, A. K., Eds.; Plenum Publishing Corp.: New York, 1975.

12. Liener, I. E. Isolation and Properties of Concanavalin A. In *Concanavalin A as a Tool*; Bittiger, H., Schnebli, H. P., Eds.; John Wiley & Sons:, 1976; pp 2.

13. Shoham, M.; Yonath, A.; Sussman, J. L.; Moult, J.; Traub, W.; Kalb, A. J. *J. Mol. Biol.* **1979**, 131, 137.

APPENDIX 1

Re-print of publication

Imaging the Molecular Dimensions and Oligomerization of Proteins at Liquid/Solid Interfaces

Mark J. Waner,† Martha Gilchrist,† Melvin Schindler,‡ and Marcos Dantus*,†

Departments of Chemistry and Biochemistry, Michigan State University, East Lansing, Michigan 48824

Received: October 2, 1997; In Final Form: December 10, 1997

Individual Concanavalin A (ConA) molecules have been imaged at the liquid/solid interface with an atomic force microscope (AFM). Three-dimensional sizing with very high resolution (<5 Å) has been obtained by a novel approach based on height distributions, which avoids the tip convolution effects which normally affect scanning probe microscopy techniques. Each height measurement correlates to a particular molecular orientation on the surface. A large number of such measurements provide a statistical ensemble of orientations. The complete height distribution reflects the three-dimensional size of the protein sample and hence its tertiary and quaternary structure. A surface adsorption and orientation model, based on a minimization of surface adsorption energy, is proposed. This model is in good agreement with the observed height distribution of Con A molecules at the liquid/solid interface. Analysis of Con A and succinylated Con A molecules on mica demonstrates that Con A dimers are the prevalent species at the liquid/solid interface. This is in contrast to the tetrameric organization of Con A normally observed in solution. The new possibilities opened by height distribution analysis on the physical characterization of biomolecules at interfaces are also discussed.

Introduction

The behavior of proteins at the solid/liquid interface has been of great interest because of the fundamental role that membranes, cytoskeletons, and other interfaces play in cells, as well as for the more practical aspects related to the development of biological assays, biocompatibility of materials, and protein processing.[1−3] The number of techniques available to quantitatively measure structure and structural changes of individual proteins at interfaces, however, is very limited. In this paper we present very high-resolution (<5 Å) three-dimensional measurements of Concanavalin A (Con A) molecules at the water/mica interface. The measurement of the molecular size allows us to determine the oligomerization state and an estimate of the surface adsorption energy for different protein orientations at the interface.

The adsorption of a protein on a solid surface is a complex process.[4,5] Current models identify two steps in the adsorption process. The first step involves the collision of the protein, which is translating by Brownian motion, with the surface. If the energetic benefit of the removal of surface adsorbed water is greater than the thermal energy of the protein molecule, then it remains on the surface; otherwise, it returns to solution.[5] Once the protein is adsorbed, the second step is thought to involve a conformational change that maximizes electrostatic interactions, local interactions between electron donor moieties in the protein with electron acceptors on the surface, and maximization of van der Waals attractions.[6] It is often assumed that this second step involves changes in the protein structure, akin to denaturation. There is now some experimental evidence that "compact" globular proteins maintain their native structure when adsorbed and that only some "soft/flexible" proteins suffer some degree of denaturation upon adsorption, especially on hydrophobic surfaces.[7] The second step, therefore, involves primarily an orientation process that is thermodynamically driven and yields proteins in the most energetically favorable orientation. Independent evidence of protein orientation at interfaces has been obtained by Lee and Saavedra.[8]

Since its recent introduction,[9] atomic force microscopy (AFM) has been extensively used for imaging biological molecules.[10−14] The AFM technique allows the examination of biological specimens without the need for chemical fixation/dehydration or growth of crystals, which is normally associated with other high-resolution techniques.[10,15,16] In fact, samples may be imaged with AFM under liquids at physiological concentrations and pH.[17] A number of groups have utilized the high resolution (~1 Å) of the AFM for examination of 2D crystallization of protein molecules.[11,18−21] Others, especially Hansma, Bustamante, and Henderson, have studied isolated proteins[17,22−28] or DNA molecules[12,29−33] and their complexes.[34] Recently, AFM has been used by the Marchant group to obtain valuable structural information for the multimeric Von Willebrand factor (MW ~ 260 kD, monomer).[25,35,37]

In the case of smaller globular proteins and other small three-dimensional structures, however, AFM has been limited by its inherent low resolution (10−20 nm) in lateral measurements. The difference between height and lateral resolution (~1 vs ~100 Å, respectively) is due to finite size of the probe tip (~10−50 nm radius of curvature) and the sample being examined[36−39] (see Figure 1A). This "convolution" effect occurs regardless of imaging conditions (vacuum, ambient, or under liquid). Some groups have proposed making sharper probe tips to reduce this problem,[36] while others have examined ways to first characterize the shape of the probe tip and then mathematically "deconvolute" the AFM data.[39] Because the height measurements obtained by AFM are unaffected by the size of the probe, resolution of 1 Å or less can be realized in this dimension.

In this study, we extend the utility of AFM to the quantitative measurement of molecular dimensions with angstrom resolution

* Corresponding author. e-mail: dantus@cem.msu.edu.
† Department of Chemistry.
‡ Department of Biochemistry.

S1089-5647(97)03221-5 CCC: $15.00
Published on Web 02/04/1998

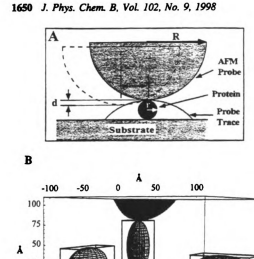


Figure 1. (A) Schematic representation of the AFM probe/protein interaction. Note that height measurements are not affected by tip convolution. There can be a minor discrepancy in the measured height (d) when the probe of radius (R) is displaced by a distance (x) from the apex of the protein, of radius (r). In our measurements this discrepancy amounts to a 0.5% error (see text). (B) Representative drawing of Concanavalin A dimers arranged in three orientations on a mica surface with an AFM probe positioned on top of one of the molecules. The orientation model assumes that the proteins may be crudely approximated as rectangular parallelepipeds and predicts their orientation on the surface according to the relative dimensions of the rectangular parallelepiped. The size of the molecules, the probe, and the mica lattice are drawn to scale.

at liquid/solid interfaces. Such measurements are of particular biological relevance since the preponderance of biological activities of enzymes and their binding properties occur on membrane or cytoskeletal surfaces. To avoid the loss of resolution that is normally introduced into AFM measurements as a result of probe tip geometry, we have devised a measurement strategy that measures the heights of an ensemble of individual protein molecules that have adsorbed to the atomically flat surface of mica (Figure 1B). The technique has been termed surface oriented molecular sizing (SOMS) because it utilizes the statistical thermodynamic distribution of orientations for adsorbed molecules to determine the three-dimensional structure of proteins. The use of the projected height of individual adsorbed protein molecules as the defining parameter for molecular dimensions makes this approach minimally dependent on probe tip geometry and utilizes the angstrom resolution of the AFM in the z direction. Indeed, it is advantageous to use a broader tip for making these measurements, thereby minimizing the applied pressure between the tip and protein and decreasing the chance of underestimating the height of a molecule due to the digitization inherent in the microscope. Furthermore, this technique is generally applicable to contact and intermittent contact modes of imaging under ambient or liquid conditions. The relative ease, speed, and resolution with which these measurements of molecular dimension may now be performed make this technique particularly useful for examining multiple conformational equilibria of adsorbed protein subunits and the changes in shape and assembly state of multi-subunit proteins resulting from oligomerization and aggregation at liquid/solid interfaces. Such changes may be systematically examined as a function of protein concentration and ligand or allosteric activator binding. Such measurements could help to determine whether single amino acid substitutions affect protein activity as a consequence of an induced change in the oligomerization state of the mutant protein or as a result of a chemical change in the active site. We will present evidence of this phenomenon in a future publication.

We present in this article a study of Con A at the liquid/solid interface. Our findings demonstrate that, at the water/mica interface, Con A exists primarily as a dimer, which is in contrast to the tetrameric form which is typically observed in solution. The dimeric structure is confirmed by measurements on succinylated Con A. We also explored the possibility of denaturation at the liquid/solid interface by comparison to measurements on sodium dodecyl sulfate (SDS) treated Con A samples. Confirmation is also obtained from measurements on ferritin protein molecules, which are found to agree well with molecular dimensions derived from images obtained with X-ray diffraction and electron microscopy. The application of a surface adsorption and orientation (SAO) model in conjunction with AFM to the study of protein structure and oligomerization at the liquid/solid interface in the presence of activators, inhibitors, and denaturing agents will be discussed.

Experimental Section

Materials. Colloidal gold particles were obtained from Sigma (St. Louis, MO) and used without further purification. Three sizes of gold particles were used for the calibration: 50 $\pm$ 8 Å (3.8×10^{13} particles/mL), 90 $\pm$ 12 Å (6.0×10^{12} particles/mL), and 180 $\pm$ 8 Å (6.7×10^{11} particles/mL). The dimensions and their deviations were determined by the manufacturer using TEM. Cationized ferritin was obtained from Polysciences (Warrington, PA) as a 20 mg/mL solution. Concanavalin A (Con A) was obtained from Boehringer-Mannheim (Indianapolis, IN) as lyophilized solid. Succinyl Concanavalin A was obtained from Sigma (St. Louis, MO) as a 95% protein lyophilisate. Protein purity was checked by SDS polyacrylamide gel electrophoresis. HEPES was obtained from Sigma (St. Louis, MO), and the $MgCl_2$ was ACS grade from Columbus Chemical Industries (Columbus, OH). The water used in preparing solutions was purified with a MilliQ Type I MilliPore water purification unit (19 MW resistivity). The concentrations of Con A and succinyl Con A stock solutions were determined using A_{280} with an extinction coefficient of 1.37 mL mg^{-1} cm^{-1}.[40]

Preparation of Samples for Height Measurements with AFM. Measurements of protein height require an atomically flat substrate (less than 5 Å corrugations). In this study we used muscovite mica obtained from Ward's Natural Science Establishment, Inc. (Rochester, NY.) as the substrate. In most cases we found a flatness better than 3 Å over areas of tens of microns squared. When freshly cleaved, mica has a net negative surface charge. Treating a freshly cleaved surface of mica with a 5 mM $MgCl_2$ solution results in the replacement of endogenous K^+ ions with Mg^{2+} making the surface more positively charged. The magnesium-treated surface has been demonstrated to enhance adsorption of macromolecules.[41] Proteins were dissolved in 10 mM HEPES solution at pH 7.3. (The concentration used in each case is specified for each protein.) Twenty-five microliters of protein solution is deposited onto the mica. Twenty minutes later, the surface is gently rinsed twice with

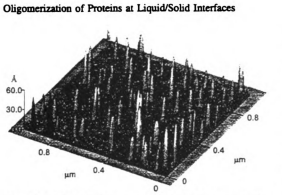


Figure 2. Typical 1 × 1 μm scan of Con A molecules on a mica substrate used to acquire height data. The sample was prepared as described in the Experimental Section. The different heights (amplified ×50) correspond to different orientations of Con A dimers (see text).

200 mL portions of MilliQ water and allowed to dry for several hours (at room temperature and 35−50% relative humidity). This preparation yields a dispersed population of individual protein molecules when imaged by AFM (Figure 2). The concentration of protein has been chosen to provide a large population of individual protein molecules and to avoid the seeding of aggregates and monolayers. Overall, the protein coverage is always kept below 1%. In the presence of a relative humidity of ~30%, it has been shown that there is a monolayer of water maintained on the mica surface which can act to hydrate and stabilize the adsorbed proteins.[42]

Atomic Force Microscopy (AFM). The microscope used for this work was an Autoprobe CP scanning probe microscope (Park Scientific Instruments), with a 5 μm high-resolution scanner. Height measurements may be obtained from AFM studies carried out by contact mode or intermittent contact mode in ambient or under liquid conditions. For this study the AFM data were recorded in contact mode under controlled ambient conditions (~40% humidity and 21° C). Under these conditions a monolayer of water exists at the mica surface,[42] and the protein molecules are fully hydrated. It was found that the protein molecules adhered to the mica well enough for measurements only when the humidity was less than 50%, probably because high humidity increases the attractive capillary forces between probe and sample.[12,24,42,43] Humidities of less than 30% were avoided to prevent protein dehydration and possible structural deformations. The effects of controlled hydration on proteins have been studied by scanning tunneling microscopy (STM).[44,45] The largest changes were found to occur as the humidity dropped below 30%. Unfortunately, those measurements did not provide quantitative structural information because the contrast mechanism (tunneling current) is highly dependent on the amount of solvent.

Sharpened Microlever cantilevers (Park Scientific Instruments) having a spring constant of approximately 0.05 N/m and an integrated silicon nitride probe with a radius of curvature of approximately 20 nm, according to the manufacturer, were used as probes. The scanner was calibrated for vertical measurements using Tobacco Mosaic Virus (cylindrical shape; diameter 18 nm) (American Type Culture Collection) and colloidal gold particles (5 nm) (Sigma, St. Louis). Horizontal measurements are calibrated using a two-dimensional grating with 1000 features per millimeter.

AFM images were acquired in constant force mode, in which the PZT scanner moves in the z direction so as to maintain a constant force between the sample and the probe tip. To determine the minimum force to achieve stable imaging with negligible or no measurement distortion, we imaged ferritin protein molecules with a range of applied forces. At low imaging forces we accurately reproduced the well-characterized dimensions (a sphere with 120 Å diameter) of this protein.[47] At higher imaging forces a deformation of ~10% was observed. We compared the force measurements on ferritin with similar ones on gold particles. As expected, the gold particles were more resistant to deformation due to applied pressure. The minimum force for stable imaging was found to occur within the electrostatic repulsive regime and was constant for all images. The total force applied between the tip and the cantilever under our imaging conditions contains a large attractive or capillary force (~10^{-9}−10^{-8} N, depending on the ambient humidity[13]) in addition to the overall repulsive force. The capillary forces between tip and sample may be minimized by imaging under liquid. These types of measurements in combination with the proposed SOMS technique may be ideally suited for the study of proteins at interfaces and are of great interest in our group.

Given that the overall accuracy of our method is limited by the roughness of the substrate to ~3−5 Å, force distortions can be neglected except for very soft and large (> 100 Å) features of proteins (not the case in the results presented in this article). In those cases we can use intermittent contact mode scanning techniques under liquid, where the tip−protein interaction forces can be further minimized.[13] Scan size for images was 2.0 × 2.0 μm and 512 × 512 pixels at a linear scan rate of 1.5 Hz, which is equivalent to 3 μm/s along the fast scan. The fairly slow linear scan rate was chosen to help minimize the lateral forces that occur between the tip and sample during the raster scanning process. The lateral forces can also be further minimized by using intermittent or noncontact modes of imaging, which are currently being investigated.

Each image of individual protein molecules contained approximately 10−50 molecules/μm^2. Individual protein molecules are measured by taking the maximum height of the molecule minus the height of the local background. Each data set, consisting of more than 1000 individual height measurements, is tabulated in 3 Å bins resulting in a height distribution histogram. The choice of bin size is based primarily on the substrate roughness. Our results are essentially independent of this choice. A combination of Gaussian functions is then used to fit the data. This fitting was done iteratively using constraints on the three parameters of a typical Gaussian function (i.e., center position, width, and amplitude). In all cases, the amplitudes of the functions were allowed to vary freely. The center positions of the functions were allowed to vary independently. The widths of the functions were constrained to be between 8 and 13 Å. The functions centered at 59 Å, having the lowest amplitude, have the greatest uncertainty, and their position was fixed while the other parameters were optimized. The centers of the Gaussian functions are taken as the measured molecular dimensions, while the area under the curves is proportional to the probability with which a given molecular orientation is measured. Consistency in the fitting parameters is checked by multiple repetitions of the experiments on different days. The results presented here conform to this strict criteria to less than 5 Å.

The final analysis consists of two different procedures: (a) determination of molecular dimensions and (b) assignment of quaternary structure (when different oligomers are observed). To assign the observed dimensions to the macromolecule, a model that takes into account the orientations of molecules at

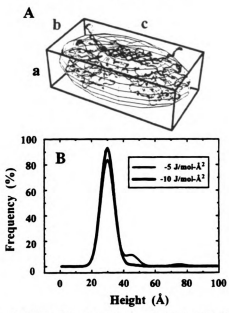


Figure 3. (A) Con A dimer (rendering based on X-ray structure) with circumscribed ellipsoid and parallelepiped of dimensions $a = 30$ Å, $b = 45$ Å, and $c = 75$ Å. (B) Calculated height distributions for parallelepipeds, of the same dimensions as Con A dimer, as predicted by the SAO model for values of ΔG_0 of -5 and -10 J/(mol Å^2). Notice that for higher protein substrate affinity proteins tend to orient so that their largest surface area lies on the mica, thereby projecting their shortest dimension perpendicular to the surface.

interfaces is required (vide infra). Protein dimensions obtained by this method are evaluated for consistency with other biophysical data available for the protein under investigation, such as molecular weight. For example, the volume of each oligomer is calculated from the assumed dimensions. Using the known molecular weight of the protein being studied, a volume is then calculated (assuming a density of ~1 g/cm^3) and compared to the volume predicted by the assumed dimensions.

Theory

In the absence of chemically specific interactions of the protein molecules with the surface, the ensemble of molecules is expected to assume a thermodynamically driven distribution of orientations. A surface adsorption and orientation (SAO) model is proposed, which treats this interfacial protein orientation as an optimization of free energy.

The model assumes each protein can be enclosed by a parallelepiped of dimensions a, b, and c, where $a \le b \le c$ (Figure 3 A). This geometrical shape is compatible with ellipsoidal models in the limit of well-rounded edges and corners. Furthermore, the measured height is independent of the degree to which the edges are rounded. A favorable interaction between the protein molecules and the hydrophilic mica surface results in orientations that maximize the surface area of contact. Therefore, a protein orientation having the bc area adsorbed on the mica (where b and c are the largest dimensions, by definition) is considered to be the most energetically favorable, yielding a maximum probability for measured heights corresponding to dimension a (see Figure 3B).

The orientation free energy after the initial adsorption step is assumed to be dependent on the protein surface in contact with the substrate. In the absence of particular surface recognition motifs, the free energy of orientation for each surface is given by

$$\Delta G^{bc}_{orient} \cong \Delta G_{bc} A_{bc} \tag{1}$$

$$\Delta G^{ac}_{orient} \cong \Delta G_{ac} A_{ac} \tag{2}$$

$$\Delta G^{ab}_{orient} \cong \Delta G_{ab} A_{ab} \tag{3}$$

where ΔG_{bc} is the free energy of orientation per unit area of the protein in contact with the substrate (in this case bc), and A is the surface area of the side identified by the subscript.

The probability of orientation such that height a, b, or c is measured depends on the different free energies of orientation, eqs 1−3. These may be estimated by a Boltzmann distribution using

$$P(a) \cong \frac{2\exp(-\Delta G^{bc}_{orient}/kT)}{Q} \tag{4}$$

$$P(b) \cong \frac{2\exp(-\Delta G^{ac}_{orient}/kT)}{Q} \tag{5}$$

$$P(c) \cong \frac{2\exp(-\Delta G^{ab}_{orient}/kT)}{Q} \tag{6}$$

where the factor of 2 arises from the symmetry of the parallelepiped, k is Boltzmann's constant, T is the temperature, and Q is the canonical partition function. These expressions are simplified by taking the ratio between these probabilities and by assuming that all orientation energies per unit area are equal (as expected in the absence of specific recognition sites in the surface or the protein) so that $\Delta G_0 \equiv \Delta G_{bc} = \Delta G_{ac} = \Delta G_{ab}$; we thus obtain

$$P(a)/P(b) = \exp\left(\frac{[A_{ac} - A_{bc}]\Delta G_0}{kT}\right) \tag{7}$$

$$P(b)/P(c) = \exp\left(\frac{[A_{ab} - A_{ac}]\Delta G_0}{kT}\right) \tag{8}$$

$$P(a)/P(c) = \exp\left(\frac{[A_{ab} - A_{bc}]\Delta G_0}{kT}\right) \tag{9}$$

where the free energies of orientation have been substituted by their corresponding expression (eqs 1, 2, and 3). Given these equations, it is possible to predict the probabilities of each protein orientation on the surface given its dimensions. It is also possible to start with the experimentally observed dimensions (a, b and c) and their probabilities (P_a, P_b, and P_c) and obtain a value for ΔG_0 using eqs 7−9. The magnitude of ΔG_0 determines the relative probabilities of measuring each of the three main dimensions of a protein. Figure 3 illustrates the theoretical height distribution for a parallelepiped, of the same dimensions as Con A dimer, for values of ΔG_0 differing by only a factor of 2.

The SAO model, therefore, provides a mathematical framework to analyze the height distribution profiles and obtain the relative amounts of each oligomeric state, e.g., monomer, dimer, tetramer (see the data analysis section). In addition to the obvious height differences for the oligomers, we observe that the probability of measuring those values varies extensively depending on the configuration. Therefore, for proteins of

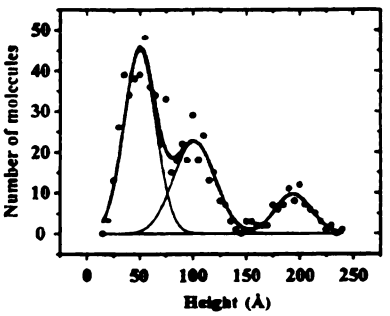


Figure 4. AFM height distribution data for colloidal gold particles adsorbed on mica. Results of the least-squares Gaussian fits to the data may be found in Table 1.

TABLE 1: AFM Height Measurements of Colloidal Gold Particles Compared to the Values Obtained Using TEM

size determined by TEM (nm)[a]	size determined by AFM (nm)	particles measured
5.0 ± 0.8	5.0 ± 0.8	353
9.0 ± 1.2	10.1 ± 1.6	207
18.0 ± 0.8	19.4 ± 2.0	87

[a] TEM analysis supplied by the manufacturer.

unknown structure, this model is useful for the determination of tertiary and quaternary structure.

Most proteins are asymmetrical or oligomeric; therefore, for a given protein, different dimensions are obtained corresponding to different protein orientations (and/or oligomers) on the substrate. Given a large enough sampling, on the order of 10^3 individual measurements, a distribution of measured heights is obtained that is characteristic of the molecular dimensions of the protein. To interpret the height distribution and decompose it into the dimensions of the protein, we use the SAO model to predict how molecules of specific dimensions adsorb to the substrate.

Results

Height Measurements of Colloidal Gold with AFM. Monodispersed spherical colloidal gold particles were initially employed to demonstrate the subnanometer height measurements of AFM that are minimally influenced by probe geometry.[46] The particles were chosen because their sizes were comparable to that of protein molecules. Twenty five microliters of sample ($10^{11}-10^{13}$ particles/mL) was deposited onto a freshly cleaved, magnesium ion-treated ~1 cm^2 piece of mica and rinsed twice with 200 mL of MilliQ water, after a 20 min incubation time. This preparation resulted in a uniform distribution of gold particles on the mica surface. The results of height measurements on these particles are presented in Figure 4 and Table 1. The three dominant peaks in the histogram are in good agreement with the TEM analysis performed by the manufacturer. For the 18 nm gold we found a greater deviation presumably because some particles had slightly different shapes (nonspherical). These types of variations are more common for larger particles. In general, the similarity between the standard deviations calculated for both TEM and SOMS measurements (Table 1) suggests that the dominant source of error in these measurements may reside in the size distribution of the particles themselves rather than the measurement technique or AFM probe geometry as discussed in Materials and Methods.

Analysis of Ferritin at the Liquid/Solid Interface. To test this approach for the determination of the macromolecular

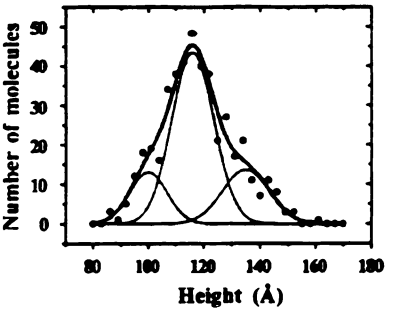


Figure 5. AFM height distribution data for ferritin. The data are best fit by three Gaussian functions centered at 97 Å, 115 Å (the intact protein), 133 Å (ferritin trimers), and 144 Å. The two assigned peaks account for 85% of the data.

dimensions of proteins, we performed measurements on ferritin for which high-resolution structures had been previously obtained by both electron microscopy[47] and X-ray crystallography.[48] Ferritin (horse spleen) has been demonstrated to be a spherical molecule comprised of 24 polypeptides (ratios of homologous H and L chains) and bound iron (Fe^{3+}).[47] Ferritin has a molecular weight of 445 kDa, and it has been shown to have a diameter of 120 Å by electron microscopy[47] and X-ray crystallography.[48] AFM measurements of samples prepared from a 20 μg/mL solution of ferritin yielded a distribution of measured heights with a dominant peak at 115 Å (Figure 5), in excellent agreement with the previously reported results for ferritin. The additional peaks in the height distribution indicate the presence of ferritin trimers (133 Å) and proteins of smaller oligomeric composition. (Such ferritin structures have been previously observed with adsorbed ferritin samples on polished Ti and Mo using STM.[49]) Our data show the presence of three main molecular species: intact ferritin (74%), trimers (11%), and some smaller species at 97 Å (9%). All other ferritin species account for ~6% of the observed values.

Analysis of Concanavalin A (Con A) at the Liquid/Solid Interface. Concanavalin A (Con A) is a lectin, obtained from the jackbean (*Canavalia ensiformis*), that binds to mannose, glucose, and glycoconjugates containing these saccharides.[50-52] The native protein is composed of four identical subunits, each with molecular weight 25 500 kDa.[52] Con A in solution is normally observed as a tetramer at pH 7.0 or above.[52] However, it was shown by Gordon and Young[53] that Con A dissociates into dimers as a consequence of dilution and ionic strength at a concentration ≤5 μg/mL and a pH ≥7.0. According to X-ray diffraction studies, the molecular dimensions of the dimer are approximately 30 Å × 45 Å × 75 Å, while those of the tetramer are approximately 60 Å × 70 Å × 70 Å (Protein DataBank, 1CN1[54]) (Figure 7).

Preparations of Con A solutions (~400 ng/mL), as described in the Experimental Section, yielded monodispersed individual protein molecules adsorbed on the mica surface (see Figure 2, for example). Analysis of several such images results in the height distribution shown in Figure 6. This distribution is the result of 2000 individual protein molecule measurements. The height distribution is best fit by four Gaussian functions centered at 18, 28, 42, and 59 Å (see Table 2 A). Inspection of the X-ray structure of the Con A dimer (see Figure 7) indicates that the observed height distribution corresponds to individual Con A dimers oriented such that their largest surface area is in contact with the mica substrate, projecting the 30 Å dimension perpendicular to the surface. This orientation is expected and can be understood on the basis of SAO model. The 28 and 42

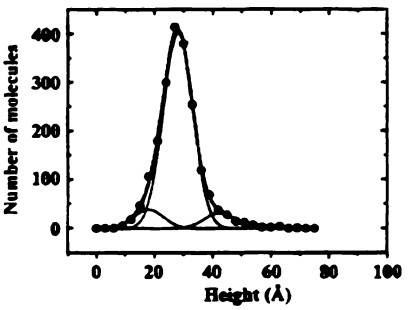


Figure 6. AFM height distribution data (2000 height measurements) for Con A The data are best fit by four Gaussian functions centered at 18, 28, 42, and 59 Å. The results obtained from this fit may be found in Table 2A.

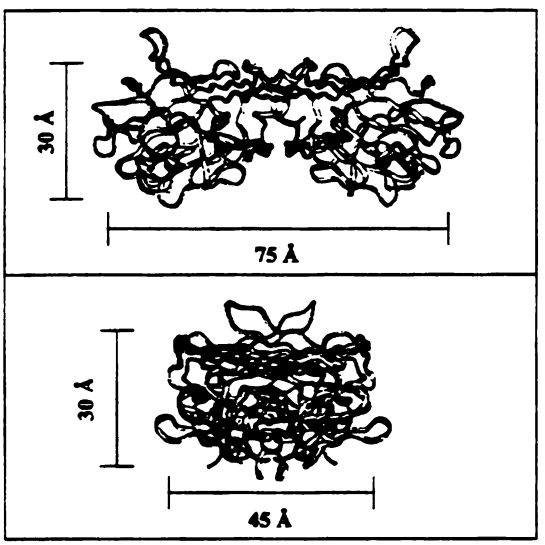


Figure 7. Ribbon rendering of the Con A dimer (Protein DataBank, 1CNI[54]). (A) Side view showing the short axis, 30 Å (which does not include the hairpin turns), and the longest axis, 75 Å for end to end. (B) Side view showing the intermediate axis measuring approximately 45 Å (neglecting the two hairpin turns at the extremes).

Å dimensions are in excellent agreement with the X-ray structure (see Figure 7). The 59 Å measurement is shorter than the 75 Å expected, assuming that individual Con A dimers project their longest axis away from the mica (0° from normal inclination). This 0° orientation is not favorable on the basis of protein structure shown in Figure 7; therefore, one would expect some angle of inclination. Our measurements indicate an inclination of about 38° from normal. We will show that the presence of the 18 Å population (~8% of the total population) can be attributed to denatured Con A in the following experiments.

To rule out the possibility that Con A was denaturing at the mica surface, two experiments were carried out. The first one involves imaging of succinylated Con A which is known to exist only as a dimer, over a wide range of pH.[55] The second experiment was to purposely denature the Con A and determine the height distribution with our technique.

Results from the succinylated Con A (1576 individual height measurements) show a very similar height distribution to that obtained from Con A (see Figure 8 and Table 2B). The three main peaks found are at 29, 45, and 59 Å, corresponding closely (within 3 Å) to those found for Con A. These results indicate

TABLE 2: Results from Height Distribution Analysis and Least-Squares Fit to Gaussian Functions[a]

	height (Å)	fwhm (Å)[b]	area[c]	probability (%)[d]
		(A) Concanavalin A		
*	18	10	489	
a	28	10	5079	91.8
b	42	10	409	7.4
c	59	10	45	0.8
		(B) Succinylated Con A		
*	19	9	1056	
a	29	13	1597	87.7
b	45	8	164	9.0
c	59	13	59	3.2
		(C) SDS-Treated Con A		
*	18	10	3679	
a	29	13	1938	81.2
b	46	12	350	14.7
c	59	12	100	4.2

[a] The values in these tables are the result of the height distribution analysis of measurements made on 2000, 1576, and 1983 individual proteins for A, B, and C, respectively. [b] These values represent the full width at half-maximum (fwhm) for each of the Gaussian functions in the fit. [c] These measurements correspond to the area under the Gaussian function and are proportional to the total number of molecules within a range of heights in Å. [d] The probability excludes the particles found to have a dimension of 18−19 Å labeled by *, which are determined to arise from the presence of denatured protein molecules. The values of ΔG_0 are the average of three measurements, using eqs 7−9 (see text). For Con A (A) and succinylated Con A (B) we obtain -9.4 ± 2.0 J/(mol Å^2) and -6.2 ± 0.2 J/(mol Å^2), respectively.

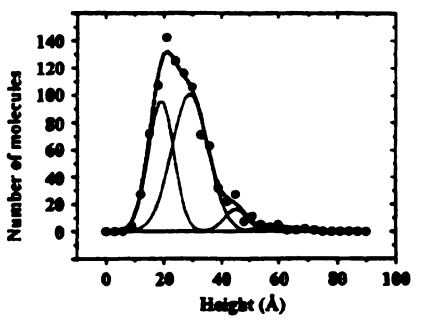


Figure 8. AFM height distribution data (1576 height measurements) for succinyl Con A. The data are best fit by four Gaussian functions centered at 19, 29, 45, and 59 Å. The results obtained from the least-squares fit may be found in Table 2B.

that Con A exists as a dimer and not a tetramer at the water/mica interface under these conditions. This result confirms the observations of Gordon and Young[53] and provides support for the capability and sensitivity of this approach for measuring changes in the oligomerization state of proteins. The only difference is an additional population of molecules with a height of 19 Å. An explanation of this population with a smaller dimension is given below.

To investigate the possibility of denaturation of Con A at the water/mica interface, we treated the protein with 2% sodium dodecyl sulfate (SDS) for 2−3 min at 90−95 °C prior to dilution and preparation of samples. The height distribution from the SDS-treated Con A samples, from 1983 individual height measurements, is shown in Figure 9 (see also Table 2C). Note that the major population is centered at 18 Å. The results indicate that ~61% of the protein molecules were denatured. The rest of the distribution shows peaks at 29, 46, and 59 Å, in excellent agreement (<5 Å) with the previous two sets of data. This indicates that some of the Con A molecules were not

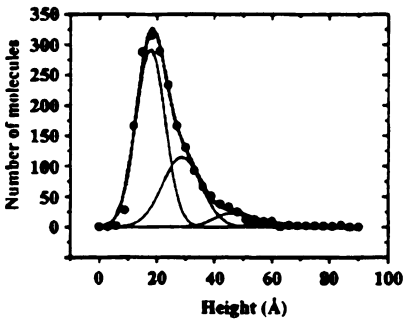


Figure 9. AFM height distribution data (1983 height measurements) for SDS-treated Con A. The data are best fit by four Gaussian functions centered at 18, 29, 46, and 59 Å. The results obtained from this least-squares fit may be found in Table 2C.

denatured or underwent refolding upon dilution of the SDS during sample preparation. This result allows us to assign the peak at 19 Å in the succinylated Con A samples to the presence of ~37% denatured protein in the sample.

Further analysis of the results by the SAO model can be used to determine the free energy of orientation per unit area, ΔG_0, for Con A and succinylated Con A on Mg^{2+}-treated mica. Using eqs 7−9 and values (height and probability) from Table 2, we calculate average values of -9.4 ± 2.0 and -6.2 ± 0.2 kJ/(mol Å^2), respectively. The uncertainty is the standard deviation of the three values calculated from eqs 7−9. Day-to-day variability of as much as 40% in the calculated value of ΔG_0 has been observed.[56] The approximate surface area of each face of the protein can be calculated from the measured dimensions. With these values we calculate the free energy of orientation for each of the three different orientations (see Table 2), the most stable of these being ~−20 kJ/mol when the Con A molecules have their largest surface oriented toward the mica. These values are larger than kT (2.5 kJ/mol), indicating this orientation is relatively stable. The free energy is small enough that distortions in the native structure are not expected. Differences in the free energy for each orientation determine the observed height distribution. The values for Con A and succinylated Con A are similar, as expected; however, we believe that there may be some small difference related to modification of amino acid side chains of Con A by the succinylation process. Succinylation involves the modification of amines in the protein with a concomitant decrease in positive charge. For Con A, an average of 10 sites per subunit are succinylated,[55] causing changes in its interaction with the mica surface.

In conclusion, we have obtained angstrom resolution measurements directly from individual proteins at the liquid/solid interface. Our measurements, 28, 42, and 59 Å, indicate that Con A adsorbs onto mica substrates as a dimer, primarily having its smallest dimension projected normal to the surface. Our measurements are in very good agreement with the X-ray structure determined by Shoham et al.[54] Although the tetrameric form of Con A is normally observed in solution, the solution from which we prepared our samples was ~400 ng/mL at a pH = 7.3, which according to Gordon and Young is expected to contain primarily the dimer species.[53] This dimeric state is conserved upon adsorption of the Con A molecules to the mica surface. The dimeric structure was confirmed with measurements on succinylated Con A. Evidence to the orientation of the Con A dimers on the mica comes from the prevalence of 28 Å measurements in the data and analysis of the different heights (42 and 59 Å) using the SAO model. The possibility of protein denaturation caused by the adsorption process was addressed by measurements on Con A treated by the denaturing agent SDS, which indicated that the degree of denaturation upon adsorption is minimal (<5 Å changes). Applying the SAO model to our measurements results in approximate values for the free energy of orientation for these molecules. We are currently investigating the effects of large changes in the protein/surface interaction on these values.

Discussion

Measurements of individual Concanavalin A and succinyl Con A molecules at the water/mica interface have determined that the Con A dimer is the prevalent structure. The measurements obtained (29 × 45 × 59 Å) for these dimers are in good agreement with the dimensions determined by X-ray diffraction studies. Further, the results for Con A, succinyl Con A, and SDS-treated ConA are self-consistent and reproducible. In addition to the structural information obtained, the proposed model also allows us to predict a value for the free energy of adsorption for Con A at the water/mica interface and the orientation of the adsorbed proteins.

We conservatively estimate that the accuracy of this technique is 5 Å or less. This estimate is related to a number of factors. First of all, there is 3−5 Å roughness in the mica substrates which we use. Second, although the ultimate resolution of the AFM in the z dimension is 1 Å or less, we are not optimizing our imaging conditions for single-point measurements of height; rather, we are trading accuracy for acquisition time. For the purposes of three-dimensional protein oligomerization state and orientation, the accuracy (<5 Å) is sufficient.

The technique presented here provides a means to measure the molecular dimensions for the complete size range of monomeric and oligomeric proteins adsorbed to a surface (natural or synthetic). These measurements may then be utilized to determine the oligomerization or aggregation state of the protein. This approach extends the nanometer resolution of AFM, in the contact or intermittent contact modes under ambient or liquid conditions, to that of angstroms by avoiding excessive image broadening of biological samples that is associated with the lateral distortions introduced by the probe tip geometry.[36−39] Probe−protein interactions, mainly lateral and vertical pressure, can cause the molecule to appear shorter.[26] With regard to protein−substrate interaction, there is an attractive interaction between the polar surface of the protein molecule and the charged mica surface, possibly deforming the molecule and causing it to flatten against the mica surface. A study using lysozyme has shown that it retains its activity when adsorbed onto mica and imaged with AFM, so these deformations may not be extensive enough to alter biological activity.[27] Our measurements confirm the conservation of native structure when compared with available X-ray data.

A thermodynamic model was derived to understand the orientation of proteins at the liquid/solid interface. Large deviations from this SAO model by adsorbed proteins would imply a preferred orientation that is presumably related to recognition between a protein domain and the surface. The extent of deviation from the orientation model may therefore be useful in defining preferred molecular orientation for oligomers or multiple equilibria for dissociated subunits of oligomers. In these cases one is able to extend the simple model presented here, by allowing each face of the protein to have a different value for ΔG_0 (eqs 1−3). In this way a quantitative value may be assigned to the energy difference associated with

each orientation. Indeed, binding studies with streptavidin and free biotin provide strong evidence for a ligand-mediated preferential orientation for the adsorption of streptavidin dimers and tetramers to the mica surface (work in progress).

An examination of the preferred orientation of the molecules also provides the opportunity to predict the organization of individual subunits within the protein oligomer and potentially evaluate the conformational flexibility of dissociated individual subunits. Measurements with angstrom resolution may be performed on functional complexes that are associated with a surface, a more common biological interaction within and between cells. This paper presents an approach that extends the analytical power and the resolution of AFM to provide direct physical characterization (both dimensions and orientation) of biomolecules at interfaces and is minimally dependent on the geometry of the probe tip. This approach provides (1) the capability to examine the concentration- and ligand-dependent properties of protein oligomerization using picomole amounts of protein, (2) offers the capability to distinguish between the possible quaternary forms of oligomers and their interconversions resulting from the binding of ligands, substrates, or allosteric effectors, and (3) determines whether adsorption of oligomeric proteins to surfaces (natural or synthetic) can modify protein assembly and aggregation state. Such measurements can be most useful in studies of enzyme regulation, DNA/RNA organization, gene regulation, protein biosynthesis and chaperonin function, bacterial attachment and adhesion, and the effects of posttranslational modification on protein organization and activity.

Acknowledgment. This research was partially funded by the REF-Protein Structure, Function and Design Award (MSU), the Center for Fundamental Materials Research (MSU), and a Camille and Henry Dreyfus New Faculty Award (M.D.). We are very grateful to Dr. Neil Bowlby for programming the data analysis software and Drs. Leslie Kuhn, Michael Garavito, John Wang, and Jack Preiss of the Department of Biochemistry for their incisive critiques. M.D. is a Beckman Young Investigator and a Packard Science and Engineering Fellow.

References and Notes

(1) Horbett, T. A.; Brash, J. L. In *Proteins at Interfaces, Physicochemical and Biochemical Studies*, Brash, J. L., Horbett, T. A., Eds.; ACS Symposium Series 343; American Chemical Society: Washington, DC, 1987; p 1.

(2) Norde, W. *Adv. Colloid Interface Sci.* **1986**, *25*, 267.

(3) Brash, J. L.; Horbett, T. A. In *Proteins at Interfaces II, Fundamentals and Applications*; Horbett, T. A., Brash, J. L., Eds.; ACS Symposium Series 602; American Chemical Society: Washington, DC, 1995; p 1.

(4) Elwing, H.; Askenda, A.; Ivarsson, B.; Nilsson, U.; Welin, S.; Lundström, I. In *Proteins at Interfaces, Physicochemical and Biochemical Studies*; Brash, J. L., Horbett, T. A., Eds.; ACS Symposium Series 343; American Chemical Society: Washington, DC, 1987; p 468.

(5) Nadarajah, A.; Lu, C. F.; Chittur, K. K. In *Proteins at Interfaces II, Fundamentals and Applications*; Horbett, T. A., Brash, J. L., Eds.; ACS Symposium Series 602; American Chemical Society: Washington, DC, 1995; p 181.

(6) van Oss, C. J.; Wu, W.; Giese, R. F. In *Proteins at Interfaces II, Fundamentals and Applications*; Horbett, T. A.; Brash, J. L., Eds.; ACS Symposium Series 602, American Chemical Society: Washington, DC, 1995; p 80.

(7) Yan, G.; Li, J.-T.; Huang, S.-C.; Caldwell, K. D. In *Proteins at Interfaces II. Fundamentals and Applications*; Horbett, T. A.; Brash, J. L., Eds.; ACS Symposium Series 602, American Chemical Society: Washington, DC, 1995; p 256.

(8) Lee, J. E.; Saavedra, S. S. In *Proteins at Interfaces II, Fundamentals and Applications*; Horbett, T. A.; Brash, J. L., Eds.; ACS Symposium Series 602; American Chemical Society: Washington, DC, 1995; p 269.

(9) Binnig, G.; Quate, C. F.; Gerber, C. *Phys. Rev. Lett.* **1986**, *56*, 930.

(10) Lal, R.; John, S. A. *Am. J. Physiol.* **1994**, *266*, C1.

(11) Butt, H.-J.; Downing, K. H.; Hansma, P. K. *Biophys. J.* **1990**, *58*, 1473.

(12) Bustamante, C.; Vesenka, J.; Tang, C. L.; Rees, W.; Guthold, M.; Keller, R. *Biochemistry* **1992**, *31*, 22.

(13) Bustamante, C.; Rivetti, C. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 395.

(14) Louder, D. R.; Parkinson, B. A. *Anal. Chem.* **1994**, *66*, 84R.

(15) Schnyder, T.; Engel, A.; Lustig, A.; Wallimann, T. *J. Biol. Chem.* **1988**, *263*, 16954.

(16) Dykstra, M. J. *A Manual of Applied Techniques for Biological Electron Microscopy*; Plenum: New York, 1993.

(17) Weisenhorn, A. L.; Drake, B.; Prater, C. B.; Gould, S. A. C.; Hansma, P. K.; Ohnesorge, F.; Egger, M.; Heyn, S. P.; Gaub, H. E. *Biophys. J.* **1990**, *58*, 1251.

(18) Schabert, F.; Hefti, A.; Goldie, K.; Stemmer, A.; Engel, A.; Meyer, E.; Overney, R.; Güntherodt, H.-J. *Ultramicroscopy* **1992**, *42B*, 1118.

(19) Schabert, F. A.; Henn, C.; Engel, A. *Science* **1995**, *268*, 92.

(20) Ohnishi, S.; Hara, M.; Furuno, T.; Sasabe, H. *Biophys. J.* **1992**, *63*, 1425.

(21) Malkin, A. J.; Kuznetsov, Y. G.; McPherson, A. *J. Struct. Biol.* **1996**, *117*, 124.

(22) Roberts, C. J.; Williams, P. M.; Davies, J.; Dawkes, A. C.; Sefton, J.; Edward, J. C.; Haymes, A. G.; Bestwick, C.; Davies, M. C.; Tendler, S. J. B. *Langmuir* **1995**, *11*, 1822.

(23) Ill, C. R.; Keivens, V. M.; Hale, J. E.; Nakamura, K. K.; Jue, R. A.; Cheng, S.; Melcher, E. D.; Drake, B. *Biophys. J.* **1993**, *64*, 919.

(24) Edstrom, R. D.; Meinke, M. H.; Yang, X.; Yang, R.; Elings, V.; Evans, D. F. *Biophys. J.* **1990**, *58*, 1437.

(25) Marchant, R. E.; Lea, A. S.; Andrade, J. D.; Bockenstedt, P. *J. Colloid Interface Sci.* **1992**, *148*, 261.

(26) Radmacher, M.; Fritz, M.; Cleveland, J. P.; Walters, D. A.; Hansma, P. K. *Langmuir* **1994**, *10*, 3809.

(27) Radmacher, M.; Fritz, M.; Hansma, H. G.; Hansma, P. K. *Science* **1994**, *265*, 1577.

(28) Henderson, R. M.; Schneider, S.; Li, Q., L.; Hornsby, D.; White, S. J.; Oberleithner, H. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8756.

(29) Benzanilla, M.; Manne, S.; Laney, D. E.; Lyubchenko, Y. L.; Hansma, H. G. *Langmuir* **1995**, *11*, 655.

(30) Hansma, H. G.; Sinsheimer, R. L.; Li, M. Q.; Hansma, P. K. *Nucl. Acids Res.* **1992**, *20*, 3585.

(31) Hansma, H. G.; Vesenka, J.; Siegerist, C.; Kelderman, G.; Morret, H.; Sinsheimer, R. L.; Elings, V.; Bustamante, C.; Hansma, P. K. *Science* **1992**, *256*, 1180.

(32) Henderson, E. *Nucl. Acids Res.* **1992**, *20*, 445.

(33) Shaiu, W.-L.; Vesenka, J.; Jondle, D.; Henderson, E.; Larson, D. D. *J. Vac. Sci. Technol. A* **1993**, *11* (4), 820.

(34) Wyman, C.; Rombel, I.; North, A. K.; Bustamante, C.; Kustu, S. *Science* **1997**, *275*, 1658.

(35) Siedlecki, C. A.; Lestini, B. J.; Kottke-Marchant, K.; Eppell, S. J.; Wilson, D. L.; Marchant, R. E. *Blood* **1996**, *88* (8), 2939.

(36) Keller, D. J.; Franke, F. S. *Surg. Sci.* **1993**, *294*, 409.

(37) Eppell, S. J.; Zypman, F. R.; Marchant, R. E. *Langmuir* **1993**, *9*, 2281.

(38) Allen, M. J.; Hud, N. V.; Balooch, M.; Tench, R. J.; Seikhaus, W. J.; Balhorn, R. *Ultramicroscopy* **1992**, *42B*, 1095.

(39) Markiewicz, P.; Goh, M. C. *Langmuir* **1994**, *10*, 5.

(40) Yariv, J.; Kalb, A. J.; Levitzski, A. *Biochim. Biophys. Acta* **1968**, *165*, 303.

(41) Apell, H. J.; Colchero, J.; Linder, A.; Marti, O. In *STM and AFM in Biology*; Marti, O., Amrein, M., Eds.; Academic Press: San Diego, CA, 1993; p 282.

(42) Hu, J.; Xiao, X.-D.; Ogletree, D. F.; Salmeron, M. *Science* **1995**, *268*, 267.

(43) Thundat, T.; Zheng, X.-Y.; Chen, G. Y.; Warmack, R. J. *Surf. Sci. Lett.* **1993**, *294*, L939.

(44) Leggett, G. J.; Davies, M. C.; Jackson, D. E.; Roberts, C. J.; Tendler, S. J. B.; Williams, P. M. *J. Phys. Chem.* **1993**, *97*, 8852.

(45) Parker, M. C.; Davies, M. C.; Tendler, S. J. B. *J. Phys. Chem.* **1995**, *99*, 16155.

(46) Height measurements of small spherical particles with AFM are slightly affected, ~1%, because of the geometry of the probe and because individual scan lines may be separated by distances comparable to the size of particles being measured. This error occurs when the end of the probe, which can be modeled as a hemisphere, is not exactly at the apex of the protein. For example, in a 2 × 2 μm square image digitized to 512 × 512 pixels, the end of the probe could be as far as 20 Å from the apex of the molecule if it is located exactly between two scan lines. In Figure 1A, the idealized shapes of the end of the probe and a spherical protein molecule are illustrated. Because of the finite number of points that are recorded, the apex of the protein will usually not be exactly under the center of the probe; therefore, the measured height of the particle will be slightly smaller. The

discrepancy, d, can be calculated as follows:

$$d = (R + r) - \sqrt{(R + r)^2 - X^2} \quad (3)$$

where R is the radius of curvature of the probe, r is the radius of a spherical protein, and X is the horizontal distance between the probe and the protein when the probe is closest to the protein (see Figure 1A). Probe sizes for our measurements had approximately a 24 nm radius of curvature, as calculated from the width and height measurements of colloidal gold particles of known dimensions.[39] For particles with a diameter of 40 Å, the maximum height deviation can be as much as 0.7 Å. We have modeled this effect by calculating the average error expected given 500 randomly placed 40 Å diameter spheres. (Spherical geometry maximizes this resolution effect.) The average error obtained from our simulation was 0.2 Å, which amounts to 0.5% error in the height measurement of the particles. This error is smaller than the natural roughness of atomically flat mica; therefore, no attempt was made to minimize it or correct for it.

(47) Massover, W. H. *Micron* **1993**, *24*, 389.
(48) Harrison, P. M. *J. Mol. Biol.* **1963**, *6*, 402.
(49) Yau, S.-T.; Zhou, Y. *Mod. Phys. Lett. B* **1995**, *9*, 187.
(50) *Concanavalin A as a Tool*; Bittiger, H., Schnebli, H. P., Eds.; John Wiley & Sons: New York, 1976.
(51) Sharon, N.; Lis, H. *Science* **1989**, *246*, 227.
(52) Reeke, G. N., Jr.; Becker, J. W.; Cunningham, B. A.; Wang, J. L.; Yahara, I.; Edelman, G. M. In *Concanavalin A*; Chowdhury, T. K., Weiss, A. K., Eds.; Plenum Publishing Corp.: New York, 1975.
(53) Gordon, J. A.; Young, R. K. *J. Biol. Chem.* **1979**, *254*, 1932.
(54) Shoham M.; Yonath, A.; Sussman, J. L.; Moult, J.; Traub, W.; Kalb, A. J. *J. Mol. Biol.* **1979**, *131*, 137.
(55) Gunther, G. R.; Wang, J. L.; Yahara, I.; Cunningham, B. A.; Edelman, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **1973**, *70* (4), 1012.
(56) The origin of this variability is currently under investigation. Some of the variation may be attributed to the amount of background noise, perhaps due to differences in the mica substrate or changes in instrument performance. Capillary forces may also play a role in this variability. Further experiments, carried out under liquid conditions, will help us make this assessment.

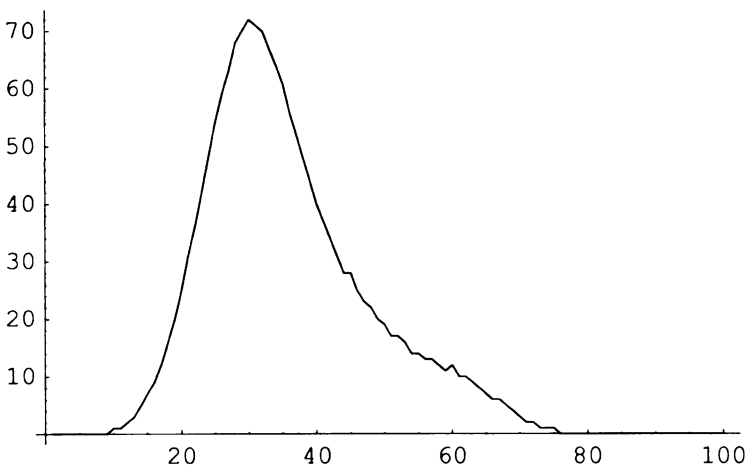
APPENDIX 2

Computer program source code

APPENDIX 2.1

Source code for Mathematica calculation of Chapter 2

```
(* Parameters *)
Nbins = 100;
SizeRange = 100; (*  in Angstroms *)
Convert = Nbins / SizeRange;

(* Set positions of the Gaussians *)
nOne = 30 Convert ;
nTwo = 45 Convert;
nThree = 60 Convert;

(* The width and amplitude of the three Gaussians *)
width = 10 Convert;
popOne = 70;
popTwo = 20;
popThree = 10;

(* The data simulated is the sum of three Gaussians *)
Table [H[n] = IntegerPart[popOne * Exp[-((n - nOne) / width) ^ 2]] +
   IntegerPart[popTwo * Exp[-((n - nTwo) / width) ^ 2]] +
   IntegerPart[popThree * Exp[-((n - nThree) / width) ^ 2]], {n, 1, Nbins}];

ListPlot[
 Table[H[n], {n, 1, Nbins}], PlotRange -> All, PlotJoined -> True]
```



- Graphics -

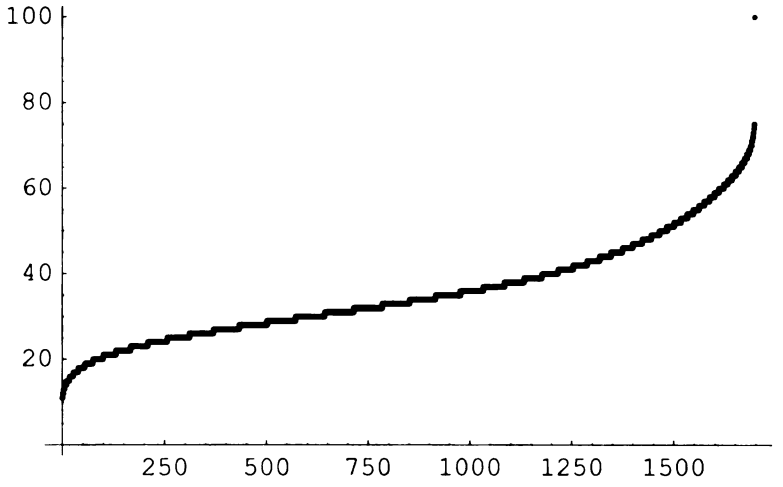
$$\text{Area} = \sum_{n=1}^{\text{Nbins}} H[n]$$

1698

```
Table[U[i] = 0, {i, 1, Area}];
For[test = 0; n = 1; i = 1, n <= Nbins,
 n++, Table[U[i] = n, {i, test, test + H[n]}]; test += H[n]]
test
```

1698

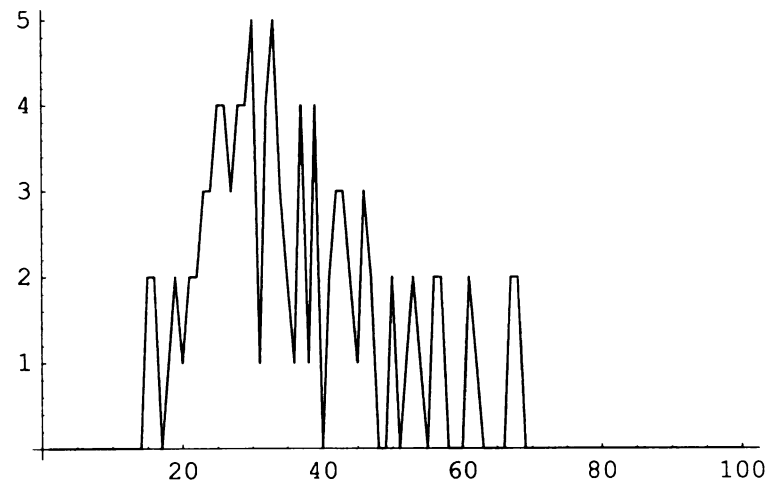
```
(* This is the distribution of all the values *)
Distrib = Table[U[i], {i, 1, Area}];
ListPlot[Distrib]
```



- Graphics -

```
Nsamples = 100;

Table[Discrete[n] = 0, {n, 1, Nbins}];
temp := Random[Integer, Area - 1] + 1

measurement = Table[Distrib[[temp]], {i, 1, Nsamples}];

For[suma = 0; j = 1, j <= Nsamples,
 j++, Discrete[measurement[[j]]] += 1; suma += 1]
suma
For[Hm = 0; j = 1, j <= Nbins, j++, If[H[j] > Hm, Hm = H[j], Hm = Hm]]
Hm
For[Dm = 0; j = 1, j <= Nbins,
 j++, If[Discrete[j] > Dm, Dm = Discrete[j], Dm = Dm]]
Dm
For[j = 1, j <= Nbins, j++,
 Dnorm[j] = (Discrete[j] / Dm); Hnorm[j] = (H[j] / Hm)]
For[sumsq = 0; j = 1, j <= Nbins, j++, sumsq += (Dnorm[j] - Hnorm[j]) ^ 2]
N[sumsq]

ListPlot[
 Table[Discrete[n], {n, 1, Nbins}], PlotRange -> All, PlotJoined -> True]
```

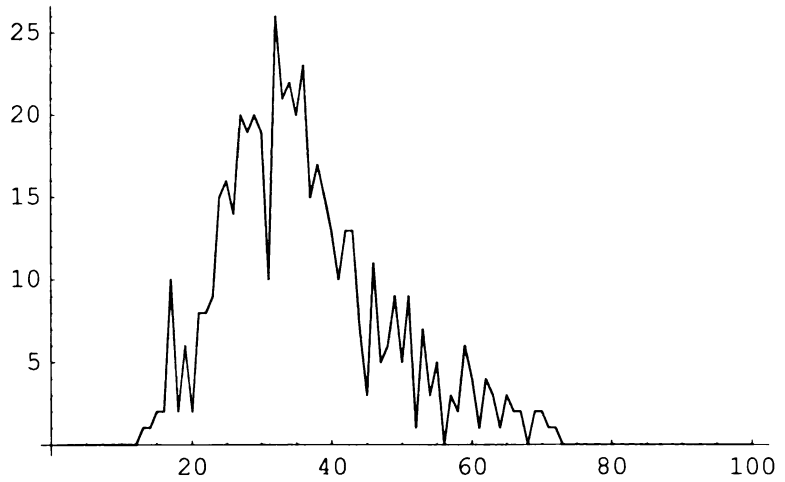
```
100

72

5

3.2102
```




```
- Graphics -
```

```
Nsamples = 500;

Table[Discrete[n] = 0, {n, 1, Nbins}];
temp := Random[Integer, Area - 1] + 1

measurement = Table[Distrib[[temp]], {i, 1, Nsamples}];

For[suma = 0; j = 1, j <= Nsamples,
 j++, Discrete[measurement[[j]]] += 1; suma += 1]
suma
For[Hm = 0; j = 1, j <= Nbins, j++, If[H[j] > Hm, Hm = H[j], Hm = Hm]]
Hm
For[Dm = 0; j = 1, j <= Nbins,
 j++, If[Discrete[j] > Dm, Dm = Discrete[j], Dm = Dm]]
Dm
For[j = 1, j <= Nbins, j++,
 Dnorm[j] = (Discrete[j] / Dm); Hnorm[j] = (H[j] / Hm)]
For[sumsq = 0; j = 1, j <= Nbins, j++, sumsq += (Dnorm[j] - Hnorm[j]) ^ 2]
N[sumsq]

ListPlot[
 Table[Discrete[n], {n, 1, Nbins}], PlotRange -> All, PlotJoined -> True]
```

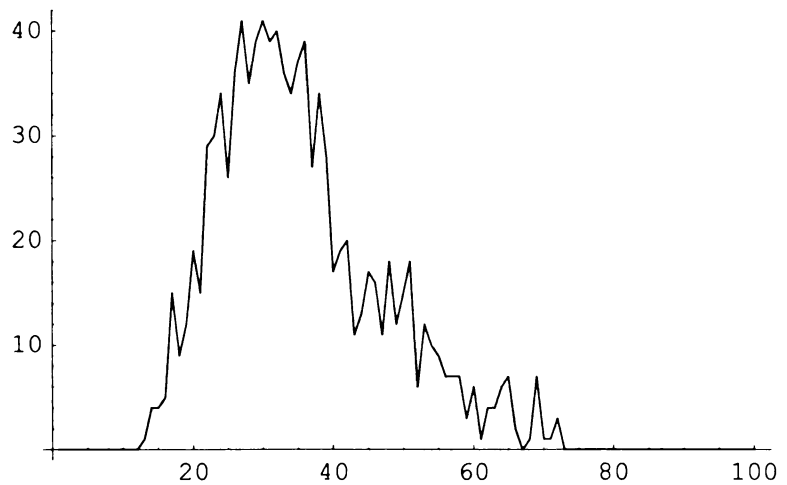
```
500

72

26

1.2382
```




```
- Graphics -
```

```
Nsamples = 1000;

Table[Discrete[n] = 0, {n, 1, Nbins}];
temp := Random[Integer, Area - 1] + 1

measurement = Table[Distrib[[temp]], {i, 1, Nsamples}];

For[suma = 0; j = 1, j <= Nsamples,
 j++, Discrete[measurement[[j]]] += 1; suma += 1]
suma
For[Hm = 0; j = 1, j <= Nbins, j++, If[H[j] > Hm, Hm = H[j], Hm = Hm]]
Hm
For[Dm = 0; j = 1, j <= Nbins,
 j++, If[Discrete[j] > Dm, Dm = Discrete[j], Dm = Dm]]
Dm
For[j = 1, j <= Nbins, j++,
 Dnorm[j] = (Discrete[j] / Dm); Hnorm[j] = (H[j] / Hm)]
For[sumsq = 0; j = 1, j <= Nbins, j++, sumsq += (Dnorm[j] - Hnorm[j]) ^ 2]
N[sumsq]

ListPlot[
 Table[Discrete[n], {n, 1, Nbins}], PlotRange -> All, PlotJoined -> True]
```

```
1000

72

41

0.4693
```




```
- Graphics -
```

APPENDIX 2.2

Source code for Image Analysis software in Visual Basic 5.0

HDF analysis program source code for Visual Basic 5.0

Main.frm Code

```
'******** HDF Analysis Visual Basic 5.0 program ********
'**************** Version 2.1 6/12/98 ****************
'*** Written by George Schoendorff and Mark J. Waner ***

'Analyzes Park Scientific Instruments AFM images to find heights of particles.

'*** Uses ScanData (supplied by Park Scientific Instruments)to open and display image.
'*** Based on hdfanal5i QuickBasic 4.0 program written by Dr. Neil Bowlby.
'*** Original Program and algorithms were formulated by:
'***  Dr. Marcos Dantus, Mark J. Waner and Dr. Neil Bowlby.

'Synopsis of program:
'  1) File is read and conversion of raw data (Integers) are converted into anstroms.
'  2) The height data is then put into a 2 angstrom bin histogram.
'  3) The Analysis routine (click button) will then find all local maxima (within
'     given radius) that are above the threshold value, but below the cutoff.
'  4) The x,y position and the height of each particle can then be saved to a text
'     file or you can go into an interactive peak selection/rejection mode (analysis.frm)
'  5) The height of a particle is determined by making a histogram of all pixel heights
'     within 40 pixels (in Y) and 1 pixel (in X) and using the most frequent height
'     as the local background

Dim Zgain, MaxPixel, MinPixel As Single
Dim DataArray(1 To 512, 1 To 512) As Single
Dim xpeaks(1 To 5000), ypeaks(1 To 5000), zpeaks(1 To 5000) As Integer
Dim peakno, radius, threshold, cutoff, MostFreq, p, p1, acceptedpeaks, size As Integer
Dim Zspan As Long
Dim fileName, defaultfilename As String  'Strings to store open and save file names
Dim xpeaks1(1 To 5000), ypeaks1(1 To 5000), zpeaks1(1 To 5000) As Integer

Public Sub About_Click()
frmAbout.Show
End Sub

Public Sub buttonGrayScale_Click()
ScanData.GrayScaleSetting
End Sub

Public Sub textMax_Change()
textCutoff.Text = textMax.Text
End Sub
```

```
Public Sub buttonMinAnal_Click()

'Do analysis to look for minima
textError.Text = ""

Dim MinExclude As Integer
MinExclude = textMinExclude.Text

'Clear the arrays to prepare for new analysis
Erase xpeaks
Erase ypeaks
Erase zpeaks
peakno = 0

'change cursor to hourglass
Form1.MousePointer = 11

'Locate the local minima
radius = textRadius.Text
    For x = radius + 1 To size - radius
      For y = radius + 1 To size - radius

        For rx = -radius To radius
        For ry = -radius To radius
          If rx = 0 Then If ry = 0 Then GoTo 16
          If DataArray(x + rx, y + ry) < DataArray(x, y) Then GoTo 20
          'jump out if there's a point (MinExclude) Angstroms higher within the radius
          If DataArray(x + rx, y + ry) > (DataArray(x, y) + MinExclude) Then GoTo 20
          If DataArray(x + rx, y + ry) = DataArray(x, y) Then If (rx <= 0) Then If (ry <=
0) Then GoTo 20
16        Next ry
        Next rx
peakno = peakno + 1
xpeaks(peakno) = x: ypeaks(peakno) = y: zpeaks(peakno) = DataArray(x, y)
If peakno > 4998 Then textError.Text = "Too many valleys. Increase radius or Min.
Exclude": peakno = 0: textNumPeaks.Text = 0: GoTo 30

20  Next y
   Next x

If peakno = 0 Then textError.Text = "No valleys detected.  Decrease radius or Min.
Exclude": peakno = 0: textNumPeaks.Text = 0: GoTo 30

30  textPeakValley.Text = "# Valleys"
   textNumPeaks.Text = peakno
```

```
buttonPeakAnalysis.Enabled = True
buttonAnalysis.Default = False
buttonPeakAnalysis.Default = True

Form1.MousePointer = 0 'Default mouse pointer
End Sub

Public Sub buttonPeakAnalysis_Click()
Form1.MousePointer = 11 'or VB hourglass

'Write all data to file in order to pass info to analysis form
Open "variable" For Output As #2
Write #2, peakno, MaxPixel, MostFreq, cutoff, threshold, radius, size
Write #2, fileName, textPeakValley.Text, textMinExclude.Text, defaultfilename
For x = 1 To size
   For y = 1 To size
   Write #2, DataArray(x, y)
   Next y
Next x
For peaks = 1 To peakno
Write #2, xpeaks(peaks), ypeaks(peaks), zpeaks(peaks)
Next peaks
Close #2

buttonPeakAnalysis.Default = False
buttonSave.Default = True

analysis.Show
Form1.MousePointer = 0 'Default mouse pointer
End Sub

Public Sub buttonSave_Click()
Form1.MousePointer = 11 'or VB hourglass
Dim SaveFileName
'Write peak data to file
SaveFileName = textSaveFileName.Text
   Open SaveFileName For Output As #1
   Write #1, fileName
   Write #1, textPeakValley.Text; acceptedpeaks
   Write #1, "maximum="; MaxPixel, "most frequent="; MostFreq, "Min. exclude=";
textMinExclude.Text
   Write #1, "cutoff="; cutoff, "threshold="; threshold, "radius="; radius
   For i = 1 To acceptedpeaks
      Write #1, xpeaks1(i), ypeaks1(i), zpeaks1(i)
   Next i
```

```
Close #1
Form1.MousePointer = 0 'Default mouse pointer
End Sub

Public Sub buttonScales_Click()

'Turn on and off display scales
If ScanData.DisplayScales Then
buttonScales.Caption = "Scales On"
ScanData.DisplayScales = False
Else
buttonScales.Caption = "Scales Off"
ScanData.DisplayScales = True
End If

End Sub


Public Sub Exit_Click()
Kill "data"
End   'Exit application
End Sub

Public Sub buttonAnalysis_Click()

textError.Text = ""

'Clear the arrays to prepare for new analysis
Erase xpeaks
Erase ypeaks
Erase zpeaks
peakno = 0

'Get peak-finding parameters
radius = textRadius.Text
cutoff = textCutoff.Text
threshold = textThreshold.Text

'Change cursor to hourglass
Form1.MousePointer = 11

'Locate the local maxima
    For x = radius + 1 To size - radius
      For y = radius + 1 To size - radius
        If DataArray(x, y) < threshold Then GoTo 20
        If DataArray(x, y) > cutoff Then GoTo 20
```

```
      For rx = -radius To radius
        For ry = -radius To radius
         If rx = 0 Then If ry = 0 Then GoTo 16
         'jump out of the loop if there's anything taller within the radius
         If DataArray(x + rx, y + ry) > DataArray(x, y) Then GoTo 20
         If DataArray(x + rx, y + ry) = DataArray(x, y) Then If (rx <= 0) Then If (ry <= 0)
Then GoTo 20
16       Next ry
      Next rx
peakno = peakno + 1
xpeaks(peakno) = x: ypeaks(peakno) = y: zpeaks(peakno) = DataArray(x, y)
'Indicate overflow error
If peakno > 4998 Then textError.Text = "Too many peaks.  Increase threshold.": peakno
= 0: textNumPeaks.Text = 0: GoTo 30
20  Next y
   Next x
'Indicate lack of particles found
If peakno = 0 Then textError.Text = "No peaks detected.  Decrease threshold.": peakno =
0: textNumPeaks.Text = 0: GoTo 30
30  textPeakValley.Text = "# peaks"
   textNumPeaks.Text = peakno

'Calculate histogram for local background around each peak found
Dim lowx, highx, lowy, highy, maxpeakheight, maxpeakhistbar, maxbin, maxzcoord,
peakheight, bin, HistBar(1 To 600) As Integer
For p = 1 To peakno
lowx = xpeaks(p) - 1:  If lowx < 1 Then lowx = 1
highx = xpeaks(p) + 1:  If highx > size Then highx = size
lowy = ypeaks(p) - 40:  If lowy < 1 Then lowy = 1
highy = ypeaks(p) + 40:  If highy > size Then highy = size
maxpeakheight = 0
For x = lowx To highx
   maxpeakhistbar = 0:  maxbin = 0:  maxzcoord = 0:  peakheight = 0
   For bin = 1 To zpeaks(p):  HistBar(bin) = 0:  Next bin
   For y = lowy To highy
      If DataArray(x, y) > maxzcoord Then maxzcoord = DataArray(x, y)
      bin = DataArray(x, y)
      If bin > zpeaks(p) Then GoTo 4020
      If bin < 1 Then GoTo 4020
      HistBar(bin) = HistBar(bin) + 1
4020 Next y
   For bin = 1 To zpeaks(p)
   If HistBar(bin) > maxpeakhistbar Then maxpeakhistbar = HistBar(bin): maxbin = bin
   Next bin
'calculate height of particle relative to local background
   peakheight = maxzcoord - maxbin
```

```
   If peakheight > maxpeakheight Then maxpeakheight = peakheight
Next x

xpeaks1(p) = xpeaks(p)
ypeaks1(p) = ypeaks(p)
zpeaks1(p) = zpeaks(p) - maxbin
Next p
acceptedpeaks = peakno

buttonPeakAnalysis.Enabled = True
buttonAnalysis.Default = False
buttonPeakAnalysis.Default = True

Form1.MousePointer = 0 'Default mouse pointer
End Sub

Public Sub Open_Click()
'CommonDialog is used to show Open and Save Dialog Boxes
'Microsoft Common Dialog Control 5.0 has to be checked in Project/Components...
'and then installed on Form
Form1.MousePointer = 11 'or VB hourglass
textPeakValley.Text = "": textNumPeaks.Text = "": textComments.Text = ""
textCutoff.Text = "": textError.Text = "": textFilename.Text = ""
textMaxPixel.Text = "": textMinPixel = "": textMostFreq.Text = ""
textSaveFileName.Text = "": textScanSize.Text = "": textThreshold.Text = ""
peakno = 0: threshold = 0: cutoff = 0: MostFreq = 0: MaxPixel = 0: MinPixel = 0
fileName = ""
Erase xpeaks: Erase ypeaks: Erase zpeaks: Erase xpeaks1: Erase ypeaks1: Erase
zpeaks1
Erase DataArray
buttonPeakAnalysis.Enabled = False
buttonAnalysis.Enabled = False
buttonMinAnal.Enabled = False
buttonSave.Enabled = False
buttonAnalysis.Default = False
CommonDialog.ShowOpen
fileName = CommonDialog.fileName
ScanData.DisplayScales = False
ScanData.LoadHDF (fileName)  'Load HDF file with name from Open File Dialog box
textFilename.Text = fileName
textScanSize.Text = Str(ScanData.YScanSize) + "x" + Str(ScanData.XScanSize)
textComments.Text = ScanData.Comments
size = ScanData.GetNumCols

'Determine default filename for peak height data file
Dim tempstr(1 To 500) As String
```

```
Dim store, store2 As Integer

For i = 1 To Len(fileName)
   tempstr(i) = Mid(fileName, i, 1)
   If tempstr(i) = "\" Then store = i
   If tempstr(i) = "." Then store2 = i
Next i
textSaveFileName.Text = Mid(fileName, store + 1, store2 - store) + "txt"
defaultfilename = textSaveFileName.Text

'Here we will take image data which is loaded into Scan Data Control and put it
'into an array
Dim DataGain As Single
Dim TempArray(1 To 512, 1 To 512) As Single
Dim MinDataAt As Single

DataGain = ScanData.DataGain
MinDataAt = 32767
MaxPixel = -32767
MinPixel = 32767

Form1.MousePointer = 11 'or VB hourglass

'Put data from column counter1-1 and row counter2-1 in HDF file
'into array (DataAt array is from 0 to 511, not 1 to 512)
'Rows & columns are transposed in DataAt !!!!!!!
If optionY.Value = True Then
For counter1 = 1 To size
   For counter2 = 1 To size
   TempArray(counter2, counter1) = (CSng(ScanData.DataAt(counter1 - 1, counter2 - 1))
* DataGain)
   If TempArray(counter2, counter1) < MinDataAt Then MinDataAt =
TempArray(counter2, counter1)
   Next counter2
Next counter1
Else
For counter1 = 1 To 512
   For counter2 = 1 To 512
   'Rows & Columns may be transposed !!!!!!!!!!!!!!!
   TempArray(counter1, counter2) = (ScanData.DataAt(counter1 - 1, counter2 - 1) *
DataGain)
   If TempArray(counter1, counter2) < MinDataAt Then MinDataAt =
TempArray(counter1, counter2)
   Next counter2
Next counter1
End If
```

```
'calibrate pixel values in angstroms and put into array
For counter1 = 1 To size
   For counter2 = 1 To size
   DataArray(counter1, counter2) = (TempArray(counter1, counter2) - MinDataAt) *
10000
   If DataArray(counter1, counter2) > MaxPixel Then MaxPixel = DataArray(counter1,
counter2)
   If DataArray(counter1, counter2) < MinPixel Then MinPixel = DataArray(counter1,
counter2)
   Next counter2
Next counter1

'Calculate histogram for entire image
Dim Sizecategory, MaxSizeCategory, TotalHistBars As Integer
Dim binsize, BarLength As Integer
Dim MaxHistBar As Long
Dim HistBar(1 To 300) As Long
MaxHistBar = 0

'define the histogram
Zspan = MaxPixel - MinPixel
TotalHistBars = 100
binsize = Zspan / TotalHistBars

For x = 1 To size
For y = 1 To size
   Sizecategory = (1 + ((DataArray(x, y) - MinPixel) / Zspan) * (TotalHistBars - 1))
   HistBar(Sizecategory) = HistBar(Sizecategory) + 1
   If HistBar(Sizecategory) > MaxHistBar Then MaxHistBar = HistBar(Sizecategory):
MaxSizeCategory = Sizecategory
Next y
Next x
MostFreq = Int(MinPixel + (MaxSizeCategory * Zspan / TotalHistBars))

'Find histogram range
Dim LastHistBar, check As Integer
check = 0
LastHistBar = 100

For i = 100 To 1 Step -1
   BarLength = Int((HistBar(i) / MaxHistBar) * 90)
   If BarLength = 0 And check = 0 Then GoTo 50
   If BarLength > 1 Then GoTo 60
40  Next i
   GoTo 70
```

```
50  LastHistBar = i
    GoTo 40

60  check = 1
    GoTo 40

70  'Draw the histogram for the entire image
Dim Width As Integer
'Width is total "# pixels" / "# relevant histogram bars" - 1 pixel spacing between bars
Width = Int((280 / LastHistBar) - 1)

pictureImageHist.Cls
pictureImageHist.DrawWidth = Width
pictureImageHist.ForeColor = 255
pictureImageHist.DrawMode = 13
For i = 1 To LastHistBar
    If HistBar(i) = 0 Then GoTo 80
    BarLength = Int((HistBar(i) / MaxHistBar) * 90)
    x = 10 + ((Width + 1) * (i - 1))
    y = 100 - BarLength
    pictureImageHist.Line (x, y)-(x, 100)


80  Next i

'Assign values to text boxes
textMaxPixel.Text = Int(MaxPixel)
textMinPixel.Text = Int(MinPixel)
textCutoff.Text = Int(MaxPixel)
textMostFreq.Text = MostFreq
textThreshold.Text = MostFreq + 9

buttonAnalysis.Enabled = True
buttonMinAnal.Enabled = True
buttonSave.Enabled = True
buttonAnalysis.Default = True

Form1.MousePointer = 0 'Default mouse pointer
End Sub

Public Sub Save_Click()
CommonDialog.ShowSave
fileName = CommonDialog.fileName
ScanData.SaveHDF (fileName)  'Saving HDF file with name from Save File Dialog box
End Sub
```

Analysis.frm Code

```
Dim DataArray(1 To 512, 1 To 512), zpeaks(1 To 5000), zpeaks1(1 To 5000) As Single
Dim MaxPixel As Single
Dim xpeaks(1 To 5000), ypeaks(1 To 5000) As Integer
Dim peakno, p, p1, acceptedpeaks, size As Integer
Dim xpeaks1(1 To 5000), ypeaks1(1 To 5000) As Integer
Dim xzoomcenter, yzoomcenter, x1, y1 As Integer
Dim HistBar(1 To 300) As Integer
Dim lowx, highx, lowy, highy, maxbin, maxpeakheight As Integer
Dim peakheight, maxzcoord, phb, bin, maxpeakhistbar As Integer
Dim bkc As Long
Dim Last As Boolean
Dim MostFreq, cutoff, threshold, radius As Integer
Dim textPeakValley, textMinExclude, fileName, defaultfilename As String

Public Sub buttonAccept_Click()

'indicate that last peak was accepted and enable back button
Last = True
buttonBack.Enabled = True

textZHeight.ForeColor = 0

xpeaks1(p1) = xpeaks(p)
ypeaks1(p1) = ypeaks(p)
zpeaks1(p1) = zpeaks(p) - maxbin
p = p + 1
p1 = p1 + 1
acceptedpeaks = acceptedpeaks + 1

If p > peakno Then
   analysis.Cls
   pictureHistogram.Cls
   Label4.Caption = "Save points or Quit?"
   Label5.Caption = "Enter Filename."
   Exit Sub
Else

xzoomcenter = xpeaks(p):  yzoomcenter = ypeaks(p)

'clear form
analysis.Cls
```

```
'Zoom in on point
Dim heightmax, heightmin, step, xcount, ycount As Integer
Dim ptcol(1 To 25, 1 To 25) As Long
heightmax = DataArray(xpeaks(p), ypeaks(p))
heightmin = heightmax
For x = xzoomcenter - 12 To xzoomcenter + 12    'Find minimum height
x1 = x
If x < 1 Then x1 = 1
If x > size Then x1 = size
For y = yzoomcenter - 12 To yzoomcenter + 12
y1 = y
If y < 1 Then y1 = 1
If y > size Then y1 = size
If DataArray(x1, y1) < heightmin Then heightmin = DataArray(x1, y1)
Next y
Next x
step = Int(255 / (heightmax - heightmin))
xcount = 1:  ycount = 1
For x = xzoomcenter - 12 To xzoomcenter + 12
x1 = x
If x < 1 Then x1 = 1
If x > size Then x1 = size
For y = yzoomcenter - 12 To yzoomcenter + 12
y1 = y
If y < 1 Then y1 = 1
If y > size Then y1 = size
Color = Int((DataArray(x1, y1) - heightmin) * step)
If Color > 255 Then Color = 255
ptcol(xcount, ycount) = RGB(Color, Color, Color)
'Set all points >= the peak to green
If DataArray(x1, y1) >= DataArray(xpeaks(p), ypeaks(p)) Then ptcol(xcount, ycount) =
RGB(0, 255, 0)
ycount = ycount + 1
Next y
ycount = 1
xcount = xcount + 1
Next x

'Set the peak to red
ptcol(13, 13) = 255

Dim xconv, yconv As Single
xconv = 3#
For px = 1 To 25
yconv = 3#
For py = 1 To 25
```

```
'analysis.PSet (xconv, yconv), ptcol(px, py)
analysis.PSet (xconv, yconv), ptcol(px, py)
yconv = CSng(yconv + 5)
Next py
xconv = CSng(xconv + 5)
Next px

'Calculate histogram for local background around each peak found
lowx = xpeaks(p) - 1:  If lowx < 1 Then lowx = 1
highx = xpeaks(p) + 1:  If highx > size Then highx = size
lowy = ypeaks(p) - 40:  If lowy < 1 Then lowy = 1
highy = ypeaks(p) + 40:  If highy > size Then highy = size
maxpeakheight = 0
pictureHistogram.ForeColor = 255
pictureHistogram.DrawMode = 13
pictureHistogram.DrawWidth = 1
pictureHistogram.Cls

For x = lowx To highx
   maxpeakhistbar = 0:  maxbin = 0:  maxzcoord = 0:  peakheight = 0
   For bin = 1 To zpeaks(p):  HistBar(bin) = 0:  Next bin
   For y = lowy To highy
      If DataArray(x, y) > maxzcoord Then maxzcoord = DataArray(x, y)
      bin = DataArray(x, y)
      If bin > zpeaks(p) Then GoTo 4020
      If bin < 1 Then GoTo 4020
      HistBar(bin) = HistBar(bin) + 1
      phb = HistBar(bin)
      If phb > 15 Then phb = 15
      pictureHistogram.Line (bin * 5, 141)-((bin + 1) * 5, 141 - phb * 9), , BF
4020 Next y
   For bin = 1 To zpeaks(p)
   If HistBar(bin) > maxpeakhistbar Then maxpeakhistbar = HistBar(bin): maxbin = bin
   Next bin
   peakheight = maxzcoord - maxbin
   If peakheight > maxpeakheight Then maxpeakheight = peakheight
Next x
pictureHistogram.ForeColor = RGB(0, 0, 255)
pictureHistogram.Line (maxbin * 5, 141)-((maxbin + 1) * 5, 0), , BF

textCurrentPeak.Text = p
textAcceptedPeaks.Text = acceptedpeaks
If DataArray(xpeaks(p), ypeaks(p)) - heightmin <= 5 Then textZHeight.ForeColor =
RGB(255, 0, 0)
textZHeight.Text = Int(DataArray(xpeaks(p), ypeaks(p)) - maxbin)
```

```
Form1.MousePointer = 0 'Default mouse pointer
End If
End Sub

Private Sub buttonBack_Click()
textZHeight.ForeColor = 0

If Last = False Then GoTo 100
p1 = p1 - 1
acceptedpeaks = acceptedpeaks - 1
xpeaks1(p1) = 0
ypeaks1(p1) = 0
zpeaks1(p1) = 0

100 p = p - 1
buttonBack.Enabled = False

xzoomcenter = xpeaks(p):  yzoomcenter = ypeaks(p)

'clear form
analysis.Cls

'Zoom in on point
Dim heightmax, heightmin, step, xcount, ycount As Integer
Dim ptcol(1 To 25, 1 To 25) As Long
heightmax = DataArray(xpeaks(p), ypeaks(p))
heightmin = heightmax
For x = xzoomcenter - 12 To xzoomcenter + 12    'Find minimum height
x1 = x
If x < 1 Then x1 = 1
If x > size Then x1 = size
For y = yzoomcenter - 12 To yzoomcenter + 12
y1 = y
If y < 1 Then y1 = 1
If y > size Then y1 = size
If DataArray(x1, y1) < heightmin Then heightmin = DataArray(x1, y1)
Next y
Next x
step = Int(255 / (heightmax - heightmin))
xcount = 1:  ycount = 1
For x = xzoomcenter - 12 To xzoomcenter + 12
x1 = x
If x < 1 Then x1 = 1
If x > size Then x1 = size
For y = yzoomcenter - 12 To yzoomcenter + 12
y1 = y
```

```
If y < 1 Then y1 = 1
If y > size Then y1 = size
Color = Int((DataArray(x1, y1) - heightmin) * step)
If Color > 255 Then Color = 255
ptcol(xcount, ycount) = RGB(Color, Color, Color)
'Set all points >= the peak to green
If DataArray(x1, y1) >= DataArray(xpeaks(p), ypeaks(p)) Then ptcol(xcount, ycount) =
RGB(0, 255, 0)
ycount = ycount + 1
Next y
ycount = 1
xcount = xcount + 1
Next x

'Set the peak to red
ptcol(13, 13) = 255

Dim xconv, yconv As Single
xconv = 3#
For px = 1 To 25
yconv = 3#
For py = 1 To 25
analysis.PSet (xconv, yconv), ptcol(px, py)
yconv = CSng(yconv + 5)
Next py
xconv = CSng(xconv + 5)
Next px

'Calculate histogram for local background around each peak found
lowx = xpeaks(p) - 1:  If lowx < 1 Then lowx = 1
highx = xpeaks(p) + 1:  If highx > size Then highx = size
lowy = ypeaks(p) - 40:  If lowy < 1 Then lowy = 1
highy = ypeaks(p) + 40:  If highy > size Then highy = size
maxpeakheight = 0
pictureHistogram.ForeColor = 255
pictureHistogram.DrawMode = 13
pictureHistogram.DrawWidth = 1
pictureHistogram.Cls

For x = lowx To highx
   maxpeakhistbar = 0:  maxbin = 0:  maxzcoord = 0:  peakheight = 0
   For bin = 1 To zpeaks(p):  HistBar(bin) = 0:  Next bin
   For y = lowy To highy
      If DataArray(x, y) > maxzcoord Then maxzcoord = DataArray(x, y)
      bin = DataArray(x, y)
      If bin > zpeaks(p) Then GoTo 4020
```

```
      If bin < 1 Then GoTo 4020
      HistBar(bin) = HistBar(bin) + 1
      phb = HistBar(bin)
      If phb > 15 Then phb = 15
      pictureHistogram.Line (bin * 5, 141)-((bin + 1) * 5, 141 - phb * 9), , BF
4020 Next y
   For bin = 1 To zpeaks(p)
   If HistBar(bin) > maxpeakhistbar Then maxpeakhistbar = HistBar(bin): maxbin = bin
   Next bin
   peakheight = maxzcoord - maxbin
   If peakheight > maxpeakheight Then maxpeakheight = peakheight
Next x
pictureHistogram.ForeColor = RGB(0, 0, 255)
pictureHistogram.Line (maxbin * 5, 141)-((maxbin + 1) * 5, 0), , BF

textCurrentPeak.Text = p
textAcceptedPeaks.Text = acceptedpeaks
If DataArray(xpeaks(p), ypeaks(p)) - heightmin <= 5 Then textZHeight.ForeColor =
RGB(255, 0, 0)
textZHeight.Text = Int(DataArray(xpeaks(p), ypeaks(p)) - maxbin)

Form1.MousePointer = 0 'Default mouse pointer

End Sub

Public Sub buttonQuit_Click()
Unload Me
End Sub

Public Sub buttonReject_Click()
'indicate Last peak was reject for back routine and enable back button
Last = False
buttonBack.Enabled = True

textZHeight.ForeColor = 0

p = p + 1
If p > peakno Then
   analysis.Cls
   pictureHistogram.Cls
   Label4.Caption = "Save points or Quit?"
   Label5.Caption = "Enter Filename."
   Exit Sub
Else

xzoomcenter = xpeaks(p):  yzoomcenter = ypeaks(p)
```

```
'clear form
analysis.Cls

'Zoom in on point
Dim heightmax, heightmin, step, xcount, ycount As Integer
Dim ptcol(1 To 25, 1 To 25) As Long
heightmax = DataArray(xpeaks(p), ypeaks(p))
heightmin = heightmax
For x = xzoomcenter - 12 To xzoomcenter + 12    'Find minimum height
x1 = x
If x < 1 Then x1 = 1
If x > size Then x1 = size
For y = yzoomcenter - 12 To yzoomcenter + 12
y1 = y
If y < 1 Then y1 = 1
If y > size Then y1 = size
If DataArray(x1, y1) < heightmin Then heightmin = DataArray(x1, y1)
Next y
Next x
step = Int(255 / (heightmax - heightmin))
xcount = 1:  ycount = 1
For x = xzoomcenter - 12 To xzoomcenter + 12
x1 = x
If x < 1 Then x1 = 1
If x > size Then x1 = size
For y = yzoomcenter - 12 To yzoomcenter + 12
y1 = y
If y < 1 Then y1 = 1
If y > size Then y1 = size
Color = Int((DataArray(x1, y1) - heightmin) * step)
If Color > 255 Then Color = 255
ptcol(xcount, ycount) = RGB(Color, Color, Color)
'Set all points >= the peak to green
If DataArray(x1, y1) >= DataArray(xpeaks(p), ypeaks(p)) Then ptcol(xcount, ycount) =
RGB(0, 255, 0)
ycount = ycount + 1
Next y
ycount = 1
xcount = xcount + 1
Next x

'Set the peak to red
ptcol(13, 13) = 255

Dim xconv, yconv As Single
```

```
xconv = 3#
For px = 1 To 25
yconv = 3#
For py = 1 To 25
analysis.PSet (xconv, yconv), ptcol(px, py)
yconv = CSng(yconv + 5)
Next py
xconv = CSng(xconv + 5)
Next px

'Calculate histogram for local background around each peak found
lowx = xpeaks(p) - 1:  If lowx < 1 Then lowx = 1
highx = xpeaks(p) + 1:  If highx > size Then highx = size
lowy = ypeaks(p) - 40:  If lowy < 1 Then lowy = 1
highy = ypeaks(p) + 40:  If highy > size Then highy = size
maxpeakheight = 0
pictureHistogram.ForeColor = 255
pictureHistogram.DrawMode = 13
pictureHistogram.DrawWidth = 1
pictureHistogram.Cls

For x = lowx To highx
   maxpeakhistbar = 0:  maxbin = 0:  maxzcoord = 0:  peakheight = 0
   For bin = 1 To zpeaks(p):  HistBar(bin) = 0:  Next bin
   For y = lowy To highy
      If DataArray(x, y) > maxzcoord Then maxzcoord = DataArray(x, y)
      bin = DataArray(x, y)
      If bin > zpeaks(p) Then GoTo 4020
      If bin < 1 Then GoTo 4020
      HistBar(bin) = HistBar(bin) + 1
      phb = HistBar(bin)
      If phb > 15 Then phb = 15
      pictureHistogram.Line (bin * 5, 141)-((bin + 1) * 5, 141 - phb * 9), , BF
4020 Next y
   For bin = 1 To zpeaks(p)
   If HistBar(bin) > maxpeakhistbar Then maxpeakhistbar = HistBar(bin): maxbin = bin
   Next bin
   peakheight = maxzcoord - maxbin
   If peakheight > maxpeakheight Then maxpeakheight = peakheight
Next x
pictureHistogram.ForeColor = RGB(0, 0, 255)
pictureHistogram.Line (maxbin * 5, 141)-((maxbin + 1) * 5, 0), , BF

textCurrentPeak.Text = p
textAcceptedPeaks.Text = acceptedpeaks
```

```
If DataArray(xpeaks(p), ypeaks(p)) - heightmin <= 5 Then textZHeight.ForeColor =
RGB(255, 0, 0)
textZHeight.Text = Int(DataArray(xpeaks(p), ypeaks(p)) - maxbin)

Form1.MousePointer = 0 'Default mouse pointer
End If
End Sub

Public Sub buttonSave_Click()
Form1.MousePointer = 11 'or VB hourglass
Dim SaveFileName
Dim peakcount As Integer
'Write peak data to file
SaveFileName = textSaveFileName.Text
   Open SaveFileName For Output As #15
   Write #15, fileName
   Write #15, textPeakValley; acceptedpeaks
   Write #15, "maximum="; MaxPixel, "most frequent="; MostFreq, "Min. exclude=";
textMinExclude
   Write #15, "cutoff="; cutoff, "threshold="; threshold, "radius="; radius
   For peakcount = 1 To acceptedpeaks
      Write #15, xpeaks1(peakcount), ypeaks1(peakcount), zpeaks1(peakcount)
   Next peakcount
   Close #15
Form1.MousePointer = 0 'Default mouse pointer
Unload Me
End Sub

Private Sub Form_Activate()

buttonBack.Enabled = False

textZHeight.ForeColor = 0

acceptedpeaks = 0
p = 1
p1 = 1
textSaveFileName.Text = defaultfilename
xzoomcenter = xpeaks(p):  yzoomcenter = ypeaks(p)

'clear form
analysis.Cls

'Zoom in on point
Dim heightmax, heightmin, step, xcount, ycount As Integer
Dim ptcol(1 To 25, 1 To 25) As Long
```

```
heightmax = DataArray(xpeaks(p), ypeaks(p))
heightmin = heightmax
For x = xzoomcenter - 12 To xzoomcenter + 12    'Find minimum height
x1 = x
If x < 1 Then x1 = 1
If x > size Then x1 = size
For y = yzoomcenter - 12 To yzoomcenter + 12
y1 = y
If y < 1 Then y1 = 1
If y > size Then y1 = size
If DataArray(x1, y1) < heightmin Then heightmin = DataArray(x1, y1)
Next y
Next x
step = Int(255 / (heightmax - heightmin))
xcount = 1:  ycount = 1
For x = xzoomcenter - 12 To xzoomcenter + 12
x1 = x
If x < 1 Then x1 = 1
If x > size Then x1 = size
For y = yzoomcenter - 12 To yzoomcenter + 12
y1 = y
If y < 1 Then y1 = 1
If y > size Then y1 = size
Color = Int((DataArray(x1, y1) - heightmin) * step)
If Color > 255 Then Color = 255
ptcol(xcount, ycount) = RGB(Color, Color, Color)
'Set all points >= the peak to green
If DataArray(x1, y1) >= DataArray(xpeaks(p), ypeaks(p)) Then ptcol(xcount, ycount) =
RGB(0, 255, 0)
ycount = ycount + 1
Next y
ycount = 1
xcount = xcount + 1
Next x

'Set the peak to red
ptcol(13, 13) = 255

Dim xconv, yconv As Single
xconv = 3#
For px = 1 To 25
yconv = 3#
For py = 1 To 25
analysis.PSet (xconv, yconv), ptcol(px, py)
yconv = CSng(yconv + 5)
Next py
```

```
xconv = CSng(xconv + 5)
Next px

'Calculate histogram for local background around each peak found
lowx = xpeaks(p) - 1:  If lowx < 1 Then lowx = 1
highx = xpeaks(p) + 1:  If highx > size Then highx = size
lowy = ypeaks(p) - 40:  If lowy < 1 Then lowy = 1
highy = ypeaks(p) + 40:  If highy > size Then highy = size
maxpeakheight = 0
pictureHistogram.ForeColor = 255
pictureHistogram.DrawMode = 13
pictureHistogram.DrawWidth = 1
pictureHistogram.Cls

For x = lowx To highx
   maxpeakhistbar = 0:  maxbin = 0:  maxzcoord = 0:  peakheight = 0
   For bin = 1 To zpeaks(p):  HistBar(bin) = 0:  Next bin
   For y = lowy To highy
      If DataArray(x, y) > maxzcoord Then maxzcoord = DataArray(x, y)
      bin = DataArray(x, y)
      If bin > zpeaks(p) Then GoTo 4020
      If bin < 1 Then GoTo 4020
      HistBar(bin) = HistBar(bin) + 1
      phb = HistBar(bin)
      If phb > 15 Then phb = 15
      pictureHistogram.Line (bin * 5, 141)-((bin + 1) * 5, 141 - phb * 9), , BF
4020 Next y
   For bin = 1 To zpeaks(p)
   If HistBar(bin) > maxpeakhistbar Then maxpeakhistbar = HistBar(bin): maxbin = bin
   Next bin
   peakheight = maxzcoord - maxbin
   If peakheight > maxpeakheight Then maxpeakheight = peakheight
Next x
pictureHistogram.ForeColor = RGB(0, 0, 255)
pictureHistogram.Line (maxbin * 5, 141)-((maxbin + 1) * 5, 0), , BF

textCurrentPeak.Text = p
textAcceptedPeaks.Text = acceptedpeaks
If DataArray(xpeaks(p), ypeaks(p)) - heightmin <= 5 Then textZHeight.ForeColor =
RGB(255, 0, 0)
textZHeight.Text = Int(DataArray(xpeaks(p), ypeaks(p)) - maxbin)

Form1.MousePointer = 0 'Default mouse pointer
End Sub

Public Sub Form_Load()
```

```
Open "variable" For Input As #3
Input #3, peakno, MaxPixel, MostFreq, cutoff, threshold, radius, size
Input #3, fileName, textPeakValley, textMinExclude, defaultfilename
For x = 1 To size
   For y = 1 To size
   Input #3, DataArray(x, y)
   Next y
Next x
For peaks = 1 To peakno
Input #3, xpeaks(peaks), ypeaks(peaks), zpeaks(peaks)
Next peaks
Close #3
Kill "variable"

Form1.MousePointer = 0 'Default mouse pointer
End Sub
```